JOURNAL OF AEROSPACE COMPUTING, INFORMATION, AND COMMUNICATION
Vol. 5, September 2008

Unbiased Kalman Consensus Algorithm

Mehdi Alighanbari* and Jonathan P. How†
Massachusetts Institute of Technology, Cambridge, MA 02139

DOI: 10.2514/1.34226

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 are shown to converge for strongly connected networks. It is demonstrated in this paper, however, that these previous techniques can result in biased estimates that deviate from the centralized solution, if it had been computed. An extension to the basic algorithm is presented to ensure the Kalman filter converges to an unbiased estimate. The proof of convergence for this new 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 examples for different network structures.

I. Introduction

COORDINATED planning for a group of agents has been given significant attention in recent research [1–7]. This includes work on various planning architectures, such as distributed [1,5], hierarchic [3,4], and centralized [2,6,7]. 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 straightforward because the central planner has access to all information. This approach is often not practical, however, owing to communication limits, robustness issues, and poor scalability [1,5]. 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 [5].

Several different algorithms have been developed in the literature for agents to reach consensus [8–16] for a wide range of static and dynamic communication structures. In particular, a recent paper by Ren et al. [15] uses the well-known Kalman filtering approach to develop the Kalman consensus algorithm (KCA) for both continuous and discrete updates and presents numerical examples and analytical proofs to show their convergence.

The objective of this paper is to extend the algorithm developed by Ren et al. [15] not only to ensure its convergence for the general form of communication networks, but also to ensure that the algorithm converges to the desired value. In the KCA the desired value is the value that is achieved if a centralized Kalman filter was applied to the initial information of the agents. We show, both by simulation and analytical proofs, that the new extended algorithm always converges to the desired value.

Received 23 August 2007; accepted for publication 30 May 2008.
* Research Assistant, Department of Aeronautics and Astronautics, mehdi@alum.mit.edu
† Professor, Department of Aeronautics and Astronautics, jhow@mit.edu, Associate Fellow AIAA

The main contribution of this paper is developing a KCA that gives an unbiased estimate of the desired value for static and dynamic communication networks. The proof of convergence of the new unbiased decentralized kalman consensus (UDKC) algorithm to this unbiased estimate is then provided for both static and dynamic communication networks. As the desired value in Kalman consensus is essentially a weighted average of the initial information of the agents, the proposed algorithm can also be used to achieve a general weighted average for the very general form of communication networks. Previous research has shown that the weighted average can only be achieved for the special case of strongly connected balanced networks. Another contribution of this paper is showing that these constraints on the network can be relaxed and the proposed algorithm still reaches the desired weighted average.

Section II provides some background on the consensus problem and Sec. III formulates the KCA and discusses the convergence properties. The new extension to the KCA is formulated in Sec. IV and more examples are given to show its convergence to an unbiased estimate. Finally, the proof of convergence to an unbiased estimate for static and dynamic communication structure is given.

II. Consensus Problem

This section presents the consensus problem statement and discusses some common algorithms for this problem [8–16].

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 among 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

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

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 [17] and a balanced network is defined as a network where for any node $\mathcal{A}_i$, its outflow equals its inflow.

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

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

where $\alpha_{ij}(t) \geqslant 0$ 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 [8]. 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

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

Several methods such as fixed coefficients, Vicsek model, gossip algorithm, and Kalman filtering have been proposed to pick values for the matrix $\boldsymbol{A}$ [9,15]. In the Kalman filtering approach, the coefficients, a_{ij}, are chosen to account for the uncertainty each agent has in its information. Section A summarizes the Kalman filter formulation of consensus problem from Ren et al. [15]. Simulations are then presented to show that the performance of this algorithm strongly depends on the structure of the communication network. An extension to this algorithm is proposed in Sec. IV that is shown to work for more general communication networks.

III. Kalman Consensus Formulation

This section provides a brief summary of work presented by Ren et al. [15], 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 $w \sim (0, Q)$

$$x^*(t+1) = x^*(t) + w(t)$$

The measurements for agents $\mathcal{A}_i$ at time t are the information that it receives from other agents

$$z_i(t) = \begin{bmatrix} g_{i1}(t)x_1(t) \\ \vdots \\ g_{in}(t)x_n(t) \end{bmatrix} \tag{4}$$

where $g_{ij}(t) = 1$ 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' initial estimation errors, $(x_i(0) - x^*)$, are uncorrelated, $E[(x_i(0) - x^*)(x_j(0) - x^*)^T] = 0, i \neq j$ and by defining

$$P_i(0) = E[(x_i(0) - x^*)(x_i(0) - x^*)^T]$$

then the discrete-time KCA for agent i can be written as [15]

$$P_i(t+1) = \left\{ [P_i(t) + Q(t)]^{-1} + \sum_{j=1, j\neq i}^{n} g_{ij}(t) \left[P_j(t)\right]^{-1} \right\}^{-1} \tag{5}$$

$$x_i(t+1) = x_i(t) + P_i(t+1) \sum_{j=1, j\neq i}^{n} \left\{ g_{ij}(t) \left[P_j(t)\right]^{-1} \left[x_j(t) - x_i(t)\right] \right\}$$

