

DISTRIBUTED TRACKING IN MULTI-HOP NETWORKS

Murat Alanyali and Venkatesh Saligrama

Department of Electrical and Computer Engineering
Boston University, Boston MA 02215
Email: {alanyali,srv}@bu.edu

ABSTRACT

We describe distributed tracking of a linear dynamical system via networked sensors. The networked sensors communicate with each other by means of a multi-hop protocol over a communication network. MMSE-optimal solution is Kalman filtering when measurements are available centrally, but new methods are required to account for communication constraints. We derive optimal algorithms to deal with arbitrary network topology and then extend these results to account for communication delays and packet losses. The proposed techniques differ from related techniques proposed in two important aspects: a) there is no designated leader/fusion node and each sensor attempts to optimally track the system trajectory based on its local observations and time-dependent information available from other sensors in the network; b) the message computation at each sensor is structurally identical, where the computed message from each sensor is the innovation in the state conditioned on all the information available upto that time at each sensor. Consequently, the sensor network can be queried at any time and at any node to obtain optimal estimates for the state of the dynamical system.

1. INTRODUCTION

We consider a collection of sensors that collectively track a linear system that is driven by noise. The measurement model is also linear, hence the optimal MMSE centralized solution is given by the canonical Kalman filter. Here we consider decentralized tracking subject to a multi-hop communication model among the

sensors. A special case of this scenario was studied in [2, 4] for a completely connected topology with no communication delay. They show that the centralized estimate can be constructed at a designated node if nodes send their local estimates along with a correction term that can be computed locally. Recently, we have been made aware of the work in [6], which describes approximate distributed Kalman filtering for arbitrary networks with no communication delays. However, the underlying philosophy of our scheme deviates significantly from the existing schemes:

(a) similar to [6] there is no designated leader/fusion node and the idea is to reach a consensus at each sensor node;

(b) However, [6] uses message passing to compute the least squares solution of the current state based only on current data and does not incorporate correlations of the current state with the past data. This is suboptimal from an energy efficient perspective. We build on [2, 4] and incorporate correlations in our message passing scheme for arbitrary networks. Our scheme has the pleasing feature of *transmitting only the local innovations*.

(c) [6] is suboptimal in the information sense, i.e., computed estimate at any stage is not necessarily equal to the centralized estimate.

Large collections of sensors call for multi-hop communications, whose implications need to be clarified as to what should be sent on each link towards a designated sensor. Furthermore, if communication delays are significant then the optimal centralized estimate that respects such delays is different from sensor to sensor. We derive optimal algorithms to deal with arbitrary network topologies, wherein at each stage a sensor computes the optimal estimate conditioned on locally available information and the new information

This work was supported by PECASE grant no. N00014-02-100362, NSF grants CCF-0430983, CNS-0435353, ANI-0238397.

in the form of innovations received from adjacent sensors. We show that when no delays are present this scheme converges to the centralized solution. We next extend this scheme to account for packet losses. Finally, we derive distributed algorithms to deal with communication delays for some special network topologies. With delays consensus is not the correct notion (since sensors with superior SNRs will always have better estimates). Consequently, the best one can hope for is to have each sensor attempt to optimally track the system trajectory based on its local observations and time-dependent information available from other sensors in the network. Consequently, the sensor network can be queried at any time and at any node to obtain optimal estimates for the state of the dynamical system.

2. PROBLEM STATEMENT

Let \mathbf{R}^n denote n -dimensional real vectors, and denote by $\mathbf{M}^{n \times n}$ the set of symmetric positive-definite matrices of dimension $n \times n$. We consider the discrete-time system

$$X_{t+1} = AX_t + W_t, \quad X_0 \sim N(0, \Sigma_0),$$

where A is a stable $n \times n$ matrix and $W = (W_t : t = 0, 1, 2, \dots)$ is an IID sequence such that $W_t \sim N(0, \Sigma_W)$, independently of X_0 .

