THE INSTITUTE FOR SYSTEMS RESEARCH

ISR TECHNICAL REPORT 2010-5

Consensus-Based Linear Distributed Filtering

Ion Matei and John S. Baras

Consensus-Based Linear Distributed Filtering

Ion Matei and John S. Baras

Abstract

We address the consensus-based distributed linear filtering problem, where a discrete time, linear stochastic process is observed by a network of sensors. We assume that the consensus weights are known and we first provide sufficient conditions under which the stochastic process is detectable, i.e. for a specific choice of consensus weights there exists a set of filtering gains such that the dynamics of the estimation errors (without noise) is asymptotically stable. Next, we provide a distributed, sub-optimal filtering scheme based on minimizing an upper bound on a quadratic filtering cost. In the stationary case, we provide sufficient conditions under which this scheme converges; conditions expressed in terms of the convergence properties of a set of coupled Riccati equations. We continue with presenting a connection between the consensus-based distributed linear filter and the optimal linear filter of a Markovian jump linear system, appropriately defined. More specifically, we show that if the Markovian jump linear system is (mean square) detectable, then the stochastic process is detectable under the consensus-based distributed linear filtering scheme. We also show that the optimal gains of a linear filter for estimating the state of a Markovian jump linear system appropriately defined can be seen as an approximation of the optimal gains of the consensus-based linear filter.

I. Introduction

Sensor networks have broad applications in surveillance and monitoring of an environment, collaborative processing of information, and gathering scientific data from spatially distributed sources for environmental modeling and protection. A fundamental problem in sensor networks is developing distributed algorithms for the state estimation of a process of interest. Generically, a process is observed by a group of (mobile) sensors organized in a network. The goal of each sensor is to computed accurate state estimates. The distributed filtering (estimation) problem has received a lot of attention during the past thirty years. An important contribution was brought by Borkar and Varaiya [2], who address the distributed estimation problem of a random variable by a group of sensors. The particularity of their formulation is that both estimates and measurements are shared among neighboring sensors. The authors show that if the sensors form a communication ring, through which information is exchanged infinitely often, then the estimates converge asymptotically to the same value, i.e. they asymptotically agree. An extension of the results in reference [2] is given in [7]. The recent technological advances in mobile sensor networks have re-ignited the interest for the distributed estimation problem. Most papers focusing on distributed estimation propose different mechanisms for combining the Kalman filter with a consensus filter in order to ensure that the estimates asymptotically converge to the same value, schemes which will be henceforth called consensus based distributed filtering (estimation) algorithms. In [8] and [9], several algorithms based on the idea mentioned above are introduced. In [3], the authors study the interaction between the consensus matrix, the number of messages exchanged per sampling time, and the Kalman gain for scalar systems. It is shown that optimizing the consensus matrix for fastest convergence and using the centralized optimal gain is not necessarily the optimal strategy if the number of exchanged messages per sampling time is small. In [11], the weights are adaptively updated to minimize the variance of the estimation error. Both the estimation and the parameter optimization are performed in a distributed manner. The authors derive an upper bound of the error variance in each node which decreases with the number of neighboring nodes.

In this note we address the consensus-based distributed linear filtering problem as well. We assume that each agent updates its (local) estimate in two steps. In the first step, an update is produced using a Luenberger observer type of filter. In the second step, called *consensus step*, every sensor computes a convex combination between its local update and the updates received from the neighboring sensors. Our focus is *not* on designing the consensus weights, but on designing the *filter gains*. For given consensus weights, we will first give sufficient conditions for the existence of filter gains such that the dynamics of the estimation errors (without noise) is asymptotically stable. These sufficient conditions are also expressible in terms of the feasibility of a set of linear matrix inequalities. Next, we present a distributed (in the sense that each sensor uses only information available within its neighborhood), sub-optimal filtering algorithm, valid for time varying topologies as well, resulted from minimizing an upper bound on a quadratic cost expressed in terms of the covariances matrices of the estimation errors. In the case where the matrices defining the stochastic process and the consensus weights are time invariant, we present sufficient conditions such that the aforementioned distributed algorithm produces filter gains which converge and ensure the stability of the dynamics of the covariances matrices of the estimation errors. We will also present a connection between the consensus-based linear filter and the linear filtering of a Markovian jump linear system appropriately defined. More precisely, we show that if the aforementioned Markovian jump linear system is (mean square) detectable then the stochastic process is detectable as well under the consensus-based distributed linear filtering scheme. Finally we show that the optimal gains of a linear filter for the state estimation of the Markovian jump linear system can be viewed as an approximation of the optimal gains of the consensus-based distributed linear filtering strategy.

Paper structure: In Section II we describe the problems addressed in this note. Section III introduces the sufficient conditions for detectability under the consensus-based linear filtering scheme together with a test expressed in terms of the feasibility of a set of linear matrix inequalities. In Section IV we present a sub-optimal distributed consensus based linear filtering scheme with quantifiable performances. Section V makes a connection between the consensus-based distributed linear filtering algorithm and the linear filtering scheme for a Markovian jump linear system.

