DISTRIBUTED KALMAN FILTERS IN SENSOR NETWORKS: BIPARTITE FUSION GRAPHS

Usman A. Khan and José M. F. Moura

Carnegie Mellon University
Department of Electrical and Computer Engineering
Pittsburgh, PA 15213
{ukhan, moura}@ece.cmu.edu

ABSTRACT

We study the distributed Kalman filter in sensor networks where multiple sensors collaborate to achieve a common objective. Our motivation is to distribute the global model that comes from the state-space representation of a sparse and localized large-scale system into reduced **coupled** sensor-based models. We implement local Kalman filters on these reduced models, by approximating the Gaussian error process of the Kalman filter to be Gauss-Markov, ensuring that each sensor is involved only in reduced-order computations and local communication. We propose a generalized distributed Jacobi algorithm to compute global matrix inversion, locally, in an iterative fashion. We employ bipartite fusion graphs in order to fuse the shared observations and shared estimates across the local models.

Index Terms— Large-scale systems, Sparse matrices, Distributed algorithms, Matrix inversion, Kalman filtering

1. INTRODUCTION

Recent technological advances in solid-state electronics and wireless communication have made it possible to monitor very large-scale dynamical systems, e.g., power grid, weather forecast systems, and earthquake tracking systems, using sensor networks. These geographically distributed sensors take measurements of the variables pertinent to the system. These measurements, in addition to be employed in other tasks, are used for state estimation. State estimation is essential to these dynamical systems for control, tracking and navigational purposes. We develop a distributed Kalman filter for the multisensor large-scale dynamical systems with localized and sparse structure.

With sensor networks the observations of the field of interest are distributed across different sensors. All these observations are to be incorporated in the implementation of the Kalman filter to ensure optimal performance. Collecting these observations at a single location (fusion center) implements a centralized Kalman filter. The fusion center then communicates the estimates back to the sensors. In large-scale dynamical systems, the centralized Kalman filter is impractical because it requires: (i) long-distance communication since the sensors span a large geographical area; and (ii) high computation because the state-space models coming from such large-scale systems have high-dimensional state vectors. Furthermore, a centralized scheme has the disadvantages of large latency and a single point of failure.

This work was partially supported by the DARPA DSO Advanced Computing and Mathematics Program Integrated Sensing and Processing (ISP) Initiative under ARO grant # DAAD 19-02-1-0180, by NSF under grants # ECS-0225449 and # CNS-0428404, and by an IBM Faculty Award.

To reduce the inordinate communication requirements, Kalman filters have been implemented by replicating the global model at each sensor. This is still computationally expensive involving nth order matrices and vectors (where n is the dimension of the large-scale system). Much of the existing research addresses this problem of reducing the communication burden, but, by replicating the nth order global dynamics at each sensor, which, in general, requires $O(n^3)$ computations. In this paper, we present a solution that reduces the communication as well as the computation requirements. In our approach, we distribute the global dynamics into reduced-order local dynamics. This reduces the computational burden at each sensor, since the computations are of the order of the local reduced-order models, n_l, where $n_l \ll n$. We devise local Kalman filters that are efficient for large-scale systems and avoid at each sensor the drawbacks, extensive computation and inordinate communication, of the centralized implementation.

Decentralizing Kalman filter dates back to [2, 3, 4], requiring all-to-all communication networks and nth order replicated models. A decentralized Kalman filter where the observations are fused using local communication and iterative consensus filters is provided in [5]. Distributing the communication requirements in the presence of uncertain communication links and packet losses is addressed in [6], where the problem of target tracking is considered, see also [7]. The problem of target tracking requires a few state variables, e.g., velocity and acceleration, and has an inherent structure of decoupled dynamics in the case of multiple targets. All these decentralized schemes reduce the communication burden but replicate the nth order Kalman filter at each sensor with $O(n^3)$ computational requirements, a practically infeasibility when n is large.

Kalman filters with reduced models were addressed in [8]. This work was further extended in [9]. In their work, the reduced models at each sensor are decoupled, forcing the model matrix to be block diagonal. Furthermore, the network topology is either fully connected, [8], or is close to fully connected, [9], requiring long distance links that are expensive. Willsky et. al. [10], have also addressed the problem of combining estimates from the subsystems of a global system, but in their implementation the local sub-systems are decoupled, making the sub-systems independent Markov processes. This solution does not address the problem we consider here of coupled large-scale dynamical systems.

