Distributed Kalman Estimation with Decoupled Local Filters

Damián Marelli^{a,b}, Tianju Sui^c, Minyue Fu^d

^aSchool of Automation, Guangdong University of Technology, Guangzhou, China.

^b French-Argentinean International Center for Information and Systems Sciences, National Scientific and Technical Research Council, Rosario 2000, Argentina.

^cSchool of Control Science and Engineering, Dalian University of Technology, Dalian, China.

^dSchool of Electrical Engineering and Computing, University of Newcastle, Callaghan, NSW 2308, Australia.

Abstract

We study a distributed Kalman filtering problem in which a number of nodes cooperate without central coordination to estimate a common state based on local measurements and data received from neighbors. This is typically done by running a local filter at each node using information obtained through some procedure for fusing data across the network. A common problem with existing methods is that the outcome of local filters at each time step depends on the data fused at the previous step. We propose an alternative approach to eliminate this error propagation. The proposed local filters are guaranteed to be stable under some mild conditions on certain global structural data, and their fusion yields the centralized Kalman estimate. The main feature of the new approach is that fusion errors introduced at a given time step do not carry over to subsequent steps. This offers advantages in many situations including when a global estimate in only needed at a rate slower than that of measurements or when there are network interruptions. If the global structural data can be fused correctly asymptotically, the stability of local filters is equivalent to that of the centralized Kalman filter. Otherwise, we provide conditions to guarantee stability and bound the resulting estimation error. Numerical experiments are given to show the advantage of our method over other existing alternatives.

Key words: Kalman filters, networked control systems, sensor networks, estimation theory, statistical analysis, stability analysis.

1 Introduction

A networked system consists in a collection of nodes (or sub-systems), connected via a communication network, executing certain processing task [1]. The processing is called distributed if it is carried out by a cooperative strategy among nodes without central coordination [2]. The design of distributed methods aims at minimizing the amount of computation and communication required by each node, as well as making these requirements scalable in the number of nodes. Distributed methods are available for parameter estimation [3,4], Kalman filtering [5], control [6,7], optimization [8], etc. A Kalman filter gives the optimal maximum a posteriori estimation of the state for linear systems with Gaussian noises. This is done by alternating two steps called prediction and update. A major division among distributed Kalman filtering methods is based on whether all nodes estimate the full system state [9], or each node only estimates a subset of the state variables [10,11,12,13,14,15]. This work concerns with methods of the first type. Generally speaking, all methods of this type assume that nodes know the state transition equation. This permits that the prediction step is locally executed at each node. The challenge then consists in how to distributedly execute the update step. Most available methods do so by making use of the information form of the Kalman filter. This requires computing two quantities called the *infor*mation vector and information matrix, the former involving the fusion of measured signals at different nodes and the latter involving the fusion of structural data of the sub-systems. We broadly classify the available methods in two categories.

Email addresses: damian.marelli@newcastle.edu.au (Damián Marelli), suitj@mail.dlut.edu.cn (Tianju Sui), minyue.fu@newcastle.edu.au (Minyue Fu).

¹ This work was supported by the Argentinean Agency for Scientific and Technological Promotion (PICT- 201-0985) and by the National Natural Science Foundation of China (Grant Nos. 61633014, 61803101 and U1701264).

In the first category the information vector and matrix are formed by adding, using different communication schemes, partial components from all nodes of the network. An early method was proposed in [16], which requires full connectivity among all nodes. This restriction was overcome in [17] by using dynamic consensus [18] to fuse information across the network. The same method was refined in [19] by using different consensus stages for fusing information vectors and matrices. In [20] accuracy was improved, at the expense of extra communication, by adding consensus sub-iterations between every two sample times. A variant of this method was proposed [21] and analyzed in [22], where two parallel consensus stages are run for each, information vectors and matrices. A different variant was proposed in [23], where a particular kind of dynamic consensus was used guaranteeing convergence on the time-varying information vectors if certain assumptions are met. In [24,25], fusion of information vectors was done by representing them using a state-space model and estimated them using a distributed Kalman filter of the second category. In [26] and [27] the fusion scheme was complemented by using the covariance intersection method [28] to fuse the outcomes of the prediction steps from each node. Finally, in [29] fusion was done by using a message passing algorithm, rather than a form of consensus, with the advantage of finite-time convergence in the case of an acyclic communication network graph.

In the second category, the fusion of information vectors and matrices used in methods of the first category is complemented by fusion of Kalman estimates. This approach was proposed in [30]. Its optimal design was studied in [31] and its performance analyzed in [32]. A recent improvement of this method was proposed in [33], by using dynamic consensus to fuse information vectors and matrices. In [34] the design was approached by proposing a particular structure, with free parameters, which are optimized to minimize the estimation error. A similar approach was later considered in [35] using a more general structure and setup. Finally, in [36], fusion of information vectors is eliminated and only Kalman estimates are fused.

Broadly speaking, all methods from the two categories described above require carrying out two kinds of data fusion. The first one aims to fuse information associated with the parameters of the measurement equation, typically to form the global information/covariance matrix. We refer to this as *structural data fusion*. The second one aims to fuse information associated to the measurements locally acquired at each node, typically to form the global information vector. We refer to this as *signal fusion*. In the case of time-invariant measurement equations, structural data fusion needs to be done once, possibly during initialization. Also, even in the time-varying case, the rate of change of this data is typically slow, and can be easily tracked using dynamic consensus with limited communications. In these cases, struc-

tural data fusion can be done with negligible error. On the other hand, the change of measurements across time steps is typically much faster than that of the measurement equation. This requires a signal fusion stage with more communications.

A common property of the all the available methods described above is that information fusion needs to be carried out at every Kalman update step, because its result is needed for the subsequent Kalman prediction and update steps. Due to the large communication demands associated to this stage, it is often done approximately. The resulting approximation error then propagates, in the sense that it affects subsequent steps. This *signal fusion error propagation* leads to deviations between the estimates produced by the centralized Kalman filter and those of their distributed counterparts, which accumulate across time steps.

To overcome signal fusion error propagation, in this work we propose an alternative method which avoids this drawback. In the proposed method, each node runs a local estimator which does not require signal fusion. Obviously, none of these local estimators can produce the global Kalman estimate, since they only use local measurement information. However, they have the property that the global Kalman estimate is obtained by fusing their local estimates. In this way, the proposed scheme avoids the aforementioned signal fusion error propagation problem. For this reason, it is in our view a proper generalization of a Kalman filter to a distributed setting. Apart from avoiding the accuracy problems resulting from signal fusion error propagation, the proposed scheme is advantageous in applications where a global estimate is required at a rate slower than the one at which measurements are acquired. This is because information fusion needs only be done at the slower rate. Also, in the case of unreliable communications, where fusion cannot be done during certain periods, the proposed scheme immediately recovers without errors after communications resume.

An additional property of the proposed method is that, provided that structural data fusion is accurately done, the stability of each local estimator is equivalent to that of the centralized Kalman filter. However, stability can be lost if the structural data fusion is done with significant errors. We do a stability analysis in which we provide a bound on the structural data fusion error that guarantees stability. We also bound the difference between the distributed state estimate and the centralized Kalman estimate due to both, structural data and signal fusion errors.

The rest of the paper is organized as follows. In Section 2 we describe the research problem. In Section 3 we give an overview of the available approaches for distributed Kalman filtering and point out their common

drawback that motivates our work. In Section 4 we introduce the proposed distributed Kalman filtering scheme addressing the aforementioned drawback. In Section 5 we present our stability and accuracy analysis results and in Section 6 we derive their proofs. In Section 8 we give experimental evidence of our claims. Concluding remarks are given in Section 9. For ease of readability, the proofs of some auxiliary results appear in the Appendix.

2 Problem description

Notation 1 For a vector x, ||x|| denotes its 2-norm and for a matrix X, ||X|| denotes its operator norm. We use $\mathbb{S}_N(\mathbb{R}) \subset \mathbb{R}^{N \times N}$ to denote the set of real symmetric $N \times N$ matrices, and $\mathbb{P}_N(\mathbb{R}) \subset \mathbb{R}^{N \times N}$ to denote the set of real positive definite $N \times N$ matrices. Also, $\operatorname{col}(x_1, \dots, x_I)$ denotes the column vector formed by stacking the symbols (either vectors or matrices) x_i , $i = 1, \dots, I$, and $\operatorname{diag}(x_1, \dots, x_I)$ denotes the diagonal matrix with the same symbols on its main diagonal. We use $\mathbf{1}_N$ to denote the N-dimensional column vector filled with ones, \mathbf{I}_N to denote the N-dimensional identity matrix and \otimes to denote the Kronecker product. For a symbol Ξ^i we use the handy notation $\Xi^{-i} \triangleq (\Xi^i)^{-1}$ and $\Xi^{i\top} \triangleq (\Xi^i)^{\top}$.

Consider a random vector sequence described by the following recursions

$$x_t = Ax_{t-1} + w_t, \tag{1}$$

where $\mathbb{R}^N \ni x_0 \sim \mathcal{N}(\mu, P)$ and $w_t \sim \mathcal{N}(0, Q)$, with $P, Q \in \mathbb{P}_N(\mathbb{R})$. We assume that we have I nodes acquiring measurements from x_t . In order to model moving nodes, we assume that their associated measurement equations are time-varying, i.e., at time step t, node i measures

$$y_t^i = C_t^i x_t + v_t^i, (2)$$

with $v_t^i \sim \mathcal{N}(0, R_t^i)$, $R_t^i \in \mathbb{P}_M(\mathbb{R})$. We assume that the set $\{x_0, w_t, v_t^i : t \in \mathbb{N}, i = 1, \cdots, I\}$ is statistically mutually independent.