We consider tracking the sequence $(X_t : t = 0, 1, 2, \dots)$ based on measurements taken by a collection V of sensors. The measurement of sensor $v \in V$ taken at time slot t is denoted by $Y_t(v) \in \mathbf{R}^m$ and it satisfies

$$Y_t(v) = C_t(v)X_t + U_t(v), \quad v \in V,$$

where $C_t(v)$ is an $m \times n$ matrix, and $(U_t(v) : t = 0, 1, 2, \dots)$ is an IID sequence such that $U_t(v) \sim N(0, \Sigma_U)$. In particular if all measurements are immediately available to a central processor then the MMSE estimator is a Kalman filter. Specifically, the MMSE estimate $X_{t|t} = E[X_t | Y_\tau(v) : v \in V, \tau \leq t]$ of X_t based on $Y_\tau(v)$, $v \in V, \tau \leq t$, satisfies

$$X_{t|t} = X_{t|t-1} + P_{t|t} \sum_{v \in V} C_t^T(v) \Sigma_U^{-1} (Y_t(v) - C_t(v) X_{t|t-1}) \quad (1)$$

where $P_{t|t} = E(X_t - X_{t|t})(X_t - X_{t|t})^T$ is the conditional error covariance matrix at time t and is given by the recursion:

$$P_{t|t}^{-1} = P_{t|t-1}^{-1} + \sum_{v \in V} C_t^T(v) \Sigma_U^{-1} C_t(v)$$

These steps are commonly referred to as the measurement update steps. To complete the Kalman filter updates we require the so called prediction steps, which are given by:

$$X_{t|t} = AX_{t|t-1}, \quad P_{t+1|t} = AP_{t|t}A^T + \Sigma_W \quad (2)$$

Our objective in this paper is to address the tracking problem when sensors are networked through a communication infrastructure represented by a directed graph $G = (V, E)$, in which each edge $(v, v') \in E$ indicates a directed communication link from sensor v to sensor v' . Let $F = [I\{(v, v') \in E\}]_{V \times V}$ be the connectivity matrix of G . We will assume that G is strongly connected and will maintain a self-loop $(v, v) \in E$ at each sensor $v \in V$, so that F is indecomposable and aperiodic.

3. FAST NETWORK/SLOW SENSING

Here we assume that the network is significantly faster than the dynamics of the system. In this case we can assume that the system time is frozen and the network has infinite time to share the data. We analyze the following message passing scheme in this context. Each sensor updates its state based on past information and the new message received from its neighboring sensors. It then sends out the innovation in the state information to its neighbors. In addition the updated error covariance matrix is also transmitted. The latter is particularly important, whenever the global observation model is not locally available at each sensor.

The general idea for fusion can be described by considering the two sensor case. Consider a local sensor that is required to transmit its information to the fusion sensor. The state update for the centralized kalman filter is given by Equation 1. For the local sensor, the state update, $X_{t|t}^l$ is given by,

$$X_{t|t}^l = E(X_t | Y_l(\tau), \tau \leq t) = X_{t|t-1}^l + P_{t|t}^l C_t^T(l) \Sigma_U^{-1} Z_l(t)$$

where, $Y_l(\tau)$, $P_{t|t}^l$, $Z_l(t) = Y_l(t) - C_t(l)X_{t|t-1}^l$, $C_t(l)$ are the local sensor data, local error covariance, local

innovations and the local observation matrix respectively. It follows that,

$$\begin{aligned} C_t^T(l)\Sigma_U^{-1}Y_l(t) &= (P_{t|t}^l)^{-1}(X_{t|t}^l \\ &\quad - (I - P_{t|t}^l C_t^T(l)\Sigma_U^{-1}C_t(l))X_{t|t-1}^l) \end{aligned}$$

Let $Y_f(\tau)$ be the data available at the fusion sensor. The centralized update can be simplified as follows:

$$\begin{aligned} X_{t|t} &= X_{t|t-1} + P_{t|t}C_t^T(f)\Sigma_U^{-1}(Y_f(t) - C_t(f)X_{t|t-1}) \\ &\quad + P_{t|t}C_t^T(l)\Sigma_U^{-1}(Y_l(t) - C_t(l)X_{t|t-1}) \\ &= X_{t|t-1} + P_{t|t}C_t^T(f)\Sigma_U^{-1}(Y_f(t) - C_t(f)X_{t|t-1}) \\ &\quad + P_{t|t}(P_{t|t}^l)^{-1}(X_{t|t}^l - X_{t|t-1}^l) \\ &\quad - P_{t|t}C_t^T(l)\Sigma_U^{-1}C_t(l)(X_{t|t-1} - X_{t|t-1}^l) \end{aligned}$$

Therefore, the centralized Kalman filter can be written in terms of the local Kalman filter estimates. This is the key insight we employ in generalizing the scheme to general graphs. We state this fact as a lemma next:

Lemma 3.1 *Suppose, $(X_{t|t}^1, P_{t|t}^1)$ $(X_{t|t}^2, P_{t|t}^2)$ are two local sensor state and error covariance streams respectively and $P_{t|t}$ is the centralized error covariance, it follows that the fused sensor update is given by:*

$$\begin{aligned} X_{t|t} &= K_0 X_{t|t-1} + P_{t|t} \sum_{j=1}^2 (P_{t|t}^j)^{-1} (X_{t|t}^j - X_{t|t-1}^j) \\ &\quad - P_{t|t} \sum_{j=1}^2 C_t^T(j)\Sigma_U^{-1}C_t(j)(X_{t|t-1} - X_{t|t-1}^j) \quad (3) \end{aligned}$$

To describe the general case let $N(v)$ denote the set of sensors such that there exists a directed(one-hop communication) edge from v to each $v' \in N(v)$ in G . Suppose for the moment that the first t updates of (1) have already been executed and the network has achieved a consensus for state estimate $X_{t-1|t-1}$, $X_{t|t-1}$ and error covariance $P_{t-1|t-1}$, $P_{t|t-1}$. This implies that the last term in Equation 3 is identically zero. Therefore, only the differences between measured and the predicted, i.e., $(X_{t|t}^j - X_{t|t-1}^j)$ needs to be communicated. To this end, let $m_k(v)$ be the k th state message update and $\Delta P_{t|t}^v(k)$ be the error covariance update, sent by sensor v to each neighboring sensor in $N(v)$. The updates for error covariance transmitted, $\Delta P_{t|t}^v(k)$ and

the local error covariance update, $P_{t|t}^v(k)$ are:

$$\begin{aligned} \Delta P_v(k+1) &= \frac{1}{|N(v')|} \sum_{u \in N(v)} \Delta P_u(k) \\ \Delta P_{t|t}^v(0) &= C_t(v)\Sigma_U^{-1}C_t^T(v) \\ \left(P_{t|t}^v(k+1)\right)^{-1} &= P_{t|t-1}^{-1} + |V|\Delta P_v(k+1) \end{aligned}$$

The message update transmitted and the local state update is given by:

$$\begin{aligned} m_{k+1}(v) &= \frac{1}{|N(v')|} \sum_{u \in N(v)} P_{t|t}^v(k+1) \left(P_{t|t}^u(k)\right)^{-1} m_k(u) \\ m_0(v) &= (X_{t|t}^v(0) - X_{t|t-1}) \\ &= P_{t|t}^v(0)C_t(v)\Sigma_U^{-1}(Y_t(v) - C_t(v)X_{t|t-1}) \end{aligned}$$

$$X_{t|t}^v(k+1) = (I + |V|\Delta P_v(k+1))X_{t|t-1} + |V|m_{k+1}(v)$$

Theorem 3.1 *Consider the decentralized updating scheme presented above. Suppose, $X_{t|t}$, $P_{t|t}$ is the centralized state estimate and the error covariance matrix respectively conditioned on sensor data upto time t for all of the sensors. It follows that,*

$$\lim_{k \rightarrow \infty} X_{t|t}^v(k) = X_{t|t}, \quad \lim_{k \rightarrow \infty} P_{t|t}^v(k) = P_{t|t} \quad \forall v \in V$$

4. PACKET LOSSES

Our aim in this section is to account for the following two effects: First, messages may be corrupted and lost due to imperfections in point-to-point communication. Although link layer protocols would provide some relief against this issue, robustness of network operation against message losses needs to be addressed, especially if the physical communication medium is wireless. Secondly, one can imagine situations where some sensors operate on a slower time-scale than others, thereby slowing down the network under the lock-step message-passing algorithm outlined above. This limitation may be overcome if each sensor contributes to the collaborative effort at its own time-scale. In both cases described above the network operation is asynchronous in the sense that not all links are necessarily active at each round of the algorithm.

