SCALABLE DISTRIBUTED KALMAN FILTERING THROUGH CONSENSUS

Shrut Kirti and Anna Scaglione

School of Electrical and Computer Engineering
Cornell University, Ithaca, NY 14853
Email: sk482@cornell.edu, anna@ece.cornell.edu

ABSTRACT

Kalman filtering is a classical problem of significant interest in the context of a distributed application for wireless sensor networks. In this paper we consider a specific algorithm for distributed Kalman filtering proposed recently by Olfati-Saber [1] and present a scalable wireless communication architecture suited for implementation in sensor networks. The proposed architecture uses a data driven average consensus framework. This allows us to explicitly characterize the delay vs. estimate accuracy tradeoff in filtering. By exploiting the structure of the distributed filtering computations, we derive an optimal communication resource allocation policy for minimizing the component-wise state estimation error. Furthermore, our architecture is scalable in terms of the network size n. We provide simulation results demonstrating the performance of our architecture.

***Index Terms*—** Kalman filtering, distributed algorithms, average consensus

1. INTRODUCTION

A fundamental problem in sensor networks is distributed detection and estimation. A practical solution to this can lead to significant applications in areas such as the distributed monitoring and control of dynamical systems. One of the most computationally efficient and mathematically elegant algorithms for the state estimation of dynamical systems, in a centralized setting, is the Kalman filter. There are several works in literature that propose decentralized versions of the Kalman filter [2, 3, 4, 5].

Recently, [1] proposed a promising algorithm for distributed Kalman filtering (DKF) using average consensus. Specifically, it showed that the information form of the classical centralized Kalman filter can be decomposed into a totally distributed version in which each node i in the sensor network can compute the complete system state, $\mathbf{x}(k)$, using local system observations $\mathbf{z}_i(k)$ and low cost near-neighbor communication. Furthermore, when near-neighbor communication occurs with infinite precision, the local state estimates $\hat{\mathbf{x}}(k)$ obtained using the DKF algorithm are the same as the state estimate $\hat{\mathbf{x}}(k)$ obtained using the centralized Kalman filter.

1.1. Motivation

Unfortunately, the communication cost of the average consensus algorithm does not scale well with the network size n when point-to-point wireless communication is used because of transmission scheduling related congestion. This is particularly so for networks with high connectivity (Fig. 1).

Consider, for instance, random access scheduling which is time slotted for simplicity. Each node i independently transmits its packet with probability $p = 0.1$ in any slot. A node is deemed active in slot l if it transmits in that slot. Neighbor j successfully receives a packet from active node i if-

$$SINR_{ij} = \frac{l(d_{ij})|h_{ij}[l]|^2}{\frac{N_o}{P_t} + \sum_{k \in \mathcal{M}_l \setminus \{i\}} l(d_{ik})|h_{ik}[l]|^2} \cdot b_j[l] > \tau \quad (1)$$

where P_t is finite transmit power, $l(d_{ij})$ is the path loss between node pair (i, j), $h_{ij}[l]$ is the Rayleigh fading channel in slot l, N_o is AWGN channel noise variance such that SNR = $15dB$, $\mathcal{M}_l$ is the set of all nodes active in slot l, and τ is some threshold. $b_j[l]$ is an indicator variable which is 1 if node j is active in slot l and 0 otherwise. This accounts for the half-duplex constraint.

Toss n nodes randomly in a unit square. Nodes i and j are neighbors if $d_{ij} < r$ where r is the radius of connectivity. Ignoring the overhead of acknowledgment packets, we determine the number of transmission slots required for each node to successfully receive packets from all its neighbors. This constitutes one iteration of the average consensus algorithm. Fig. 1 shows that the communication cost in terms of the required slots per consensus iteration increases dramatically with both r and n. Note that r influences the algebraic connectivity of the network and there is a tradeoff between the connectivity and the average number of consensus iterations required for convergence of the algorithm [6].