Nodes are communicated via a consensus network. We assume that between every two consecutive time steps t and t+1, there are K communication cycles. In order to model a time-varying connection topology, at time $t \in \mathbb{N}$ and cycle $k \in \{1, \dots, K\}$, node i can send messages to its neighbors $\mathcal{N}_{t,k}^i \subseteq \{1, \dots, I\}$. The communication link from node i to node $j \in \mathcal{N}_{t,k}^i$ has gain $w_{t,k}^{j,i}$. The gains are such that the communication graph is undirected, i.e., $w_{t,k}^{i,j} = w_{t,k}^{j,i}$. We also assume that the adjacency matrix $W_{t,k} = \begin{bmatrix} w_{t,k}^{i,j} \end{bmatrix}_{i,j=1}^{I}$ satisfies $\lim_{K \to \infty} \lambda_2 \left(W_{t,K} \times \cdots \times W_{t,1} \right) = 0$,

where
$$\lambda_2(X)$$
 denotes the algebraic connectivity of matrix X , i.e., the second largest eigenvalue. This guarantees that, for any $x_{t,0} = \begin{bmatrix} x_{t,1}^i, \cdots, x_{t,1}^I \end{bmatrix}^\top \in \mathbb{R}^I$, the sequence generated by $x_{t,k} = W_{t,k}x_{t,k-1}$ satisfies

$$\lim_{t \to \infty} x_{t,k} = \mathbf{1}_I \otimes \frac{1}{I} \sum_{i=1}^I x_0^i.$$

Writing (2) in block form we obtain

k

$$y_t = C_t x_t + v_t, \tag{3}$$

where $v_t \sim \mathcal{N}(0, R_t)$ and

$$y_t = \operatorname{col} \left(y_t^1, \cdots, y_t^I \right),$$

$$v_t = \operatorname{col} \left(v_t^1, \cdots, v_t^I \right),$$

$$C_t = \operatorname{col} \left(C_t^1, \cdots, C_t^I \right),$$

$$R_t = \operatorname{diag} \left(R_t^1, \cdots, R_t^I \right).$$

A research challenge consists in deriving a distributed method for running a Kalman filter on the system (1)-(3). As mentioned in Section 1, a number of method are available for doing so. In Section 3 we give an overview of these methods and point out their common drawback. In Section 4 we propose a method which avoids this drawback.

3 Overview of available distributed methods

In this section we briefly summarize available approaches for distributed Kalman filtering. Let

$$x_{t+1|t}^{i} = A x_{t|t}^{i}, (4)$$

$$\Sigma_{t+1|t}^{i} = A \Sigma_{t|t}^{i} A^{\top} + Q, \qquad (5)$$

$$x_{t|t} = \Sigma_{t|t} \left(\Sigma_{t|t-1}^{-1} x_{t|t-1} + C_t^{\top} R_t^{-1} y_t \right), \quad (6)$$

$$\Sigma_{t|t} = \left(\Sigma_{t|t-1}^{-1} + C_t^\top R_t^{-1} C_t\right)^{-1}, \tag{7}$$

denote the centralized Kalman filter equations, where the update step is expressed in information form, and $x_{t|s}^i$ and $\Sigma_{t|s}^i$ denote the approximations obtained at node *i*. All methods assume that the number *I* of nodes is known at each node. Notice that it is possible to compute *I* in a distributed manner using the method proposed in [37]. They also assume that all nodes know *A* and *Q* and the initial values $x_{0|0} = \mu$ and $\Sigma_{0|0} = P$. Then, at time step *t*, given an update estimate/covariance pair $x_{t|t}^i, \Sigma_{t|t}^i$, the Kalman prediction step can be carried out at each node using (4)-(6). The different methods differ in how the Kalman update step is carried out. In Section 3.1 we describe how this is done in the two method categories mentioned in Section 1. Carrying out this step requires some form of data fusion across nodes. In Section 3.2 we describe the most common options used for doing so. Finally, in Section 3.3 we comment on a common limitation of all available approaches.

3.1 Distributed Kalman update step

Let

$$\Psi_t = C_t^{\top} R_t^{-1} C_t = \sum_{i=1}^{I} C_t^{i\top} R_t^{-i} C_t^i, \qquad (8)$$

$$\psi_t = C_t^{\top} R_t^{-1} y_t = \sum_{i=1}^{I} C_t^{i\top} R_t^{-i} y_t^i.$$
(9)

We refer to ψ_t and Ψ_t as the (global) signal and structural data, respectively. In view of (9) and (8), the signal ψ_t and structural data Ψ_t can be made available at each node using some kind of data fusion. The fusion stage yields, at each node *i*, estimates ψ_t^i and Ψ_t^i of ψ_t and Ψ_t , respectively. The different available methods depend on how, using ψ_t^i and Ψ_t^i , the update step (6)-(7) is computed at each node. We describe below how this is done in the aforementioned two categories:

3.1.1 Consensus on global signal data

Using again any form of data fusion, an approximation ψ_t^i of ψ_t can be obtained at each node *i*. Using this approximation, in [20,21,22,23,26,27] $x_{t|t}^i$ is obtained using (6), i.e.,

$$x_{t|t}^{i} = \Sigma_{t|t}^{i} \left(\Sigma_{t|t-1}^{-i} x_{t|t-1}^{i} + \psi_{t}^{i} \right).$$
 (10)

Alternatively, the Kalman gain

$$K_t = \Sigma_{t|t} C_t^\top R_t^{-1},$$

is used in [17,19] to compute

$$\begin{aligned} x_{t|t}^{i} &= x_{t|t-1}^{i} + K_{t} \left(y_{t} - C_{t} x_{t|t-1}^{i} \right) \\ &= x_{t|t-1}^{i} + \Sigma_{t|t} \left(C_{t}^{\top} R_{t}^{-1} y_{t} - C_{t}^{\top} R_{t}^{-1} C_{t} x_{t|t-1}^{i} \right) \\ &= x_{t|t-1}^{i} + \Sigma_{t|t} \left(\psi_{t} - \Psi_{t} x_{t|t-1}^{i} \right) \\ &\simeq x_{t|t-1}^{i} + \Sigma_{t|t}^{i} \left(\psi_{t}^{i} - \Psi_{t}^{i} x_{t|t-1}^{i} \right). \end{aligned}$$
(11)

3.1.2 Consensus on global signal data and estimates

In order to help the estimates in (11) to converge to a common value, in [30,31,32,33], an extra term penalizing

inter-node mismatches is added. This leads to

$$x_{t|t}^{i} = x_{t|t-1}^{i} + \Sigma_{t|t}^{i} \left(\psi_{t}^{i} - \Psi_{t}^{i} x_{t|t-1}^{i} \right) + D_{t} \sum_{j \in \mathcal{N}_{i}} \left(x_{t|t-1}^{j} - x_{t|t-1}^{i} \right), \qquad (12)$$

where matrix D_t is a free parameter that needs to be designed. In particular, the choice $D_t = I - \sum_{t|t}^i \Psi_t^i$ is implicitly made in [33].

3.2 Information fusion using consensus

In this section we describe the different data fusion methods used in the distributed Kalman filtering literature. These methods apply to the fusion of both, global signal data ψ_t and global structural data Ψ_t . We describe then for fusing signal data. Its application to the fusion of structural data is straightforward.

3.2.1 Local fusion of neighbor data

In [24,25,30,31,32,35,36], ψ_t^i is built by using only data from neighbor nodes. More precisely, they assume that K = 1, i.e., there is a single communication cycle between consecutive time steps. Let

$$\mathring{\psi}_t^i = \left(C_t^i\right)^\top \left(R_t^i\right)^{-1} y_t^i. \tag{13}$$

Then

$$\psi_t^i = I \sum_{j=1}^I w_{t,1}^{i,j} \dot{\psi}_t^i.$$

3.2.2 Global fusion using consensus

In [20,21,22,34], the fusion is done using K > 1 consensus iterations, i.e., they run the following recursions

$$\psi_{t,k}^{i} = \sum_{j=1}^{I} w_{t,k}^{i,j} \psi_{t,k-1}^{j}, \qquad (14)$$

initialized by $\psi_{t,0}^j = I \dot{\psi}_t^i$. The fused data is then $\psi_t^i = \psi_{t,K}^i$, i.e., the one yield after K cycles.

3.2.3 Local fusion using dynamic consensus

In [17,19,23,33], fusion is done using dynamic consensus. More precisely, they assume K = 1 and the fused local data ψ_t^i is computed by modifying its previous value ψ_{t-1}^i with an update term, i.e.,

$$\psi_t^i = \sum_{j=1}^{I} w_{t,1}^{i,j} \left[\psi_{t-1}^j + I \dot{\psi}_t^j - I \dot{\psi}_{t-1}^j \right].$$

3.2.4 Global fusion using dynamic consensus

The advantage of dynamic consensus is that it leads to an approximation error $\psi_t - \psi_t^i$ that decreases as so does the rate of change of ψ_t . Also, the advantage of using K > 1 consensus iterations is that it also permits reducing this error, at the expense of extra communications. These two advantages can be readily combined to increase accuracy as follows

$$\psi_{t,k}^{i} = \sum_{j=1}^{I} w_{t,k}^{i,j} \psi_{t,k-1}^{j}, \qquad (15)$$

initialized by

$$\psi_{t,0}^{j} = \psi_{t-1,K}^{j} + I \dot{\psi}_{t}^{j} - I \dot{\psi}_{t-1}^{j}, \qquad (16)$$

The fused data is then $\psi_t^i = \psi_{t,K}^i$. This is the fusion method that we use in this work.

3.3 Common drawback of all available methods

The methods described above require running two fusion stages for computing (8) and (9). The first one computes the global structural data Ψ_t . Since Ψ_t is structural data, it is often time-invariant or its change from one time step to the next one in typically slow. In the former case, it can be readily computed during initialization using some fusion mechanism. Otherwise, we can track its slow evolution using dynamic consensus with a relatively small number K of consensus iterations. On the other hand, the second consensus stage computes ψ_t . Since this quantity depends on the measurements y_t , its change across time steps is typically much faster than that of Ψ_t . This requires using consensus with a larger value of K to make an accurate estimate ψ_t^i of ψ_t available at each node. A common feature of the available methods described above is that the signal fusion error incurred in the estimation ψ_t^i is carried over to the next time step. This requires that the estimation of ψ_t is accurately done at each time step, using a large number of consensus iterations K, even if an estimate $x_{t|t}$ is not required at that step. In the next section we propose an alternative distributed method which avoids this drawback.

