2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

On the stability and optimality of distributed Kalman filters with finite-time data fusion*

Usman A. Khan‡, Ali Jadbabaie†

***Abstract*— In this paper, we consider distributed estimation for discrete-time, linear systems, with *finite-time data fusion of agent measurements* between each time-step of the dynamics. Prior work in this context is related to average-consensus, where either the data fusion is implemented for an infinite time (in general) to reach average-consensus, or under restricted observability requirements (one-step and/or local), whereas, our results hold under the broadest observability conditions (n-step global observability, where n is the dimension of the dynamics).**

We show that after the finite-time data fusion on agent measurements, the observation map at each agent is a linear combination of the local observation maps. We then show that this new observation map is observable (if the data is fused for a sufficient number of iterations that we lower bound) resulting in a stable distributed estimator that can be implemented using semi-definite programming at each agent. We further characterize the performance of such distributed estimators by comparing the positive-definiteness of their corresponding information matrices. The centralized and distributed performance gap, although cannot be written in closed form, can be computed using the infinite horizon Kalman gain of each filter. Finally, we consider special cases under which the performance of these distributed estimators is equal to the performance of the centralized Kalman filter.

I. Introduction

The advent of sensor and social networks has directed much of the recent research interest to the distributed estimation and control problems for dynamical systems whose observation are distributed over a network of agents. Much work has focussed on distributed estimation with measurement fusion algorithms, where either the measurement data is fused for an infinite time (in general) between each step of the dynamics to reach average-consensus [1]–[4], or when restricted models for observability are used [5]–[7]. However, there are practical limitations in reaching a consensus as infinite iterations of any data fusion algorithm are infeasible to implement, and in cases, when infinite iterations are not required, the observability conditions have been simplified to address the finite-time data fusion. Relevant work in this context is available in [1]–[3], [5] and references therein; see also the foundation work in [8]–[10].

The restrictions posed on the data fusion algorithms limit the generality of the observability conditions. For instance, References [5]–[7], [11] consider distributed estimation with single iteration of data fusion, but they require one-step global observability (collection of all the observation matrices is full rank). On the other hand, References [1]–[4] consider n-step global observability (collection of all the observation matrices along with the system matrix is full rank), but the data fusion algorithms must reach consensus, which, in general, requires infinite iterations. Note that References [2], [6], [7] also consider single-step data fusion but with a restricted notion of observability, i.e., local observability at each agent. Specifically, References [2], [6], [11] consider scalar-state dynamical systems with a scalar local observation at each agent; thus assuming that each agent is observable. On the contrary, our work considers vector state-space and assumes n-step global observability, with only a finite iterations of the data fusion algorithm. Furthermore, we do not require any agent to be locally observable.

The distributed estimator we propose is based on *finite-time data fusion of agent measurements* and does not require any relaxation of the centralized observability condition. We show that the finite-time data fusion results into a new observation map at each agent that is a linear combination of the local observation maps. We show that these local observation maps guarantee observability at each agent when the data fusion is implemented for a sufficient number of iterations (greater than or equal to the primitivity index of the underlying agent communication graph). Other techniques with fusion of the state estimates at each agent have also been considered, see, e.g., [5], [12] and the references therein. Fusion on state estimates requires designing a block-diagonal Kalman gain matrix that is equivalent to solving related (Linear Matrix Inequalities) LMIs under structural constraints, whose solution is not guaranteed even if the system is centrally observable. The main contribution of this paper is to provide a lower bound on the number of data fusion iterations that guarantees local observability at each agent. This local observability results in a block-diagonal Kalman gain that can be computed locally at each agent.

In addition to showing the stability of the local error processes, we provide a method for comparing the steady state solution of the Riccati equations corresponding to the distributed and the centralized estimator. We show that the performance of the distributed estimator with finite-time data fusion is always upper bounded (in the information sense, or lower bounded in the error sense) by the performance of the centralized Kalman filter. Although this result is intuitive but has not been explicitly shown in the context of distributed estimation with finite data fusion iterations. Furthermore, we show that the performance of the distributed

*This research was partially supported by the following grants: ONR MURI N000140810747, NSF Career, and AFOSR's Complex Networks Program.