We consider next the message passing algorithm above in the case when communication links are imperfect in that a transmitted message can be lost. The evolution of messages can then be represented as

$$m_{k+1} = Q_k Q_{k-1} \cdots Q_0 m_0, \quad (4)$$

where $Q_k = (I + F_k)D_k$ such that $F_k = [f_{ij}(k)]$ is a binary matrix and $D_k = [d_{ij}(k)]$ is a diagonal matrix with

$$d_{jj}(k) = \left(1 + \sum_i f_{ij}(k)\right)^{-1},$$

so that in particular columns of Q_k are probability vectors. Given sensors i, j we shall say that link $i \rightarrow j$ is *functional* in round k if sensor j receives a message from sensor i in that round. Entries of F_k are then interpreted as

$$f_{ij}(k) = I\{\text{link } j \rightarrow i \text{ is functional at round } k\}.$$

Hence the system (4) describes the evolution of local messages when each transmitted message is normalized by the number of outgoing functional links (i.e., the number of receivers of the message) in the same round.

Theorem 4.1 *Suppose that the matrices $(F_k : k \geq 1)$ are IID, and that $E[F_1]$ is irreducible. Then for $v \in V$ a consensus to the centralized Kalman Filter state estimate and the corresponding error covariance is achieved by the distributed tracking algorithm described above.*

5. COMMUNICATION DELAYS

Our next task is to deal with communication delays. We first focus on a completely connected network for simplicity of exposition and describe implementation on a general network later in the section. In this setup we assume that a transmitted estimate arrives at the destination nodes with a delay equal to unit time. Thus, if a message is transmitted at time t , its reception is completed by another sensor at time $t + 1$. We shall denote the message transmitted by sensor v at time t by $m_t(v)$, and define it as an encoding of the optimal prediction of the state at time t based on data available

at the originating sensor subject to communication delays. Namely,

$$m_t(v) = (\xi_t(v), \tilde{P}_t(v))$$

where

$$\xi_t(v) = \mu_t(v) - X_{t|t-1} \quad (5)$$

$$\mu_t(v) = E\left(X(k) \mid Y^k(v), Y^{k-1}(-v)\right) \quad (6)$$

$$\tilde{P}_t(v) = E\left((X_t - \mu_t(v))(X_t - \mu_t(v))^T\right)$$

with $Y^k(v) = (Y_\tau(v) : \tau \leq k)$ and $Y^{k-1}(-v) = (Y_\tau(u) : \tau < k, u \neq v)$.

In a completely connected network this message is simultaneously received by all sensors at time $t + 1$. The same message transmission scheme is adopted by all sensors, though it is clear that messages differ from sensor to sensor since they are adapted to different filtrations. At time $t + 1$ each sensor v constructs its next message, $m_{t+1}(v)$, based on the received messages $m_t(u)$, $u \neq v$, as well as the local measurement $Y_{t+1}(v)$ taken after the last transmitted message. We specify this construction as follows. Note that

$$\mu_{t+1}(v) = X_{t+1|t} + \tilde{K}_{t+1}(v) (Y_{t+1}(v) - C_{t+1}(v)X_{t+1|t}) \quad (7)$$

where $\tilde{K}_{t+1}(v) = \tilde{P}_{t+1}(v)C_{t+1}(v)^T \Sigma_U^{-1}$. The conditional error covariance $\tilde{P}_{t+1}(v)$ is given by

$$\tilde{P}_{t+1}^{-1}(v) = P_{t+1|t}^{-1} + C_{t+1}(v)^T \Sigma_U^{-1} C_{t+1}(v), \quad (8)$$

where $P_{t+1|t}$ can be constructed from $P_{t|t-1}$ and received information $\tilde{P}_t(u)$, $u \neq v$, via the recursion (2) and the representation

$$P_{t|t} = \left(\sum_{u \in V} \tilde{P}_t^{-1}(u) - (V-1)P_{t|t-1}^{-1} \right)^{-1}. \quad (9)$$

Note that the last term in (8) is local information at sensor v . Manipulation of equality (1), as outlined in Section 3, yields that

$$X_{t|t} = X_{t|t-1} + P_{t|t} \sum_{u \in V} \tilde{P}_t^{-1}(u) \xi_t(u); \quad (10)$$

therefore $X_{t+1|t}$ can be computed based on $X_{t|t-1}$ and $m_t(u)$, $u \neq v$ via (2), in turn $\mu_{t+1}(v)$ is calculated via (7). We collect these observations in the following theorem.

Theorem 5.1 *When initiated with*

$$\begin{aligned} X_{0|0} &= 0 & P_{0|-1} &= \Sigma_0 \\ \xi_0(v) &= 0 & \tilde{P}_0^{-1}(v) &= 0, \quad v \in V, \end{aligned}$$

the algorithm described above maintains equality (6) at each step k of the recursion (7).