4 Proposed distributed method

In this section we describe the proposed distributed Kalman filtering method. The covariance prediction and update steps are carried out using (5) and (7), as in the methods described in Section 3. For the state estimate, suppose that

$$x_{t-1|t-1} = \sum_{i=1}^{I} \xi_{t-1|t-1}^{i},$$

for some $\xi_{t-1|t-1}^i$, $i = 1, \dots, I$, which are only known at node *i*. We then have

$$\begin{aligned} x_{t|t} &= Ax_{t-1|t-1} + K_t \left(y_t - C_t Ax_{t-1|t-1} \right) \\ &= \left(I - K_t C_t \right) Ax_{t-1|t-1} + K_t y_t \\ &= \sum_{i=1}^{I} \left[\left(I - \Phi_t \right) A\xi_{t-1|t-1}^i + K_t^i y_t^i \right], \end{aligned}$$

where

Letting $K_t^{\top} = \left[\left(K_t^1 \right)^{\top}, \cdots, \left(K_t^I \right)^{\top} \right]$, where for each $i = 1, \cdots, I$, the number of columns of K_t^i equals the dimension of y_t^i , and

 $\Phi_t = K_t C_t.$

$$\xi_{t|t}^{i} = (I - \Phi_{t}) A \xi_{t|t-1}^{i} + K_{t}^{i} y_{t}^{i}, \qquad (17)$$

we obtain

$$x_{t|t} = \sum_{i=1}^{I} \xi_{t|t}^{i}.$$
 (18)

The above means that if we could distributedly compute the structural data Φ_t and K_t^i , then each node could run the local filter (17) without needing to exchange information with its neighbors unless an estimate of $x_{t|t}$ is needed at time step t. We address the distributed computation of Φ_t and K_t^i below.

From the information form of the Kalman filter, we have

$$K_t = \Sigma_{t|t} C_t^\top R_t^{-1}.$$

Hence, K_t^i can be readily computed at each node using

$$K_t^i = \Sigma_{t|t} \left(C_t^i \right)^\top \left(R_t^i \right)^{-1}.$$
(19)

Also,

$$\Sigma_{t|t} = (I - \Phi_t) \Sigma_{t|t-1},$$

leading to

$$\Phi_{t} = I - \Sigma_{t|t} \Sigma_{t|t-1}^{-1} = I - \Sigma_{t|t} \left(\Sigma_{t|t}^{-1} - \Psi_{t} \right) = \Sigma_{t|t} \Psi_{t}.$$
(20)

Hence, Φ_t can be locally computed at each node provided an estimate of Ψ_t is available.

The resulting method then requires a fusion stage to compute an estimate of Ψ_t at each node, and another one for computing $x_{t|t}$ using (18). As we mentioned, we do fusion using the dynamic consensus procedure (15)-(16). We use K_{Ψ} and K_x to denote the number of consensus iterations used to compute Ψ_t and $x_{t|t}$, respectively. The resulting method is summarized in Algorithm 1. **Algorithm 1** Proposed distributed Kalman filtering algorithm.

Initialization: We assume that, for each $i \in \{1, \dots, I\}$, node *i* knows *I*, *A*, *Q* and *P*. Set

$$\xi_{0|0}^{i} = \mu, \quad \Sigma_{0|0} = P \quad \text{and} \quad \Psi_{0}^{i} = 0.$$

Main iterations: At each $t \in \mathbb{N}$, we assume that node i knows C_t^i , R_t^i and y_t^i . (1) **Prediction:**

$$\Sigma_{t|t-1}^{i} = A \Sigma_{t-1|t-1}^{i} A^{\top} + Q, \qquad (21)$$

$$\xi_{t|t-1}^i = A\xi_{t-1|t-1}^i. \tag{22}$$

(2) Structural data fusion: For $k = 1, \dots, K_{\Psi}$, run

$$\Psi_{t,k}^{i} = \sum_{j=1}^{I} w_{t,k}^{i,j} \Psi_{t,k-1}^{j}, \qquad (23)$$

initialized by

$$\Psi_{t,0}^{j} = \Psi_{t-1}^{j} + I \mathring{\Psi}_{t}^{j} - I \mathring{\Psi}_{t-1}^{j},$$

where

$$\mathring{\Psi}_t^j = C_t^{j\top} R_t^{-j} C_t^j.$$
(24)

Upon completion set $\Psi_t^i = \Psi_{t,K\Psi}^i$. In the timeinvariant case, run this step only at t = 1.

(3) Update:

$$\Sigma_{t|t}^{i} = \left(\Sigma_{t|t-1}^{-i} + \Psi_{t}^{i}\right)^{-1}, \qquad (25)$$

$$K_t^i = \Sigma_{t|t}^i C_t^{i+} R_t^{-i}, (26)$$

$$\Phi_t^i = \Sigma_{t|t}^i \Psi_t^i, \tag{27}$$

$$\xi_{t|t}^{i} = \left(I - \Phi_{t}^{i}\right)\xi_{t|t-1}^{i} + K_{t}^{i}y_{t}^{i}.$$
 (28)

(4) **Signal fusion:** If an estimate is required at t, then for $k = 1, \dots, K_x$, run

$$x_{t|t,k}^{i} = \sum_{j=1}^{I} w_{t,k}^{i,j} x_{t|t,k-1}^{j}, \qquad (29)$$

initialized by

$$x_{t|t,0}^{j} = x_{s|s}^{j} + I\xi_{t|t}^{j} - I\xi_{s|s}^{j}$$

where s is the previous time an estimate was required. Upon completion set $\xi^i_{t|t} = x^i_{t|t} = x^i_{t|t,K_x}$.

Remark 2 In many applications, the structural data Ψ_t is typically either time-invariant or changes slowly with time in comparison with $x_{t|t}$. We then typically use K_{Ψ} much smaller than K_x . A smaller K_{Ψ} reduces the complexity of the algorithm while allowing us to keep track of slow changes of the structural data.

Remark 3 Let M_i denote dimension of the measurement vector y_t^i at node *i* and $L_i = \max{\{M_i, N\}}$. The complexity of Algorithm 1 is as follows: Each prediction/update step requires $O(L_i^2N)$ multiplications, each structural data fusion stage requires $O(L_i^2M_i)$ and each signal fusion stage $O(N^2)$.

Clearly, if structural data and signal fusions are done without errors, every time t signal fusion occurs, the estimate $x_{t|t}^i$ produced at each node i equals the centralized Kalman estimate $x_{t|t}$. The question then naturally arises as to whether the linear maps $(y_t)_{t\in\mathbb{N}} \mapsto (d_t^i)_{t\in\mathbb{N}}$, where $d_t^i = \xi_{t|t}^i - \sum_{j=1}^I \xi_{t|t}^j$, are stable for each $i = 1, \dots, I$. This is guaranteed by setting $\xi_{t|t}^i = x_{t|t}^i$ at the end of every signal fusion step. This requires running signal fusion steps on a regular basis. However, if the dynamics of local filters (22), (28) (equivalently (17)) are stable, this requirements can be dropped. The following result gives conditions guaranteeing this.

Theorem 4 If $\Psi_t^i = \Psi_t$, for all $i = 1, \dots, I$ and $t \in \mathbb{N}$ (*i.e.*, structural data fusion is done without errors), and the centralized Kalman filter $(y_t)_{t\in\mathbb{N}} \mapsto (x_{t|t})_{t\in\mathbb{N}}$ is stable, then the local filters $(y_t^i)_{t\in\mathbb{N}} \mapsto (\xi_{t|t}^i)_{t\in\mathbb{N}}$ are stable.

PROOF. Since $\Psi_t^i = \Psi_t$, it follows from (25), (27) and (21) that $\Phi_t^i = \Phi_t$. Therefore, from (26), (28) and (22), that

$$\xi_{t|t}^{i} = (I - \Phi_{t}) A \xi_{t-1|t-1}^{i} + \Sigma_{t|t} C_{t}^{i\top} R_{t}^{-i} y_{t}^{i}.$$

Hence, the dynamics of each local filter are determined by the matrix $(I - \Phi_t) A$. The result then follows since this is also the matrix that determines the dynamics of the centralized Kalman filter. \Box

On the other hand, if errors are introduced at the structural data fusion stage, they will affect local filter dynamics by introducing errors in the recursions (21)-(28). This in turn raises a question about which error tolerance can be allowed at the structural data fusion stage so as to preserve the stability of the local filters $(y_t^i)_{t\in\mathbb{N}} \mapsto (\xi_{t|t}^i)_{t\in\mathbb{N}}$, as well as that of the mismatch map $(y_t)_{t\in\mathbb{N}} \mapsto (x_{t|t}^i - x_{t|t})_{t\in\mathbb{N}}$ between the estimates produced at each node and the centralized Kalman one. We address these two questions in the next section.

5 Stability and accuracy analysis

In this section we study the accuracy requirements in the structural data fusion stage to guarantee the stability of local filters. We also derive a bound on the mismatch between the estimates $x_{t|t}^i$ produced at each node and the centralized Kalman estimate $x_{t|t}$, caused by errors introduced at both fusion stages.

Notation 5 We use $\tilde{\Psi}_t^i = \Psi_t^i - \Psi_t$ to denote the error introduced at each node by the structural data fusion stage. We use the same notation, e.g., $\tilde{\Sigma}_{t|s}^i = \Sigma_{t|s}^i - \Sigma_{t|s}$ and $\tilde{\Phi}_t^i = \Phi_t^i - \Phi_t$ for the resulting errors introduced in the values of $\Sigma_{t|s}^i$ and Φ_t^i , respectively. We also use $\xi_{t|s}^i$ to denote the value of $\xi_{t|s}^i$ that would result if no errors were introduced at the structural data fusion stage, and define $\tilde{\xi}_{t|s}^i = \xi_{t|s}^i - \xi_{t|s}^i$.

Let

$$\bar{\tilde{\psi}} = \sup_{\substack{t \in \mathbb{N} \\ 1 \leq i \leq I}} \left\| \tilde{\Psi}_t^i \right\|,$$

be a bound on the error introduced at all structural data fusion stages. Our first result states a sufficient condition on $\bar{\psi}$ to guarantee the stability of all local filters.

Theorem 6 Let $\bar{\sigma} = \sup_{t \in \mathbb{N}} \left\| \Sigma_{t|t} \right\|$ and

$$\bar{\gamma} = \sup_{t \in N} \sum_{s=1}^{t} \| (I - \Phi_{t-1}) A \times \dots \times (I - \Phi_s) A \|$$

$$\beta = \operatorname{sol}_b \left\{ b + \log \bar{\sigma} \|A\|^2 \|Q^{-1}\| b = 0 \right\}.$$

If the centralized Kalman filter $(y_t)_{t\in\mathbb{N}} \mapsto (x_{t|t})_{t\in\mathbb{N}}$ is stable, and

$$\bar{\tilde{\psi}} \leq \min\left\{\bar{\sigma}^{-1}\left[1 - \exp\left(-\frac{\beta}{\sqrt{N}}\right)\right], \bar{\gamma}^{-1} \|A\|^{-1}\right\},$$
(30)
then, the local filters $\left(y_t^i\right)_{t\in\mathbb{N}} \mapsto \left(\xi_{t|t}^i\right)_{t\in\mathbb{N}}$ are stable.