‡Department of Electrical and Computer Engineering, Tufts University, khan@ece.tufts.edu. †Department of Electrical and Systems Engineering, University of Pennsylvania, jadbabai@seas.upenn.edu.

estimator is optimal under the available information, i.e., a better estimator cannot be obtained with the given set of information. We further consider special cases where the distributed estimators guarantee the performance of the centralized Kalman filter. In this context, the approach we use is novel and provides interesting insights into scenarios where finite data fusion iterations are, in fact, optimal. The performance analysis carried out in this paper can also be extended to study the role of information fusion in Riccati equations that we consider elsewhere.

We now describe the rest of the paper. Section II provides the preliminaries and sets the notation, whereas, Section III provides the distributed estimator with finite-time data fusion and considers the stability analysis. We then provide the optimality (performance) analysis in Section IV. We provide simulation results in Section V, and, finally, Section VI concludes the paper.

II. Preliminaries

Consider the following discrete-time dynamical system,

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

where $x_k \in \mathbb{R}^n$ is the state vector, and v_k is the system noise distributed as Gaussian, $\mathcal{N}(0, Q)$. Consider the following observation model

$$\begin{aligned} y_k &= Cx_k + r_k, && (2) \\ \triangleq \begin{bmatrix} y_k^1 \\ \vdots \\ y_k^N \end{bmatrix} &= \begin{bmatrix} C_1 \\ \vdots \\ C_n \end{bmatrix} x_k + \begin{bmatrix} r_k^1 \\ \vdots \\ r_k^N \end{bmatrix} && (3) \end{aligned}$$

where $y_k^i \in \mathbb{R}^m$ and $C \triangleq [C_1^T, \ldots, \ldots C_N^T]^T$ is the observation matrix[1]. We assume that the noise at each agent r_k^i is Gaussian,[2] $\mathcal{N}(0, I_m)$. We assume that the above system is observable, i.e., the pair (A, C) is observable (this is the n-step global observability that we mentioned before).

Network connectivity: The interactions among the agents are modeled with an undirected graph, $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, \ldots, N\}$ is the set of vertices and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of ordered pairs describing the interconnections among the agents. The neighborhood at the ith agent is defined as $\mathcal{N}_i \triangleq \{i\} \cup \{j \mid (i, j) \in \mathcal{E}\}$. For details on graph-theoretic concepts, see [13].

A. Centralized Kalman filter

The optimal centralized state estimate, $\widehat{x}^c_{k+1|k+1}$, at time $k+1$, given all the agent observations up to time, $k+1$, is the centralized Kalman filter [14], [15]:

$$\widehat{x}^c_{k+1|k+1} = A\widehat{x}^c_{k|k} + K^c_{k+1}(y_{k+1} - CA\widehat{x}^c_{k|k}), \tag{4}$$

[1]Here, we assume that each agent has the same number, m, of observations, this assumption is made for simplification of the discussion and does not lose generality.

[2]Note that the assumption of having an I_m noise covariance can be made without the loss of generality. This is because if the observation noise covariance were $R_i > 0$, then a transformation of the local observations with $R_i^{-1/2}$ does not lose any information in y_k^i, and results into a covariance of I_m. In such cases, we only require $\mathbb{E}[r_k^i r_k^{jT}] = 0, \forall i \neq j$ for our results to hold.

where K^c_{k+1} is the centralized Kalman gain that is computed as a function of the centralized estimator and predictor error covariances. Clearly, if the system is (A, C) observable, the error dynamics in the Kalman filter are stable.

In the following, we give a distributed observer that is based on data fusion between each innovation update. This data fusion only requires finite iterations, unlike consensus-based estimators [1]. We split the analysis of the proposed distributed estimator into two categories: first we address stability, and then we address optimality.

III. Distributed Kalman filter: Stability

Consider the following data fusion algorithm at agent i:

$$d^i_{k+1}(t) = \sum_{j \in \mathcal{N}_i} w_{ij} d^j_{k+1}(t-1), \tag{5}$$

with $d^i_{k+1}(0) = C_i^T y^i_{k+1}$. It follows that

$$\begin{aligned} d^i_{k+1}(t) &= \sum_{j=1}^{N} [W^t]_{ij} C_j^T y^j_{k+1}, \\ &= \sum_{j=1}^{N} [W^t]_{ij} C_j^T C_j x_{k+1} + \sum_{j=1}^{N} [W^t]_{ij} C_j^T r^j_{k+1}, \end{aligned}$$

