An Unbiased Kalman Consensus Algorithm

Mehdi Alighanbari and Jonathan P. How

Aerospace Controls Laboratory
Massachusetts Institute of Technology
{mehdi_a, jhow}@mit.edu

Abstract— **This paper investigates the consensus problem for a team of agents with inconsistent information, which is a core component for many proposed distributed planning schemes. Kalman filtering approaches to the consensus problem have been proposed, and they were shown to converge for strongly connected networks. However, it is demonstrated in this paper that these previous techniques can result in biased estimates that deviate from the centralized solution, if it had been computed. A modification to the basic algorithm is presented to ensure the Kalman Filter converges to an unbiased estimate. The proof of convergence for this Modified Distributed Kalman Consensus algorithm to the unbiased estimate is then provided for both static and dynamic communication networks. These results are demonstrated in simulation using several simple examples.**

I. INTRODUCTION

Coordinated planning for a group of agents has been given significant attention in recent research [1], [2], [3]. This includes work on various planning architectures, such as distributed [3], hierarchic [1], and centralized [2]. In a centralized planning scheme, all of the agents communicate with a central agent to report their information and new measurements. The central planner gathers this available information to produce coordinated plans for all agents, which are then redistributed to the team. Note that generating a coordinated plan using a centralized approach can be computationally intensive, but otherwise it is relatively straight forward because the central planner has access to all information. However, this approach is often not practical due to communication limits, robustness issues, and poor scalability [3]. Thus attention has also focused on distributed planning approaches, but this process is complicated by the extent to which the agents must share their information to develop coordinated plans. This complexity can be a result of dynamic or risky environments or strong coupling between tasks, such as tight timing constraints. One proposed approach to coordinated distributed planning is to have the agents share their information to reach *consensus* and then plan independently [3]. Several different algorithms have been developed in the literature for agents to reach consensus [4]–[10] for a wide range of static and dynamic communication structures. In particular, a recent paper by Ren *et al.* [5] uses the well known Kalman filtering approach to develop the Kalman Consensus Algorithm (KCA) for both continuous and discrete updates and presents numerical examples to show their convergence.

In this paper we will compare the result of the KCA with the centralized algorithm in which estimation is ac- complished by aggregating the information in a centralized fashion using the appropriate Kalman gains. It will be shown that, although the KCA converges for strongly connected networks, its convergence to the true centralized estimate is only guaranteed for a special type of balanced communication network in which, at each time step, the outflows of all agents are equal. In particular, examples are given to show that a slight change in the special network structure considered in Ref. [5] leads to a deviation in the converged estimate away from the true centralized value. The main contribution of this paper is to extend the Kalman Consensus Algorithm so that it gives unbiased estimates for any strongly connected communication network. The proof of convergence of the new *Modified Decentralized Kalman Consensus* (MDKC) algorithm to this unbiased estimate is then provided for both static and dynamic communication networks.

The paper outline is as follows: Section II provides some background on the consensus problem and Section III formulates the Kalman Consensus algorithm and discusses the convergence properties. The new Modified Kalman Consensus algorithm is formulated in Section IV and examples are given to show its convergence to an unbiased estimate. Finally, we give the proof of convergence to an unbiased estimate for static and dynamic communication structure.

II. CONSENSUS PROBLEM

This section presents the consensus problem statement and discusses common solution algorithms [4]–[10].

A. Problem Statement

Suppose there are n agents $\mathcal{A} = \{\mathcal{A}_1, \ldots, \mathcal{A}_n\}$ with inconsistent information and let x_i be the information associated with agent i. The objective is for the agents to communicate this information amongst themselves to reach consensus, which means that all of the agents have the same information ($x_i = x_j, \; \forall i, j \in \{1, \ldots, n\}$).

To simplify the notation in this paper, we assume that the information is a scalar value, but the results can be easily extended to the case of a vector of information.

The communication pattern at any time t can be described in terms of a directed graph $\mathbb{G}(t) = (\mathcal{A}, \mathcal{E}(t))$, where $(\mathcal{A}_i, \mathcal{A}_j) \in \mathcal{E}(t)$ if and only if there is a unidirectional information exchange link from $\mathcal{A}_i$ to $\mathcal{A}_j$ at time t. Here we assume that there is a link from each agent to itself, $(\mathcal{A}_i, \mathcal{A}_i) \in \mathcal{E}(t), \, \forall \, i, t$. The adjacency matrix $G(t) = [g_{ij}(t)]$

of a graph $\mathbb{G}(t)$ is defined as

$$g_{ij}(t) = \begin{cases} 1 & \text{if} \quad (\mathcal{A}_j, \mathcal{A}_i) \in \mathcal{E}(t) \\ 0 & \text{if} \quad (\mathcal{A}_j, \mathcal{A}_i) \notin \mathcal{E}(t) \end{cases} \quad (1)$$

and a directed path from $\mathcal{A}_i$ to $\mathcal{A}_j$ is a sequence of ordered links (edges) in $\mathcal{E}$ of the form $(\mathcal{A}_i, \mathcal{A}_{i_1}), (\mathcal{A}_{i_1}, \mathcal{A}_{i_2}), \ldots, (\mathcal{A}_{i_r}, \mathcal{A}_j)$. A directed graph $\mathbb{G}$ is called strongly connected if there is a directed path from any node to all other nodes.