1.2. Contribution

In this paper we propose a data driven consensus based physical layer cooperative wireless communication architecture for DKF whose communication cost is scalable. Each data driven DKF iteration takes a fixed $|\mathcal{Q}|\mathcal{B}^{-1}$ *time* to complete regardless of the network size. $|\mathcal{Q}|$ is the cardinality of the set of quantizer centroids indicating the precision with which nodes communicate and $\mathcal{B}$ is the bandwidth of the node transmissions. It provides an intuitive explanation of the tradeoff between the accuracy achieved in the state estimate and the communication cost of achieving it. We also provide a communication resource allocation policy for our filtering architecture.

The paper is organized as follows. Section 2 describes the system model and assumptions. Section 3 summarizes the DKF decomposition and delineates the data driven consensus architecture for it. In Section 4 we derive a bandwidth allocation policy to minimize the error in the state estimate. Finally, we present simulation results in Section 5 showing the performance of our architecture and resource allocation policy.

This work is supported in part by the grant NSF CNS 05-24695.

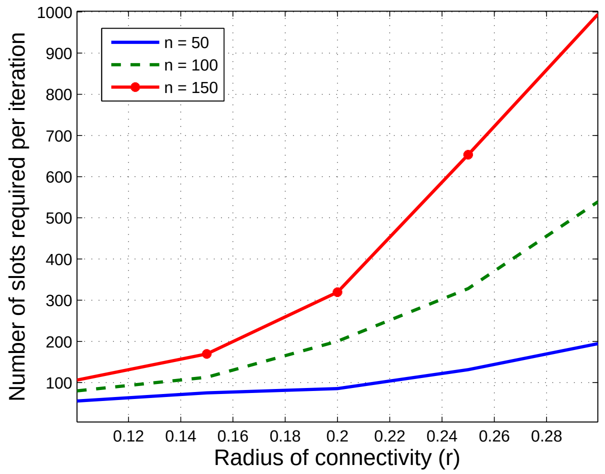

Fig. 1. **Network congestion** increases rapidly with the network size n and radius of connectivity r when average consensus is implemented using point-to-point wireless communication in a slotted ALOHA-like scheme.

2. SYSTEM MODEL

Consider a linear dynamical system with the following state-space model-

$$\mathbf{x}_{k+1} = A_k\mathbf{x}_k + B_k\mathbf{w}_k \tag{2}$$

where $\mathbf{x}_k \in \mathbf{R}^m$ is the system state at time step k, $\mathbf{w}_k$ is Gaussian noise such that $\mathbb{E}\{\mathbf{w}_k\mathbf{w}_l^T\} = Q_k\delta[k-l]$, and A_k, B_k are known $\mathbf{R}^{m\times m}$ matrices. We assume that A_k is such that the norm of $\mathbf{x}_k$ remains within a pre-specified range $\forall k$. The problem objective is that each node i in a network of n nodes should distributedly estimate the state $\hat{\mathbf{x}}_k$.

Each node i makes noisy observations of the state- $\mathbf{z}_k^i = H_k^i\mathbf{x}_k + \mathbf{v}_k^i$, where H_k^i is a known $\mathbf{R}^{p\times m}$ matrix, v_k^i is Gaussian noise with $\mathbb{E}\{\mathbf{v}_k^i\mathbf{v}_l^{iT}\} = R_k^i\delta[k-l]$ and $\mathbb{E}\{\mathbf{v}_k^i\mathbf{v}_k^{jT}\} = R_k^i\delta[i-j]$, and $\mathbf{z}_k^i \in \mathbf{R}^p$, $p \leq m$.

The wireless channel model for inter-node communications assumes the following: $(a1)$ the channel is broadcast with flat Rayleigh fading. The channel gains for any node pair i, j at time slot l are circularly symmetric samples $h_{ij}[l] \overset{iid}{\sim} \mathcal{CN}(0, \sigma_{ij}^2)$ where $\sigma_{ij}^2 = \mathcal{K}(d^* + d_{ij})^{-\alpha}$, α is the path loss exponent, d_{ij} is the distance between the nodes, and d^* and $\mathcal{K}$, related by $\mathcal{K} = (\frac{1}{d^*})^{-\alpha}$, are modeling parameters that take into account the carrier frequency, the scattering environment and antennae gains. $(a2)$ The additive noise at node i is $t_i[l] \overset{iid}{\sim} \mathcal{CN}(0, N_0)$. $(a3)$ The nodes are half-duplex. Denoting the samples of the discrete-time complex base-band equivalent transmit and receive signal models of node i by $s_i[l]$ and $r_i[l]$ respectively we have-

