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 technique with a number of potential distributed applications in sensor networks. In this paper we consider a specific algorithm for distributed Kalman filtering proposed recently by Olfati-Saber [1]. We design a communication access protocol for wireless sensor networks that is tailored to converge rapidly to the desired estimate and provides scalable error performance as number of sensors increases. 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. We provide simulation results demonstrating the performance of our architecture.

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

1. INTRODUCTION AND SYSTEM MODEL

A fundamental problem in sensor networks is distributed detection and estimation. A practical solution can have a great impact in supporting distributed monitoring operations 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. More specifically, consider a linear dynamical system with the following state-space model-

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

where $\mathbf{x}_k \in \mathbb{R}^m$ is the system state at time step k, $\mathbf{w}_k$ is Gaussian noise such that $\mathbb{E}\{\mathbf{w}_k\mathbf{w}_l^T\} = G_k\delta[k-l]$, and A_k, B_k are known matrices $\in \mathbb{R}^{m\times m}$. 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 estimate, in a distributed fashion, 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 matrix $\in \mathbb{R}^{p\times m}$, 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 $\mathbf{z}_k^i \in \mathbb{R}^p$, $p \le m$. For simplicity, in the remainder of the paper we will focus on the case where H^i and R^i are time invariant.

In [1] it was shown that:

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

Theorem 1.1 *Let* $\mathbf{y}_k^i = (H^i)^T(R^i)^{-1}\mathbf{z}_k^i$. *If the two following quantities are available at all nodes:*

$$\mathbf{y}_k = \frac{1}{n}\sum_{i=1}^{n}\mathbf{y}_k^i\,, \qquad S = \frac{1}{n}\sum_{i=1}^{n}(H^i)^T(R^i)^{-1}H^i,$$

then each node can compute its state estimate $\hat{\mathbf{x}}_{k|k}$ *through the following local micro Kalman filter (μ-KF) iterations -*

$$\begin{aligned} M_k &= ((nP_{k|k-1})^{-1} + S)^{-1} & (2)\\ \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + M_k(\mathbf{y}_k - S\hat{\mathbf{x}}_{k|k-1}) & (3)\\ P_{k+1|k} &= A_kM_kA_k^T + B_k(nG_k)B_k^T & (4)\\ \hat{\mathbf{x}}_{k+1|k} &= A_k\hat{\mathbf{x}}_{k|k} & (5)\end{aligned}$$

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

To distribute $\mathbf{y}_k$ and S, [1] proposed replacing $\mathbf{y}_k$ and S respectively with the ith node estimates $\hat{\mathbf{y}}_k(i)$ and $\hat{S}(i)$ obtained with low-pass and high-pass average consensus filters (see [1] for more details). We denote the state estimates obtained this way by $\hat{\mathbf{x}}_{k|k}(i)$.

While Theorem 1.1 [1], provides the basic decomposition needed for solving the state estimation problem via network gossiping, a few problems stand in the way of the practical application of this algorithm: (1) more often than not in practical problems the observations and/or the state vector evolve based on non linear dynamics; (2) the work lacks a good physical model for the network communications, which are assumed to be instantaneous, perfectly synchronized and with infinite precision. This is clearly impossible in the bandwidth and energy constrained environment of a wireless sensor network. While the extension to non linear observers is left for future work, the gap we fill here is dealing with the communication problem. We propose a more realistic modeling and efficient implementation of the average consensus gossiping to support the DKF, based on our recent study on multiple access for wireless average consensus [6].

2. PROPOSED APPROACH

First, we modify the algorithm slightly to allow for an optimized communication resource allocation. Since the matrix S is assumed to be constant, we compute $\hat{S}(i)$ before hand and with much higher accuracy, and assume in the following that $\hat{S}(i) \approx S$. We note in equation (3) that the error made on the term $(\hat{\mathbf{y}}_k(i) - S\hat{\mathbf{x}}_{k|k-1})$ is filtered through the symmetric matrix M_k. Because of symmetry, the eigenvalue decomposition $M_k = U_k\Lambda_kU_k^T$, where U_k is some unitary matrix and $\Lambda_k = diag(\lambda_k^1, \ldots, \lambda_k^m)$, exists.