Remark 7 Theorem 6 states that, if the error tolerance $\bar{\psi}$ of structural data fusion is smaller than the threshold given in (30), the stability of local filters is equivalent to that of the centralized Kalman filter. Notice that, if measurement equations are time-invariant, so is the structural data, i.e., $\Psi_t = \Psi$, for all $t \in \mathbb{N}$. Hence, arbitrarily accurate structural data fusion can be guaranteed, either during an initialization phase, or asymptotically at running time. In this case, stability of local filters is simply

equivalent to that of the centralized Kalman filter. Also notice that the required boundness of $\bar{\gamma}_t$ is equivalent to the stability of the centralized Kalman filter.

Let $\mathbf{x}_{t|t} = \mathbf{1}_I \otimes x_{t|t}$ denote a vector with I copies of the centralized Kalman estimate $x_{t|t}$, and $\hat{\mathbf{x}}_{t|t} =$ $\operatorname{col}\left(x_{t|t}^1, \cdots, x_{t|t}^I\right)$ denote the vector of estimates produced by each node. Let also $\check{\mathbf{x}}_{t|t} = \mathbf{1}_I \otimes \check{x}_{t|t}$, where

$$\check{x}_{t|t} = \sum_{i=1}^{I} \xi_{t|t}^{i},$$

denotes the estimate that would be obtained at all nodes if no error were introduced a the signal fusion stage. Our second result bounds the covariance of the error $\tilde{\mathbf{x}}_{t|t} \triangleq \mathbf{x}_{t|t} - \hat{\mathbf{x}}_{t|t}$. This bound depends on two terms. The first one depends on the error $\bar{\psi}$ introduced at the structural data fusion stage and the second one depends on the error $\|\tilde{\mathbf{x}}_{t|t} - \hat{\mathbf{x}}_{t|t}\|$ introduced at the signal fusion stage.

Lemma 8 Let $\bar{v}\left(\tilde{\psi}\right) = \sqrt{N} \left|\log\left(1 - \bar{\sigma}\tilde{\psi}\right)\right|$. If (30) holds, then the following equation has at least one solution

$$x = \frac{\bar{\sigma} \|A\|^2 x}{\bar{\sigma} \|A\|^2 + \|Q^{-1}\|^{-1} e^{-x}} + \bar{v} \left(\bar{\psi}\right).$$
(31)

Theorem 9 Let $\mathfrak{y}_t^i = \operatorname{col}\left(\breve{\xi}_{t|t-1}^i, \mathring{\psi}_t^i\right)$ with $\mathring{\psi}_t^i$ given by (13) and

$$\bar{\mathfrak{y}} = \sup_{\substack{t \in \mathbb{N} \\ 1 \leq i \leq I}} \left\| \mathcal{E} \left\{ \mathfrak{y}_t^i \mathfrak{y}_t^{i\top} \right\} \right\|^{1/2}.$$
(32)

Suppose that (30) holds, and let $\bar{\delta}\left(\bar{\psi}\right)$ denote the smallest solution of (31). Then

$$\left\| \mathcal{E}\left\{ \tilde{\mathbf{x}}_{t|t} \tilde{\mathbf{x}}_{t|t}^{\top} \right\} \right\|^{1/2} \le \mathcal{E}\left(\tilde{\psi} \right) + \mathcal{E}\left\{ \left\| \check{\mathbf{x}}_{t|t} - \hat{\mathbf{x}}_{t|t} \right\|^2 \right\}^{1/2},$$
(33)

where

$$\mathscr{E}\left(\bar{\tilde{\psi}}\right) = NI\left(\frac{\bar{\gamma}\bar{\mathfrak{y}}}{1-\bar{\gamma}\bar{\tilde{\psi}}}\right)^2 \left(\bar{\phi}^2\left(\bar{\tilde{\psi}}\right) + \bar{\tilde{\sigma}}^2\left(\bar{\tilde{\psi}}\right)\right),$$

with

$$\bar{\phi}\left(\bar{\tilde{\psi}}\right) = \left[\left(e^{\bar{\delta}\left(\bar{\tilde{\psi}}\right)} - 1\right)\bar{\psi} + e^{\bar{\delta}\left(\bar{\tilde{\psi}}\right)}\bar{\tilde{\psi}}\right]\bar{\sigma}$$
$$\bar{\tilde{\sigma}}\left(\bar{\tilde{\psi}}\right) = \left(e^{\bar{\delta}\left(\bar{\tilde{\psi}}\right)} - 1\right)\bar{\sigma},$$

and $\bar{\psi} = \sup_{t \in \mathbb{N}} \|\Psi_t\|.$

Remark 10 The above result is stated in terms of the bound $\overline{\mathfrak{y}}$. We give in Section 7 details on how to compute this bound.

Remark 11 Notice that the first term $\mathscr{E}\left(\bar{\psi}\right)$ in (33)

depends only on the bound $\tilde{\psi}$ of the structural data fusion error, and the second term is the signal fusion error at each sample time t. These two errors are determined by the numbers K_{Ψ} and K_x of consensus iterations used on each fusion stage. Notice also that, for $\mathscr{E}\left(\tilde{\psi}\right)$ to be bounded, so need to be $\bar{\eta}$ and $\bar{\psi}$.

6 Proofs of the main results

The proofs of Theorems 6 and 9 are given in Section 6.4. Deriving these results requires certain mathematical background, which is introduced in three preceding sections. In Section 6.1 we introduce a Riemannian metric on the differentiable manifold $\mathbb{P}_N(\mathbb{R})$ of positive-definite matrices and state its properties. In Section 6.2 we introduce a convenient algebraic structure on random vectors, namely, a Hibert C^{*}-module. Finally, in Section 6.3 we use this structure to characterize the output covariance of a perturbed linear time-varying (LTV) system.

6.1 A Riemannian metric on $\mathbb{P}_N(\mathbb{R})$

For a given $N \in \mathbb{N}$, the set $\mathbb{P}_N(\mathbb{R})$ of positive-definite matrices can be considered as a differentiable manifold inside $\mathbb{R}^{N \times N}$. We define the following map $\delta : \mathbb{P}_N(\mathbb{R}) \times \mathbb{P}_N(\mathbb{R}) \to [0, \infty)$:

Definition 12 [38, Chapter 6] For $P, Q \in \mathbb{P}_N(\mathbb{R})$ we define

$$\delta(P,Q) = \left\| \log Q^{-1/2} P Q^{-1/2} \right\|_{\mathrm{F}}.$$

It is shown in [38, Chapter 6] that the map δ is a Riemannian metric on $\mathbb{P}_N(\mathbb{R})$. This metric enjoys the properties given in the following proposition, whose proof appears in the appendix.

Proposition 13 For $P, Q \in \mathbb{P}_N(\mathbb{R})$ and $R \in \mathbb{S}_N(\mathbb{R})$:

- (1) $\delta(P^{-1}, Q^{-1}) = \delta(P, Q);$
- (2) for any $W \in \mathcal{P}^M(\mathbb{R})$ and $M \times N$ matrix B, we have

$$\delta\left(W + BPB^{\top}, W + BQB^{\top}\right) \le \frac{\alpha}{\alpha + \beta}\delta\left(P, Q\right)$$

where $\alpha = \max \{ \|BPB^{\top}\|, \|BQB^{\top}\| \}$ and $\beta = \|W^{-1}\|^{-1}$;

- (3) $||P Q|| \le (e^{\delta(P,Q)} 1) \min \{||P||, ||Q||\}.$
- (4) If $\|P^{-1}\| \|R\| < 1$, then

$$\delta(P, P+R) \le \sqrt{N} \left| \log \left(1 - \|P^{-1}\| \|R\| \right) \right|.$$

6.2 A Hilbert C^* -module of random vectors

A Hilbert C^{*}-module is an algebraic structure that offers an elegant and compact way to work with random vectors and their covariance matrices. In this section we very briefly introduce the concepts needed for our analysis. A general treatment of Hilbert C^{*}-modules can be found in [39], and its application to covariance matrices in [40, Section 3.3.1].

Let x and y be $N\text{-dimensional real random vectors. We define the following <math display="inline">\mathbb{R}^{N\times N}\text{-valued inner product}$

$$\left\langle x,y\right\rangle _{\star}=\mathcal{E}\left\{ xy^{\top}\right\} .$$

This inner product induces the following norm on N-dimensional random vectors

$$||x||_{\star} = ||\langle x, x \rangle_{\star}||^{1/2}$$

It is shown in [39, Chapter 1] that $||x||_{\star}$ is indeed a norm. This norm enjoys the following additional property, whose proof appears in the appendix:

Lemma 14 Let x and y be random vectors of the same dimension. Then

$$\left\| \left\langle x, y \right\rangle_{\star} \right\| \le \left\| x \right\|_{\star} \left\| y \right\|_{\star}.$$

6.3 Output covariance of perturbed LTV systems

In this section we use the Hilbert C^{\star}-module structure described in Section 6.2 to bound the output covariance of a perturbed LTV system. Consider the following LTV system

$$x_t = A_{t-1}x_{t-1} + u_t, (34)$$

$$x_0 = 0, \tag{35}$$

with u_t being a possibly colored and non-stationary vector random process. Suppose we have a perturbed version $\hat{A}_t = A_t + \tilde{A}_t$ of the sequence A_t and let \hat{x}_t denote the sequence generated by (34)-(35) when A_t is replaced \hat{A}_t . The following lemma gives a bound on the norm $\|\hat{x}_t\|_{\star}$, in terms of the the non perturbed sequence A_t and a measure of the perturbation $\tilde{A}_t = \hat{A}_t - A_t$. Its proof appears in the appendix.

Lemma 15 Let $A = (A_t)_{t \in \mathbb{Z}}$ and $\hat{A} = (\hat{A}_t)_{t \in \mathbb{Z}}$ be two sequences of square matrices of the same dimension and $\tilde{A}_t = \hat{A}_t - A_t$. Let

$$\hat{x}_t = \hat{A}_{t-1}\hat{x}_{t-1} + u_t,$$

 $\hat{x}_0 = 0.$

Let also $\mu = \sup_{t \in \mathbb{N}} \left\| \tilde{A}_t \right\|, \, \bar{u} = \max_{1 \le s \le t} \left\| u_s \right\|_{\star} \text{ and }$

$$\gamma = \sup_{t \in Z} \sum_{s=1}^{t} \|A_{t-1} \times \dots \times A_s\|.$$

If $\mu\gamma < 1$, then

$$\|\hat{x}_t\|_\star \le \frac{\gamma}{1-\gamma\mu}\bar{u}.$$

6.4 Proofs of the main results

In this section we give the proofs of Theorems 6 and 9. We arrive to them through a sequence of lemmas, whose proofs appear in the appendix. The first lemma gives a bound of the difference between the ideal predicted covariance $\Sigma_{t|t}$ and its approximation $\Sigma_{t|t}^{i}$ at node *i*. This difference is measured using the Riemannian metric δ introduced in Section 6.1.