B. Consensus Algorithm

If the information, x_i of agent $\mathcal{A}_i$, is updated in discrete time steps using the data communicated from the other agents, then the update law can be written as

$$x_i(t+1) = x_i(t) + \sum_{j=1}^{N} \alpha_{ij}(t) g_{ij}(t)(x_j(t) - x_i(t)) \quad (2)$$

where $\alpha_{ij}(t) \geq 0$ represents the relative effect of information of agent $\mathcal{A}_j$ on the information of agent $\mathcal{A}_i$. The parameter $\alpha_{ij}(t)$ can be interpreted as the relative confidence that agent $\mathcal{A}_i$ and $\mathcal{A}_j$ have that their information variables are correct. Equation (2) can also be written in matrix form as $\mathbf{x}(t+1) = A(t)\mathbf{x}(t)$, where $\mathbf{x}(t) = [x_1(t), \ldots, x_n(t)]^T$, and the $n \times n$ matrix $A(t) = [a_{ij}(t)]$ is given by

$$a_{ij}(t) \begin{cases} \geq 0 & \text{if} \quad g_{ij}(t) = 1 \\ = 0 & \text{if} \quad g_{ij}(t) = 0 \end{cases} \quad (3)$$

Several methods such as *Fixed Coefficients, Vicsek Model, Gossip Algorithm and Kalman Filtering* have been proposed to pick values for the matrix A [10], [5]. In the Kalman filtering approach, the coefficients, a_{ij}, are chosen to account for the uncertainty each agent has in its information. Section III-A summarizes the Kalman filter formulation of consensus problem from Ref. [5]. Simulations are then presented to show that the performance of this algorithm strongly depends on the structure of communication network. An extension to this algorithm is proposed in Section IV that is shown to work for more general communication networks.

III. KALMAN CONSENSUS FORMULATION

This section provides a brief summary of Ref. [5] which uses Kalman filtering concepts to formulate the consensus problem for a multi-agent system with static information.

A. Kalman Consensus

Suppose at time t, $x_i(t)$ represents the information (perception) of agent $\mathcal{A}_i$ about a parameter with the true value x^*. This constant true value is modeled as the state, $x^*(t)$, of a system with trivial dynamics and a zero-mean disturbance input, $x^*(t+1) = x^*(t) + w(t)$. The measurements for agents $\mathcal{A}_i$ at time t are the information that it receives from other agents,

$$z_i(t) = \begin{bmatrix} g_{i1}(t)[x_1(t) + v_{i1}(t)] \\ \vdots \\ g_{in}(t)[x_n(t) + v_{in}(t)] \end{bmatrix} \quad (4)$$

where $v_{ij}(t)$ represents the communication noise and $g_{ij}(t) = 1$ if there is a communication link at time t from

agent $\mathcal{A}_j$ to $\mathcal{A}_i$, and 0 otherwise. Assuming that the agents' estimation errors, $(x_i(t) - x^*)$, are uncorrelated,

$$E[(x_i(t) - x^*)(x_j(t) - x^*)^T] = 0 \quad , \quad i \neq j$$

and by defining $\Omega_{ij}(t) = E[v_{ij}(t)v_{ij}(t)^T]$ and

$$P_i(t) = E[(x_i(t) - x^*)(x_i(t) - x^*)^T]$$

then the discrete-time Kalman Consensus algorithm for agent i can be written as [5]

$$\begin{aligned} P_i(t+1) &= \{[P_i(t) + Q(t)]^{-1} \quad (5) \\ &+ \sum_{j=1, j \neq i}^{n} g_{ij}(t)[P_j(t) + \Omega_{ij}(t)]^{-1}\}^{-1} \end{aligned}$$

$$\begin{aligned} x_i(t+1) &= x_i(t) + P_i(t+1) \sum_{j=1, j \neq i}^{n} \{g_{ij}(t) \\ & [P_j(t) + \Omega_{ij}(t)]^{-1}[x_j(t) + v_{ij}(t+1) - x_i(t)]\} \end{aligned}$$

Since it is assumed that $g_{ii} = 1$, then to make the formulation identical to the one in Ref. [5], i is excluded from the summations ($j \neq i$) in the above equations. Equations (5) are applied recursively until all the agents converge in their information or, equivalently, consensus is reached ($t = 1, \ldots, T_C$). T_C is the time that consensus is reached.

Ref. [5] shows that under certain conditions the proposed Kalman consensus algorithm converges and the converged value is based on the confidence of each agent about the information. The following sections analyze the performance of this algorithm for different network structures and modifications are proposed to improve the convergence properties.