1-4244-1484-9/08/$25.00 ©2008 IEEE

Rather than computing an estimate of $\mathbf{y}_k$, our state updates are:

$$\hat{\mathbf{x}}_{k|k}(i) = \hat{\mathbf{x}}_{k|k-1}(i) + U_k\Lambda_k(\boldsymbol{\theta}_k(t,i) - U_k^T S\hat{\mathbf{x}}_{k|k-1}(i)) \quad (6)$$

$$\hat{\mathbf{x}}_{k+1|k}(i) = A_k\hat{\mathbf{x}}_{k|k}(i), \quad (7)$$

where $\boldsymbol{\theta}_k(t,i)$ is obtained through the iterative average consensus protocol explained next, initialized with the state $\boldsymbol{\theta}_k(0,i) = U_k^T(H^i)^T(R^i)^{-1}\mathbf{z}_k^i$. For $t \gg 1$ this provides an estimate of:

$$\boldsymbol{\theta}_k = \frac{1}{n}\sum_{i=1}^{n} U_k^T(H^i)^T(R^i)^{-1}\mathbf{z}_k^i \quad (8)$$

with an error $\mathbf{e}_k(t,i)$ that is a function of the transmission resources:

$$\boldsymbol{\theta}_k(t,i) = \boldsymbol{\theta}_k + \mathbf{e}_k(t,i). \quad (9)$$

The reason for our change in (6) is that the error added due to the communication constraints has independent entries in addition to being independent with respect to k (c.f. Section 3.1). Let the additional error incurred by computing (6) instead of the μ-KF be called:

$$\mathbf{v}_{k|k}(i) = \hat{\mathbf{x}}_{k|k} - \hat{\mathbf{x}}_{k|k}(i)\ ,\ \mathbf{v}_{k|k-1}(i) = \hat{\mathbf{x}}_{k|k-1} - \hat{\mathbf{x}}_{k|k-1}(i).$$

It is not difficult to see that $\mathbf{v}_{k|k-1}(i) = A_{k-1}\mathbf{v}_{k-1|k-1}(i)$ and that the error injected into the μ-KF estimate is such that:

$$E\{||\mathbf{v}_{k|k}(i)||^2\} = E\{||(I - M_kS)A_{k-1}\mathbf{v}_{k-1|k-1}(i)||^2\} + E\{||\Lambda_k\mathbf{e}_k(t,i)||^2\}. \quad (10)$$

Our next objective is to indicate how to compute $\boldsymbol{\theta}_k(t,i)$ and how to minimize the average error term $\frac{1}{n}\sum_{i=1}^{n} E\{||\Lambda_k\mathbf{e}_k(t,i)||^2\}$ under the communication constraints.

3. THE COMMUNICATION PROTOCOL

In this section we discuss the communication protocol that supports the acquisition of $\boldsymbol{\theta}_k(t,i)$. The estimation of S is done in an analogous fashion, by performing a similar protocol on all the upper diagonal entries (main diagonal included) of the matrix, using the upper diagonal of $S_i(0) = (H^i)^T(R^i)^{-1}H^i$ as the initial state.

Given a network with a symmetric adjacency matrix whose elements are $a_{ij}(t) \geq 0\ \forall i,j$, and $a_{ij}(t) > 0$ iff nodes i and j are neighbors, the canonical update for a infinite precision synchronous discrete time average consensus protocol is:

$$\boldsymbol{\theta}_k(t+1,i) = \boldsymbol{\theta}_k(t,i) + \sum_{j=1}^{n} a_{ij}(t)(\boldsymbol{\theta}_k(t,j) - \boldsymbol{\theta}_k(t,i)), \quad (11)$$

where the update term $\mathbf{u}_k(t,i) = \sum_{j=1}^{n} a_{ij}(t)(\boldsymbol{\theta}_k(t,j) - \boldsymbol{\theta}_k(t,i))$ is obtained with near-neighbor communications. Define the matrix

$$\{W(t)\}_{ij} = \delta_{ij} - \sum_{j'=1}^{n} a_{ij'}(t) + a_{ij}(t). \quad (12)$$

With initial states $\boldsymbol{\theta}_k(0,i) = U_k^T(H^i)^T(R^i)^{-1}\mathbf{z}_k^i$, the convergence $\boldsymbol{\theta}_k(t,i) \to \boldsymbol{\theta}_k$ in (8) is guaranteed iff $W(t)\mathbf{1} = \mathbf{1}$, $\mathbf{1}^T W(t) = \mathbf{1}^T$, and spectral radius $\rho(W(t) - J) < 1$ where with $J = \frac{1}{n}\mathbf{1}\mathbf{1}^T$. For simplicity, in the following we will assume $W(t) = W$ to be constant. Stacking the vectors $\boldsymbol{\theta}_k(t,i)$ in a vector $\boldsymbol{\theta}_k(t) \in \mathbb{R}^{mn}$:

$$\boldsymbol{\theta}_k(t) = (\boldsymbol{\theta}_k^T(t,1), \ldots \boldsymbol{\theta}_k^T(t,n))^T, \quad (13)$$

we see that the consensus recursion can be written in matrix form as: $\boldsymbol{\theta}_k(t+1) = (W \otimes I)\boldsymbol{\theta}_k(t)$, where $\otimes$ stands for the matrix Kronecker product.