As it is assumed that $g_{ii} = 1$, then to make the formulation similar to the one in [15], i is excluded from the summations $(j \neq i)$ in the above equations. Equation (5) is applied recursively until all the agents converge in their information or, equivalently, consensus is reached $(t = 1, \ldots, T_{\text{consensus}})$. Note that, although $P_i(0)$ represents the initial covariance of $x_i(0)$, the values $P_i(t);\ t > 0$ need not have the same interpretation; they are just weights used in the algorithm that are modified using the covariance update procedure of the Kalman filter.

Reference [15] shows that under certain conditions the proposed KCA 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 distributed algorithms. As the centralized solution is achieved in one iteration $(T_{\text{consensus}} = 1)$ and the decentralized solution is solved over multiple iterations $(T_{\text{consensus}} > 1)$, some assumptions are necessary to enable a comparison between the two algorithms. In particular, as the process noise is added in each iteration and the centralized solution is done in one step, consistent comparisons can only be done if the process noise is zero $(w(t) = 0;\ \forall\ t)$. 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 algorithm that will be developed in the next sections. Under these assumptions, the centralized solution using the Kalman filter is

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

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

C. Example

The meet-for-dinner example [15] is used in the current paper 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 he is uncertain about the time of dinner. A centralized solution to this problem is to have a conference call and decide on the time by some kind of averaging on their preferences. As 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 KCA from Sec. A is used to solve this problem for $n = 10$ agents. Figure 1 shows the output of this algorithm for the two cases presented in [15], demonstrating that the results obtained are consistent. These simulations use a special case of a balanced communication network in which each agent communicates with exactly one other agent so that

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

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

In Fig. 1a, the initial states and the initial variances are uniformly assigned (case 1). In Fig. 1b, 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 more weight on the final estimate (case 2). To evaluate the performance of this algorithm, the results are compared to the true estimate, $\bar{x}$, calculated from the centralized algorithm in Eq. (6). The results in Table 1 clearly show that the solution to the decentralized algorithm in Eq. (5) is identical to the true centralized estimate.

As noted, these cases assume the special case of the communication networks in Eq. (7). To investigate the performance of the decentralized algorithm in more general cases, similar examples were used with slightly different communication networks. The graphs associated with these new architectures are still strongly connected, but the assumption in Eq. (7) is relaxed. This is accomplished using the original graphs of cases 1 and 2 with four extra

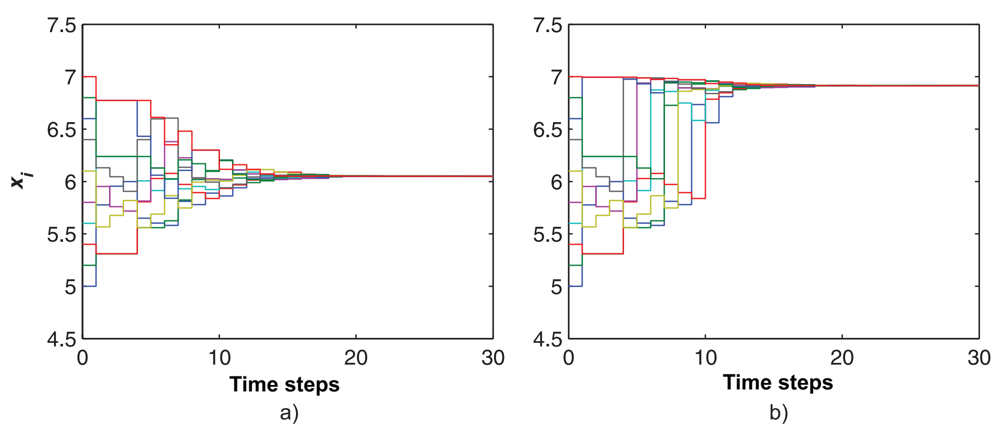

Fig. 1 The result of KCA for cases 1 and 2, demonstrating consistency with the results in [15]: a) no Leader; b) Leader.

Table 1 Comparing the results of different algorithms

Algorithm	Case 1	Case 2	Case 3	Case 4
Centralized	6.0433	6.9142	6.0433	6.9142
Kalman consensus	6.0433	6.9142	5.6598	6.2516
UDKC	6.0433	6.9142	6.0433	6.9142

links added to the original graph. The results are presented in Table 1 (cases 3, 4). For these cases, the solution of the decentralized algorithm of Eq. (5) deviates from the true estimate, $\bar{x}$, obtained from the centralized solution. The KCA always converges to a value that respects the certainty of each agent about the information, but these results show that in cases for which the network does not satisfy the condition of Eq. (7), the consensus value can be biased and deviate from the centralized solution.