Remark 5.1 *The pair $(X_{t|t-1}, P_{t|t-1})$ is an internal state for the sensor at time t : $m_t(v)$ is determined by $(X_{t|t-1}, P_{t|t-1})$, together with the local information $Y_t(v)$ and $C_t(v)^T \Sigma_U^{-1} C_t(v)$, whereas $(X_{t+1|t}, P_{t+1|t})$ is determined by new information obtained at the end of time t .*

Remark 5.2 *The algorithm guarantees that each sensor v dynamically constructs $\mu_t(v)$, i.e. the optimal estimate of X_t that can be obtained subject to the delay constraints imposed by the network. In general, the error covariances $\tilde{P}_t(v)$ associated with these estimates are different due to non-identical observation models. These covariance matrices are also constructed locally; in turn the network can be queried to identify a sensor with a highest quality estimate.*

Remark 5.3 *One can imagine other message passing algorithms that lead to the same conclusions obtained above. In particular, with appropriate modification of the processing of messages, Theorem 5.1 holds true if $\tilde{P}_t(v)$ is replaced by $C_t(v)^T \Sigma_U^{-1} C_t(v)$ in each message $m_t(v)$. Depending on the dimensions of the state and the observation variables one algorithm may be more favorable than the other, particularly for applications in which link capacities are limited.*

6. MULTI-HOP TOPOLOGIES

In this section we relax the assumption of the completely connected network topology, but assume that each link in the network can send $K \geq 1$ messages between two consecutive measurements by sensors. If K is large than flooding the network with messages that are defined in the previous section might be a viable solution to arrive at a network-wide consensus, however the present focus is on efficient message passing algorithms that would work with the smallest possible value of K . The diameter of the network graph is a trivial lower bound for such K , and we give here an

algorithm for a linear topology that is compatible with this lower bound. We give an extension of the algorithm can to a mesh topology.

We continue to indicate the measurement time-scale by the letter t and the messaging time-scale by the letter k . In this respect, the time slot between consecutive measurements is assumed to be divided into D sub-slots, and $m_{t,k}(v)$ denotes the message transmitted by sensor v in the k th sub-slot of slot t .

Consider first the linear network topology depicted by Figure 1, and assume that $(X_{t|t-1}, P_{t|t-1})$ is known by the sensors at time t . We give next an algorithm that brings all sensors to an information state that allows computation of $(X_{t+1|t}, P_{t+1|t})$ after V messaging rounds. This goal is achieved via two types of messages, namely $m_{t,k}^+(v)$ and $m_{t,k}^-(v)$, with the understanding that messages with superscript $+$ (resp. $-$) proceed from left to right (resp. right to left). Specifically, we define the messages $m_{t+1,k}^\mp(v)$ as follows:

$$m_{t+1,k}^\mp(v) = (\xi_{t+1,k}^\mp(v), \tilde{P}_{t+1,k}^\mp(v))$$

where

$$\begin{aligned} \tilde{P}_{t+1,k+1}^\mp(v) &= \tilde{P}_{t+1,k}^\mp(v \pm 1) + \tilde{P}_t^{-1}(v) \\ \xi_{t+1,k+1}^\mp(v) &= \xi_{t+1,k}^\mp(v \pm 1) + \tilde{P}_t^{-1}(v) \xi_t(v), \end{aligned}$$

with the understanding that both $\xi_{t+1,k+1}^\mp(v) = 0$ and $\tilde{P}_{t+1,k+1}^\mp(v) = 0$ for $v = -1, V + 1$.

Let $Q_{t+1,k}(v)$ be defined via the equality

$$\begin{aligned} Q_{t+1,k}^{-1}(v) &= \left(\tilde{P}_{t+1,k}^+(v-1) + \tilde{P}_{t+1,k}^-(v+1) \right)^{-1} \\ &\quad - (V-1) P_{t|t-1}^{-1}. \end{aligned}$$

It can be verified by inspection that $Q_{t+1,k}(v) = P_{t|t}$ for $k \geq \max\{v, V-v\}$; in turn the following theorem holds.

Theorem 6.1 *For $k \geq V$*

$$\begin{aligned} P_{t+1|t} &= A Q_{t+1,k}(v) A^T + \Sigma_W \\ X_{t|t} &= X_{t|t-1} \\ &\quad + Q_{t+1,k}(v) \left(\xi_{t+1,k}^+(v-1) + \xi_{t+1,k}^-(v+1) \right) \end{aligned}$$

for all $v \in V$.