Two aspects are missing from the conventional description of consensus in a wireless network: (1) the communication constraints mandate quantizing the state variables exchanged as follows: $\boldsymbol{\theta}_k(t,j) \mapsto \bar{\boldsymbol{\theta}}_k(t,j) \in \mathcal{L}_k = \{\mathbf{q}_1(k), \ldots, \mathbf{q}_L(k)\}$, where $\mathcal{L}_k$ is an m dimensional lattice; and (2) the nodes share the communication medium and require a multiple access coding policy to deal with multi-user inference.

3.1. Quantization

The residual error in consensus is:

$$\mathbf{e}_k(t) = (\mathbf{e}_k^T(t,1), \ldots \mathbf{e}_k^T(t,n))^T = \boldsymbol{\theta}_k(t) - (\mathbf{1} \otimes \boldsymbol{\theta}_k). \quad (14)$$

We need to characterize how the choice of the number of quantizer levels affect the error variance (10). In what follows we assume:

a.0 Quantization: All nodes use the same quantizer, with codes $\mathcal{L}_k = \mathcal{Q}_{1k} \times \ldots \times \mathcal{Q}_{mk}$ that form a square lattice; and denoting by Q_{pk} the number of levels in the one dimensional uniform lattice $\mathcal{Q}_{pk}$, $L = \prod_{p=1}^{m} Q_{pk}$. We assume that the range of the pth quantizer is $C\sigma_{pk}$ where $\sigma_{pk} = \max_i \sqrt{VAR[\{\boldsymbol{\theta}_k(0,i)\}_p]}$ and C is a positive constant that renders clipping errors statistically negligible. The resulting quantization error $\boldsymbol{\epsilon}_k(t,i) = \boldsymbol{\theta}_k(t,i) - \bar{\boldsymbol{\theta}}_k(t,i)$ can be assumed to be uncorrelated from state to state, approximately uniform, and with:

$$VAR[\{\boldsymbol{\epsilon}_k(t,i)\}_p] = \frac{C^2}{12}\frac{\sigma_{pk}^2}{Q_{pk}^2};\ \ E\{\boldsymbol{\epsilon}_k(t)\boldsymbol{\epsilon}_k^T(t)\} = (I \otimes \Sigma_k),$$

where $\Sigma_k = diag(\frac{\sigma_{1k}^2}{Q_{1k}^2}, \ldots, \frac{\sigma_{mk}^2}{Q_{mk}^2})$.

For consensus with finite precision, the vector update is $\boldsymbol{\theta}_k(t+1) = (W \otimes I)\bar{\boldsymbol{\theta}}_k(t) = (W \otimes I)(\boldsymbol{\theta}_k(t) + \boldsymbol{\epsilon}_k(t))$. Therefore, the error includes two terms, the first due to the incomplete convergence and the second due to the finite precision $\mathbf{e}_k(t) = ((W - J) \otimes I)\boldsymbol{\theta}_k(t-1) + (W \otimes I)\boldsymbol{\epsilon}_k(t-1)$. As a function of the initial condition $\boldsymbol{\theta}_k(0)$ we can now calculate:

$$E\{||(I \otimes \Lambda_k)\mathbf{e}_k(t)||^2\} = E\{||((W - J)^t \otimes \Lambda_k)\boldsymbol{\theta}_k(0)||^2\} + \sum_{t'=1}^{t} ||W^{t'}||^2 \sum_{p=1}^{m} (\lambda_k^p)^2 \frac{\sigma_{pk}^2}{Q_{pk}^2}, \quad (15)$$

where $||W^{t'}||^2 = trace(W^{2t'})$.