We present a fully distributed Kalman filter, implemented on sensor-based reduced models and distributed observations. In order to achieve this, we distribute the state-space models coming from the large-scale dynamical systems, into reduced *coupled* sensor-based models at each sensor[1]. These reduced models exploit the localized

[1] We assume that the state-space model cannot be decoupled under any model preserving transformation of the state vector.

and sparse structure of the system dynamics[2]. Local Kalman filters are then implemented on the reduced models. Each local Kalman filter computes local variables, a subset of the global variables (state, observations, error covariances) required in its centralized counterpart. Hence the computations required at the sensors are significantly reduced. Coupling between the reduced models is preserved and global performance is achieved by exchanging messages across different sensors using only local communication.

We distribute the Information filter form of the Kalman filter[3]. In the process, we assume information matrices in the Information filter to be L-banded (we refer to a matrix as an L-banded matrix ($L \geq 0$), if the elements outside the band defined by the Lth upper and Lth lower diagonal are zero.) This assumption is equivalent to forcing the Gaussian error processes to be Gauss-Markovian and is optimal in Kullback-Leibler or maximum entropy sense [12]. This assumption helps us in making the computations and communications to be local. It will be shown that without this assumption a distributed implementation is not possible, requiring either global communication or computing global variables. Simulations in [13] show that the centralized Information filter with L-banded approximations is virtually indistinguishable from the exact centralized Information filter.

We provide the global model and centralized Information filter in sections 2 and 3, respectively. We will discuss the sensor-based reduced models in section 4. The local Information filters are divided in section 5 and section 6, with results and conclusions in section 7.

2. GLOBAL MODEL

We assume that the dynamical system follows an n-dimensional state equation

$$\mathbf{x}_{k+1} = \mathbf{F}\mathbf{x}_k + \mathbf{G}\mathbf{u}_k, \tag{1}$$

where k is the discrete time index, $\mathbf{x}_k \in \mathbb{R}^n$ is the state vector, $\mathbf{F} \in \mathbb{R}^{n\times n}$ is a sparse localized model matrix, $\mathbf{u}_k \in \mathbb{R}^m$ is the state noise vector and $\mathbf{G} \in \mathbb{R}^{n\times m}$ is the state noise matrix, and $\mathbf{x}_0$ are the initial conditions such that $\mathbf{x}_0 \sim \mathcal{N}(\overline{\mathbf{x}}_0, \mathbf{\Sigma}_0)$. We assume that the random field (1), is monitored by N sensors. Observations at sensor l are,

$$\mathbf{y}_k^{(l)} = \mathbf{H}_l\mathbf{x}_k + \mathbf{w}_k^{(l)}, \tag{2}$$

where $\mathbf{H}_l \in \mathbb{R}^{p_l\times n}$ is the local observation matrix, and $\mathbf{w}^{(l)} \in \mathbb{R}^{p_l}$ is the local observation noise with the covariance matrix, $\mathbf{R}_l$. We can get the global observation model by stacking the observations at each sensor in a global observation vector, $\mathbf{y}_k \in \mathbb{R}^{p\times n}$ given by

$$\mathbf{y}_k = \mathbf{H}\mathbf{x}_k + \mathbf{w}_k, \tag{3}$$

where we have

$$\mathbf{y}_k = \begin{bmatrix} \mathbf{y}_k^{(1)} \\ \vdots \\ \mathbf{y}_k^{(N)} \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_N \end{bmatrix}, \quad \mathbf{w}_k = \begin{bmatrix} \mathbf{w}_k^{(1)} \\ \vdots \\ \mathbf{w}^{(N)} \end{bmatrix}. \tag{4}$$

We adopt the standard assumptions on the statistical characteristics of the white noise sequences, $\{\mathbf{u}_k\}_{k\geq 0}$, and $\{\mathbf{w}_k\}_{k\geq 0}$, with

$$E[\mathbf{u}_k\mathbf{u}_j^H] = \mathbf{Q}\delta_{ij} \text{ and } E[\mathbf{w}_k\mathbf{w}_j^H] = \mathbf{R}\delta_{ij}. \tag{5}$$

We also note that $\mathbf{R} = \text{blockdiag}[\mathbf{R}_1, \ldots, \mathbf{R}_N]$.