$$r_i[l] = \begin{cases} \sum_{j=1}^n h_{ij}[l]s_j[l] + t_i[l], & \text{if } s_i[l] = 0; \\ 0, & \text{if } s_i[l] \neq 0. \end{cases} \tag{3}$$

Finally, $(a4)$ we assume that the nodes are synchronized and (a5) all nodes have a fixed transmit power constraint P_t.

3. SCALABLE ARCHITECTURE FOR DKF

First we connect distributed Kalman filtering and average consensus. Then we outline the physical layer architecture for filtering.

3.1. Distributed Kalman Filter

Begin by considering the standard recursions of a centralized Kalman filter in the information form [7]. These recursions can be decomposed into a distributed form leading to the following theorem which gives the DKF algorithm-

Theorem 3.1 *[1] Assume that each node i solves two average consensus problems - $S_k = \frac{1}{n}\sum_{i=1}^n (H_k^i)^T(R_k^i)^{-1}H_k^i$ and $\mathbf{y}_k = \frac{1}{n}\sum_{i=1}^n \mathbf{y}_k^i = \frac{1}{n}\sum_{i=1}^n (H_k^i)^T(R_k^i)^{-1}\mathbf{z}_k^i$ - then each node can update its state estimate $\hat{\mathbf{x}}_{k|k}$ by computing the following micro Kalman filters (μ-KF) iterations locally-*

$$\begin{aligned} M_k &= ((nP_{k|k-1})^{-1} + S_k)^{-1} && (4)\\ \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + M_k(\mathbf{y}_k - S_k\hat{\mathbf{x}}_{k|k-1}) && (5)\\ P_{k+1|k} &= A_kM_kA_k^T + B_k(nQ_k)B_k^T && (6)\\ \hat{\mathbf{x}}_{k+1|k} &= A_k\hat{\mathbf{x}}_{k|k} && (7)\end{aligned}$$

where $\hat{\mathbf{x}}_{1|0} = \mathbf{x}_0$ and $P_0 = I_m$.

While the decomposition in Theorem 3.1 is exact, the accuracy of its estimates is contingent on the precision with which the consensus terms $\mathbf{y}_k$ and S_k are computed.

3.2. Data Driven Average Consensus

For simplicity, we consider the case where H_k^i and R_k^i are time invariant. Then, $S = \frac{1}{n}\sum_{i=1}^n (H^i)^T(R^i)^{-1}H^i$ needs to computed only once at the start of the filtering. S can be computed using average consensus. Initialize node states as $\boldsymbol{\theta}_i(0) = S_i$. Then the iterations-

$$\boldsymbol{\theta}_i(k+1) = \boldsymbol{\theta}_i(k) + \epsilon\mathbf{u}_i(k), \tag{8}$$

$$\mathbf{u}_i(k) = \sum_{j\in\mathcal{N}_i} a_{ij}(\boldsymbol{\theta}_j(k) - \boldsymbol{\theta}_i(k)) \tag{9}$$

ensure that $\lim_{k\to\infty}\boldsymbol{\theta}_i(k) \to S, \ \forall i$ [6], where ϵ is a step size parameter, a_{ij} are entries of the network adjacency matrix, and $\mathcal{N}_i$ is the neighbor set of node i. Since communications are bandwidth and power limited, we cannot have infinite precision. In the packet switched network described before, the precision is at best $O(2^{-2\log(1+\tau)\mathcal{B}P})$, where P is packet duration. Lengthening the packet has an exponential improvement on the precision (which is good); unfortunately, the delay in consensus depends on the congestion and is negligibly affected by the packet duration.