where $[W^t]_{ij}$ is the ij-th element of W^t, and $W = \{w_{ij}\}$. With t steps of data fusion (5), we get $d^i_{k+1}(t)$, which is a linear combination of the agents' observations. Let $W_i(t)$ denote the $mN \times mN$ diagonal matrix with $[W^t]_{ij} I_m$ as its main diagonal elements (with i fixed and $j = 1, \ldots, N$). After t steps of (5), the observation matrix for agent i's observation model, $d^i_{k+1}(t)$, can be compactly written as

$$\overline{C}_i(t) \triangleq C^T W_i(t) C = \sum_{j=1}^{N} [W^t]_{ij} C_j^T C_j, \tag{6}$$

where the observation noise covariance in $d^i_{k+1}(t)$ is

$$\overline{R}_i(t) \triangleq C^T W_i^2(t) C = \sum_{j=1}^{N} [W^t]^2_{ij} C_j^T C_j, \tag{7}$$

Using a similar *finite-time* data fusion algorithm to (5), we may compute $\overline{C}_i(t)$ (and $\overline{R}_i(t)$) at each agent i, which is a scaled linear combination of the local (transformed) observation matrices (and local weighted covariances). At each agent i, this scaling depends on the ith row of the weight matrix evolved after t steps, i.e., on $[W^t]_{ij}$. Note that this data fusion has to be carried out only once at the start of the algorithm as $\overline{C}_i(t)$ is independent of k.

We now consider the following distributed estimator.

$$\widehat{x}^i_{k+1|k+1} = A\widehat{x}^i_{k|k} + K^i_{k+1}(d^i_{k+1}(t) - \overline{C}_i(t) A\widehat{x}^i_{k|k}). \tag{8}$$

The error in the above estimator is given by

$$\begin{aligned} e^i_{k+1|k+1} &\triangleq x_{k+1} - \widehat{x}^i_{k+1|k+1}, \\ &= Ax_k + v_k - A\widehat{x}^i_{k|k} \\ &- K^i_{k+1}(\overline{C}_i(t)x_{k+1} - \overline{C}_i(t)A\widehat{x}^i_{k|k}) \\ &- K^i_{k+1}\sum_{j=1}^{N}[W^t]_{ij}C_j^T r^j_{k+1}, \\ &= \left(A - K^i_{k+1}\overline{C}_i(t)A\right)e^i_{k|k} + \eta^i_{k+1}, \end{aligned} \tag{9}$$

where

$$\eta^i_{k+1} = (I_n - K^i_{k+1}\overline{C}_i(t))v_k - K^i_{k+1}\sum_{j=1}^{N}[W^t]_{ij}C_j^T r^j_{k+1}.$$

A. Infinite-time data fusion

In this section, we show that, if we replace $\overline{C}_i(t)$ by C^TC, then the distributed estimator in (8) has the same stability notion and performance as the centralized estimator in (4). In other words, this is equivalent to saying that the two observation models, y_k and C^Ty_k, have the same stability notion and performance. We will prove this in Theorem 2 in Section IV. In the following lemma, we show that the stability of the observation model y_k implies the stability of the observation model C^Ty_k, and vice versa.

Lemma 1: The pair (A, C) is observable if and only if the pair (A, C^TC) is observable.

The above lemma is straightforward to prove. For instance, notice that the row-space of C^TC is the same as the row-space of C; thus, the row-space of the observability matrix for the pair, (A, C^TC), is the same as the row-space of the observability matrix for the pair, (A, C), and observability of the two pairs are equivalent. That the performance is also the same will be shown in Theorem 2.

One method to obtain C^TC from $\overline{C}_i(t)$ is through average-consensus as is shown in [1]. Choosing the weight matrix, $W = \{w_{ij}\}$, in the data fusion algorithm (5), to be stochastic (along with some assumptions on the underlying agent communication graph, $\mathcal{G}$), such that $\lim_{t\to\infty} W^t = 11^T/N$, guarantees

$$\lim_{t\to\infty}\overline{C}_i(t) = \frac{1}{N}C^TC. \tag{10}$$

Hence, if we allow infinite consensus iterations between each successive steps of the dynamics (between each k and $k+1$), then the local observation matrices become C^TC/N and choosing $K^i_{k+1} = NK^c_{k+1}$ at each agent suffices for the distributed estimator in (8) to have the same performance as the centralized Kalman filter (4).