3. CENTRALIZED INFORMATION FILTER

The estimator and the predictor in the Information filter domain are $\widehat{\mathbf{z}}_{k|k} = \mathbf{Z}_{k|k}\widehat{\mathbf{x}}_{k|k}$, and $\widehat{\mathbf{z}}_{k|k-1} = \mathbf{Z}_{k|k-1}\widehat{\mathbf{x}}_{k|k-1}$, respectively, where the Information matrices, $\mathbf{Z}_{k|k}$, and $\mathbf{Z}_{k|k-1}$, are the inverses of the estimation error covariance matrix, $\mathbf{S}_{k|k}$, and prediction error covariance matrix, $\mathbf{\Pi}_{k|k-1}$, respectively. Let the nth-dimensional global observation variables be $\mathbf{i}_k = \mathbf{H}^T\mathbf{R}^{-1}\mathbf{y}_k$ and $\mathbf{I}_k = \mathbf{H}^T\mathbf{R}^{-1}\mathbf{H}$ and the nth-dimensional local observation variables be $\mathbf{i}_{l,k} = \mathbf{H}_l^T\mathbf{R}_l^{-1}\mathbf{y}_k^{(l)}$ and $\mathbf{I}_{l,k} = \mathbf{H}_l^T\mathbf{R}_l^{-1}\mathbf{H}_l$. It can be shown that [9]

$$\mathbf{i}_k = \sum_{l=1}^{N} \mathbf{H}_l^T\mathbf{R}_l^{-1}\mathbf{y}_k^{(l)} = \sum_{l=1}^{N} \mathbf{i}_{l,k} \tag{6}$$

$$\mathbf{I}_k = \sum_{l=1}^{N} \mathbf{H}_l^T\mathbf{R}_l^{-1}\mathbf{H}_l = \sum_{l=1}^{N} \mathbf{I}_{l,k} \tag{7}$$

The centralized Information filter equations for the global model (1) and (3) contain initial conditions, a filter step, and a prediction step. The *initial conditions* are $\mathbf{Z}_{0|-1} = \mathbf{\Sigma}_0^{-1}$ and $\widehat{\mathbf{z}}_{0|-1} = \mathbf{Z}_{0|-1}\overline{\mathbf{x}}_0$. The *filter step* of the centralized Information filter is given by

$$\mathbf{Z}_{k|k} = \mathbf{Z}_{k|k-1} + \sum_{l=1}^{N} \mathbf{I}_{l,k}, \tag{8a}$$

$$\widehat{\mathbf{z}}_{k|k} = \widehat{\mathbf{z}}_{k|k-1} + \sum_{l=1}^{N} \mathbf{i}_{l,k}. \tag{8b}$$

The *prediction step* of the centralized Information filter is given by

$$\mathbf{Z}_{k|k-1} = \mathbf{\Pi}_{k|k-1}^{-1} = (\mathbf{F}\mathbf{Z}_{k-1|k-1}^{-1}\mathbf{F}^T + \mathbf{G}\mathbf{Q}\mathbf{G}^T)^{-1}, \tag{9a}$$

$$\widehat{\mathbf{z}}_{k|k-1} = \mathbf{Z}_{k|k-1}\left(\mathbf{F}\mathbf{Z}_{k-1|k-1}^{-1}\widehat{\mathbf{z}}_{k-1|k-1}\right). \tag{9b}$$

4. SENSOR-BASED REDUCED MODELS

In this section, we present the model distribution in order to obtain reduced sensor-based models at each sensor. At the lth sensor, we choose an $n_l \times n$ selection matrix, $\mathbf{T}_l$, see also [9], such that it chooses the n_l local states in the local state vector, $\mathbf{x}_k^{(l)}$, from the global state vector, $\mathbf{x}_k$,

$$\mathbf{x}_k^{(l)} = \mathbf{T}_l\mathbf{x}_k. \tag{10}$$

The choice of the selection matrix, $\mathbf{T}_l$, is such that, if a sensor l observes a linear combination of states through the local observation matrix, $\mathbf{H}_l$, all these states are included in the local state vector, $\mathbf{x}_k^{(l)}$.

[2]Sparse dynamical systems span a large variety of interesting applications, e.g., image reconstruction problem where the pixel values depend on the neighboring pixels, random fields obtained by discretizing PDEs, power grids. Localized structure on the global dynamics refers to systems where the correlations among the states farther apart in the state vector decay rapidly. A broad range of sparse systems, not exhibiting the localized structure, can be converted into sparse banded systems with highly localized structures, using parallelized iterative solvers, e.g., Reverse Cuthill Mckee (RCM) reordering, [11].

[3]Information filters [9], are algebraically equivalent to the Kalman filter [1]. In Information filters the information matrices (inverse of the error covariance matrices) are iterated at each time step.