We now show how these iterations can be transformed into the data driven form [8]. Note we will henceforth average S component-wise. The idea behind the data driven approach is to devote channels to data types rather than to nodes. For a given set $\mathcal{Q} = \{q_l, \ l = 0, \ldots, |\mathcal{Q}|-1\}$ where q_l represent quantizer centroids, the quantization policy adopted is $\tilde{\theta}_i(k) = \arg\min_{q_l}|\theta_i(k) - q_l|$. Rewrite (9) in terms of the quantized state variables-

$$u_i(k) = \sum_{j\in\mathcal{N}_i}\int_{-\infty}^{\infty} a_{ij}(k)(q - \theta_i(k))\delta(q - \theta_j(k))dq \tag{10}$$

to get $\tilde{u}_i(k) = \sum_{l=0}^{|\mathcal{Q}|-1}(q_l - \theta_i(k))V_{ki}(q_l)$, where

$$V_{ki}(q_l) \triangleq \sum_{j=1}^n a_{ij}(k)\delta[q_l - \tilde{\theta}_j(k)] \tag{11}$$

and $\delta(\cdot)$ and $\delta[\cdot]$ are the Dirac and Kronecker delta functions respectively.

We state a scheme for approximating (11). Parse a generalized poly-phase component of the transmit signal of node i, $s_i[k]$ into $|\mathcal{Q}|$ slots indexed by $l = 0, \ldots, |\mathcal{Q}| - 1$. Let the discrete-time complex base-band signal transmitted by any node i in channel access slot l within iteration k be-

$$s_i\left[l + (k-1)|\mathcal{Q}|\right] = e^{j\phi_i[l]}\delta[q_l - \tilde{\theta}_i(k)], \tag{12}$$

where $e^{j\phi_i[l]}$ introduces a uniform random phase offset in $[0, 2\pi]$ and is picked independently by each node. Then we can approximate $V_{ki}(q_l)$ from the cooperatively generated signal as-

$$\hat{V}_{ki}(q_l) = \begin{cases} |r_i\left[l + |\mathcal{Q}|(k-1)\right]|^2 - N_0, & \text{if } q_l \neq \tilde{\theta}_i(k) \\ 0, & \text{else.} \end{cases} \tag{13}$$

where $r_i\left[l + |\mathcal{Q}|(k-1)\right]$ is defined in (3). This leads to-

$$\hat{u}_i(k) = \sum_{l=0}^{|\mathcal{Q}|-1} (q_l - \tilde{\theta}_i(k))\hat{V}_{ki}(q_l). \tag{14}$$

It is clear from equations (13) and (14) that the nodes do not decode the signals they receive. They use the magnitude of the received signal energy, after subtracting the noise power, to compute the consensus update directly.

The main motivation behind this strategy is that the cost of receiving the update in (14) in terms of channel uses is exactly $|Q|$. Assuming that each channel use requires a time $\mathcal{B}^{-1}$, the delay of each iteration becomes $|Q|\mathcal{B}^{-1}$, irrespective of the number of nodes participating which is an improvement from Fig. 1.

The following lemma provides intuition on why the performance of the data driven algorithm is equivalent to that of a quantized average consensus scheme of the form (9).

Lemma 3.2 *Under* $(a1) - (a5)$, $\mathbb{E}\{\hat{\mathbf{u}}(k)\} = \tilde{\mathbf{u}}(k)$.

Proof From (14), $\mathbb{E}\{\hat{u}_i(k)|\boldsymbol{\theta}(k)\} = \sum_{l=0}^{|Q|-1}(q_l - \tilde{\theta}_i(k))(\lambda_l - N_o)$ where $\lambda_l = \sum_{j=1}^{n} \sigma_{ij}^2 \delta[q_l - \tilde{\theta}_j(k)] + N_o$. If we choose $a_{ij} = \sigma_{ij}^2, \ \forall i \neq j$ and 0 otherwise, then $V_{ki}(q_l) = \lambda_l - N_o = \sum_{j=1}^{n} \sigma_{ij}^2 \delta[q_l - \tilde{\theta}_j(k)]$. This gives us $\mathbb{E}\{\hat{u}_i(k)\} = \tilde{u}_i(k)$ by removing the conditioning. ■

Remark *Convergence*: By modeling the data driven algorithm as a finite markov chain we have shown [9] that it is guaranteed to converge to the quantized true average under the conditions that $|\mathcal{Q}| < \infty$ and $N_o = 0$.