Notations and Abbreviations: We represent the property of positive (semi-positive) definiteness of a symmetric matrix A, by $A \succ 0$ ($A \succeq 0$). By convention, we say that a symmetric matrix A is *negative definite* (*semi-definite*) if $-A \succ 0$ ($-A \succeq 0$) and we denote this by $A \prec 0$ ($A \preceq 0$). By $A \succ B$ we understand that $A - B$ is positive definite. Given a set of square matrices $\{A_i\}_{i=1}^N$, by $diag(A_i, i = 1 \dots n)$ we understand the block diagonal matrix which contains the matrices A_i's on the main diagonal. We use the abbreviations CBDLF, MJLS and LMI for consensus-based linear filter(ing), Markovian jump linear system and linear matrix inequality, respectively.

Remark 1.1: Given a positive integer N, a set of vectors $\{x_i\}_{i=1}^N$, a set of non-negative scalars $\{p_i\}_{i=1}^N$ summing up to one and a positive definite matrix Q, the following inequality holds

$$\left(\sum_{i=1}^N p_i x_i\right)' Q \left(\sum_{i=1}^N p_i x_i\right) \leq \sum_{i=1}^N p_i x_i' Q x_i.$$

Remark 1.2: Given a positive integer N, a set of vectors $\{x_i\}_{i=1}^N$, a set of matrices $\{A_i\}_{i=1}^N$ and a set of non-negative scalars $\{p_i\}_{i=1}^N$ summing up to one, the following holds

$$\left(\sum_{i=1}^N p_i A_i x_i\right)\left(\sum_{i=1}^N p_i A_i x_i\right)' \preceq \sum_{i=1}^N p_i A_i x_i x_i' A_i'. \tag{1}$$

II. Problem formulation

We consider a stochastic process modeled by a discrete-time linear dynamic equation

$$x(k+1) = A(k)x(k) + w(k),\ x(0) = x_0, \tag{2}$$

where $x(k) \in \mathbb{R}^n$ is the state vector and $w(k) \in \mathbb{R}^n$ is a driving noise, assumed Gaussian with zero mean and (possibly time varying) covariance matrix $\Sigma_w(k)$. The initial condition x_0 is assumed to be Gaussian with mean μ_0 and covariance matrix Σ_0. The state of the process is observed by a network of N sensors indexed by i, whose sensing models are given by

$$y_i(k) = C_i(k)x(k) + v_i(k),\ i = 1 \dots N, \tag{3}$$

where $y_i(k) \in \mathbb{R}^{r_i}$ is the observation made by sensor i and $v_i(k) \in \mathbb{R}^{r_i}$ is the measurement noise, assumed Gaussian with zero mean and (possibly time varying) covariance matrix $\Sigma_{v_i}(k)$. We assume that the matrices $\{\Sigma_{v_i}(k)\}_{i=1}^N$ and $\Sigma_w(k)$ are positive definite for $k \geq 0$ and that the initial state x_0, the noises $v_i(k)$ and $w(k)$ are independent for all $k \geq 0$. For later reference we also define $\Sigma_{v_i}^{1/2}(k)$, $\Sigma_w^{1/2}(k)$, where $\Sigma_{v_i}(k) \triangleq \Sigma_{v_i}^{1/2}(k)\Sigma_{v_i}^{1/2}(k)'$ and $\Sigma_w(k) \triangleq \Sigma_w^{1/2}(k)\Sigma_w^{1/2}(k)'$.

The set of sensors form a communication network whose topology is modeled by a directed graph that describes the information exchanged among agents. The goal of the agents is to (locally) compute estimates of the state of the process (2).

Let $\hat{x}_i(k)$ denote the state estimate computed by sensor i and let $\epsilon_i(k)$ denote the estimation error, i.e. $\epsilon_i(k) \triangleq x(k) - \hat{x}_i(k)$. The covariance matrix of the estimation error of sensor i is denoted by $\Sigma_i(k) \triangleq E[\epsilon_i(k)\epsilon_i(k)']$, with $\Sigma_i(0) = \Sigma_0$.

The sensors update their estimates in two steps. In the first step, an intermediate estimate, denoted by $\varphi_i(k)$, is produced using a Luenberger observer filter

$$\varphi_i(k) = A(k)\hat{x}_i(k) + L_i(k)(y_i(k) - C_i(k)\hat{x}_i(k)),\ i = 1 \dots N, \tag{4}$$

where $L_i(k)$ is the *filter gain*.

In the second step, the new state estimate of sensor i is generated by a convex combination between $\varphi_i(k)$ and all other intermediate estimates within its communication neighborhood, i.e.

$$\hat{x}_i(k+1) = \sum_{j=1}^{N} p_{ij}(k)\varphi_j(k),\ i = 1 \dots N, \tag{5}$$

where $p_{ij}(k)$ are non-negative scalars summing up to one ($\sum_{j=1}^{N} p_{ij}(k) = 1$), and $p_{ij}(k) = 0$ if no link from j to i exists at time k. Having $p_{ij}(k)$ dependent on time accounts for a possibly time varying communication topology.