The next section extends this algorithm to eliminate this bias and to guarantee convergence to the true centralized estimate, $\bar{x}$, for the general case of communication networks.

IV. Unbiased Decentralized Kalman Consensus

This section extends the Kalman consensus formulation of Eq. (5) to achieve the desired unbiased solution, which is the solution to the centralized algorithm presented in Eq. (6). The new extended algorithm generates the true centralized estimate, $\bar{x}$, using a decentralized estimator for any form of communication networks.

The main idea is to scale the accuracy of the agents by their outflow, which gives the UDKC algorithm. For agent $\mathcal{A}_i$ at time $t+1$, the solution is given by

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

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

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

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

To show the unbiased convergence of the UDKC algorithm, the four cases of the meet-for-dinner problem in Sec. C were resolved using this new approach. The results for the four cases are presented in Table 1. As shown, in all four cases the UDKC algorithm converges to the true estimates (the results of the centralized algorithm). The following remarks provide further details on the UDKC algorithm.

1) Both the original KCA and new UDKC formulations presented here differ from the previously developed weighted average consensus algorithms [18] in the sense that these algorithms not only update the information in each iteration, but also update the weights (Ps) that are used in the formulation. This additional update (Eqs. (5) and (8)) enables the UDKC algorithm to converge to the desired weighted average for a very general class of communication networks, while the previous form of consensus algorithm (Eq. (2)), where only the information itself is updated at each iteration [18], was limited to a special kind of strongly connected balanced network.
2) Reference [18] introduces an alternative form of the consensus algorithm that has some apparent similarities to the UDKC formulation introduced in this paper. The form of the consensus algorithm in [18] is as follows

$$\dot{x}_i = \frac{1}{|N_i|} \sum_{j\in N_i} (x_j - x_i) \tag{10}$$

where $N_i = \{j \in \mathcal{A} : (i, j) \in \mathcal{E}\}$ is the list of neighbors of agent $\mathcal{A}_i$. Note that in the notation of [18], if $(i, j) \in \mathcal{E}$ then there is a link from $\mathcal{A}_i$ to $\mathcal{A}_j$ but the information flow is from $\mathcal{A}_j$ to $\mathcal{A}_i$. Therefore, although $|N_i|$ is defined as the outdegree of agent $\mathcal{A}_i$, it is essentially the inflow of agent $\mathcal{A}_i$ in our formulation. Thus the consensus formulation of Eq. (10) has a scaling factor that is equal to the inflow of the receiving agent, $\mathcal{A}_i$. Note however, that the scaling factor in the UDKC algorithm (the coefficient μ in Eq. (8)) is the outflow of the sending agent, $\mathcal{A}_j$. This clarifies the key differences between UDKC and the method introduced in [18].

3) The scaling introduced in UDKC (the coefficient μ in Eq. (8)) does not change the topology of the network to make it a balanced network. The implicit effect of μ is essentially making the outflows of all agents equal to 1 and has no effect on the inflow of the agents. Thus the resulting network will not necessarily be a balanced network and therefore the results presented in [11] and [18] for balanced networks can not be used to prove the convergence of the UDKC algorithm to the desired weighted average. Figure 2 shows a simple network that is neither balanced (Inflow(1) = 1, Outflow(1) = 2) nor are its outflows equal (Outflow(1) = 2, Outflow(2) = 1). The adjacency matrix for this network is

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} \tag{11}$$

and applying the scaling μ defined in Eq. (9) gives

$$\begin{bmatrix} 0 & 0 & 1 \\ 0.5 & 0 & 0 \\ 0.5 & 1 & 0 \end{bmatrix} \tag{12}$$

which has the same outflow for all the nodes, but is still imbalanced (Inflow(2) = 0.5, Outflow(2) = 1).

To show why the outflow scaling results in convergence to the desired solution, a simple example is presented here. Based on the Kalman filter, the relative weights given to each estimate should be relative to the accuracy of the estimates, P_is (see Eq. (6)). The formulation in Eq. (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 Eq. (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 [15], this bias does not occur as the outflows are all equal to one.

Figure 3 presents a simple example to illustrate the problem with the KCA of Eq. (5). There are three agents with $[x_1(0)\ x_2(0)\ x_3(0)] = [\ 4 \quad 5 \quad 6\ ]$ 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. As all agents have the same initial accuracy, the centralized solution is the average of the initial estimates, $\bar{x} = 5$. Figure 3 shows four steps of the KCA 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

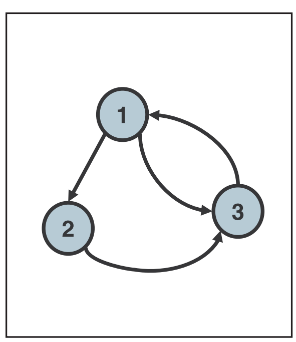

Fig. 2 A simple imbalanced network with unequal outflows.