3.3. Data Driven Dynamic Consensus

Once S is known to all nodes, they can start the local micro Kalman filter iterations which involve computing $\mathbf{y}_k$ through dynamic consensus using a low pass consensus filter of the form [10]-

$$\boldsymbol{\theta}_i(k+1) = \boldsymbol{\theta}_i(k) + \epsilon \mathbf{u}_i(k), \tag{15}$$

$$\mathbf{u}_i(k) = \sum_{j \in \mathcal{N}_i} a_{ij}(\boldsymbol{\theta}_j(k) - \boldsymbol{\theta}_i(k)) + \sum_{j \in \mathcal{N}_i \cup \{i\}} a_{ij}(\mathbf{y}_k^j - \mathbf{y}_k^i) \tag{16}$$

where $\boldsymbol{\theta}_i(k)$ denotes a node's estimate of $\mathbf{y}_k$.

Since $\mathbf{y}_k^i$ are vectors in $\mathbf{R}^m$, we solve m component-wise scalar consensus problems. In the jth consensus problem $\theta_i(k)$ tracks $[\mathbf{y}_k]_j \ i = 1, \ldots, m$. This can be extended to the data driven form in a straightforward manner. The only changes to the data driven algorithm are- the network information term:

$$V_{ki}(q_l) \triangleq \sum_{j=1}^{n} a_{ij}(k) \left[\delta[q_l - \tilde{\theta}_j(k)] + \delta[q_l - \tilde{[\mathbf{y}_k^j]_j}]\right] \tag{17}$$

and the transmitted discrete-time complex base-band signal:

$$s_i\left[l + (k-1)|\mathcal{Q}|\right] = e^{j\phi_i[l]} \left[\delta[q_l - \tilde{\theta}_i(k)] + \delta[q_l - \tilde{[\mathbf{y}_k^j]_j}]\right]. \tag{18}$$

4. RESOURCE ALLOCATION FOR DATA DRIVEN DYNAMIC CONSENSUS

From the structure of equation (5) it is clear that if we average $\mathbf{y}_k$ component-wise, errors made in the averaging of one component of $\mathbf{y}_k$ due to the data driven algorithm will affect the accuracy of the values of other components of $\hat{\mathbf{x}}_{k|k}$. So how should we select what precision i.e. $|Q_j|$ to use for averaging each component such that we get the best estimate $\hat{\mathbf{x}}_{k|k}$?

Let $\hat{\mathbf{x}}_{k|k}^{AC}$ denote the state estimate obtained at a node when finite precision data driven consensus is used to compute $\mathbf{y}_k$. Let us define the error covariance matrix $E_{\hat{\mathbf{x}}_k} = \mathbb{E}\{(\hat{\mathbf{x}}_{k|k} - \hat{\mathbf{x}}_{k|k}^{AC})(\hat{\mathbf{x}}_{k|k} - \hat{\mathbf{x}}_{k|k}^{AC})^T\}$. Our objective is to design a resource allocation policy based on the criterion-

$$\min \ Tr(E_{\hat{\mathbf{x}}_k}) \ subj. \ to \ |Q_T| = \sum_{j=1}^{m} |Q_j| \tag{19}$$

where $\log_2 |Q_T|$ and $\log_2 |Q_j|$ indicate the total number of bits available for representing $\mathbf{y}_k^i$ and $[\mathbf{y}_k^i]_j$ respectively. The solution to this is provided by the following lemma.

Lemma 4.1 *The optimal solution for (19) is-*

$$|Q_j| = \frac{(\lambda_k^j)^{\frac{2}{3}}|Q_T|}{\sum_{j=1}^{m}(\lambda_k^j)^{\frac{2}{3}}}, \ \ j = 1, \ldots, m \tag{20}$$

where λ_k^j *denote the eigenvalues of* M_k. *Therefore this is the optimal resource allocation policy.*

Proof See Appendix.