Lemma 16 If $\left\|\tilde{\Psi}_{t}^{i}\right\| < \left\|\Sigma_{t|t}\right\|^{-1}$, for all $t \in \mathbb{N}$ and $i \in \{1, \dots, I\}$, then

$$\begin{split} &\delta\left(\Sigma_{t|t}^{i}, \Sigma_{t|t}\right) \\ &\leq \frac{\|A\|^{2} \|\Sigma_{t-1|t-1}\| \delta\left(\Sigma_{t-1|t-1}^{i}, \Sigma_{t-1|t-1}\right)}{\|A\|^{2} \|\Sigma_{t-1|t-1}\| + \|Q^{-1}\|^{-1} e^{-\delta\left(\Sigma_{t-1|t-1}, \Sigma_{t-1|t-1}^{i}\right)} \\ &+ \sqrt{N} \left|\log\left(1 - \|\Sigma_{t|t}\| \left\|\tilde{\Psi}_{t}^{i}\right\|\right)\right|. \end{split}$$

The next lemma characterizes the approximation error $\tilde{\xi}^i_{t|t}$ as the output of a perturbed LTV system.

Lemma 17 For all $t \in \mathbb{N}$ and $i \in \{1, \dots, I\}$,

$$\tilde{\xi}^{i}_{t|t} = \left(I - \Phi_{t} - \tilde{\Phi}^{i}_{t}\right) A \tilde{\xi}^{i}_{t-1|t-1} + \left[\tilde{\Phi}^{i}_{t}, \tilde{\Sigma}^{i}_{t|t}\right] \mathfrak{y}^{i}_{t}, \quad (36)$$

where \mathfrak{y}_t^i is defined as in Theorem 9 and

$$\begin{split} \left\| \tilde{\Phi}_{t}^{i} \right\| &\leq \left[\left(e^{\delta \left(\Sigma_{t|t}^{i}, \Sigma_{t|t} \right)} - 1 \right) \left\| \Psi_{t} \right\| \right. \\ &+ e^{\delta \left(\Sigma_{t|t}^{i}, \Sigma_{t|t} \right)} \left\| \tilde{\Psi}_{t}^{i} \right\| \right] \left\| \Sigma_{t|t} \right\|, \\ \left\| \tilde{\Sigma}_{t|t}^{i} \right\| &\leq \left(e^{\delta \left(\Sigma_{t|t}^{i}, \Sigma_{t|t} \right)} - 1 \right) \left\| \Sigma_{t|t} \right\|. \end{split}$$

The following lemma gives a bound on the norm $\left\|\tilde{\xi}_{t|t}^{i}\right\|_{\star}$ of the approximation error $\tilde{\xi}_{t|t}^{i}$ at each node.

Lemma 18 If (30) holds, then

$$\left\|\tilde{\xi}_{t|t}^{i}\right\|_{\star} \leq \frac{\bar{\gamma}\bar{\mathfrak{y}}}{1-\bar{\gamma}\bar{\tilde{\psi}}}\sqrt{\bar{\phi}^{2}\left(\bar{\tilde{\psi}}\right) + \bar{\tilde{\sigma}}^{2}\left(\bar{\tilde{\psi}}\right)}.$$
 (37)

We can now give the proofs of our main results.

PROOF. [of Theorem 6] This is an immediate consequence of Lemma 18. $\hfill \Box$

PROOF. [of Lemma 8] Let

$$a = \frac{\left\|Q^{-1}\right\|^{-1}}{\bar{\sigma}\left\|A\right\|^{2}}, \qquad b = \bar{\upsilon}\left(\bar{\psi}\right),$$

and

$$f(x) = \frac{a}{b}(x-b), \qquad g(x) = e^x.$$

Equation (31) can then be rewritten as

$$f(x) = g(x). \tag{38}$$

Since f is affine and g convex, (38) has either zero, one or two solutions. In order for it to have a single solution, we must have

$$\frac{a}{b} = f'(x) = g'(x) = e^x.$$

Replacing the above into (38) we obtain

$$\log \frac{a}{b} = b$$

or equivalently, $b = \beta$. It then follows that (31) has at least one solution if $b \leq \beta$. It is straightforward to verify that the latter is implied by (30) and the result follows. **PROOF.** [of Theorem 9] We have

$$\begin{aligned} \left\| \mathcal{E} \left\{ \tilde{\mathbf{x}}_{t|t} \tilde{\mathbf{x}}_{t|t}^{\top} \right\} \right\|^{1/2} \\ &= \left\| \mathcal{E} \left\{ \left(\mathbf{x}_{t|t} - \hat{\mathbf{x}}_{t|t} \right) \left(\mathbf{x}_{t|t} - \hat{\mathbf{x}}_{t|t} \right)^{\top} \right\} \right\|^{1/2} \\ &\leq \mathcal{E} \left\{ \operatorname{Tr} \left\{ \left(\mathbf{x}_{t|t} - \hat{\mathbf{x}}_{t|t} \right) \left(\mathbf{x}_{t|t} - \hat{\mathbf{x}}_{t|t} \right)^{\top} \right\} \right\}^{1/2} \\ &= \mathcal{E} \left\{ \sum_{i=1}^{I} \left\| x_{t|t} - \hat{x}_{t|t}^{i} \right\|^{2} \right\}^{1/2} \\ &= \mathcal{E} \left\{ \sum_{i=1}^{I} \left\| x_{t|t} - \tilde{x}_{t|t} + \check{x}_{t|t} - \hat{x}_{t|t}^{i} \right\|^{2} \right\}^{1/2} \\ &\leq \sqrt{I} \mathcal{E} \left\{ \left\| x_{t|t} - \check{x}_{t|t} \right\|^{2} \right\}^{1/2} + \mathcal{E} \left\{ \sum_{i=1}^{I} \left\| \check{x}_{t|t} - \hat{x}_{t|t}^{i} \right\|^{2} \right\}^{1/2}. \end{aligned}$$
(39)

Now, using Lemma 18,

$$\mathcal{E}\left\{\left\|x_{t|t} - \check{x}_{t|t}\right\|^{2}\right\} = \mathcal{E}\left\{\left\|\sum_{i=1}^{I}\xi_{t|t} - \hat{\xi}_{t|t}^{i}\right\|^{2}\right\}$$

$$\leq \sum_{i=1}^{I}\mathcal{E}\left\{\left\|\tilde{\xi}_{t|t}^{i}\right\|^{2}\right\} = \sum_{i=1}^{I}\mathcal{E}\left\{\operatorname{Tr}\left\{\tilde{\xi}_{t|t}^{i}\tilde{\xi}_{t|t}^{i\top}\right\}\right\}$$

$$\leq N\sum_{i=1}^{I}\left\|\mathcal{E}\left\{\tilde{\xi}_{t|t}^{i}\tilde{\xi}_{t|t}^{i\top}\right\}\right\| = N\sum_{i=1}^{I}\left\|\tilde{\xi}_{t|t}^{i}\right\|_{\star}^{2}$$

$$\leq NI\left(\frac{\bar{\gamma}\bar{\mathfrak{y}}}{1 - \bar{\gamma}\bar{\tilde{\psi}}}\right)^{2}\left(\bar{\phi}^{2}\left(\bar{\tilde{\psi}}\right) + \bar{\sigma}^{2}\left(\bar{\tilde{\psi}}\right)\right). \quad (40)$$

The result then follows by putting (40) into (39), and noticing that

$$\mathcal{E}\left\{\sum_{i=1}^{I} \left\| \check{x}_{t|t} - \hat{x}_{t|t}^{i} \right\|^{2}\right\} = \mathcal{E}\left\{ \left\| \check{\mathbf{x}}_{t|t} - \hat{\mathbf{x}}_{t|t} \right\|^{2}\right\}.$$

7 About computing $\bar{\mathfrak{y}}$

Our first step consists in characterizing \mathfrak{y}_t^i as the output of a state-space model. This is done by defining

$$\mathfrak{x}_t^i = \begin{bmatrix} x_t \\ raket{i}_{t|t} \end{bmatrix}$$
 and $\mathfrak{e}_t^i = \begin{bmatrix} w_t \\ v_t^i \end{bmatrix}$.

We can then write

$$\mathbf{\mathfrak{x}}_t^i = F_t^i \mathbf{\mathfrak{x}}_{t-1}^i + G_t^i \mathbf{\mathfrak{e}}_t^i, \tag{41}$$

$$\mathfrak{y}_t^i = H_t^i \mathfrak{x}_t^i + E_t^i \mathfrak{e}_t^i, \tag{42}$$

with

$$\begin{split} F_t^i &= \begin{bmatrix} A & 0 \\ K_t^i C_t^i A & (I - \Phi_t) A \end{bmatrix}, \\ G_t^i &= \begin{bmatrix} I & 0 \\ K_t^i C_t^i & K_t^i \end{bmatrix}, \ H_t^i &= \begin{bmatrix} 0 & A \\ \mathring{\Psi}_t^i & 0 \end{bmatrix}, \ E_t^i &= \begin{bmatrix} 0 & 0 \\ 0 & C_t^{i\top} R_t^{-i} \end{bmatrix}, \end{split}$$

and $\mathring{\Psi}_t^i$ given by (24).

Using the model (41)-(42) we obtain the covariance of \mathfrak{y}_t^i as follows

$$\begin{split} \mathcal{E}\left\{\mathfrak{x}_{t}^{i}\mathfrak{x}_{t}^{i\top}\right\} &= \Pi_{0,t}^{i} \begin{bmatrix} P_{0} & 0\\ 0 & 0 \end{bmatrix} \Pi_{0,t}^{i\top} \\ &+ \sum_{s=1}^{t} \Pi_{s,t}^{i} G_{s}^{i} \begin{bmatrix} Q & 0\\ 0 & R_{t}^{i} \end{bmatrix} G_{s}^{i\top} \Pi_{s,t}^{i\top}, \\ \mathcal{E}\left\{\mathfrak{y}_{t}^{i}\mathfrak{y}_{t}^{i\top}\right\} &= H_{t}^{i} \mathcal{E}\left\{\mathfrak{x}_{t}^{i}\mathfrak{x}_{t}^{i\top}\right\} H_{t}^{i\top} + E_{t}^{i} \begin{bmatrix} Q & 0\\ 0 & R_{t}^{i} \end{bmatrix} E_{t}^{i\top}, \end{split}$$

where $\Pi_{s,t}^i = F_t \times \cdots \times F_{s+1}$. We can then readily compute the bound $\bar{\mathfrak{y}}$ by putting the above into (32).