For example, if we have a 5-dimensional system with sensor l having the local observation matrix,

$$\mathbf{H}_l = \begin{bmatrix} h_{11} & h_{12} & 0 & 0 & 0 \\ 0 & h_{22} & h_{23} & 0 & 0 \end{bmatrix}, \tag{11}$$

the selection matrix is,

$$\mathbf{T}_l = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}. \tag{12}$$

A detailed model distribution procedure for obtaining reduced order sensor-based models using a graph theoretic approach is provided in [14]. The reduced model at sensor l involves $\mathbf{x}_k^{(l)}$ as a local state vector. Putting (10) in (1) and defining $\mathbf{F}_l = \mathbf{T}_l\mathbf{F}$ and $\mathbf{G}_l = \mathbf{T}_l\mathbf{G}$, we have

$$\mathbf{x}_{k+1}^{(l)} = \mathbf{F}_l\mathbf{x}_k + \mathbf{G}_l\mathbf{u}_k \tag{13}$$

We partition $\mathbf{F}_l$ into a reduced model matrix, $\mathbf{F}^{(l)}$, which corresponds to the reduced state vector, $\mathbf{x}_k^{(l)}$, and an input matrix, $\mathbf{D}^{(l)}$, which corresponds to the rest of the states not included in the reduced state vector, $\mathbf{x}_k^{(l)}$ (these states that are not included will be the inputs to the reduced model so that we can preserve the coupling and in turn the global model). Notice that, since the model matrix, $\mathbf{F}$, is sparse and localized, most of the columns in $\mathbf{D}^{(l)}$ will be zero and we retain only its non zero columns, with their corresponding states in an input vector, $\mathbf{d}_k^{(l)}$. We also retain only the state noise sources relevant to the reduced model. Now the reduced model at sensor l becomes

$$\mathbf{x}_{k+1}^{(l)} = \mathbf{F}^{(l)}\mathbf{x}_k^{(l)} + \mathbf{D}^{(l)}\mathbf{d}_k^{(l)} + \mathbf{G}^{(l)}\mathbf{u}_k^{(l)} \tag{14}$$

$$\mathbf{y}_k^{(l)} = \mathbf{H}^{(l)}\mathbf{x}_k^{(l)} + \mathbf{w}_k^{(l)} \tag{15}$$

Note that the reduced local observation matrix, $\mathbf{H}^{(l)} \in \mathbb{R}^{p_l \times n_l}$ is different from the local observation matrix, $\mathbf{H}_l \in \mathbb{R}^{p_l \times n}$. The term $\mathbf{D}^{(l)}\mathbf{d}_k^{(l)}$ in (14), arises because the local model at sensor l is coupled to the local models at those neighboring sensor (recall $\mathbf{F}$ is sparse and localized), which model the states, in $\mathbf{d}_k^{(l)}$, in their local model. If we ignore this term, we reduce our model to decoupled local subsystems as in [10, 9]. We do not ignore this coupling and require it to be communicated from the neighboring sensors. Because $\mathbf{d}_k^{(l)}$ is not available, we use $\widehat{\mathbf{d}}_{k|k}^{(l)}$ as inputs, which is communicated from the neighboring sensors that are modeling the states in the vector $\mathbf{d}_k$, in their reduced models.

The local Information filters are now based on (14) and (15). The local Information filters contain a local filter step and local prediction step, divided in the next sections. The local filter step (section 5) requires observation fusion and estimate fusion because reduced models across different sensors may have overlapping states. The fusion is carried out with the help of bipartite fusion graphs, section 5.1. The local prediction step (section 6) requires a global variable because the reduced models are coupled; we avoid computing this global variable by using iterative generalized distributed Jacobi algorithms.

5. LOCAL FILTER STEP

Define n_lth-dimensional reduced local observation variables $\mathbf{i}_k^{(l)} = (\mathbf{H}^{(l)})^T\mathbf{R}_l^{-1}\mathbf{y}_k^{(l)}$ and $\mathbf{I}_k^{(l)} = (\mathbf{H}^{(l)})^T\mathbf{R}_l^{-1}\mathbf{H}^{(l)}$. Then the local filter step is given by

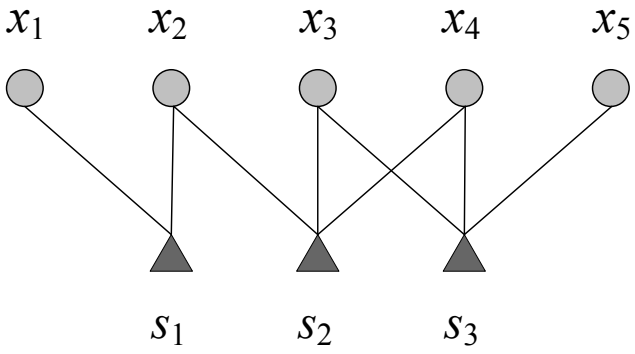

Fig. 1. Bipartite fusion graph for a 5-dimensional system with the global and local observation matrices of equation (17).