3.2. Scalable multiple access

The motivation behind our multiple-access method is the realization that the trade-off between the connectivity of the graph and speed of convergence of average consensus, renders gossiping protocols with traditional wireless packet switching unscalable in bandwidth constrained environments. By studying the effect of the algebraic connectivity $\rho(W - J)$ as a function of the nodal degrees $\sum_{j=1}^{n} a_{ij}$ it is possible to draw the following conclusions: (1) if one limits the number of neighbors, by lowering the transmit power and encouraging spectral reuse, the protocol slows down; (2) if one allows the number of neighbors to grow as the network size n increases by increasing the transmission power, the speed of convergence tends to saturate, while congestion slows down actual the transmissions. The power control problem makes the protocol less robust and less attractive.

Consider quantized consensus. The resulting quantized update $\bar{\mathbf{u}}_k(t,i) = \sum_{j=1}^{n} a_{ij}(t)(\bar{\boldsymbol{\theta}}_k(t,j) - \bar{\boldsymbol{\theta}}_k(t,i))$ can be decomposed as:

$$\bar{\mathbf{u}}_k(t,i) = \sum_{l=1}^{L} (\mathbf{q}_l(k) - \bar{\boldsymbol{\theta}}_k(t,i))\mathbf{m}_k(t,i,l) \quad (16)$$

where, using an indicator function $\delta(x) = 1$ iff $x = 0$ and 0 else:

$$\{\mathbf{m}_k(t,i,l)\}_p = \sum_{j=1}^{n} a_{ij}(t)\delta(\{\mathbf{q}_l(k) - \bar{\boldsymbol{\theta}}_k(t,i)\}_p). \quad (17)$$

This last equation is the corner point of our multiple access technique: in fact, our additional observation is that the state information can be embedded in a code that delivers at each node the message $\mathbf{m}_k(t,i,l)$ directly and collaboratively. Next, we review our novel multiple access protocol for average consensus [6] for coding $\mathbf{m}_k(t,i,l)$. Let us review first the assumptions we make on the physical communication layer. Let $\tau \in \mathbb{R}$ indicate the continuous time variable.
a.1 Signal Space – Time is slotted in slots of duration $T = 1$. The RF signals transmitted belong to a signal space of dimension $Q_k \propto BT$ where B is the single-sided bandwidth allotted for communication. We denote by $\{c_1(\tau)\}_{l=1}^{Q_k}$, the base-band complex equivalent orthonormal basis chosen to span the signal space. The signal transmitted by node i in the tth time slot is:

$$S_i(\tau,t) = \sum_{l=1}^{Q_k} s_i(t,l)c_l(\tau - tT) \quad (18)$$

where $s_i(t) = (s_i(t,1),\ldots,s_i(t,Q_k))^T$ is the vector of coordinates of the transmit signal with respect to the basis $\{c_1(\tau)\}_{l=1}^{Q_k}$.
a.2 Node Power constraint – Each node has a per iteration power constraint $P = \sum_{l=1}^{Q_k} |s_i(t,l)|^2$.
a.3 Incoherent channel: Fading + AWGN – Each received signal $R_i(\tau,t)$ is affected by an independent additive white Gaussian noise process $\mathrm{W}_i(\tau)$ with noise spectral density $N_0/2$. The channel is broadcast. Its distortion on $c_i(\tau)$ can be captured by a single independent fading coefficient, changing over the slots as block fading, denoted by $h_{ij}(t) \sim \mathcal{CN}(0,\alpha_{ij})$, where α_{ij} is the average path-loss (large scale fading). Reciprocity holds on average, *i.e.* $\alpha_{ij} = \alpha_{ji}$.
a.4 Half-duplex channel: If node i has $s_i(t,l) \neq 0$ for $l \in \mathcal{S} \subseteq [1,Q_k]$, node i cannot sense any code transmitted in the sub-space spanned by $\{c_l(\tau)\}_{l\in\mathcal{S}}$.

Based on (**a.1-a.4**), a sufficient statistic for the received signal is $r_i(t,l) = n^{-1/2} < R_i(\tau,t), c_l(\tau - lT) >$, for $l \in [1,Q_k]$:

$$r_i(t,l) = \begin{cases} \frac{1}{\sqrt{n}} \sum_{j=1}^{n} h_{ji}(t)s_j(t,l) + w_i(t,l), & \text{if } s_i(t,l) = 0; \\ 0, & \text{else.} \end{cases} \quad (19)$$

Given $s_j(t,l)$, $r_i(t,l) \sim \mathcal{CN}(0, \frac{1}{n}\sum_{j\neq i} \alpha_{ij}|s_j(t,l)|^2 + \frac{N_0}{n})$.