Clearly, there are several limitations of reaching consensus. Mainly, the process is asymptotic, thus, requires infinite iterations between any two time-steps of the dynamics. This requirement, although has pedagogical contributions, is infeasible to implement in any practical situation. We now explore the distributed estimator of (8), when the data fusion between any two steps of the dynamics is carried out for a finite time.

B. Finite-time data fusion

We now provide the main result of this section, i.e., the distributed estimator in (8) results into a stable estimator for any $t \geq \tau(\mathcal{G})$ along with some restrictions on the weight matrix, W, where $\tau(\mathcal{G})$ is the index of primitivity of the underlying agent communication graph, $\mathcal{G}$. To proceed further, we first characterize the observability of the observation matrix, $\sum_{j=1}^N z_{ij}C_j^TC_j$, when $z_{ij} > 0$ for all i, j, in the following lemma.

Lemma 2: A dynamical system is (A, C^TC)-observable if and only if it is $(A, \sum_{j=1}^N z_{ij}C_j^TC_j)$-observable[3] for strictly positive z_{ij}'s.

Proof: Let $D_c = \sum_{j=1}^N C_j^TC_j$. Recall that the pair (A, D_c) is observable if and only if the observability Gramian,

$$\mathcal{O} = \left[(D_c)^T,\ (D_cA)^T, \ldots, (D_cA^{n-1})^T\right]^T \tag{11}$$

is full rank. That the observability Gramian is full rank is true if and only if the following matrix

$$\begin{aligned} \mathcal{O}^T\mathcal{O} &= \sum_{k=0}^{n-1}(A^k)^TD_c^TD_cA^k, & (12) \\ &= \sum_{k=0}^{n-1}(A^k)^T\sum_{j=1}^{N}C_j^TC_j\sum_{j=1}^{N}C_j^TC_jA^k, & (13) \end{aligned}$$

is strictly positive-definite. Now note that $\mathcal{O}^T\mathcal{O}$ is strictly positive definite if and only if

$$\sum_{k=0}^{n-1}(A^k)^T\sum_{j=1}^{N}z_{ij}C_j^TC_j\sum_{j=1}^{N}z_{ij}C_j^TC_jA^k, \tag{14}$$

is strictly positive definite for $z_{ij} > 0$ for all i, j, which concludes the proof. ■

The proof of this lemma is straightforward and we provide it here for the sake of completeness. With the help of the above lemma, we provide the following theorem.

Theorem 1: Let the agent communication graph, $\mathcal{G}$, be irreducible, and let τ be its primitivity index. For each i, let $w_{ij} > 0,\ \forall\ j \in \mathcal{N}_i$, then $\exists K^i$ at each i, such that the distributed estimator in (8) is stable for any $t \geq \tau$.

Proof: Since $\mathcal{G}$ is irreducible, and for each i, $w_{ij} > 0,\ \forall\ j \in \mathcal{N}_i$, the matrix W is irreducible and primitive. Hence, for any $t \geq \tau$, W^t is a strictly positive matrix, i.e., each element of W^t is strictly positive, and, at any agent i, $\overline{C}_i(t) = \sum_{j=1}^N w_{ij}C_j^TC_j$ is a linear combination of $C_j^TC_j$'s with strictly positive weights. Hence, $(A, \overline{C}_i(t))$ is observable at each i from Lemma 2, and there exists a gain matrix K^i at each i that stabilizes the error in (8). ■

We may implement the following procedure at each agent i, to compute a stabilizing gain, K^i, that guarantees a stable error process, i.e., $\rho(A - K^i\overline{C}_iA) < 1$. From Lyapunov theory [16], we know that $\rho(A - K^i\overline{C}_iA) < 1$, if and only

[3] Note that $C^TC = \sum_{j=1}^N C_j^TC_j$.

if there exists a symmetric positive-definite matrix, $P > 0$, such that

$$\begin{aligned}(A - K^i\overline{C}_iA)^T P(A - K^i\overline{C}_iA) - P &< 0,\\ \Rightarrow (PA - Y^i\overline{C}_iA)^T P^{-1}(PA - Y^i\overline{C}_iA) - P &< 0,\end{aligned}$$

where $Y^i = PK^i$. Writing the above equation in its Schur complement form, we can solve the following Linear Matrix Inequality (LMI) [17],

$$\begin{bmatrix} P & (PA - Y^i\overline{C}_iA)^T \\ (PA - Y^i\overline{C}_iA) & P \end{bmatrix} > 0, \qquad (15)$$

for P and Y^i using semi-definite programming (SDP) [18], and compute $K^i = P^{-1}Y^i$.