B. Centralized Kalman Consensus

The centralized Kalman estimator for the consensus problem is formulated in this section to be used as a benchmark to evaluate different decentralized algorithms. Since the centralized solution is achieved in one iteration ($T_C = 1$) and the decentralized solution is solved over multiple iterations ($T_C > 1$), some assumptions are necessary to enable a comparison between the two algorithms. In particular since the communication noise and process noise are added in each iteration and the centralized solution is done in one step, consistent comparisons can only be done if the sensor/process noise are zero ($v_{ij}(t) = 0$, $w(t) = 0$; $\forall \ i, j, t$). These assumptions are made solely to enable a comparison of different algorithms with the benchmark (centralized), and they do not impose any limitations on the modified distributed algorithms that will be developed in the next sections. Under these assumptions, the centralized solution using Kalman filter is

$$\bar{P} = \left\{ \sum_{i=1}^{n} [P_i(0)]^{-1} \right\}^{-1} \quad (6)$$

$$\bar{x} = \bar{P} \sum_{i=1}^{n} \left\{ [P_i(0)]^{-1} x_i(0) \right\}$$

which is calculated in one step (compared to the iterative decentralized KCA).

TABLE I
COMPARING THE RESULTS OF DIFFERENT ALGORITHMS.

Algorithm	Case 1	2	3	4	5	6
Centralized	6.043	6.914	6.043	6.914	6.043	6.914
KCA [5]	6.043	6.914	5.659	6.251	5.839	6.215
MDKC	6.043	6.914	6.043	6.914	6.116	6.901

C. Example

The *meet-for-dinner* example [5] is used here as a benchmark to compare the performance (accuracy) of different algorithms. In this problem, a group of friends decide to meet for dinner, but fail to specify a precise time to meet. On the afternoon of the dinner appointment, each individual realizes that they are uncertain about the time of dinner. A centralized solution to this problem would be to have a conference call and decide on the time by some kind of averaging on their preferences. Since the conference call is not always possible, a decentralized solution is required. In the decentralized solution, individuals contact each other (call, leave messages) and iterate to converge to a time (reach consensus). Here the Kalman Consensus algorithm from Section III-A is used to solve this problem with $n = 10$ agents. The Kalman Consensus algorithm in (5) is applied to the two cases of the meet-for-dinner example in Ref. [5]. In Case 1, the initial states and the initial variances are uniformly assigned in the ranges of $[5 \quad 7]$ and $[0.8 \quad 1.2]$ respectively. In Case 2, the variance of the agent with initial data $x_i(0) = 7$ (leader) is given an initial variance of $P_i(0) = 0.001$, which is significantly lower than the other agents, and therefore has a higher weight on the final estimate. The simulation results given in Table I are consistent with the results in Ref. [5].

To evaluate the performance of the various algorithms, the simulation results are compared to the true estimate, $\bar{x}$, calculated from the centralized algorithm in (6). A comparison of the estimates in Table I clearly shows that the solution to the decentralized algorithm in (5) is identical to the true centralized estimate for Cases 1 and 2. These two simulations use a special case of a balanced communication network in which each agent communicates with exactly one other agent, so that

$$\text{Inflow}(\mathcal{A}_i) = \text{Outflow}(\mathcal{A}_i) = 1, \quad \forall \ \mathcal{A}_i \in \mathcal{A} \qquad (7)$$

where $\text{Inflow}(\mathcal{A}_i)$ is the number of links of the form $(\mathcal{A}_j, \mathcal{A}_i) \in \mathcal{E}$ and $\text{Outflow}(\mathcal{A}_i)$ is the number of links of the form $(\mathcal{A}_i, \mathcal{A}_j) \in \mathcal{E}$.

To investigate the performance of the decentralized algorithm in more general configurations, similar examples were used with slightly different communication networks. The graphs associated with these new architectures are still strongly connected, but the assumption in (7) is relaxed, which is accomplished by adding four extra links to the original graphs of Cases 1 and 2. The results are presented in Table I as Cases 3 and 4. For these cases, the solution of decentralized algorithm of (5) deviates from the true estimate, $\bar{x}$, resulting from the centralized solution. These
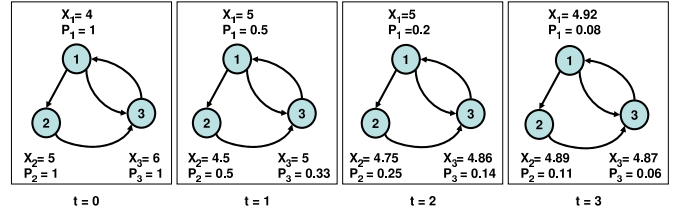

Fig. 1. An example to show the bias of the Decentralized Kalman Consensus Algorithm, $x_i(t)$ and $P_i(t)$ are the estimate and its accuracy of agent $\mathcal{A}_i$ at time t.

results confirm that the Kalman Consensus algorithm always converges to a value that respects the certainty of each agent about the information, but shows that, in cases for which the network does not satisfy the special condition of (7), the consensus value can be biased and deviate from the centralized solution.