The proof provides intuition on why our chosen optimization problem is reasonable. The policy states that greater precision should be used for those components $[\mathbf{y}_k^i]_j$ for which λ_k^j is large.

5. NUMERICAL RESULTS

We simulate a network with n nodes which track the position of an object moving in roughly a planar circle. The parameters of the dynamical system (2) are - $A = [1 \ -\Delta; \Delta \ 1]$ where $\Delta = 0.02$ is the time-step, $B = I_2$, $H_i = I_2$, $Q = 0.1$, $R_i \in [0.1, 0.45]$, $P_0 = I_2$, and $x_0 = [0, -1]^T$. The initial estimate of each node is x_0.

Fig. 2 shows the estimates of the object's position at all the nodes at different time steps as circles. Fig. 3 compares the MSE in the state estimates for DFK with equal resource allocation and with the optimal resource allocation policy. It highlights the estimate accuracy gained through the allocation policy. Interestingly, the quality of nodes's estimates changes with variations in the original process being tracked.

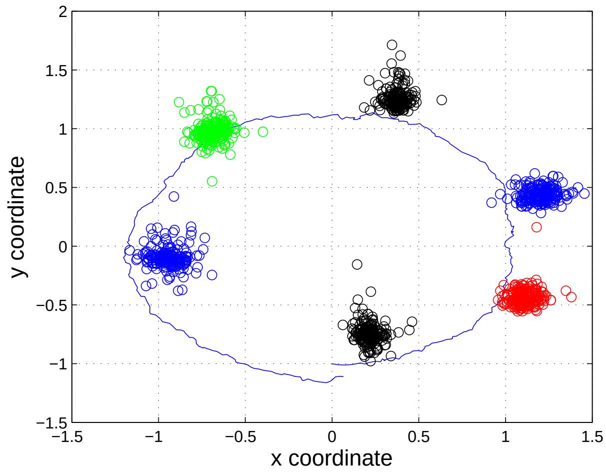

Fig. 2. Node estimates $\hat{\mathbf{x}}_{k|k}$ of the position of the object at different times. Network size $n = 200$.

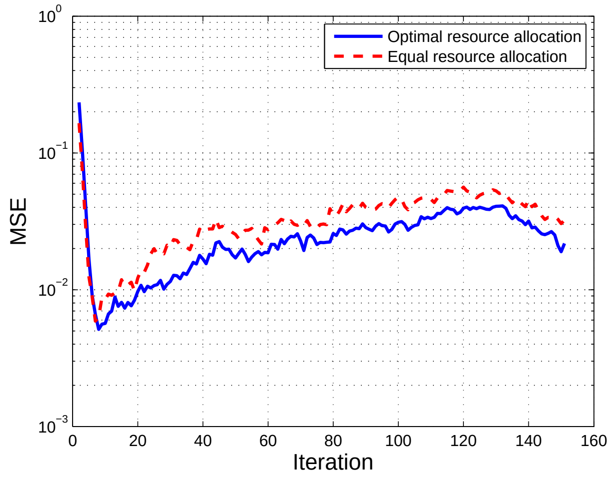

Fig. 3. MSE comparison of filtering with and without resource allocation policy. Network size $n = 50$.

6. CONCLUSION

In this paper, we presented a wireless communication architecture for distributed Kalman filtering based on average consensus in the context of sensor networks. In our architecture, nodes schedule their transmissions according to the data they possess. This leads to a fixed communication cost independent of the network size n. We provided a strategy for allocating communication resources for data driven dynamic consensus which minimizes the error across components in the state estimate $\hat{\mathbf{x}}_{k|k}$.