8 Numerical experiments

In this section we evaluate the performance of our method. For comparison we use one method form each of the two categories described in Section 1. For the first category we consider the method proposed in [20]. We refer to it as Algorithm A. For the second category we consider the method recently proposed in [33], which we refer to as Algorithm B.

For evaluation we use a randomly generated timeinvariant system of order N = 10. Matrices A and Qhave spectral radii $\rho(A) = 0.999$ and $\rho(Q) = 1$, respectively. Also, measurements are one-dimensional, i.e., M = 1, with $C^i \sim \mathcal{N}(0, \mathbf{I}_N)$ and $R^i = 10r^2 + 0.1$, with $r \sim \mathcal{N}(0, 1)$. Nodes are connected via a time-invariant network with ring topology, whose gains are given by

$$w_{t,k}^{i,j} = \begin{cases} 0.5, & i = j, \\ 0.25, & | \mod(i - j + 1, I) - 1| = 1, \\ 0, & \text{otherwise.} \end{cases}$$

This results in an algebraic connectivity of $\lambda_2 = 0.9891$.

As performance index we use the estimation mismatch error defined as

$$e^{2} = \frac{1}{T} \sum_{t=1}^{T} e_{t}^{2}$$
 with $e_{t}^{2} = \frac{1}{I} \sum_{i=1}^{I} \left\| x_{t|t}^{i} - x_{t|t} \right\|^{2}$,



Fig. 1. Error vs number of consensus iterations for structural data fusion.



Fig. 2. Error vs number of consensus iterations for signal fusion.

where $x_{t|t}$ denotes the centralized Kalman estimate.

In the first experiment we evaluate the performance when errors appear in the structural and signal fusion stages. In Figure 1 we show the effect produced by an approximation error in fusing structural data. To this end we use $K_x = 100$ consensus iterations for signal fusion and show the mismatch error as a function of the number K_{Ψ} of consensus cycles used for structural fusion. We see that Algorithm B and the proposed one performs similarly, with a noticeable advantage over Algorithm A. In Figure 2 we use $K_{\Psi} = 100$ iterations for structural fusion and show the mismatch error as a function of the number K_x of consensus iterations used for signal fusion. We again see that Algorithm B and the proposed one performs similarly, with certain advantage over Algorithm A for large values of K_x . We conclude that, when there are no network interruptions, the proposed algorithm performs similarly to the best available ones.

As mentioned, the advantage of the proposed method is that errors in signal fusion do not carry over across time steps. This can be seen in Figure 3, where we simulate a network interruption from sample times t = 20 to t = 25. We use $K_x = K_{\Psi} = 100$. We see that, while the proposed algorithm gives an accurate estimate as soon as connectivity is restored, Algorithms A and B require several time steps to do so. In Figure 4 we show the performance of the algorithms when network availability follows a symmetric Gilbert-Elliott model [41,42] with transition probability p = 0.05. We see how, while the



Fig. 3. Performance under a network interruption.



Fig. 4. Performance under network interruptions following a Gilbert-Elliott model.



Fig. 5. Error vs transition probability p.

proposed algorithm is always able to produce an accurate estimate as soon as network connectivity is available, Algorithms A and B are not able to produce accurate estimates during certain long time periods.

In Figure 5 we show the mismatch error, as a function of the transition probability p. We see that, in this case, the proposed algorithm has a significant advantage over its rivals.

9 Conclusion

We proposed a novel approach for distributed Kalman filtering. The essential difference with existing approaches is that, provided certain global structural data is available at each node, local filters do not require data fusion across the network. The latter is only needed when a global estimation is required. Hence, errors produced by inaccurate fusion do not carry over across time steps. This is advantageous in a number of situations where fusion is not needed or cannot be made at each time step. If global structural data is exactly known at each node, the stability of local filters is equivalent to that of the centralized Kalman filter. Otherwise, we give conditions to guarantee stability and bound the estimation error induced by inaccurate global structural data fusion. We also present numerical experiments showing the advantage of our method over other available alternatives.

A Proofs

PROOF. [of Proposition 13] Let $\sigma_n(X)$ and $\lambda_n(X)$ denote the singular values and eigenvalues of matrix X, respectively. We have

$$\delta(P,Q) = \left\| \log Q^{-1/2} P Q^{-1/2} \right\|_{\mathrm{F}}$$
$$= \sqrt{\sum_{n=1}^{N} \sigma_n^2 \left(\log Q^{-1/2} P Q^{-1/2} \right)}$$
$$= \sqrt{\sum_{n=1}^{N} \lambda_n^2 \left(\log Q^{-1/2} P Q^{-1/2} \right)}$$
$$= \sqrt{\sum_{n=1}^{N} \log^2 \lambda_n \left(Q^{-1/2} P Q^{-1/2} \right)}$$
$$= \sqrt{\sum_{n=1}^{N} \log^2 \lambda_n \left(P Q^{-1} \right)}.$$

Then, $\delta(P,Q)$ equals the distance defined in [43, Definition 1.4]. Hence, Properties 1 and 3 follow from [43], and Property 2 follows from [44, Proposition 6].

For Property 4 we have

$$\delta(P+R,P) = \sqrt{\sum_{n=1}^{N} \log^2 \lambda_n \left(P^{-1/2}(P+R)P^{-1/2}\right)}$$
$$= \sqrt{\sum_{n=1}^{N} \log^2 \lambda_n \left(I + P^{-1/2}RP^{-1/2}\right)}$$
$$= \sqrt{\sum_{n=1}^{N} \log^2 \left(1 + \lambda_n \left(P^{-1/2}RP^{-1/2}\right)\right)}$$

Then

$$\delta(P+R,P) \le \sqrt{\sum_{n=1}^{N} \log^2 \left(1 - \left|\lambda_n \left(P^{-1/2}RP^{-1/2}\right)\right|\right)} \le \sqrt{N} \max_n \left|\log \left(1 - \left|\lambda_n \left(P^{-1/2}RP^{-1/2}\right)\right|\right)\right| = \sqrt{N} \left|\log \left(1 - \max_n \left|\lambda_n \left(P^{-1/2}RP^{-1/2}\right)\right|\right)\right| \le \sqrt{N} \left|\log \left(1 - \left\|P^{-1/2}RP^{-1/2}\right\|\right)\right| \le \sqrt{N} \left|\log \left(1 - \left\|P^{-1/2}RP^{-1/2}\right\|\right)\right| \le \sqrt{N} \left|\log \left(1 - \left\|P^{-1}\right\|\|R\|\right)\right|.$$

PROOF. [of Lemma 14] From [39, Proposition 1.1],

$$\mathcal{E}\left\{xy^{\top}\right\}\mathcal{E}\left\{yx^{\top}\right\} \leq \left\|\mathcal{E}\left\{yy^{\top}\right\}\right\|\mathcal{E}\left\{xx^{\top}\right\}.$$

Then

$$\begin{aligned} \left\| \mathcal{E} \left\{ xy^{\top} \right\} \right\|^2 &= \left\| \mathcal{E} \left\{ xy^{\top} \right\} \mathcal{E} \left\{ yx^{\top} \right\} \right\| \\ &\leq \left\| \mathcal{E} \left\{ xx^{\top} \right\} \right\| \left\| \mathcal{E} \left\{ yy^{\top} \right\} \right\| \end{aligned}$$

and the result follows.

PROOF. [of Lemma 15] We have

$$\hat{x}_{t} = \hat{A}_{t-1}\hat{x}_{t-1} + u_{t}$$

$$= A_{t-1}\hat{x}_{t-1} + \left(\hat{A}_{t-1} - A_{t-1}\right)\hat{x}_{t-1} + u_{t}$$

$$= \sum_{s=1}^{t} \prod_{t,s} \left[\tilde{A}_{s-1}\hat{x}_{s-1} + u_{s}\right],$$

where $\Pi_{t,s} = A_{t-1} \times \cdots \times A_s$. Then

$$\mathcal{E}\left\{x_{t}x_{t}^{\top}\right\} = R_{t}^{(1)} + R_{t}^{(2)} + R_{t}^{(3)} + R_{t}^{(4)},$$

with

$$R_{t}^{(1)} = \sum_{s,r=1}^{t} \Pi_{t,s} \tilde{A}_{s-1} \mathcal{E} \left\{ \hat{x}_{s-1} \hat{x}_{r-1}^{\top} \right\} \tilde{A}_{r-1}^{\top} \Pi_{t,r}^{\top}$$
$$R_{t}^{(2)} = \sum_{s,r=1}^{t} \Pi_{t,s} \tilde{A}_{s-1} \mathcal{E} \left\{ \hat{x}_{s-1} u_{r}^{\top} \right\} \Pi_{t,r}^{\top},$$
$$R_{t}^{(3)} = \left(R_{t}^{(2)} \right)^{\top},$$
$$R_{t}^{(4)} = \sum_{s,r=1}^{t} \Pi_{t,s} \mathcal{E} \left\{ u_{s} u_{r}^{\top} \right\} \Pi_{t,r}^{\top}.$$

Let $\nu_{t-1} = \max_{1 \le s \le t-1} \|\hat{x}_t\|_{\star}$. Then

$$\begin{aligned} & \left\| R_{t}^{(1)} \right\| \\ \leq & \sum_{s,r=1}^{t} \left\| \Pi_{t,s} \right\| \left\| \tilde{A}_{s-1} \right\| \left\| \langle x_{r-1}^{\top}, x_{s-1} \rangle_{\star} \right\| \left\| \tilde{A}_{r-1}^{\top} \right\| \left\| \Pi_{t,r}^{\top} \right\| \\ \leq & \mu^{2} \sum_{s,r=1}^{t} \left\| \Pi_{t,s} \right\| \left\| \Pi_{t,r}^{\top} \right\| \left\| x_{s-1} \right\|_{\star} \left\| x_{r-1} \right\|_{\star} \\ \leq & \mu^{2} \left(\sum_{s=1}^{t} \left\| \Pi_{t,s} \right\| \right)^{2} \nu_{t-1}^{2} \leq & \mu^{2} \gamma^{2} \nu_{t-1}^{2}. \end{aligned}$$