IV. MODIFIED KALMAN CONSENSUS ALGORITHM

This section addresses the cause of the inaccuracy in the Kalman Consensus Algorithm (Table I, cases 3 and 4) and presents a modification to the original algorithm to fix this problem. The new algorithm generates the true centralized estimate, $\bar{x}$, using a decentralized estimator for a general, strongly connected network. The problem with the original formulation in (5) is a result of the weight that is given to the estimate of each agent $\mathcal{A}_i$. Based on the Kalman filter, the relative weights given to each estimate should be relative to the accuracy of the estimates, P_i's (see (6)). The formulation in (5) uses the same idea, but these weights are further scaled by the outflow of the agents. This means that if agent $\mathcal{A}_i$ and $\mathcal{A}_j$ have exactly the same accuracy, $P_i = P_j$, but in addition the outflow of agent $\mathcal{A}_i$ is greater than the outflow of agent $\mathcal{A}_j$, then using (5) causes the information of agent $\mathcal{A}_i$ to be treated as if it is more accurate than information of $\mathcal{A}_j$ (or the effective value of P_i is less than P_j), which creates a bias in the converged estimate. Obviously, for the special balanced networks considered in the simulations of Ref. [5], this bias does not occur.

Figure 1 presents a simple example to illustrate the problem with the Kalman Consensus algorithm of (5). There are 3 agents with $[x_1(0) \ x_2(0) \ x_3(0)] = [4 \ 5 \ 6]$ and $(P_i(0) = 1, \ i \in \{1, 2, 3\})$. As shown in the figure, the outflows of agents 2 and 3 are both one, but it is two for agent 1. Since all agents have the same initial accuracy, the centralized solution is the average of the initial estimates, $\bar{x} = 5$. Figure 1 shows four steps of the Kalman consensus algorithm for this example. At time $t = 3$ all of the estimates are less than 5, and the final converged estimate is 4.89, which is different from the centralized estimate. Note also that the deviation of the final value from the correct estimate is towards the initial value of agent 3, which has the larger outflow. This bias is essentially the result of an imbalanced network in which information of agents with different outflows are accounted for in the estimation with different weights.

In order to eliminate the bias, weights should be modified to cancel the effect of different outflows. The following

presents a modification to the Kalman Consensus algorithm that eliminates this bias. The main idea is to scale the accuracy of the agents by their outflow, which gives the *Modified Decentralized Kalman Consensus* (MDKC) algorithm. For agent $\mathcal{A}_i$ at time $t + 1$, the solution is given by

$$P_i(t + 1) = \{[P_i(t) + Q(t)]^{-1} +$$
$$\sum_{j=1}^{n}(g_{ij}(t)[\mu_j(t)P_j(t) + \Omega_{ij}(t)]^{-1})\}^{-1} \quad (8)$$

$$x_i(t + 1) = x_i(t) + P_i(t + 1)\sum_{j=1}^{n}\{g_{ij}(t) \cdot$$
$$[\mu_j(t)P_j(t) + \Omega_{ij}(t)]^{-1}(x_j(t) + v_{ij}(t) - x_i(t))\}$$

where $\mu_j(t)$ is the scaling factor associated with agent $\mathcal{A}_j$ and,

$$\mu_j(t) = \sum_{k=1,\ k \neq j}^{n} g_{kj}(t) \quad (9)$$

To show the performance (accuracy) of the MDKC algorithm, the four cases of the meet-for-dinner problem in Section III-C were re-solved using this new approach. The results for the four cases are presented in Table I. As shown, in all four cases the MDKC algorithm converges to the true estimates (the results of the centralized algorithm). The following sections present the proof of convergence of the MDKC algorithm to the true centralized estimate.

As discussed, a fair comparison of the centralized solution with any decentralized solution is only possible if the noise is eliminated from the formulation. However, as shown in (5) and (8), both of the KCA and MDKC algorithms can perform consensus when there is process, measurement, and communication noise. To compare the two algorithms with noise, Cases 3 and 4 were repeated with additive communication noise ($v_{ij}(t)$ in (4)). The noise was assumed to have a normal distribution with $\Omega = 0.1^2$. The results are shown in Table I (Cases 5 and 6). Note that because of the noise, the agents do not reach complete consensus and the values shown in Table I for Cases 5 and 6 are the average of agents' final values. The results of these two cases clearly show that MDKC converges to a solution that is much closer to the centralized solution than KCA.

A. Information Form of MDKC

The information form of Kalman filtering is used to prove that the MDKC algorithm converges to the true centralized estimate, $\bar{x}$, in (6). The Information filter is an equivalent form of the Kalman filter that simplifies the measurement update, but complicates the propagation [11]. It is typically used in systems with a large measurement vector, such as sensor fusion problems [12]. Since the propagation part of the Kalman filter is absent (or very simple) in the consensus problem, the information form of the filter also simplifies the formulation of that problem. The following briefly presents the information form of the Kalman consensus problem. To be consistent with the example in Section III-C, it is assumed that the process and measurement noise are zero. To write the