IV. Distributed Kalman filter: Optimality

In this section, we will provide a quantitative comparison between the performance of the centralized Kalman filter (4) and the distributed Kalman filter with finite-time data fusion (8). We will also consider cases when the *performance of the distributed estimator is the same as the centralized performance*. To this end, we note that the observation set of the centralized Kalman filter is given by the pair (C, R), where $R > 0$ is the observation noise covariance, which is block-diagonal as we assume that the observation noise is independent among the agents. As we mentioned before, we can assume $R = I_{mN}$, without loss of generality. Thus, at any time k, the information content[4], $\mathcal{I}_c$, associated to the centralized observation set, (C, I_{mN}), is given by

$$\mathcal{I}_c = C^T R^{-1} C = C^T C. \qquad (16)$$

Similarly, as shown in (6) and (7), the observation set of the distributed estimator is given by $(\overline{C}_i(t), \overline{R}_i(t))$, and the information content, at any time k, in the distributed filter is

$$\mathcal{I}_d = (C^T W_i(t)C)(C^T W_i^2(t)C)^\dagger (C^T W_i(t)C), \qquad (17)$$

where '$\dagger$' denotes the Moore-Penrose pseudo-inverse [20]. We replace the inverse with a pseudo-inverse as $C^T W_i^2(t)C$ is not necessarily invertible. Let $Z^c_{k|k}$ and $Z^c_{k|k-1}$ be the information matrices of the centralized Kalman filter, which are the inverses of the error covariances, $S^c_{k|k}$ and $S^c_{k|k-1}$, respectively[5]. The following lemma shows that the comparison of the information matrices coming from two different observation sets is governed by the their information content.

Lemma 3: Let (C_1, I) and (C_2, I) be two observation sets corresponding to the dynamical system in (1), such that (A, C_1) and (A, C_2) are both observable. If $Z^1_{0|0} = Z^2_{0|0}$ and $\mathcal{I}_1 \geq \mathcal{I}_2$, then $Z^1_{k|k} \geq Z^2_{k|k}$, $\forall\ k$.

Proof: We will prove this lemma by induction. The base case ($k = 0$) holds trivially. To show for time k, we assume the following is true for time $k - 1$,

$$Z^1_{k-1|k-1} \geq Z^2_{k-1|k-1}. \qquad (18)$$

[4]Note that for static linear parameter estimation, the information content is the inverse of the Cramer-Rao lower bound [19].

[5]The superscript 'c' denotes the centralized variables, whereas, the superscripts 'd' will denote the variables in the distributed filter.

Then,

$$(Z^1_{k-1|k-1})^{-1} \leq (Z^2_{k-1|k-1})^{-1},$$

or,

$$A(Z^1_{k-1|k-1})^{-1}A^T + Q \leq A(Z^2_{k-1|k-1})^{-1}A^T + Q,$$

or,

$$\begin{aligned}&(A(Z^1_{k-1|k-1})^{-1}A^T + Q)^{-1}\\ \geq\ &(A(Z^2_{k-1|k-1})^{-1}A^T + Q)^{-}1,\end{aligned}$$

or,

$$\begin{aligned}&(A(Z^1_{k-1|k-1})^{-1}A^T + Q)^{-1} + C_1^T C_1\\ \geq\ &(A(Z^2_{k-1|k-1})^{-1}A^T + Q)^{-}1 + C_2^T C_2,\end{aligned}$$

since $\mathcal{I}_1 \geq \mathcal{I}_2$, and the proof follows by realizing that

$$Z^1_{k|k} = (A(Z^1_{k-1|k-1})^{-1}A^T + Q)^{-1} + C_1^T C_1.$$

■

The above lemma shows that the performance of two different Kalman filter (both estimating the same system (1)) with two different but *observable* observation sets, can be compared by comparing their associated information content, $\mathcal{I}$. Note that since both the observation sets are observable, $\lim_{k\to\infty} Z^1_{k|k}$ and $\lim_{k\to\infty} Z^2_{k|k}$, exist and are strictly positive definite. Since both of these sequences converge (to a unique solution if we further assume $(A, Q^{1/2})$ to be stabilizable [16]), and one always stays below the other, the asymptotic performance of $\lim_{k\to\infty} Z^2_{k|k}$ will always be bounded above by $\lim_{k\to\infty} Z^1_{k|k}$.