$$\mathbf{Z}_{k|k}^{(l)} = \mathbf{Z}_{k|k-1}^{(l)} + \mathbf{I}_{f,k}^{(l)}, \tag{16a}$$

$$\widehat{\mathbf{z}}_{k|k}^{(l)} = \widehat{\mathbf{z}}_{k|k-1}^{(l)} + \mathbf{i}_{f,k}^{(l)}, \tag{16b}$$

where the fused observation variables $\mathbf{I}_{f,k}^{(l)}$ and $\mathbf{I}_{f,k}^{(l)}$ are discussed in section 5.1.1.We go from the estimates in the Information filter domain, $\widehat{\mathbf{z}}_{k|k}^{(l)}$ to the estimates in the Kalman filter domain, $\widehat{\mathbf{x}}_{k|k}^{(l)}$, by solving a linear system of equations $\mathbf{Z}_{k|k}\widehat{\mathbf{x}}_{k|k} = \widehat{\mathbf{z}}_{k|k}$, which requires long distance communication and extensive computations for arbitrary estimation information matrices, $\mathbf{Z}_{k|k}$. If we approximate $\mathbf{Z}_{k|k}$ to be an L-banded matrix $\forall k$, a solution with only local communication and computations with local variables is possible, using an iterative distributed Jacobi algorithm for vectors (DJV), see also [15].

5.1. Bipartite Fusion Graphs

After the model distribution step introduced in section 4, several sensors may share states in their reduced models (14). This is equivalent to saying that the local state vectors, $\mathbf{x}_k^{(l)}$, for all the sensors may overlap and hence the sensor-based reduced models share the overlapped states. Recall that the choice of the selection matrix, $\mathbf{T}_l$, is based on the local observation matrices, $\mathbf{H}_l$. So the sensors having shared states have different observations of the shared states. The observations corresponding to the shared states need to be fused in order to guarantee global performance. Since each sensor implements a separate local Information filter, the estimates corresponding to the shared states also need to be fused across the sensors containing those states. We implement the fusion procedure with the help of bipartite fusion graphs, introduced below.

We present a simple example illustrating the bipartite fusion graphs. We can easily extend this illustration for the case of higher order state-vectors and a large number of sensors. Consider a 5-dimensional system observed by $N = 3$ sensors, where the global observation matrix, $\mathbf{H}$, composed by local observation matrices $\mathbf{H}_l$, see equation (4), is given by

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \mathbf{H}_3 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & 0 & 0 & 0 \\ 0 & h_{22} & h_{23} & h_{24} & 0 \\ 0 & 0 & h_{23} & h_{24} & h_{25} \end{bmatrix}. \tag{17}$$

A bipartite fusion graph, $\mathcal{B} = [X \cup S, E]$, where $X = \{x_i\}_{i=1,\ldots,n}$ is the state-set and $S = \{s_j\}_{j=1,\ldots,N}$ is the sensor-set, consists of

the partitioned vertex set, $X \cup S$, and an interconnection matrix, E. The structure of the interconnection matrix, E, is imposed by the global observation matrix, $\mathbf{H}$, in the following way. The sensor, s_i, is connected to the state variable, x_j, if s_i observes the state variable, x_j. In other words, we have an edge between the sensor, s_i, and the state variable, x_j, if the local observation matrix, $\mathbf{H}_i$, at sensor s_i, contains a non-zero in its jth column. The bipartite fusion graph, $\mathcal{B}$, for the global observation matrix, $\mathbf{H}$ in (17), is shown in figure 1.

The bipartite fusion graphs provides a natural way of selecting the sensors required in the fusion corresponding to each state. For example, figure 1 suggests that: for state x_1, no observation fusion or estimate fusion is required since it is observed and hence modeled only at sensor s_1; for state x_2, observation fusion and estimate fusion is required on the sensors s_1 and s_2; and so on. In this way, we *distribute* the global observation into local observation fusion for each state, where only the information from neighboring sensors is required.

We provide some notation for the next subsections. Let $\mathcal{G}$ be the sensor communication graph; this entails the sensor communication pattern. For each state, x_j, let $\mathcal{G}_j$ be the induced subgraph of $\mathcal{G}$, such that the vertices of $\mathcal{G}_j$ are all the sensors connected to the state, x_j, in the bipartite fusion graph, $\mathcal{B}$. It is obvious that, to properly carry out the fusion procedure, we require G and $G_j \forall j$ to be connected. We present observation fusion in subsection 5.1.1 and estimate fusion in subsection 5.1.2.