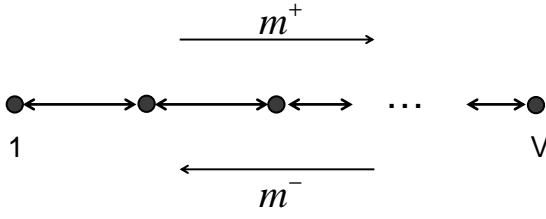


Fig. 1. Message passing on a linear topology. Messages with superscript + (resp. -) are sent to node with larger (resp. smaller) index, and are constructed as indicated in the text.

The methodology can be extended to the mesh topology of Figure 2 by defining 8 message types

$$m_{t+1,k}^D(v) = (\xi_{t+1,k}^D(v), \tilde{P}_{t+1,k}^D(v)),$$

$D \in \{E, W, S, N, NE, NW, SE, SW\}$, in the following fashion: For $D \in \{NE, NW, SE, SW\}$ let $n_D(v)$ denote the neighbor of sensor v in the exact opposite direction indicated by D (e.g. $n_{NE}(v)$ is the south-west neighbor, etc.), and let

$$m_{t+1,k+1}^D(v) = m_{t+1,k}^D(n_D(v)) + \left(\tilde{P}_t^{-1}(v) \xi_t(v), \tilde{P}_t^{-1}(v) \right)$$

for such D . For $D \in \{E, W, S, N\}$ let $n_D(v)$ denote the neighbors of v that are in the three directions that are opposite D (e.g. $n_E(v)$ consists of the west, the north-west, and the south-west neighbors of v), and let

$$m_{t+1,k+1}^D(v) = \sum_{u \in n_D(v)} m_{t+1,k}^D(u) + \left(\tilde{P}_t^{-1}(v) \xi_t(v), \tilde{P}_t^{-1}(v) \right).$$

All summations above are understood to be component-wise.

Theorem 6.2 Let $o(u, v)$ denote the orientation of sensor v with respect to its neighbor u . If k is larger than the diameter of the network then

$$P_{t|t}^{-1} = \left(\sum_{u \in N(v)} \tilde{P}_{t+1,k}^{o(u,v)}(u) \right)^{-1} - (V-1)P_{t|t-1}^{-1}$$

$$X_{t|t} = X_{t|t-1} + P_{t|t} \sum_{u \in N(v)} \xi_{t+1,k}^{o(u,v)}(u)$$

for all $v \in V$.

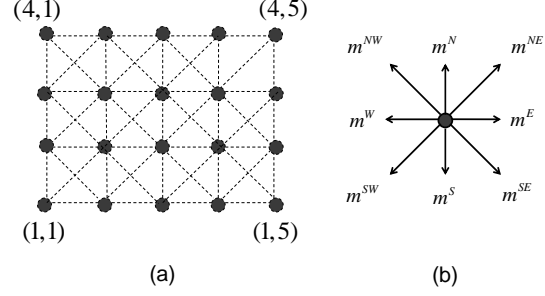


Fig. 2. a) A mesh topology for the communication graph. b) Messages sent by each (interior) node.

7. REFERENCES

- [1] R.A Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.
- [2] J. Speyer, "Computation and transmission requirements for a decentralized linear-quadratic-Gaussian control problem," *IEEE Transactions on Automatic Control*, vol. 24, no. 2, pp. 266–269, 1979.
- [3] V. Saligrama, M. Alanyali, and O. Savas, "Asynchronous Distributed Detection in Sensor Networks," (BU Technical Report 2005-03), submitted for publication to *IEEE Transactions on Signal Processing*, 2005.
- [4] A. Willsky, M. G. Bello, D. A. Castanon, B. C. Levy, G. C. Verghese, "Combining and Updating of Local Estimates and Regional Maps Along Sets of One Dimensional Tracks," *IEEE Transactions on Automatic Control*, vol. 24, no. 4, pp. 799–813, 1982.
- [5] J. Wolfowitz, "Products of indecomposable, aperiodic, stochastic matrices," *Proceedings of the American Mathematical Society*, vol. 14, pp. 733–737, 1963.
- [6] D. Spanos, R. Olfati-Saber, R. Murray, "Distributed Kalman Filtering in Sensor Networks with Quantifiable Performance," In *Proceedings of Information Processing in Sensor Networks*, April 2005