A. Comparison between centralized and distributed filters: Special cases

We now show some relevant comparisons between the information content of the centralized and distributed estimators.

Lemma 4: The following are true:

(i) We have
$$\mathcal{I}(C^TC, C^TC) = \mathcal{I}(C, I); \qquad (19)$$

(ii) If $W_i(t) = wI_{mN}$ for $w \neq 0$, then
$$\mathcal{I}(C, I) = \mathcal{I}(C^T W_i(t)C, C^T W_i^2(t)C); \qquad (20)$$

(iii) If C is invertible and $t \geq \tau$, then
$$\mathcal{I}(C, I) = \mathcal{I}(C^T W_i(t)C, C^T W_i^2(t)C); \qquad (21)$$

Proof:

(i) Note that

$$\begin{aligned}\mathcal{I}(C^TC, C^TC) &= C^TC(C^TC)^\dagger C^TC,\\ &= C^TC,\\ &= \mathcal{I}(C, I),\end{aligned}$$

by the properties of the Moore-Penrose pseudo-inverse [20].

(ii) We can write

$$\begin{aligned} & \mathcal{I}(C^TW_i(t)C, C^TW_i(t)^2C) \\ = \;& wC^TC(w^2C^TC)^\dagger wC^TC, \\ = \;& C^TC, \\ = \;& \mathcal{I}(C,I), \end{aligned}$$

by the properties of the Moore-Penrose pseudo-inverse [20].

(iii) Finally, note that if $t \geq \tau$, then $W_i(t)$ is invertible. This is because $W_i(t)$ is a diagonal matrix whose diagonal elements are given by $[W^t]_{ij} I_m$ and $t \geq \tau$ ensures that $[W^t]_{ij} > 0$ for each j. We thus have,

$$\begin{aligned} & \mathcal{I}(C^TW_i(t)C, C^TW_i(t)^2C) \\ = \;& C^TW_i(t)C(C^TW_i(t)^2C)^\dagger C^TW_i(t)C, \\ = \;& C^TW_i(t)CC^{-1}W_i(t)^{-2}C^{-T}C^TW_i(t)C, \\ = \;& C^TC, \\ = \;& \mathcal{I}(C,I), \end{aligned}$$

by the properties of the Moore-Penrose pseudo-inverse [20].

■

Remarks:

(i) Note that $\mathcal{I}(C^TC, C^TC)$ is the information content of the observation model, C^Ty_k, as it has the observation matrix, C^TC, and the observation noise covariance, C^TC. Hence, part (i) of Lemma 4 shows that the two observation models, y_k and C^Ty_k, possess the same information.

(ii) Part (ii) of Lemma 4 shows that if the ith row of W^t is such that it can be written as $w \neq 0$ times 1^T, then the performance of the distributed estimator is the same as the performance of the centralized estimator. One method to obtain the ith row as $w1^T$ is via the average-consensus [1], which results into $w = 1/N$ as $t \rightarrow \infty$.

(iii) Part (iii) of Lemma 4 shows that if C is invertible ($n \times n$ square), then a distributed estimator with $t \geq \tau$ will result into the centralized performance. This is a very important result as it shows that a Kalman filter with finite-time data fusion can provide the exact same performance as the Kalman filter with infinite consensus iterations (when C is invertible).

Finally, we now revisit the distributed filter with infinite-time data fusion (average-consensus) in the following theorem.

Theorem 2: The distributed estimator in (8) with infinite average-consensus iterations between each step of the dynamics (between each k and $k+1$) is observable when the centralized estimator (4) is observable. Furthermore, the two filters give the exact same performance.

Proof: The observability part is already established in Lemma 1. The information content is compared in Lemma 4-(i), and the theorem follows from Lemma 3. ■

B. Performance gap between the centralized and the distributed estimator

We now quantify the performance gap between the centralized Kalman filter (4) and the distributed estimator (8) when $t \geq \tau$ in the following lemma.