5.1.1. Observation Fusion

With the help of the above discussion and (6), we establish the fusion of the local observation variables. Let the entries of the $n_l \times 1$ reduced observation vector, $\mathbf{i}_k^{(l)}$, at sensor l, be subscripted by the n_l state variables modeled at sensor l. In the context of our example system in (17) and figure 1, we have

$$\mathbf{i}_k^{(1)} = \begin{bmatrix} i_{k,x_1}^{(1)} \\ i_{k,x_2}^{(1)} \end{bmatrix}, \quad \mathbf{i}_k^{(2)} = \begin{bmatrix} i_{k,x_2}^{(2)} \\ i_{k,x_3}^{(2)} \\ i_{k,x_4}^{(2)} \end{bmatrix}, \quad \mathbf{i}_k^{(3)} = \begin{bmatrix} i_{k,x_3}^{(3)} \\ i_{k,x_4}^{(3)} \\ i_{k,x_5}^{(3)} \end{bmatrix}. \tag{18}$$

For each state x_j, the observation fusion is carried out on the sensors attached to this state in the bipartite fusion graph, $\mathcal{B}$. The fused observation vector, for instance at sensor 2 denoted by $\mathbf{i}_{f,k}^{(2)}$, is given by

$$\mathbf{i}_{f,k}^{(2)} = \begin{bmatrix} i_{k,x_2}^{(2)} + i_{k,x_2}^{(1)} \\ i_{k,x_3}^{(2)} + i_{k,x_3}^{(3)} \\ i_{k,x_4}^{(2)} + i_{k,x_4}^{(3)} \end{bmatrix}. \tag{19}$$

Generalizing to the arbitrary sensor l, we may write the entry, $i_{f,k,x_j}^{(l)}$, corresponding to x_j in the fused observation vector, $\mathbf{i}_{f,k}^{(l)}$, as

$$i_{f,k,x_j}^{(l)} = \sum_{s \in \mathcal{G}_j} i_{k,x_j}^{(s)}, \tag{20}$$

where $i_{k,x_j}^{(s)}$ is the entry corresponding to x_j in the reduced observation vector at sensor s, $\mathbf{i}_k^{(s)}$.

Since the communication network on $\mathcal{G}_j$ will not be, in general, all-to-all, an iterative weighted averaging algorithm [16] can be used to compute the fusion in (20) over arbitrarily connected communication networks with only local communication. A similar procedure on the pairs of state variables and their associated subgraphs, $\mathcal{G}_{jm}$, can be implemented to fuse the reduced observation matrices, $\mathbf{I}_k^{(l)}$.

5.1.2. Estimate Fusion

For each state x_j, each sensor $s \in \mathcal{G}_j$ has an estimate, $\widehat{x}_{j,k|k}^{(s)}$. If $\mathcal{G}_j$ contains more than one sensor, there are multiple correlated estimates of the same state, which should be fused in order to obtain an estimate with smaller variance, for vector extensions, see e.g., [10]. At each sensor $s \in \mathcal{G}_j$, let $\pi_j^{(s)}$ be the variance of the jth state estimate, $\widehat{x}_{j,k|k}^{(s)}$, where $\pi_j^{(s)}$ is a diagonal element in the local estimation error covariance matrix, $\mathbf{S}_{k|k}^{(s)}$, at sensor s. We fuse the estimates using the parallel fusion of estimates formula, which can be derived by using Lagrange multipliers,

$$\widehat{x}_{j,k|k} = \left(\sum_{s \in \mathcal{G}_j} \left(\pi_j^{(s)} \right)^{-1} \right)^{-1} \left(\sum_{s \in \mathcal{G}_j} \left(\pi_j^{(s)} \right)^{-1} \widehat{x}_{j,k|k}^{(s)} \right). \tag{21}$$

Both sums in (21) can be carried out using the weighted averaging algorithm [16].