Also

$$\begin{split} \left\| R_{t}^{(2)} \right\| &= \left\| R_{t}^{(3)} \right\| \\ &\leq \sum_{s,r=1}^{t} \left\| \Pi_{t,s} \right\| \left\| \tilde{A}_{s-1} \right\| \left\| \left\langle u_{r}^{\top}, x_{s-1} \right\rangle_{\star} \right\| \left\| \Pi_{t,r}^{\top} \right\| \\ &\leq \mu \sum_{s,r=1}^{t} \left\| \Pi_{B} \left(t, s \right) \right\| \left\| x_{s-1} \right\|_{\star} \left\| u_{s} \right\|_{\star} \left\| \Pi_{B}^{\top} \left(t, r \right) \right\| \\ &\leq \mu \gamma^{2} \bar{u} \nu_{t-1}, \end{split}$$

and

$$\left\| R_t^{(4)} \right\| \le \sum_{s,r=1}^t \left\| \Pi_{t,s} \right\| \left\| u_s \right\|_{\star}^2 \left\| \Pi_{t,r}^{\top} \right\| \le \gamma^2 \bar{u}^2.$$

We then obtain

$$\begin{aligned} \|x_t\|_{\star} &= \left\| \mathcal{E} \left\{ x_t x_t^{\top} \right\} \right\|^{1/2} \\ &\leq \sqrt{\left\| R_t^{(1)} \right\| + 2 \left\| R_t^{(2)} \right\| + \left\| R_t^{(4)} \right\|} = \gamma \mu \nu_{t-1} + \gamma \bar{u}. \end{aligned}$$

Since $x_0 = 0$, it follows that $\nu_t \leq \gamma \mu \nu_{t-1} + \gamma \bar{u}$. Hence,

$$\nu_t \le \frac{\gamma \bar{u}}{1 - \gamma \mu},$$

and the result follows.

PROOF. [of Lemma 16] We have

$$\delta\left(\Sigma_{t|t}^{i}, \Sigma_{t|t}\right) = \delta\left(\Sigma_{t|t-1}^{-i} + \Psi_{t}^{i}, \Sigma_{t|t-1}^{-1} + \Psi_{t}\right) = \delta\left(\Sigma_{t|t-1}^{-i} + \Psi_{t}^{i}, \Sigma_{t|t-1}^{-1} + \Psi_{t}^{i}\right) + \delta\left(\Sigma_{t|t-1}^{-1} + \Psi_{t}^{i}, \Sigma_{t|t-1}^{-1} + \Psi_{t}\right). \quad (A.1)$$

Since $\left\|\tilde{\Psi}_{t}^{i}\right\| \leq \left\|\Sigma_{t|t}\right\|^{-1}$, we have from Proposition 13 4 that

$$\delta \left(\Sigma_{t|t-1}^{-1} + \Psi_{t}^{i}, \Sigma_{t|t-1}^{-1} + \Psi_{t} \right) \\ = \delta \left(\Sigma_{t|t-1}^{-1} + \Psi_{t} + \tilde{\Psi}_{t}^{i}, \Sigma_{t|t-1}^{-1} + \Psi_{t} \right) \\ = \delta \left(\Sigma_{t|t}^{-1} + \tilde{\Psi}_{t}^{i}, \Sigma_{t|t}^{-1} \right) \\ \leq \sqrt{N} \left| \log \left(1 - \left\| \Sigma_{t|t} \right\| \left\| \tilde{\Psi}_{t}^{i} \right\| \right) \right|.$$
(A.2)

 Also

$$\delta\left(\Sigma_{t|t-1}^{-i} + \Psi_{t}^{i}, \Sigma_{t|t-1}^{-1} + \Psi_{t}^{i}\right) \\ \leq \delta\left(\Sigma_{t|t-1}^{-i}, \Sigma_{t|t-1}^{-1}\right) \\ = \delta\left(\Sigma_{t|t-1}^{i}, \Sigma_{t|t-1}\right) \\ = \delta\left(A\Sigma_{t-1|t-1}^{i}A^{\top} + Q, A\Sigma_{t-1|t-1}A^{\top} + Q\right) \\ \leq \lambda_{t}\delta\left(\Sigma_{t-1|t-1}^{i}, \Sigma_{t-1|t-1}\right), \qquad (A.3)$$

with

$$\begin{split} \lambda_t &= \frac{\alpha_t}{\alpha_t + \beta_t}, \\ \alpha_t &= \max\left\{ \left\| A \Sigma_{t-1|t-1}^i A^\top \right\|, \left\| A \Sigma_{t-1|t-1} A^\top \right\| \right\}, \\ \beta_t &= \left\| Q^{-1} \right\|^{-1}. \end{split}$$

Now

$$\alpha_{t} \leq \|A\|^{2} \max\left\{\left\|\Sigma_{t-1|t-1}^{i}\right\|, \left\|\Sigma_{t-1|t-1}\right\|\right\}$$
$$\leq \|A\|^{2} \left(\left\|\Sigma_{t-1|t-1}\right\| + \left\|\Sigma_{t-1|t-1} - \Sigma_{t-1|t-1}^{i}\right\|\right)$$
$$\leq \|A\|^{2} \left\|\Sigma_{t-1|t-1}\right\| e^{\delta\left(\Sigma_{t-1|t-1}, \Sigma_{t-1|t-1}^{i}\right)}$$

We then have

$$\lambda_{t} \leq \frac{\|A\|^{2} \|\Sigma_{t-1|t-1}\|}{\|A\|^{2} \|\Sigma_{t-1|t-1}\| + \|Q^{-1}\|^{-1} e^{-\delta \left(\Sigma_{t-1|t-1}, \Sigma_{t-1|t-1}^{i}\right)}}.$$
(A.4)
The result then follows by putting (A.4) into (A.3)

and the resulting equation, together with (A.2) into (A.1).

PROOF. [of Lemma 17] From (28)-(22), we have

$$\begin{split} \breve{\xi}_{t|t}^{i} &= (I - \Phi_{t}) \, A \breve{\xi}_{t-1|t-1}^{i} + \Sigma_{t|t} C_{t}^{i\top} R_{t}^{-i} y_{t}^{i}, \\ \xi_{t|t}^{i} &= \left(I - \Phi_{t}^{i}\right) A \xi_{t-1|t-1}^{i} + \Sigma_{t|t}^{i} C_{t}^{i\top} R_{t}^{-i} y_{t}^{i}. \end{split}$$

Then

$$\begin{split} \tilde{\xi}_{t|t}^{i} &= \left(I - \Phi_{t}^{i}\right) A \xi_{t-1|t-1}^{i} - \left(I - \Phi_{t}\right) A \breve{\xi}_{t-1|t-1}^{i} \\ &+ \left[\Sigma_{t|t}^{i} - \Sigma_{t|t}\right] C_{t}^{i\top} R_{t}^{-i} y_{t}^{i} \\ &= \left(I - \Phi_{t} - \tilde{\Phi}_{t}^{i}\right) A \breve{\xi}_{t-1|t-1}^{i} - \tilde{\Phi}_{t} \breve{\xi}_{t|t-1}^{i} + \tilde{\Sigma}_{t|t}^{i} \mathring{\psi}_{t}^{i} \\ &= \left(I - \Phi_{t} - \tilde{\Phi}_{t}^{i}\right) A \breve{\xi}_{t-1|t-1}^{i} + \left[\tilde{\Phi}_{t}^{i}, \tilde{\Sigma}_{t|t}^{i}\right] \mathfrak{y}_{t}^{i}, \end{split}$$

where

$$\tilde{\Phi}^i_t = \Sigma^i_{t|t} \Psi^i_t - \Sigma_{t|t} \Psi_t = \tilde{\Sigma}^i_{t|t} \Psi_t + \Sigma_{t|t} \tilde{\Psi}^i_t + \tilde{\Sigma}^i_{t|t} \tilde{\Psi}^i_t.$$

Now

$$\left\|\tilde{\Sigma}_{t|t}^{i}\right\| \leq \left(e^{\delta\left(\Sigma_{t|t}^{i},\Sigma_{t|t}\right)} - 1\right) \left\|\Sigma_{t|t}\right\|.$$

Hence

$$\begin{split} \left\| \tilde{\Phi}_{t}^{i} \right\| &\leq \left\| \tilde{\Sigma}_{t|t}^{i} \right\| \left\| \Psi_{t} \right\| + \left\| \Sigma_{t|t} \right\| \left\| \tilde{\Psi}_{t}^{i} \right\| + \left\| \tilde{\Sigma}_{t|t}^{i} \right\| \left\| \tilde{\Psi}_{t}^{i} \right\| \\ &\leq \left[\left(e^{\delta \left(\Sigma_{t|t}^{i}, \Sigma_{t|t} \right)} - 1 \right) \left\| \Psi_{t} \right\| + \right. \\ &+ \left. e^{\delta \left(\Sigma_{t|t}^{i}, \Sigma_{t|t} \right)} \left\| \tilde{\Psi}_{t}^{i} \right\| \right] \left\| \Sigma_{t|t} \right\|. \end{split}$$

PROOF. [of Lemma 18] It follows from (30) that $\tilde{\psi} < \bar{\sigma}^{-1}$, which in turn implies the condition of Lemma 16. From the latter we then obtain

$$\delta\left(\Sigma_{t|t}^{i}, \Sigma_{t|t}\right) \leq \frac{\delta\left(\Sigma_{t-1|t-1}^{i}, \Sigma_{t-1|t-1}\right)}{1 + \frac{\|Q^{-1}\|^{-1}}{\bar{\sigma}\|A\|^{2}}} e^{-\delta\left(\Sigma_{t-1|t-1}^{i}, \Sigma_{t-1|t-1}\right)} + \bar{\upsilon}\left(\bar{\tilde{\psi}}\right). \quad (A.5)$$

Since $\delta\left(\Sigma_{1|1}^{i}, \Sigma_{1|1}\right) = 0$, we have from Lemma 8 that the iterations (A.5) converge to $\bar{\delta}\left(\bar{\psi}\right)$, and

$$\delta\left(\Sigma_{t|t}^{i}, \Sigma_{t|t}\right) \leq \bar{\delta}\left(\bar{\tilde{\psi}}\right). \tag{A.6}$$