Lemma 5: Let $Z^c_{k|k}$ and $Z^d_{k|k}$ be the information matrices of the centralized Kalman filter (4) and the distributed estimator (8) with $t \geq \tau$, respectively. Let the underlying agent communication graph, $\mathcal{G}$ be irreducible, and, for each i, let $w_{ij} > 0, \; \forall \; j \in \mathcal{N}_i$, then $\forall \; k$,

$$Z^c_{k|k} \geq Z^d_{k|k}. \tag{22}$$

Proof: To prove the lemma, it suffices to show that

$$\mathcal{I}_c \geq \mathcal{I}_d,$$

form Lemma 3, or

$$C^TC - C^TW_i(t)C(C^TW_i(t)^2C)^\dagger C^TW_i(t)C \;\; \geq \;\; 0,$$

or,

$$C^T(I - W_i(t)C(C^TW_i(t)^2C)^\dagger C^TW_i(t))C \;\; \geq \;\; 0,$$

or,

$$I - W_i(t)C(C^TW_i(t)^2C)^\dagger C^TW_i(t) \;\; \geq \;\; 0.$$

To this end, let $X = W_i(t)C$, and assume $X = U\Sigma V^T$ be its singular value decomposition. We can write the above equation as

$$\begin{aligned} I - X(X^TX)^\dagger X &= I - XX^\dagger, \\ &= I - U\Sigma\Sigma^\dagger U^T, \\ &= I - U_1U_1^T, \end{aligned} \tag{23}$$

where U_1 is a matrix that contains the first r columns of U and $r = \text{rank}(\Sigma) = \text{rank}(X) = \text{rank}(C)$, since $W_i(t)$ is invertible. Since $I - U_1U_1^T$ is a projection on the null space of $U_1U_1^T$, it is positive semi-definite with r unit eigenvalues, and the lemma follows. ■

In summary, the main contribution of the Lemma 4 and Lemma 5 is to explicitly characterize when the distributed estimator with $t \geq \tau$ gives the centralized performance. In addition, we have shown that distributed estimator can achieve the centralized performance even when consensus is not reached, i.e., data fusion is not implemented for infinite time. In other cases, the performance of the distributed estimator will be upper bounded by the performance in the centralized case Lemma 5. The distributed estimator (8) can thus be implemented as a Kalman filter with t steps of data fusion. If $t \geq \tau$, then the distributed estimator is stable and also optimal within the given observation set $(C^TW_i(t)C, C^TW_i^2(t)C)$, i.e., a better estimator cannot be obtained with at least t steps of data fusion on agent measurements with fixed data fusion rule. However, optimizing the data fusion rule (w_{ij}'s at each agent) may result in better performance. The exact performance gap, although cannot be computed in closed form, can be quantified using the infinite horizon Kalman gains, K^c_∞ and K^i_∞, and their associated asymptotic information matrices, $\lim_{k\rightarrow\infty} Z^c_{k|k}$ and $\lim_{k\rightarrow\infty} Z^i_{k|k}$.

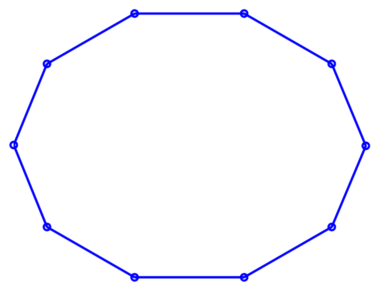

Fig. 1. A circulant graph with $N = 10$ agents.

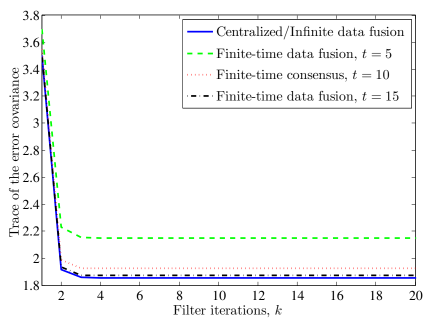

Fig. 2. Performance comparison of the centralized Kalman filter with the distributed Kalman filter with finite-time data fusion.

V. Simulations

We now present simulations to illustrate some of the results in this paper. We randomly choose a 5×5 system matrix, A, whose largest absolute eigenvalue comes out to be 1.43. We consider a network of $N = 10$ agents connected in a circle as shown in Fig. 1, such that each agent has one observation of the state vector, x_k, i.e., the local observation matrix $C_i \in \mathbb{R}^{1\times 5}$, and hence the global observation matrix $C \in \mathbb{R}^{10\times 5}$. We randomly choose the elements in $C_i, \ \forall\ i$, such that $C_i(1,5) = 0, \ \forall\ i$. It is straightforward to note that C^TC will have rank at most 4, and thus the system is not observable in one time-step. Furthermore, the pair A and C is chosen such that (A, C) is n-step observable ($n = 5$) but (A, C_i) is not observable for any i.