6. LOCAL PREDICTION STEP

We address the computation of the local prediction information matrix, $\mathbf{Z}_{k|k-1}^{(l)}$, first. Following (9a), it can be shown that the local prediction error covariance matrix, $\mathbf{\Pi}_{k|k-1}^{(l)}$, is given by

$$\mathbf{\Pi}_{k|k-1}^{(l)} = \mathbf{F}_l \mathbf{Z}_{k-1|k-1}^{-1} \mathbf{F}_l^T + \mathbf{G}^{(l)} \mathbf{Q}^{(l)} \mathbf{G}^{(l)T}. \tag{22}$$

We need to go from a local matrix, $\mathbf{Z}_{k-1|k-1}^{(l)}$, resulting from the local filter step (16), to a global matrix, $\mathbf{Z}_{k-1|k-1}^{-1}$, in order to compute $\mathbf{\Pi}_{k|k-1}^{(l)}$ from (22). This problem cannot be solved for arbitrary symmetric matrices, $\mathbf{Z}_{k-1|k-1}$, since it will require long distance communication and an $n \times n$ matrix inverse, infeasible to implement at sensors. To achieve this, we use a generalized distributed Jacobi (GDJ) algorithm for banded matrix inversion presented elsewhere, which iterates on local matrices and requires local communication. A distributed Jacobi algorithm to solve a single linear system of equations with similar banded structure is presented in [15]. The GDJ we propose here solves matrix inversion (n coupled linear systems of equations), with only local computation and local communication using L-banded theorems from [13].

We need to go from $\mathbf{\Pi}_{k|k-1}^{(l)}$ (a submatrix in $\mathbf{\Pi}_{k|k-1}$) to $\mathbf{Z}_{k|k-1}^{(l)}$ (a submatrix in $\mathbf{Z}_{k|k-1}$), where $\mathbf{\Pi}_{k|k-1} = \mathbf{Z}_{k|k-1}^{-1}$. A solution with only local communication and computations with local matrices is possible, if we approximate the prediction information matrix, $\mathbf{Z}_{k|k-1}$, to be an L-banded matrix $\forall k$, using the L-banded inversion theorem in [12].

The second part is to calculate the local transformed predicted estimate, $\widehat{\mathbf{z}}_{k|k-1}^{(l)}$. Following (9b) and the transformations in section 3, it can be shown that

$$\widehat{\mathbf{z}}_{k|k-1}^{(l)} = \mathbf{T}_l \widehat{\mathbf{z}}_{k|k-1} = \mathbf{T}_l \mathbf{Z}_{k|k-1} \mathbf{F} \widehat{\mathbf{x}}_{k-1|k-1}, \tag{23}$$

where the product $\mathbf{T}_l \mathbf{Z}_{k|k-1}$ contains the n_l rows of the prediction information matrix, $\mathbf{Z}_{k|k-1}$, corresponding to the locally modeled states at sensor l. Here we recall that $\mathbf{Z}_{k|k-1}$ is an L-banded matrix, which helps us in computing (23) locally. Notice that if we do not use this assumption than the computation in (23) will require a linear combination of arbitrary estimated states in $\widehat{\mathbf{x}}_{k-1|k-1}$, which will require long distance communication. Depending on the value of L,

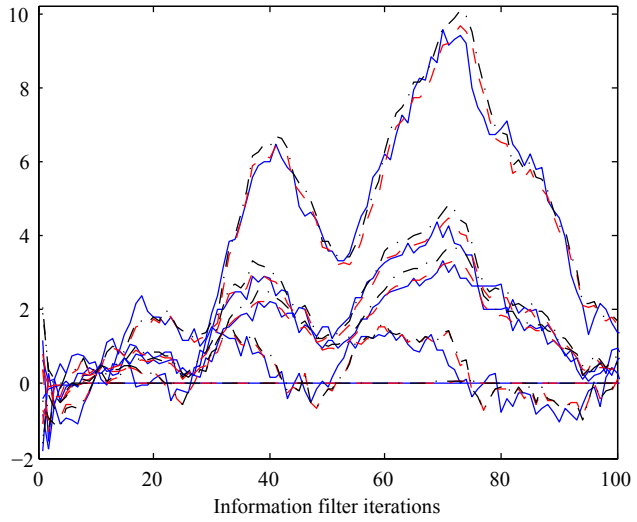

Fig. 2. Simulation Results: We show the original state variables (solid/blue) and their estimates using the the proposed scheme (dash-dot/black), which are virtually indistinguishable from the optimal estimates using the centralized Information filter (dashed/red).

$\mathbf{T}_l\mathbf{Z}_{k|k-1}$ will pick a linear combination of the entries in the vector $\mathbf{F}\hat{\mathbf{x}}_{k-1|k-1}$. Since the model matrix, $\mathbf{F}$, has a localized/spare structure, the communication required will always be local. The communication might be multi-hop and will depend on the value of L.