Using (A.6) in Lemma 17 we obtain

$$\begin{split} \left\| \tilde{\Phi}_t^i \right\| &\leq \left[\left(e^{\bar{\delta}\left(\bar{\psi}\right)} - 1 \right) \bar{\psi} + e^{\bar{\delta}\left(\bar{\psi}\right)} \bar{\psi} \right] \bar{\sigma} = \bar{\phi} \left(\bar{\psi} \right), \\ \left\| \tilde{\Sigma}_{t|t}^i \right\| &\leq \left(e^{\bar{\delta}\left(\bar{\psi}\right)} - 1 \right) \bar{\sigma} = \bar{\sigma} \left(\bar{\psi} \right). \end{split}$$

Let $\mathfrak{u}_t^i = \left[\tilde{\Phi}_t^i, \; \tilde{\Sigma}_{t|t}^i \right] \mathfrak{y}_t^i$. We have

$$\mathcal{E}\left\{\mathfrak{u}_{t}^{i}\mathfrak{u}_{t}^{i\top}\right\} = \left[\left.\tilde{\Phi}_{t}^{i}, \right.\tilde{\Sigma}_{t|t}^{i}\right]\mathcal{E}\left\{\mathfrak{y}_{t}^{i}\mathfrak{y}_{t}^{i\top}\right\}\left[\left.\begin{array}{c}\tilde{\Phi}_{t}^{i}\\ \tilde{\Sigma}_{t|t}^{i}\end{array}\right].$$

It then follows that

$$\begin{split} \left\| \mathcal{E} \left\{ \mathfrak{u}_{t}^{i}\mathfrak{u}_{t}^{i\top} \right\} \right\| &\leq \left\| \mathcal{E} \left\{ \mathfrak{y}_{t}^{i}\mathfrak{y}_{t}^{i\top} \right\} \right\| \left(\left\| \tilde{\Phi}_{t}^{i} \right\|^{2} + \left\| \tilde{\Sigma}_{t|t}^{i} \right\|^{2} \right) \\ &\leq \bar{\mathfrak{y}}^{2} \left(\bar{\phi}^{2} \left(\bar{\tilde{\psi}} \right) + \bar{\sigma}^{2} \left(\bar{\tilde{\psi}} \right) \right), \\ \text{or } \left\| \mathfrak{u}_{t}^{i} \right\|_{\star} &\leq \bar{\mathfrak{y}} \sqrt{\bar{\phi}^{2} \left(\bar{\tilde{\psi}} \right) + \bar{\sigma}^{2} \left(\bar{\tilde{\psi}} \right)}. \end{split}$$

Equation (36) defines a perturbed linear system with $-\tilde{\Phi}_t^i A$ being the perturbation of the nominal statetransition matrix $(I - \Phi_t) A$ and \mathfrak{u}_t^i being the input. Since $\bar{\psi} \bar{\gamma} ||A|| < 1$, we can apply Lemma 15 to this perturbed system to obtain (37).

References

- [1] Francesco Bullo. *Lectures on network systems*. Kindle Direct Publishing, 2019.
- [2] Z. Luo, M. Gastpar, J. Liu, and A. Swami. Distributed signal processing in sensor networks. *IEEE Signal Process Mag*, 23(4):14–15, 2006.
- [3] J. Xiao, A. Ribeiro, Z. Luo, and G. Giannakis. Distributed compression-estimation using wireless sensor networks. *IEEE Signal Process Mag*, 23(4):27–41, 2006.
- [4] Damián Marelli and Minyue Fu. Distributed weighted least-squares estimation with fast convergence for large-scale systems. *Automatica*, 51:27–39, 2015.
- [5] Alejandro Ribeiro, Ioannis D Schizas, Stergios I Roumeliotis, and Georgios Giannakis. Kalman filtering in wireless sensor networks. *IEEE Control Syst Mag*, 30(2):66–86, 2010.
- [6] P. Massioni and M. Verhaegen. Distributed control for identical dynamically coupled systems: A decomposition approach. *IEEE Trans Autom Control*, 54(1):124–135, 2009.
- [7] Raffaello D'Andrea and Geir E Dullerud. Distributed control design for spatially interconnected systems. *IEEE Trans Autom Control*, 48(9):1478–1495, 2003.
- [8] T. Yang, X. Yi, J. Wu, Y. Yuan, D. Wu, Z. Meng, Y. Hong, H. Wang, Z. Lin, and K. Johansson. A survey of distributed optimization. *Annu Rev Control*, 2019.
- [9] X. Li, H. Caimou, and H. Haoji. Distributed filter with consensus strategies for sensor networks. J Appl Math, 2013, 2013.
- [10] T. Zhou. Coordinated one-step optimal distributed state prediction for a networked dynamical system. *IEEE Trans Autom Control*, 58(11):2756–2771, 2013.
- [11] T. Zhou. On the controllability and observability of networked dynamic systems. *Automatica*, 52:63–75, 2015.
- [12] M. Farina, G. Ferrari-Trecate, and R. Scattolini. Movinghorizon partition-based state estimation of large-scale systems. *Automatica*, 46(5):910–918, 2010.

- [13] U. Khan and J. Moura. Distributing the Kalman filter for large-scale systems. *IEEE Trans Signal Process*, 56(10):4919– 4935, 2008.
- [14] A. Haber and M. Verhaegen. Moving horizon estimation for large-scale interconnected systems. *IEEE Trans Autom Control*, 58(11):2834–2847, 2013.
- [15] Y. Sun, M. Fu, B. Wang, H. Zhang, and D. Marelli. Dynamic state estimation for power networks using distributed map technique. *Automatica*, 73:27–37, 2016.
- [16] B. Rao and H. Durrant-Whyte. Fully decentralised algorithm for multisensor Kalman filtering. *IEE Proceedings-D*, 138(5):413–420, 1991.
- [17] D. Spanos, R. Olfati-Saber, and R. Murray. Approximate distributed Kalman filtering in sensor networks with quantifiable performance. In *International Symposium on IPSN*, pages 133–139. Ieee, 2005.
- [18] R. Olfati-Saber and J. Shamma. Consensus filters for sensor networks and distributed sensor fusion. In *IEEE Conf Decis Control*, pages 6698–6703. IEEE, 2005.
- [19] R. Olfati-Saber. Distributed Kalman filter with embedded consensus filters. In *IEEE Conf Decis Control*, pages 8179– 8184. IEEE, 2005.
- [20] David W Casbeer and Randy Beard. Distributed information filtering using consensus filters. In Amer Contr Conf, pages 1882–1887. IEEE, 2009.
- [21] G. Battistelli, L. Chisci, G. Mugnai, A. Farina, and A. Graziano. Consensus-based linear and nonlinear filtering. *IEEE Trans Autom Control*, 60(5):1410–1415, 2014.
- [22] Giorgio Battistelli and Luigi Chisci. Stability of consensus extended Kalman filter for distributed state estimation. *Automatica*, 68:169–178, 2016.
- [23] H. Bai, R. Freeman, and K. Lynch. Distributed Kalman filtering using the internal model average consensus estimator. In *Amer Contr Conf*, pages 1500–1505. IEEE, 2011.
- [24] S. Das and J.F. Moura. Distributed Kalman filtering with dynamic observations consensus. *IEEE Trans Signal Process*, 63(17):4458–4473, 2015.
- [25] S. Das and J. Moura. Consensus+ innovationsdistributed Kalman filter with optimized gains. *IEEE Trans Signal Process*, 65(2):467–481, 2016.
- [26] J. Hu, L. Xie, and C. Zhang. Diffusion Kalman filtering based on covariance intersection. *IEEE Trans Signal Process*, 60(2):891–902, 2011.
- [27] S. Wang and W. Ren. On the convergence conditions of distributed dynamic state estimation using sensor networks: A unified framework. *IEEE Trans Control Syst Technol*, 26(4):1300–1316, 2017.
- [28] S. Julier and J. Uhlmann. General decentralized data fusion with covariance intersection. *Handbook of multisensor data* fusion: theory and practice, pages 319–344, 2009.
- [29] Z. Wu, M. Fu, Y. Xu, and R. Lu. A distributed Kalman filtering algorithm with fast finite-time convergence for sensor networks. *Automatica*, 95:63–72, 2018.
- [30] R. Olfati-Saber. Distributed Kalman filtering for sensor networks. In *IEEE Conf Decis Control*, pages 5492–5498. IEEE, 2007.
- [31] R. Olfati-Saber. Kalman-consensus filter: Optimality, stability, and performance. In *IEEE Conf Decis Control*, pages 7036–7042. IEEE, 2009.
- [32] U. Khan, S. Kar, A. Jadbabaie, and J. Moura. On connectivity, observability, and stability in distributed

estimation. In *IEEE Conf Decis Control*, pages 6639–6644. IEEE, 2010.

- [33] C. Li, H. Dong, J. Li, and F. Wang. Distributed Kalman filtering for sensor network with balanced topology. *Systems Control Lett*, 131:104500, 2019.
- [34] R. Carli, A. Chiuso, L. Schenato, and S. Zampieri. Distributed Kalman filtering based on consensus strategies. *IEEE J Sel Areas Commun*, 26(4):622–633, 2008.
- [35] Ion Matei and John S Baras. Consensus-based linear distributed filtering. Automatica, 48(8):1776–1782, 2012.
- [36] F. Cattivelli and A. Sayed. Diffusion strategies for distributed Kalman filtering and smoothing. *IEEE Trans Autom Control*, 55(9):2069–2084, 2010.
- [37] I. Shames, T. Charalambous, C. Hadjicostis, and M. Johansson. Distributed network size estimation and average degree estimation and control in networks isomorphic to directed graphs. In Annual Allerton Conf. on Comm., Control, and Comp., pages 1885–1892. IEEE, 2012.
- [38] Rajendra Bhatia. Positive definite matrices, volume 24. Princeton university press, 2009.
- [39] E.C. Lance. Hilbert C*-modules: a toolkit for operator algebraists, volume 210. Cambridge University Press, 1995.
- [40] Thomas Kailath, Ali H Sayed, and Babak Hassibi. *Linear estimation*. Prentice Hall, 2000.
- [41] E. N. Gilbert. Capacity of a burst-noise channel. Bell Syst. Tech. J, 39(9):1253–1265, 1960.
- [42] E. O. Elliott. Estimates of error rates for codes on burstnoise channels. Bell Syst. Tech. J, 42(9):1977–1997, 1963.
- [43] P. Bougerol. Kalman filtering with random coefficients and contractions. SIAM J Control Optim, 31(4):942–959, 1993.
- [44] Tianju Sui, Damián Edgardo Marelli, Minyue Fu, and Renquan Lu. Accuracy analysis for distributed weighted least-squares estimation in finite steps and loopy networks. *Automatica*, 97:82–91, 2018.