Channel codes: We fix $Q_k = \sum_{p=1}^{m} Q_{pk}$ and let $l' = \sum_{p'=1}^{p-1} Q_{p'k} + l$. We propose choosing the coefficients $s_i(t,l')$ as follows:

$$\boldsymbol{\theta}_k(t,i) \mapsto \bar{\boldsymbol{\theta}}_k(t,i) \mapsto s_i(t,l') = e^{j\phi_l}\delta(q_{l'}(k) - \{\bar{\boldsymbol{\theta}}_k(t,i)\}_p). \quad (20)$$

where $q_{l'}(k) \in \{\mathcal{Q}_{1k},\ldots,\mathcal{Q}_{mk}\}$ is $q_{l'}(k) = \{\mathbf{q}_l(k)\}_p$ and $\phi_l \sim \mathcal{U}[0,2\pi)$. With these codes, setting the adjacency matrix coefficients to be equal to $a_{ij} = \alpha_{ij}/n$, the Maximum Likelihood (ML) estimate of $\{\mathbf{m}_k(t,i,l)\}_p$ in (17) is:

$$\{\hat{\mathbf{m}}_k(t,i,l)\}_{p=1}^{m} = |r_i(t,l')|^2 - \frac{N_0}{n}. \quad (21)$$

The following two lemmas apply, which we state here without proof:

Lemma 3.1 *For any given initial state $\boldsymbol{\theta}_k(0)$:*

$$E\{\hat{\mathbf{m}}_k(t,i,l)\} = \mathbf{m}_k(t,i,l).$$

Therefore, the multiple access coding method proposed on average tends to the same result as quantized consensus.

Lemma 3.2 *A consensus state is $\boldsymbol{\theta}_k(t) = (\mathbf{1}\otimes\mathbf{c})$. In the limit as $n \to \infty$ consensus states are absorbing, and the protocol converges to consensus with probability one.*

With our access policy, the extra error term that affects the quantized consensus the update in (16) in each consensus iteration:

$$\boldsymbol{\epsilon}'_k(t,i) = \sum_{l=1}^{L} (\mathbf{q}_l(k) - \bar{\boldsymbol{\theta}}_k(t,i))(\mathbf{m}_k(t,i,l) - E\{\hat{\mathbf{m}}_k(t,i,l)\}), \quad (22)$$

tends to vanish in the mean square sense if $n \gg 1$ due to the law of large numbers.

3.3. Optimal resource allocation

When the number of nodes is large the noise becomes negligible and the ML estimate $\hat{\mathbf{m}}_k(t,i,l) \to \mathbf{m}_k(t,i,l)$ in the mean square sense. In these conditions the brunt of the error that is fed into the average consensus iteration is due to the quantization and the effect of (22) can be neglected.

Note that, since $Q_k = \Theta(BT)$ the cost in time and bandwidth associated with the transmission of our channel codes in (20) is equal to the sum of the number of levels, i.e. $Q_k = \sum_{p=1}^{m} Q_{pk}$. Interestingly, this cost is unaffected by how large n is!

With $\lambda_2(W)$ representing the second largest eigenvalue of W (the largest one being 1) the error can be bounded as follows:

$$\frac{1}{n}E\{||(I\otimes\Lambda_k)\mathbf{e}_k(t)||^2\} =\leq \lambda_2^{2t}(W)E\{||(I\otimes\Lambda_k)\boldsymbol{\theta}_k(0)||^2\} + t\sum_{p=1}^{m}(\lambda_k^p)^2\frac{\sigma_{pk}^2}{Q_{pk}^2}, \quad (23)$$

and the communication constraint only affects the magnitude of the term $\sum_{p=1}^{m}(\lambda_k^p)^2\frac{\sigma_{pk}^2}{Q_{pk}^2}$. Hence the resource allocation optimization problem becomes:

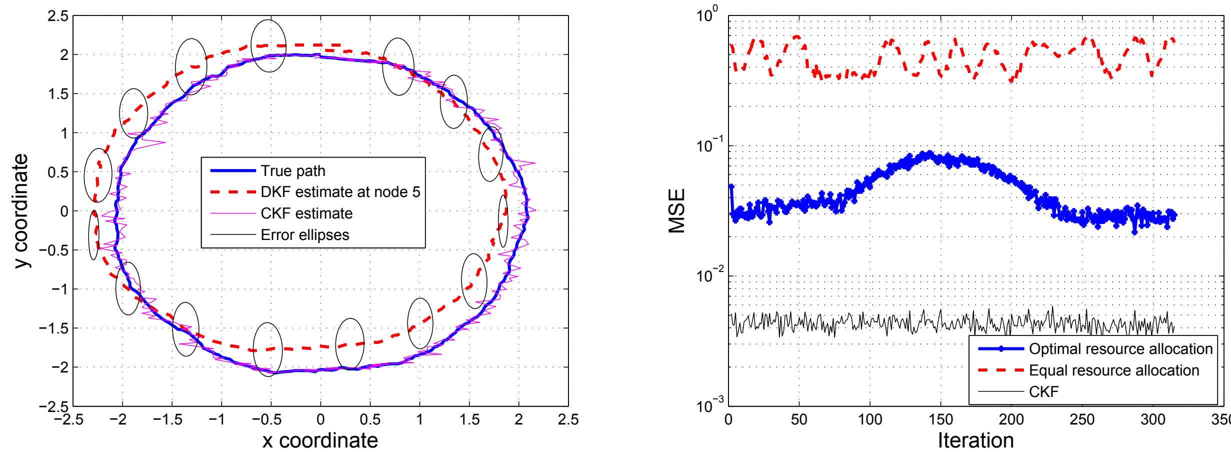
$$\min_{Q_{pk}} \sum_{p=1}^{m}(\lambda_k^p)^2\frac{\sigma_{pk}^2}{Q_{pk}^2} \;\; s.t. \;\; \sum_{p=1}^{m} Q_{pk} = Q_k = \Theta(BT). \quad (24)$$

Solving this using Lagrange multipliers gives us the following lemma:

Lemma 3.3 *The resource allocation policy is given by -*

$$Q_{pk} = \frac{Q_k(\lambda_k^p)^{\frac{2}{3}}\sigma_{pk}^{\frac{1}{3}}}{\sum_{p'=1}^{m}(\lambda_k^{p'})^{\frac{2}{3}}\sigma_{p'k}^{\frac{1}{3}}}, \;\; p = 1,\ldots,m \quad (25)$$

where $\sigma_{pk}^2 = VAR(\{U_k^T(H^i)^T(R^i)^{-1}\mathbf{z}_k^i\}_p)$ and $M_k = U_k\Lambda_kU_k$ with $\Lambda_k = diag(\lambda_k^1,\ldots,\lambda_k^m)$.



(a) (b)

Fig. 1. (a) *Error ellipses* for the $\hat{\mathbf{x}}_{k|k}$ at node 5 are shown at every 20 time steps. This is representative of the error ellipses at other nodes. (b) *MSE:* performance of the data driven DKF algorithm is compared with and without resource allocation. MSE of the CKF is also shown as a baseline.

4. NUMERICAL RESULTS

We simulated a network with 25 nodes all of which track the position of an object moving in a rough planar circle. The nodes are located randomly in a unit square plane through a uniform distribution. The parameters of the linear dynamical system (1) are - $A = [1\ -\Delta; \Delta\ 1]$ where $\Delta = 0.02$ is the time-step, $B = H_i = P_0 = I_2$, $G = 0.01$, R_i is diagonal with entries $\in [0.01, 0.4]$, and $x_0 = [0, 2]^T$. The initial estimate of each node is x_0. The resource allocation policy uses $Q_k = 200$ and $\mathbb{E}\{U_k^T \mathbf{y}_k^i \mathbf{y}_k^{iT} U_k\} = U_k R_i^{-1}((G+2)R_i^{-1} + I_2)U_k^T$. Each DKF iteration uses $t = 5$ data driven consensus iterations to compute $\boldsymbol{\theta}_k(t, i)$ at SNR $30dB$.

Fig. 1(a) shows the path estimated by node 5 (picked arbitrarily) and its accompanying 99% confidence region error ellipses. The error ellipses were obtained through monte carlo simulations using 300 trials. This numerically shows that the proposed distributed filter does not diverge. Furthermore, the path estimated by the distributed filter compares favorably to the estimate obtained by the central KF.

Fig. 1(b) compares the MSE of the distributed filter and shows that the performance is improved by utilizing the proposed resource allocation policy. The MSE is compared with a naive scheme which simply allocates equal precision to each component of the vector being averaged. Interestingly, the quality of nodes's estimates changes with variations in the process being tracked.

5. 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 consensus which minimizes the error across components in the state estimate $\hat{\mathbf{x}}_{k|k}$.

6. 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] 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*.