7. RESULTS AND COCLUSIONS

We simulate a 5 dimensional system monitored by 3 sensors. We implement the *proposed scheme*, with $L = 1$-banded approximation, and compare its performance with the centralized Information filter estimates, figure 2. The original state variables are shown as solid lines (blue). The optimal estimates, computed from the centralized Information filter, are shown as dashed lines (red). The estimates using the proposed scheme are shown as dash-dot lines (black), which are virtually indistinguishable from the centralized estimates (red/dashed). It can also be shown that the local L-banded Information filters and the centralized L-banded Information filters have the same performance. To conclude, we comment on the complexity of the distributed Kalman filter presented in the paper.

The computational complexity of the centralized Information filter is $O(n^3k)$ and for the Information filter with distributed observations in [5] is $O((n^3 + f(n)t_c)k)$, where t_c is the number of iterations required for the consensus algorithm [5] (with complexity, say $O(f(n))$) to converge. The computational complexity for the proposed scheme is $O((n_l^3 + n_l^3 t_{J_1} + n_l^2 t_{J_2} + n_l^2 t_w)k)$, where t_{J_1}, t_{J_2} and t_w are the iterations required for DJV, GDJ, and weighted averaging algorithm to converge, respectively. Even for the toy-example simulated here, where $n = 5$ and $n_1 = 3, n_2 = 3$ and $n_3 = 2$, the computational advantage at the sensors is evident. In this example, typical values for t_{J_1}, t_{J_2} and t_w (we used local degree weights for observation fusion in [16] with a sensor communication network $1 \leftrightarrow 2 \leftrightarrow 3$, other techniques e.g., semi-definite programming [16] can be used to decrease t_w significantly) are $6, 7$ and 9 respectively. The convergence rate of the iterative algorithms can be increased by optimizing the communication network topology [17].

8. REFERENCES

[1] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice Hall, Englewoods Cliff, NJ, 1979.

[2] J. L. 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, Feb 1979.

[3] H. R. Hashemipour, S. Roy, and A. J. Laub, "Decentralized structures for parallel Kalman filtering," *IEEE Transactions on Automatic Control*, vol. 33, no. 1, pp. 88–94, Jan. 1988.

[4] B.S. Rao and H. Durrant-Whyte, "Fully decentralized algorithm for multisensor Kalman filtering," in *IEEE Proceedings-Control Theory and Applications*, Sep. 1991, vol. 138, pp. 413–420.

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

[6] M. Alanyali and V. Saligrama, "Distributed Tracking in Multi- Hop Networks with Communication Delays and Packet Losses," in *IEEE Workshop on Statistical Signal Processing*, 2005, pp. 1190–1195.

[7] T. H. Chung, V. Gupta, J. W. Burdick, and R. M. Murray, "On a Decentralized Active Sensing Strategy using Mobile Sensor Platforms in a Network," in *Proc. of the IEEE Conf. on Decision and Control*, Paradise Island, Bahamas, Dec 2004.

[8] T. Berg and H. Durrant-Whyte, "Model distribution in decentralized multi-sensor data fusion," Tech. Rep., University of Oxford, 1990.

[9] A. G. O. Mutambara, *Decentralized Estimation and Control for multisensor systems*, CRC Press, Boca Raton, FL, 1998.

[10] A. Willsky, M. G. Bello, D. A. Castanon, B. C. Levy, and 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.

[11] A. George and J. Liu, *Computer Solution of Large Sparse Positive Definite Systems*, Prentice Hall, 1981.

[12] A. Kavcic and J. Moura, "Matrices with Banded Inverses: Inversion Algorithms and Factorization of Gauss-Markov Processes," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1495–1509, Jul. 2000.

[13] A. Asif and J. Moura, "Inversion of Block matrices with L-Block Banded Inverse," *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 630–642, Feb. 2005.

[14] Usman A. Khan and José M. F. Moura, "Model Distribution for Distributed Kalman Filters: A Graph Theoretic Approach," Submitted, 5 pages, Jun. 2007.

[15] V. Delouille, R. Neelamani, and R. Baraniuk, "Robust Distributed Estimation using the Embedded Subgraphs Algorithm," *IEEE Transactions on Signal Processing*, vol. 54, no. 8, pp. 2998–3010, Aug 2006.

[16] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," in *Systems and Controls Letters*, 2004, pp. 65–78.

[17] S. Aldosari and J. Moura, "Distributed Detection in Sensor Networks: Connectivity Graph and Small World Networks," in *The 39 th Asilomar Conference on Signals, Systems, and Computers*, Oct. 2005.