At each agent i, we choose constant weights, $w_{ij} = 1/3, \forall\ j \in \mathcal{N}_i$. Note that the weight matrix W is a circulant matrix with exactly 3 elements being $1/3$ and the rest zero. It is straightforward to note that $\tau = 5$ for this configuration. We run distributed filters with $t = 5,\ 10,\ 15$, and the trace of the corresponding error covariances is plotted in Fig. 2 and compared with the trace of the centralized error covariance. Note that as we have shown in the paper, for $t \geq 5$, the distributed filters have stable error, but the performance of the centralized filter is better when compared to any finite t. Another observation is that as t increases, the performance of the distributed Kalman filter with finite time data fusion improves.

VI. Conclusions

In this paper, we present a distributed Kalman filter with finite-time data fusion on agent measurements. We show that if the data fusion is carried for a sufficient number of iterations (at least equal to the primitivity index of the agent communication graph) then the distributed filter has a stable error. We then study the optimality of the distributed filters and show that the performance of the distributed filters is upper bounded (in the information sense) by the centralized performance. We then provide special cases under which the distributed filters with finite data fusion have exactly the same performance as the centralized Kalman filter (or with infinite consensus iterations).

References

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

[2] R. Carli, A. Chiuso, L. Schenato, and S. Zampieri, "Distributed Kalman filtering using consensus strategies," in *Proceedings of the 46th IEEE Conference on Decision and Control*, 2007, pp. 5486–5491.

[3] Usman A. Khan and José M. F. Moura, "Distributing the Kalman filters for large-scale systems," *IEEE Transactions on Signal Processing*, vol. 56(1), no. 10, pp. 4919–4935, Oct. 2008.

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

[5] Usman A. Khan, Soummya Kar, Ali Jadbabaie, and José M. F. Moura, "On connectivity, observability, and stability in distributed estimation," in *49th IEEE Conference on Decision and Control*, Atlanta, GA, Dec. 2010.

[6] D. Acemoglu, A. Nedic, and A. Ozdaglar, "Convergence of rule-of-thumb learning rules in social networks," in *47th IEEE Conference on Decision and Control*, Dec. 2008, pp. 1714–1720.

[7] R. Olfati-Saber, "Kalman-consensus filter : Optimality, stability, and performance," in *48th IEEE Conference on Decision and Control*, Shanghai, China, Dec. 2009, pp. 7036–7042.

[8] H. Hashemipour, S. Roy, and A. Laub, "Decentralized structures for parallel Kalman filtering," *IEEE Trans. on Automatic Control*, vol. 33, no. 1, pp. 88–94, Jan. 1988.

[9] B. Rao and H. Durrant-Whyte, "Fully decentralized algorithm for multisensor Kalman filtering," *IEE Proceedings-Control Theory and Applications*, vol. 138, pp. 413–420, Sep. 1991.

[10] A. Mutambara, *Decentralized estimation and control for multisensor systems*, CRC Press, Boca Raton, FL, 1998.

[11] A. Speranzon, C. Fischione, and K. H. Johansson, "Distributed and collaborative estimation over wireless sensor networks," in *IEEE Conference on Decision and Control*, San Diego, CA, 2006, pp. 1025–1030.

[12] Peter Alriksson and Anders Rantzer, "Distributed Kalman filtering using weighted averaging," in *Proceedings of the 17th International Symposium on Mathematical Theory of Networks and Systems*, Kyoto, Japan, July 2006.

[13] Béla Bollobás, *Modern graph theory*, Springer, New York, NY, 1998.

[14] R. Kalman, "A new approach to linear filtering and prediction problems," *Trans. of the ASME - Journal of Basic Engineering*, vol. 82, no. 2, pp. 35–45, 1960.

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

[16] G. E. Dullereud and F. Paganini, *A course in robust control theory: A convex approach*, Springer, 1999.

[17] S. Boyd, L. El-Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*, SIAM, 1994.

[18] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, New York, NY, USA, 2004.

[19] B. Anderson and J. Moore, *Optimal filtering*, Prentice Hall, Englewood Cliffs, NJ, 1979.

[20] G. Golub and C. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, MD, 1996.