Modified Kalman Consensus algorithm (8) in the information form, for agent $\mathcal{A}_i$ define

$$Y_i(t) \equiv P_i(t)^{-1} \text{ and } y_i(t) \equiv Y_i(t)x_i(t) \quad (10)$$

then, (8) can be written as

$$Y_i(t + 1) = \frac{1}{2}\left\{Y_i(t) + \sum_{j=1,j\neq i}^{n}\frac{g_{ij}(t)}{\mu_j(t)}Y_j(t)\right\} \quad (11)$$

$$y_i(t + 1) = \frac{1}{2}\left\{y_i(t) + \sum_{j=1,j\neq i}^{n}\frac{g_{ij}(t)}{\mu_j(t)}y_j(t)\right\}$$

and after each iteration (time t), for agent $\mathcal{A}_i$

$$x_i(t) = Y_i(t)^{-1}y_i(t) \quad (12)$$

Note that the expressions in (12) are scaled by a factor of $1/2$, which has no effect on the estimation, but simplifies later proofs. These equations can be written in matrix form,

$$\mathbf{Y}(t + 1) = \Psi(t)\mathbf{Y}(t) \quad (13)$$
$$\mathbf{y}(t + 1) = \Psi(t)\mathbf{y}(t) \quad (14)$$

where $\mathbf{Y}(t) = [Y_1(t), \dots, Y_n(t)]^T$, $\mathbf{y}(t) = [y_1(t), \dots, y_n(t)]^T$ and $\Psi(t) = [\psi_{ij}(t)]$ with

$$\psi_{ij}(t) = \begin{cases} \dfrac{1}{2} & \text{if } j = i \\ \dfrac{g_{ij}(t)}{2\mu_j(t)} & \text{if } j \neq i \end{cases} \quad (15)$$

A comparison of the simple linear update in (13) and (14) with the nonlinear updates of the Kalman filter (8) shows the simplicity of this information form for the consensus problem. Note that since agents iterate on communicating and updating their information before using it, the inversions in (10) and (12) do not need to be performed every iteration. At the beginning of the consensus process, each agent $\mathcal{A}_i$ transforms its initial information, $x_i(0)$, and associated accuracy, $P_i(0)$, to $y_i(0)$ and $Y_i(0)$ using (10). In each following iteration, the transformed values ($y_i(t)$, $Y_i(t)$) are communicated to other agents and are used in the update process of (12). At the end of the consensus process the state $x_i(T_C)$ can be extracted from $y_i(T_C)$ and $Y_i(T_C)$ using (12).

B. Proof of Unbiased Convergence

This section provides the results necessary to support the proof of convergence of the MDKC algorithm to an unbiased estimate in the absence of noise.

Definition 1 ([15]): A nonnegative matrix $A = [a_{ij}] \in \mathbb{C}^{n \times n}$ is called row stochastic if $\sum_{j=1}^{n} a_{ij} = 1$, $1 \leq i \leq n$ and it is called column stochastic if $\sum_{i=1}^{n} a_{ij} = 1$, $1 \leq j \leq n$. Note that if A is a row stochastic matrix, A^T is a column stochastic matrix. If we denote by $\mathbf{e} \in \mathbb{R}^n$ the vector with all components $+1$, a nonnegative matrix A is row stochastic if and only if $A\mathbf{e} = \mathbf{e}$.

Lemma 1: The matrix $\Psi(t) = [\psi_{ij}(t)]$ defined in (15) is *column stochastic.*

Proof: For any column j,

$$\sum_{i=1}^{n} \psi_{ij}(t) = \frac{1}{2}\left(1 + \sum_{i=1, i \neq j}^{n} \frac{g_{ij}(t)}{\mu_j(t)}\right)$$

Thus using (9)

$$\sum_{i=1}^{n} \psi_{ij}(t) = \frac{1}{2}\left(1 + \frac{1}{\mu_j(t)}\mu_j(t)\right) = 1 \qquad (16)$$

so Ψ is column stochastic. $\qquad\square$

Lemma 2: The directed graph associated with matrix $\Psi = [\psi_{ij}]$ defined in (15), is strongly connected.

Proof: By definition (15), $\psi_{ij} > 0$ if $g_{ij} > 0$ and $\psi_{ij} = 0$ if $g_{ij} = 0$ and therefore matrices $\Psi = [\psi_{ij}]$ and $G = [g_{ij}]$ are both adjacency matrices to the same graph, which was assumed to be strongly connected. $\qquad\square$

Theorem 1 ([14]): For any $A = [a_{ij}] \in \mathbb{C}^{n \times n}$, A is irreducible if and only if its directed graph $\mathbb{G}(A)$ is strongly connected.

Theorem 2 (Perron-Frobenius Theorem, [14]): Given any $A = [a_{ij}] \in \mathbb{R}^{n \times n}$, with $A \succeq 0$ (A is nonnegative or $a_{ij} \geq 0$) and with A irreducible, then:

i) A has a positive real eigenvalue equal to its spectral radius $\rho(A)$;

ii) to $\rho(A)$ there corresponds an eigenvector $\mathbf{v} = [v_1, v_2, \ldots, v_n]^T \succ \mathbf{0}$;

iii) $\rho(A)$ is a simple eigenvalue of A.