7. APPENDIX

Proof of Lemma 4.1- $P_{k|k-1}$ and R^i are symmetric since they are covariance matrices. Consequently, S and M_k are symmetric. The eigenvalue decomposition $M_k = U_k \Lambda_k U_k^T$ where U_k is some unitary matrix and $\Lambda_k = diag(\lambda_k^1, \dots, \lambda_k^m)$ is then guaranteed to exist. Let $\xi_{k|k-1} = U_k^T \hat{\mathbf{x}}_{k|k-1}$. Then we can rewrite (5) as a *decoupled* system of equations-

$$\xi_{k|k} = \xi_{k|k-1} + \Lambda_k (U_k^T \mathbf{y}_k - U_k^T S_k \hat{\mathbf{x}}_{k|k-1}). \quad (21)$$

Since we solve the consensus problem component-wise, (21) ensures that errors in the jth consensus problem do not affect the ith component of $\hat{\mathbf{x}}_{k|k}$, $i \neq j$. Let $E_{\xi_k} = \mathbb{E}\{(\xi_{k|k} - \xi_{k|k}^{AC})(\xi_{k|k} - \xi_{k|k}^{AC})^T\}$. By substituting (21) into E_{ξ_k} we get $E_{\xi_k} = \Lambda_k \Phi \Lambda_k^T$ where

$$\Phi = \mathbb{E}\{(U_k^T \mathbf{y}_k - U_k^T \mathbf{y}_k^{AC})(U_k^T \mathbf{y}_k - U_k^T \mathbf{y}_k^{AC})^T\} \quad (22)$$

since the only source of error is $\mathbf{y}_k$ which is computed through consensus. Φ is diagonal because each of the m components of $U_k^T \mathbf{y}_k^i$, $i = 1, \dots, n$, are averaged independently. Furthermore, over a *normalized* range, a bound on the error due to imprecision in consensus is $[U_k^T \mathbf{y}_k]_j - [U_k^T \mathbf{y}_k^{AC}]_j \leq 1/(2|Q_j|)$, giving us that $\Phi \leq diag(1/(2|Q_1|), \dots, 1/(2|Q_m|))$. By the definition of $\xi_{k|k}$ we must have that $E_{\hat{\mathbf{x}}_k} = U_k E_{\xi_k} U_k^T$. Since $Tr(\cdot)$ is invariant under a similarity transform, $Tr(E_{\hat{\mathbf{x}}_k}) = Tr(E_{\xi_k})$. Therefore, (19) can be written as-

$$\min \ \frac{1}{4} \sum_{j=1}^{m} \left(\frac{\lambda_k^j}{|Q_j|} \right)^2 \ subj.\ to\ |Q_T| = \sum_{j=1}^{m} |Q_j|. \quad (23)$$

This can be solved using Lagrangian multipliers to get the desired result. ■

8. REFERENCES

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

[2] J. Speyer, "Computation and transmission requirements for a decentralized linear-quadratic-gaussian control problem," in *Proc. IEEE Trans. Automatic Control*, vol. 24(2), Apr. 1979.

[3] B. Rao, H. F. Durrant-Whyte, and J. A. Sheen, "A fully decentralized multi-sensor system for tracking and surveillance," in *Proc. Int. Journal of Robotics Research*, vol. 12, Feb. 1993.

[4] Y. A. Vershinin, "A data fusion algorithm for multisensor systems," in *Proc. Fifth Int. Conf. on Information Fusion*.

[5] D. P. Spanos, R. Olfati-Saber, and R. Murray, "Approximate distributed kalman filtering in sensor networks with quantifiable performance," in *Proc. 4th Int. Symposium on Information Processing in Sensor Networks*, 2005, pp. 133–139.

[6] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Trans. on Automatic Control*, vol. 49(9), 2004.

[7] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Prentice Hall, 2000.

[8] S. Kirti, A. Scaglione, and R. J. Thomas, "A scalable wireless communication architecture for average consensus," in *Proc. 46th IEEE Conf. on Decision and Control*.

[9] S. Kirti, M. E. Yildiz, and A. Scaglione, "Scalable channel coding for average consensus," *submitted to IEEE JSAC sp. issue on Control and Communications*, 2007.

[10] R. Olfati-Saber, "Consensus filters for sensor networks and distributed sensor fusion," in *Proc. 44th IEEE Conference on Decision and Control*, Dec. 2005, pp. 6698 – 6703.