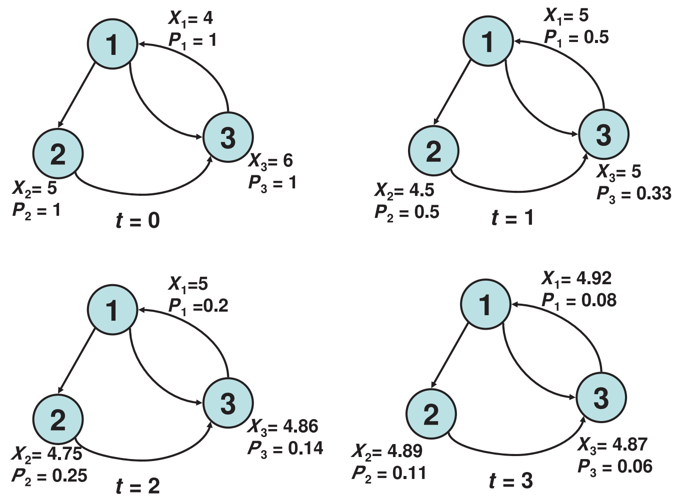

Fig. 3 An example to show the bias of the decentralized KCA, $x_i(t)$ and $P_i(t)$ are the estimate and its accuracy of agent $\mathcal{A}_i$ at time t.

estimate. Note also that the deviation of the final value from the correct estimate is towards the initial value of agent 1, which has the largest outflow. This bias is essentially the result of an imbalanced network in which information of agents with different outflows is accounted for in the estimation with different weights. To eliminate the bias, weights should be modified to cancel the effect of different outflows, which is essentially the modification that is introduced in Eq. (8).

The following sections present the proof of convergence of the UDKC algorithm to the true centralized estimate.

A. Information Form of UDKC

The information form of Kalman filtering is used to prove that the UDKC algorithm converges to the true centralized estimate, $\bar{x}$, in Eq. (6). The information filter is an equivalent form of the Kalman filter that simplifies the measurement update, but complicates the propagation [19]. It is typically used in systems with a large measurement vector, such as sensor fusion problems [20,21]. As 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 Sec. C, it is assumed that the process noise is zero. To write the UDKC (8) in the information form, for agent $\mathcal{A}_i$ define

$$Y_i(t) \equiv P_i(t)^{-1} \text{ and } y_i(t) \equiv Y_i(t)x_i(t) \tag{13}$$

then, Eq. (8) can be written as

$$Y_i(t+1) = \frac{1}{2}\left\{Y_i(t) + \sum_{j=1,j\neq i}^{n} \frac{g_{ij}(t)}{\mu_j(t)} Y_j(t)\right\} \tag{14}$$

$$y_i(t+1) = \frac{1}{2}\left\{y_i(t) + \sum_{j=1,j\neq i}^{n} \frac{g_{ij}(t)}{\mu_j(t)} y_j(t)\right\} \tag{15}$$

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

$$x_i(t) = Y_i(t)^{-1} y_i(t) \tag{16}$$

Note that the expressions in Eq. (15) 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

$$\mathbf{Y}(t+1) = \Psi(t)\mathbf{Y}(t) \tag{17}$$

$$\mathbf{y}(t+1) = \Psi(t)\mathbf{y}(t) \tag{18}$$

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

$$\psi_{ij}(t) = \begin{cases} \dfrac{1}{2} & \text{if} \quad j = i \\ \dfrac{g_{ij}(t)}{2\mu_j(t)} & \text{if} \quad j \neq i \end{cases} \tag{19}$$

A comparison of the simple linear update in Eqs. (17) and (18) with the nonlinear updates of the Kalman filter Eq. (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 Eqs. (13) and (16) 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 Eq. (13). 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 Eq. (15). At the end of the consensus process the state $x_i(T_{\text{consensus}})$ can be extracted from $y_i(T_{\text{consensus}})$ and $Y_i(T_{\text{consensus}})$ using Eq. (16).

B. Proof of Unbiased Convergence

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

Definition 1 [22].
A nonnegative matrix $A = [a_{ij}] \in \mathbb{C}^{n\times n}$ is called row stochastic if $\sum_{j=1}^{n} a_{ij} = 1$, $1 \leqslant i \leqslant n$ and it is called column stochastic if $\sum_{i=1}^{n} a_{ij} = 1$, $1 \leqslant j \leqslant n$. Note that if A is a row stochastic matrix, A^T is a column stochastic matrix.

Theorem 1 [22].
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 Eq. (19) is *column stochastic*.