Theorem 3 (Geršgorin, [13]): Let $A = [a_{ij}] \in \mathbb{C}^{n \times n}$, and let

$$R_i(A) \equiv \sum_{j=1, j \neq i}^{n} |a_{ij}|, \qquad 1 \leq i \leq n \qquad (17)$$

denote the *deleted absolute row sums* of A. Then all the eigenvalues of A are located in the union of n discs

$$\bigcup_{i=1}^{n} \{z \in \mathbb{C} : |z - a_{ii}| \leq R_i(A)\}$$

Definition 2: A nonnegative matrix $A \in \mathbb{C}^{n \times n}$ is said to be *primitive* if it is irreducible and has only one eigenvalue of maximum modulus.

Theorem 4 ([13]): If $A \in \mathbb{C}^{n \times n}$ is nonnegative and primitive, then

$$\lim_{m \to \infty} [\rho(A)^{-1}A]^m = L \succ 0$$

where $L = \mathbf{v}\mathbf{u}^T$, $A\mathbf{v} = \rho(A)\mathbf{v}$, $A^T\mathbf{u} = \rho(A)\mathbf{u}$, $\mathbf{v} \succ 0$, $\mathbf{u} \succ 0$, and $\mathbf{v}^T\mathbf{u} = 1$.

Lemma 3: For the matrix $\Psi = [\psi_{ij}]$ defined in (15),

$$\lim_{m \to \infty} \Psi^m = \mathbf{v}\mathbf{e}^T \succ 0$$

where $\mathbf{v}$ is a column vector and for the matrix C, $C \succ 0$ means that $c_{ij} > 0 \; \forall \, i, j$.

Proof: By definition $\Psi \succeq 0$ ($\psi_{ij} \geq 0$), and the directed graph associated with it is strongly connected (*Lemma 2*), so from *Theorem 1*, Ψ is irreducible. Thus Ψ has a simple eigenvalue equal to $\rho(\Psi)$ (*Theorem 2*).

Furthermore, Ψ is column stochastic (*Lemma 1*) and by definition Ψ has an eigenvalue $\lambda_1 = 1$ (*Definition 1*). Using the Geršgorin Theorem (*Theorem 3*), all of the eigenvalues of the row-stochastic matrix Ψ^T are located in the union of n disks

$$\bigcup_{i=1}^{n} \{z \in \mathbb{C} : |z - \psi_{ii}| \leq R_i(\Psi^T)\}$$

Using (15), $\psi_{ii} = 0.5, \forall \, i$, and $R_i(\Psi^T) = 0.5$ (see (17)), and thus all the eigenvalues of Ψ^T and Ψ are located in the disc $\{z \in \mathbb{C} : |z - 0.5| \leq 0.5\}$. Consequently all the eigenvalues of Ψ satisfy $|\lambda_i| \leq 1$, $\forall i$, and hence $\rho(\Psi) \leq 1$.

Since $\lambda_1 = 1$, therefore $\lambda_1 = \rho(\Psi) = 1$. As a result Ψ has only one eigenvalue of maximum modulus and therefore is *primitive* (see *Definition 2*). Finally, using *Theorem 4*,

$$\lim_{m \to \infty} [\rho(\Psi)^{-1}\Psi]^m = L \succ 0$$

where $L = \mathbf{v}\mathbf{u}^T$, $\Psi\mathbf{v} = \rho(\Psi)\mathbf{v}$, $\Psi^T\mathbf{u} = \rho(\Psi)\mathbf{u}$, $\mathbf{v} \succ 0$, $\mathbf{u} \succ 0$, and $\mathbf{v}^T\mathbf{u} = 1$. However, since $\rho(\Psi) = 1$, and using *Definition 1*, $\mathbf{u} = \mathbf{e}$, then it follows that $\lim_{m \to \infty} \Psi^m = \mathbf{v}\mathbf{e}^T \succ 0$. $\qquad\square$

With these results, we can now state the main result of the paper.

Theorem 5: For any strongly connected, time-invariant communication network, $\mathbb{G}$, and for any agent $\mathcal{A}_i$ and any initial estimate, $x_i(0)$, and variance, $P_i(0)$, the estimate, $x_i(t)$, resulting from the Modified Distributed Kalman Consensus algorithm introduced in (8) and (12), converges to the true centralized estimate, $\bar{x}$, calculated using (6), or equivalently,

$$\lim_{t \to \infty} x_i(t) \to \bar{x} \qquad \forall i \in \{1, \ldots, n\} \qquad (18)$$

Proof: The objective is to show that (18) is satisfied or equivalently, $\lim_{t \to \infty} \mathbf{x}(t) \to \bar{x}\mathbf{e}$, where $\mathbf{x} = [\, x_1, \ldots, x_n \,]^T$. Let $\mathbf{v}^\dagger$ denote the element inverse of a vector, $\mathbf{v}^\dagger = [v_1^{-1}, \ldots, v_n^{-1}]^T$. Using (12) it follows that $\lim_{t \to \infty} \mathbf{x}(t) = \lim_{t \to \infty} \mathbf{Y}^\dagger(t) \odot \mathbf{y}(t)$, where the operator $\odot$ represents the element by element multiplication. With the assumed time-invariance of the communication network, $\Psi(t) = \Psi$, and using (13) and (14)

$$\lim_{t \to \infty} \mathbf{x}(t) = \lim_{t \to \infty} \left(\Psi^t \mathbf{Y}(0)\right)^\dagger \odot \left(\Psi^t \mathbf{y}(0)\right)$$