Proof.
For any column j

$$\sum_{i=1}^{n} \psi_{ij}(t) = \frac{1}{2}\left(1 + \sum_{i=1, i\neq j}^{n} \frac{g_{ij}(t)}{\mu_j(t)}\right) = \frac{1}{2}\left(1 + \frac{1}{\mu_j(t)} \sum_{i=1, i\neq j}^{n} g_{ij}(t)\right) \tag{20}$$

Thus using Eq. (9)

$$\sum_{i=1}^{n} \psi_{ij}(t) = \frac{1}{2}\left(1 + \frac{1}{\mu_j(t)}\mu_j(t)\right) = 1 \tag{21}$$

so Ψ is column stochastic. ∎

Lemma 2.
The directed graph associated with matrix $\Psi = [\psi_{ij}]$ defined in Eq. (19), is strongly connected.

Proof.
By definition (19), $\psi_{ij} > 0$ 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. ∎

Theorem 2 [23].
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 3 (Perron–Frobenius Theorem [23]).
Given any $A = [a_{ij}] \in \mathbb{R}^{n\times n}$, with $A \succeq 0$ and with A irreducible, then

1) A has a positive real eigenvalue equal to its spectral radius $\rho(A)$;
2) to $\rho(A)$ there corresponds an eigenvector $\boldsymbol{v} = [v_1, v_2, \ldots, v_n]^T \succ \mathbf{0}$;
3) $\rho(A)$ is a simple eigenvalue of A.

Theorem 4 (Geršgorin [24]).
Let $A = [a_{ij}] \in \mathbb{C}^{n\times n}$, and let

$$R_i(A) \equiv \sum_{j=1, j\neq i}^{n} |a_{ij}|, \quad 1 \leqslant i \leqslant n \tag{22}$$

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

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

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 5 [24].
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 $L = \boldsymbol{v}\boldsymbol{u}^T$, $A\boldsymbol{v} = \rho(A)\boldsymbol{v}$, $A^T\boldsymbol{u} = \rho(A)\boldsymbol{u}$, $\boldsymbol{v} \succ 0$, $\boldsymbol{u} \succ 0$, and $\boldsymbol{v}^T\boldsymbol{u} = 1$.

Lemma 3.
For the matrix $\Psi = [\psi_{ij}]$ defined in Eq. (19),

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

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

Proof.
By definition $\Psi \succeq 0$ ($\psi_{ij} \geqslant 0$), and the directed graph associated with it is strongly connected (Lemma 2), so from Theorem 2, Ψ is irreducible. Thus Ψ has a simple eigenvalue equal to $\rho(\Psi)$ (Theorem 3).

Furthermore, Ψ is column stochastic (Lemma 1) and by definition Ψ has an eigenvalue $\lambda_1 = 1$ (Theorem 1). Using the Geršgorin Theorem (Theorem 4), all of the eigenvalues of the row-stochastic matrix Ψ^T are located in the union of n disks

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

Using Eq. (19), $\psi_{ii} = 0.5$, $\forall i$, and $R_i(\Psi^T) = 0.5$ (see Eq. (22)), and thus all the eigenvalues of Ψ^T and Ψ are located in the disc $\{z \in \mathbb{C} : |z - 0.5| \leqslant 0.5\}$. Consequently all the eigenvalues of Ψ satisfy $|\lambda_i| \leqslant 1$, $\forall i$, and hence $\rho(\Psi) \leqslant 1$.

As $\lambda_1 = 1$, 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 5,

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

where $L = \boldsymbol{v}\boldsymbol{u}^T$, $\Psi\boldsymbol{v} = \rho(\Psi)\boldsymbol{v}$, $\Psi^T\boldsymbol{u} = \rho(\Psi)\boldsymbol{u}$, $\boldsymbol{v} \succ 0$, $\boldsymbol{u} \succ 0$, and $\boldsymbol{v}^T\boldsymbol{u} = 1$. However, as $\rho(\Psi) = 1$, and using Theorem 1, $\boldsymbol{u} = \mathbf{e}$, then it follows that $\lim_{m\to\infty} \Psi^m = \boldsymbol{v}\boldsymbol{e}^T \succ 0$. ■

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

Theorem 6.
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 Eq. (8) and (15), converges to the true centralized estimate, $\bar{x}$, calculated using Eq. (6), or equivalently,

$$\lim_{t\to\infty} x_i(t) \to \bar{x} \quad \forall i \in \{1, \dots, n\} \tag{23}$$

Proof.
The objective is to show that Eq. (23) is satisfied or equivalently, $\lim_{t\to\infty} \boldsymbol{x}(t) \to \bar{x}\mathbf{e}$, where $\boldsymbol{x} = [x_1, \dots, x_n]^T$. Let $\boldsymbol{v}^\dagger$ denote the element inverse of a vector, $\boldsymbol{v}^\dagger = [v_1^{-1}, \dots, v_n^{-1}]^T$. Using Eq. (16) it follows that $\lim_{t\to\infty} \boldsymbol{x}(t) = \lim_{t\to\infty} \mathbf{Y}^\dagger(t) \odot \boldsymbol{y}(t)$, where the operator $\odot$ represents the element by element multiplication. With the assumed time-invariance of the communication network, $\Psi(t) = \Psi$, and using Eqs. (17) and (18)

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

Using Lemma 3

$$\begin{aligned}
\lim_{t\to\infty} \boldsymbol{x}(t) &= \left(\boldsymbol{v}\underbrace{\boldsymbol{e}^T\mathbf{Y}(0)}_{\texttt{scalar}}\right)^\dagger \odot \left(\boldsymbol{v}\underbrace{\boldsymbol{e}^T\boldsymbol{y}(0)}_{\texttt{scalar}}\right) \\
&= \left(\boldsymbol{e}^T\mathbf{Y}(0)\right)^{-1}\left(\boldsymbol{v}^\dagger \odot \boldsymbol{v}\right)\left(\boldsymbol{e}^T\boldsymbol{y}(0)\right)
\end{aligned}$$

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

$$\begin{aligned}
\lim_{t\to\infty} \boldsymbol{x}(t) &= \left(\boldsymbol{e}^T\mathbf{Y}(0)\right)^{-1}\left(\boldsymbol{e}^T\boldsymbol{y}(0)\right)\boldsymbol{e} \\
&= \left\{\sum_{i=1}^n Y_i(0)\right\}^{-1}\left\{\sum_{i=1}^n y_i(0)\right\}\boldsymbol{e}
\end{aligned}$$

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

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

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

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

Definition 3 [25].
A stochastic matrix A is called indecomposable and aperiodic (SIA) if

$$L = \lim_{m\to\infty} A^m$$

exists and all the rows of L are the same. Define $\delta(A)$ by

$$\delta(A) = \max_j \max_{i,k} |a_{ij} - a_{kj}|$$

Note that if the rows of A are identical, $\delta(A) = 0$, and vice versa.

Definition 4.
Let $A_1, \ldots, A_k \in \mathbb{C}^{n\times n}$. By a *word* in A_is of the length t we mean the product of t A_is with repetition permitted.

Theorem 7 [25].
Let $A_1, \ldots, A_k$ be square row-stochastic matrices of the same order such that any word in the A_is is SIA. For any $\epsilon > 0$ there exists an integer $\nu(\epsilon)$ such that any word B (in the As) of length $m \geqslant \nu(\epsilon)$ satisfies $\delta(B) < \epsilon$.

In other words, the result is that any sufficiently long word in the A_is has all its rows the same or, $\lim_{m\to\infty} A_1 A_2 \ldots A_m = \boldsymbol{e}\boldsymbol{v}^T$.

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 $D = AB$, then

$$\begin{aligned} d_{ij} &= \sum_{k=1}^{n} \underbrace{a_{ik} b_{kj}}_{\geqslant 0} \geqslant 0 \\ d_{ii} &= \sum_{k=1}^{n} a_{ik} b_{ki} = \underbrace{a_{ii} b_{ii}}_{>0} + \sum_{k=1, k\neq i}^{n} \underbrace{a_{ik} b_{ki}}_{\geqslant 0} > 0 \end{aligned}$$

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. ■

Theorem 8.
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 Eqs. (8) and (15), converges to the true centralized estimate, $\bar{x}$, calculated using Eq. (6).

Proof.
From Lemma 3, for any t, $\lim_{m\to\infty} (\Psi^T(t))^m = \boldsymbol{e}\boldsymbol{v}_t^T$, where the $\boldsymbol{v}_t$ is a column vector. Using Eq. (19) and *Lemma 1*, $\Psi^T(t)$ is row stochastic, so for any t, $\Psi^T(t)$ is SIA (see Definition 3). Then from Theorem 7

$$\lim_{t\to\infty} \Psi^T(1)\Psi^T(2)\ldots\Psi^T(t) = \boldsymbol{e}\boldsymbol{v}^T$$

for some $\boldsymbol{v}$, or equivalently,

$$\lim_{t\to\infty} \Psi(t)\Psi(t-1)\ldots\Psi(2)\Psi(1) = \boldsymbol{v}\boldsymbol{e}^T \tag{24}$$

Thus if it can be shown that $\boldsymbol{v} \succ 0$, then the proof of Theorem 8 would follow the same steps as the proof for the time-invariant case in Theorem 6. To demonstrate that $\boldsymbol{v} \succ 0$, we first show that the diagonal elements of

$$L = \lim_{t\to\infty} \Psi^T(1)\Psi^T(2)\ldots\Psi^T(t) \tag{25}$$

are positive ($L_{ii} > 0,\ \forall i$). As, by its definition in Eq. (19), $\Psi(t) \succeq 0$ 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 Eq. (25) have positive elements, $L_{ij} \geqslant 0, \forall i, j$, and strictly positive diagonal elements, $L_{ii} > 0, \forall i$, (see Lemma 4).