Using *Lemma 3*,

$$\lim_{t \to \infty} \mathbf{x}(t) = \left(\mathbf{v}\underbrace{\mathbf{e}^T\mathbf{Y}(0)}_{\text{scalar}}\right)^\dagger \odot \left(\mathbf{v}\underbrace{\mathbf{e}^T\mathbf{y}(0)}_{\text{scalar}}\right)$$

$$= \left(\mathbf{e}^T\mathbf{Y}(0)\right)^{-1}\left(\mathbf{v}^\dagger \odot \mathbf{v}\right)\left(\mathbf{e}^T\mathbf{y}(0)\right)$$

Since $\mathbf{v}^T\mathbf{e} \succ 0$ (*Lemma 3*), $\mathbf{v} \succ 0$, therefore, $\mathbf{v}^\dagger \odot \mathbf{v} = \mathbf{e}$ and

$$\lim_{t \to \infty} \mathbf{x}(t) = \left(\mathbf{e}^T\mathbf{Y}(0)\right)^{-1}\left(\mathbf{e}^T\mathbf{y}(0)\right)\mathbf{e}$$

$$= \left\{\sum_{i=1}^{n} Y_i(0)\right\}^{-1}\left\{\sum_{i=1}^{n} y_i(0)\right\}\mathbf{e}$$

Using the relationship $Y_i(0) = P_i(0)^{-1}$, it follows that

$$\lim_{t \to \infty} \mathbf{x}(t) = \left\{ \sum_{i=1}^{n} P_i(0)^{-1} \right\}^{-1} \left\{ \sum_{i=1}^{n} P_i(0)^{-1} x_i(0) \right\} \mathbf{e}$$

and then from (6), $\lim_{t \to \infty} \mathbf{x}(t) = \bar{x}\mathbf{e}$. Thus the MDKC algorithm introduced in (8) converges to the true centralized estimate, $\bar{x}$, when the strongly connected communication network is time-invariant. $\square$

In what follows we prove that the same is true for a time-varying communication network.

Definition 3 ([16]): A stochastic matrix A is called inde-composable and aperiodic (SIA) if $L = \lim_{m \to \infty} A^m$ exists and all the rows of L are the same.

Definition 4: Let $A_1, \ldots, A_k \in \mathbb{C}^{n \times n}$. By a *word* in A_i's of the length t we mean the product of t A_i's with repetition permitted.

Theorem 6 ([16]): Let $A_1, \ldots, A_k$ be square row-stochastic matrices of the same order such that any word in the A_i's is SIA. For any $\epsilon > 0$ there exists an integer $\nu(\epsilon)$ such that any word B (in the A's) of length $m \geq \nu(\epsilon)$ satisfies $\delta(B) < \epsilon$. In other words, the result is that any sufficiently long word in the A_i's has all its rows the same or, $\lim_{m \to \infty} A_1 A_2 \ldots A_m = \mathbf{e}\mathbf{v}^T$.

Lemma 4: If matrices $A_1, \ldots, A_N \in \mathbb{R}^{n \times n}$, $\forall i, A_i \succeq 0$ have strictly positive diagonal elements, then matrix $C = A_1 A_2 \ldots A_N$ has the same properties ($C \succeq 0$ and all diagonal elements of C are strictly positive).

Proof: To establish this result, it will first be shown that if matrices $A, B \succeq 0$ have strictly positive diagonal elements then $D = AB$ has the same properties. Given that $D = AB$,

$$d_{ij} = \sum_{k=1}^{n} \underbrace{a_{ik} b_{kj}}_{\geq 0} \geq 0$$

$$d_{ii} = \sum_{k=1}^{n} a_{ik} b_{ki} = \underbrace{a_{ii} b_{ii}}_{>0} + \sum_{k=1, k \neq i}^{n} \underbrace{a_{ik} b_{ki}}_{\geq 0} > 0$$

which provides the necessary result. Therefore by induction, $C = A_1, \ldots, A_N \succeq 0$ and all diagonal elements of C are strictly positive. $\square$

Theorem 7: Let $\mathbb{G}$ be any dynamic communication network, where at each time step, $\mathbb{G}(t)$ is strongly connected. Then for any agent $\mathcal{A}_i$ and any initial estimate, $x_i(0)$, and variance, $P_i(0)$, the estimate, $x_i(t)$, resulting from the Modified Distributed Kalman Consensus algorithm, introduced in (8) and (12) converges to the true centralized estimate, $\bar{x}$, calculated using (6).