Also, as $L = \boldsymbol{e}\boldsymbol{v}^T$ (see Eqs. (24) and (25)), then all of the rows of L are equal ($L_{ji} = L_{ii}, \forall i, j$). Furthermore, since $L_{ii} > 0, \forall i$ then $L_{ji} > 0, \forall i, j$, which implies that $L = \boldsymbol{e}\boldsymbol{v}^T \succ 0$ and that $\boldsymbol{v} \succ 0$. The remainder of the proof then follows the same steps as the proof for the time-invariant case in Theorem 6. ■

C. Convergence Proof for General Network Structure

In this section the UDKC is extended to be applied to the general communication networks, which means the strong connectivity assumption is relaxed and more general assumptions are made.

Assumption 1.
There exists a positive constant α such that:
(a) $a_{ii}(t) \geqslant \alpha, \forall i, t$,
(b) $a_{ij}(t) \in \{0\} \cup [\alpha, 1], \forall i, j, t$,
(c) $\sum_{j=1}^{n} a_{ij}(t) = 1, \forall i, t$.

Assumption 2 (connectivity).
The graph $(N, \bigcup_{s\geqslant t} E(s))$ is strongly connected. This assumption says that the union of the graphs from anytime to infinity is strongly connected, which means that when all the future networks are overlapped, then there is a directed graph from any node to any other node.

Assumption 3 (bounded intercommunication interval).
If i communicates to j an infinite number of times, then there is some B such that, for all t, $(i, j) \in E(t) \cup E(t+1) \cup \cdots \cup E(t+B-1)$.

Theorem 9.
Consider an infinite sequence of stochastic matrices $A(0), A(1), \ldots$, that satisfies Assumptions 1, 2, and 3. There exists a nonnegative vector v such that

$$\lim_{t\to\infty} A(t)A(t-1)A(t-2)\ldots A(1)A(0) = \boldsymbol{e}v^T$$

Proof.
See the proof in [9]. ■

Theorem 10.
Let $\mathbb{G}$ be any dynamic communication network that satisfies Assumptions 2 and 3. 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 UDKC algorithm, introduced in Eqs. (8) and (15), converges to the true centralized estimate, $\bar{x}$, calculated using Eq. (6).

Proof.
By construction $\Psi^T(t)$ has the properties of Assumption 1:
(a) $\psi_{ii}(t) \geqslant \alpha, \forall i, t$.

(b) $\psi_{ij}(t) \in \{0\} \cup [\alpha, 1], \forall i, j, t$.
(c) $\sum_{i=1}^{n} \psi_{ij}(t) = 1, \forall j, t$.

Therefore all the assumption of Theorem 9 are satisfied and therefore

$$\lim_{t\to\infty} \Psi^T(1)\Psi^T(2)\Psi^T(3)\ldots\Psi^T(t) = \boldsymbol{e}v^T$$

and therefore

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

The rest of the proof follows the proof of Theorem 6. ■

V. Conclusions

The performance of the KCA 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 is 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.

Acknowledgment

Research funded by AFOSR Grant no. FA9550-04-1-0458.

References

[1] Alighanbari, M., and How, J., "Decentralized Task Assignment for Unmanned Aerial Vehicles," *Proceedings of the IEEE European Control Conference and Conference on Decision and Control,* Institute of Electrical and Electronics Engineers, New York, NY, 2005, pp. 5668–5673.

[2] Bellingham, J., Tillerson, M., Richards, A., and How, J., "Multi-Task Allocation and Path Planning for Cooperating UAVs," in S. Butenko, R. Murphey, and P. M. Pardalos (editors)—*Cooperative Control: Models, Applications and Algorithms*, Kluwer Academic Publishers, Norwell, MA, 2003, pp. 23–41.

[3] Boskovic, J., Prasanth, R., and Mehra, R., "An Autonomous Hierarchical Control Architecture for Unmanned Aerial Vehicles," *Proceedings of the AIAA Guidance, Navigation, and Control Conference,* AIAA, Reston, VA, 2004, AIAA paper 2002-4468.

[4] Chandler, P., and Pachter, M., "Hierarchical Control for Autonomous Teams," *Proceedings of the AIAA Guidance, Navigation, and Control Conference,* AIAA, Reston, VA, 2001, AIAA paper 2001-4149.

[5] Chandler, P., "Decentralized Control for an Autonomous Team," *Proceedings of the 2nd AIAA Unmanned Unlimited Conference,* AIAA, Reston, VA, 2003, AIAA paper 2003-6571.