Proof: From *Lemma 3*, for any t, $\lim_{m \to \infty}(\Psi^T(t))^m = \mathbf{e}\mathbf{v}_t^T$, where the $\mathbf{v}_t$ is a column vector. Using *(15)* and *Lemma 1*, $\Psi^T(t)$ is row stochastic, so for any t, $\Psi^T(t)$ is SIA (see *Definition 3*). Then from *Theorem 6*, $\lim_{t \to \infty} \Psi^T(1)\Psi^T(2)\ldots \Psi^T(t) = \mathbf{e}\mathbf{v}^T$ for some $\mathbf{v}$, or equivalently,

$$\lim_{t \to \infty} \Psi(t)\Psi(t-1)\ldots \Psi(2)\Psi(1) = \mathbf{v}\mathbf{e}^T \quad (19)$$

Thus if it can be shown that $\mathbf{v} \succ 0$, then the proof of Theorem 7 would follow the same steps as the proof for

the time-invariant case in Theorem 5. To demonstrate that $\mathbf{v} \succ 0$, we first show that the diagonal elements of

$$L = \lim_{t \to \infty} \Psi^T(1)\Psi^T(2)\ldots \Psi^T(t) \quad (20)$$

are positive ($L_{ii} > 0$, $\forall i$). Since by its definition in (15), $\Psi(t) \succeq 0$ and all the diagonal elements of $\Psi(t)$ are strictly positive, then $C = \Psi^T(1)\Psi^T(2)\ldots \Psi^T(t)$ and consequently L in (20) have positive elements, $L_{ij} \geq 0$, $\forall i, j$, and strictly positive diagonal elements, $L_{ii} > 0$, $\forall i$, (see *Lemma 4*).

Also, since $L = \mathbf{e}\mathbf{v}^T$ (see (19) and (20)), then all of the rows of L are equal. Furthermore, since $L_{ii} > 0, \forall~i$ then $L_{ji} > 0, \forall i, j$, which implies that $L = \mathbf{e}\mathbf{v}^T \succ 0$ and that $\mathbf{v} \succ 0$. The remainder of the proof follows the same steps as the proof for the time-invariant case in Theorem 5. $\square$

V. CONCLUSIONS

The performance of the Kalman consensus algorithm was investigated for a team of agents with static data. It was shown that, although this algorithm converges for the general case of strongly connected communication networks, it can result in a biased estimate when the outflow of the agents are not equal. An extension to this algorithm was then presented which was shown in simulations to converge to the true centralized estimate for general strongly connected networks. This algorithm was further proved to converge to an unbiased estimate for both static and dynamic communication networks.

REFERENCES

[1] P. Chandler and M. Pachter, "Hierarchical Control for Autonomous Teams" proceedings of the AIAA *GNC*, Montreal, Canada, Aug. 2001.
[2] Yan Jin, A. Minai, and M. Polycarpou, "Cooperative real-time search and task allocation in UAV teams," IEEE *CDC*, Dec. 2003.
[3] M. Alighanbari and J. How, "Decentralized Task Assignment for Unmanned Aerial Vehicles," IEEE *ECC and CDC*, Dec. 2005.
[4] R. Olfati-Saber and R.M. Murray, "Consensus Problems in Network of Agents With Switching Topology and Time-Delay," proceeding of the IEEE *Transaction on Automatic Control*, Vol. 49, No. 9, Sept. 2004.
[5] W. Ren, R. Beard, and D. Kingston, "Multi-agent Kalman Consensus with Relative Uncertainty", IEEE *ACC*, June 2005.
[6] W. Ren, R. Beard, and E. Atkins, "A Survey of Consensus Problems in Multi-agent Coordination", IEEE *ACC*, June 2005.
[7] A. Jadbabaie, J. Lin, and A. S. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules", IEEE *Transactions on Automatic Control*, Vol. 48, No. 6, June 2003.
[8] R. Olfati-Saber, "Distributed Kalman Filter with Embedded Consenus Filter", proceeding of the *ECC and CDC*, Seville, Spain, Dec. 2005.
[9] R. Olfati-Saber, "Consensus Filters for Sensor Networks and Distributed Sensor Fusion," proceeding of the *ECC and CDC*, Seville, Spain, Dec. 2005.
[10] V. Blondel, J. Hendricks, A. Olshevsky, and J. Tsitsiklis, "Convergence in Multiagent Coordination, Consensus, and Flocking," proceeding of the *ECC and CDC*, Seville, Spain, Dec. 2005.
[11] G. Minkler, Theory and Applications of Kalman Filtering, Magellan Book Company, 1990.
[12] Y. Kim and K. Hong, "Decentralized information filter in federated form", proceeding of the IEEE *SICE Conference*, Aug. 2003.
[13] R. A. Horn and C. R. Johnson, Matrix Analysis, Cambridge University Press, 1985.
[14] R. S. Varga, Geršgorin and His Circles, Springer; 2004.
[15] E. Seneta, Non-negative Matrices, John Wiley & Sons, 1973.
[16] J. Wolfowitz, "Product of Indecomposable, Aperiodic Stochastic Matrices", Proceeding of the *American Mathematical Society*, 1963.