[6] Jin, Y., Minai, A., and Polycarpou, M., "Cooperative Real-Time Search and Task Allocation in UAV Teams," *Proceedings of the IEEE Conference on Decision and Control,* Institute of Electrical and Electronics Engineers, New York, NY, 2003, pp. 7–12.

[7] Schumacher, C., and Chandler, P., "Task Allocation for Wide Area Search Munitions," *Proceedings of the IEEE American Control Conference,* Institute of Electrical and Electronics Engineers, New York, NY, 2002, pp. 1917–1922.

[8] Beard, R. W., and Stepanyan, V., "Synchronization of Information in Distributed Multiple Vehicle Coordinated Control," *Proceedings of the IEEE Conference on Decision and Control,* Institute of Electrical and Electronics Engineers, New York, NY, 2003, pp. 2029–2034.

[9] Blondel, V., Hendricks, J., Olshevsky, A., and Tsitsiklis, J., "Convergence in Multiagent Coordination, Consensus, and Flocking," *Proceedings of the IEEE European Control Conference and Conference on Decision and Control,* Institute of Electrical and Electronics Engineers, New York, NY, 2005, pp. 2996–3000.

[10] Jadbabaie, A., Lin, J., and Morse, A. S., "Coordination of Groups of Mobile Autonomous Agents Using Nearest Neighbor Rules," *IEEE Transactions on Automatic Control,* Vol. 48, No. 6, 2003, pp. 988–1001.
doi: 10.1109/TAC.2003.812781

[11] Olfati-Saber, R., and Murray, R. M., "Consensus Problems in Network of Agents With Switching Topology and Time-Delay," *IEEE Transaction on Automatic Control,* Vol. 49, No. 9, 2004, pp. 1520–1533.
doi: 10.1109/TAC.2004.834113

[12] Olfati-Saber, R., “Distributed Kalman Filter with Embedded Consensus Filter,” *Proceedings of the IEEE European Control Conference and Conference on Decision and Control,* Institute of Electrical and Electronics Engineers, New York, NY, 2005, pp. 8179–8184.

[13] Olfati-Saber, R., “Consensus Filters for Sensor Networks and Distributed Sensor Fusion,” *Proceedings of the IEEE European Control Conference and Conference on Decision and Control,* Institute of Electrical and Electronics Engineers, New York, NY, 2005, pp. 6698–6703.

[14] Ren, W., and Beard, R., “Consensus of Information Under Dynamically Changing Interaction Topologies,” *Proceedings of the IEEE American Control Conference,* Institute of Electrical and Electronics Engineers, New York, NY, 2004, pp. 4939–4944.

[15] Ren, W., Beard, R., and Kingston, D., “Multi-Agent Kalman Consensus with Relative Uncertainty,” *Proceedings of the IEEE American Control Conference,* Institute of Electrical and Electronics Engineers, New York, NY, 2005, pp. 1865–1870.

[16] Ren, W., Beard, R., and Atkins, E., “A Survey of Consensus Problems in Multi-Agent Coordination,” *Proceedings of the IEEE American Control Conference,* Institute of Electrical and Electronics Engineers, New York, NY, 2005, pp. 1859–1864.

[17] Gibbons, A., *Algorithmic Graph Theory*, Cambridge Univ. Press, New York/London/Cambridge, England, UK, 1985.

[18] Olfati-Saber, R., Fax, J. A., and Murray, R. M., “Consensus and Cooperation in Networked Multi-Agent Systems,” *Proceedings of the IEEE,* Vol. 95, 2007, pp. 215–233.

[19] Minkler, G., *Theory and Applications of Kalman Filtering*, Magellan Book Company, Baltimore, MD, 1990.

[20] Kim, Y., and Hong, K., “Decentralized Information Filter in Federated Form,” *Proceedings of the IEEE SICE 2003 Annual Conference,* Institute of Electrical and Electronics Engineers, New York, NY, 2003, pp. 2176–2181.

[21] Kim, Y., and Hong, K., “Decentralized Sigma-Point Information Filters for Target Tracking in Collaborative Sensor Networks,” *IEEE Transactions on Acoustics, Speech, and Signal Processing,* Vol. 53, 2005, pp. 2997–3009.

[22] Seneta, E., *Non-negative Matrices*, Wiley, New York, NY, 1973.

[23] Varga, R. S., *Geršgorin and His Circles*, Springer-Verlag, Berlin/New York/Heidelberg, 2004.

[24] Horn, R. A., and Johnson, C. R., Matrix Analysis, Cambridge Univ. Press, New York/London/Cambridge, England, UK, 1985.

[25] Wolfowitz, J., “Product of Indecomposable, Aperiodic Stochastic Matrices,” *Proceedings of the American Mathematical Society,* Vol. 15, 1963, pp. 733–736.

Roy Sterritt
Associate Editor