Combining (4) and (5) we obtain the dynamic equations for the consensus based distributed filter:

$$\hat{x}_i(k+1) = \sum_{j=1}^{N} p_{ij}(k)\Big[A(k)\hat{x}_j(k) + L_j(k)\Big(y_j(k) - C_j(k)\hat{x}_j(k)\Big)\Big],\ i = 1\ldots N. \tag{6}$$

From (6) the estimation errors evolve according to

$$\epsilon_i(k+1) = \sum_{j=1}^{N} p_{ij}(k)\Big[\Big(A(k) - L_j(k)C_j(k)\Big)\epsilon_j(k) + w(k) - L_j(k)v_j(k)\Big],\ i = 1\ldots N. \tag{7}$$

We define the aggregate vectors of estimates, measurements, estimation errors, driving noise and measurements noise, respectively

$$\mathbf{\hat{x}}(k)' \triangleq (\hat{x}_1(k)', \ldots, \hat{x}_N(k)'),$$

$$\mathbf{y}(k)' \triangleq (y_1(k)', \ldots, y_N(k)'),$$

$$\epsilon(k)' \triangleq (\epsilon_1(k)', \ldots, \epsilon_N(k)'),$$

$$\mathbf{w}(k)' \triangleq (w(k)', \ldots, w(k)'),$$

$$\mathbf{v}(k)' \triangleq (v_1(k)', \ldots, v_N(k)'),$$

and the following block matrices

$$\mathcal{A}(k) \triangleq \begin{pmatrix} A(k) & O_{n\times n} & \cdots & O_{n\times n} \\ O_{n\times n} & A(k) & \cdots & O_{n\times n} \\ \vdots & \vdots & \ddots & \vdots \\ O_{n\times n} & O_{n\times n} & \cdots & A(k) \end{pmatrix} \in \mathbb{R}^{nN\times nN},$$

$$C(k) \triangleq \begin{pmatrix} C_1(k) & O_{r_2\times n} & \cdots & O_{r_N\times n} \\ O_{r_1\times n} & C_2(k) & \cdots & O_{r_N\times n} \\ \vdots & \vdots & \ddots & \vdots \\ O_{r_1\times n} & O_{r_2\times n} & \cdots & C_N(k) \end{pmatrix} \in \mathbb{R}^{r\times nN},\ \mathcal{L}(k) \triangleq \begin{pmatrix} L_1(k) & O_{n\times r_1} & \cdots & O_{n\times r_N} \\ O_{n\times r_1} & L_2(k) & \cdots & O_{n\times r_N} \\ \vdots & \vdots & \ddots & \vdots \\ O_{n\times r_1} & O_{n\times r_2} & \cdots & L_N(k) \end{pmatrix} \in \mathbb{R}^{nN\times r},$$

where $r = \sum_{i=1}^{N} r_i$. The dynamics (6) and (7) can be compactly written as

$$\mathbf{\hat{x}}(k+1) = \mathcal{P}(k)\mathcal{A}(k)\mathbf{\hat{x}}(k) + \mathcal{P}(k)\mathcal{L}(k)[\mathbf{y}(k) - C(k)\mathbf{\hat{x}}(k)], \tag{8}$$

$$\epsilon(k+1) = \mathcal{P}(k)[\mathcal{A}(k) - \mathcal{L}(k)\mathcal{C}(k)]\epsilon(k) + \mathbf{w}(k) - \mathcal{P}(k)\mathcal{L}(k)\mathbf{v}((k), \tag{9}$$

where $\mathcal{P}(k) = P(k) \otimes I$ and $P(k) = (p_{ij}(k))$ is a stochastic matrix, with rows summing up to one.

Definition 2.1: (*distributed detectability*) Assuming that $A(k)$, $\mathbf{C}(k) \triangleq \{C_i(k)\}_{i=1}^N$ and $\mathbf{p}(k) \triangleq \{p_{ij}(k)\}_{i,j=1}^N$ are time invariant, we say that the linear system (2) is *detectable* using the CBDLF scheme (6), if there exist a set of matrices $\mathbf{L} \triangleq \{L_i\}_{i=1}^N$ such that the system (7), without the noise, is asymptotically stable.

We introduce the following finite horizon quadratic filtering cost function

$$J_K(\mathbf{L}(\cdot)) = \sum_{k=0}^{K} \sum_{i=1}^{N} E[\|\epsilon_i(k)\|^2], \tag{10}$$

where by $\mathbf{L}(\cdot)$ we understand the set of matrices $\mathbf{L}(\cdot) \triangleq \{L_i(k), k = 0 \dots K-1\}_{i=1}^N$. The optimal filtering gains represent the solution of the following optimization problem

$$\mathbf{L}^*(\cdot) = \arg\min_{\mathbf{L}(\cdot)} J^K(\mathbf{L}(\cdot)). \tag{11}$$

Assuming that $A(k)$, $\mathbf{C}(k) \triangleq \{C_i(k)\}_{i=1}^N$, $\Sigma_w(k)$, $\Sigma_v(k) \triangleq \{\Sigma_{v_i}(k)\}$ and $\mathbf{p}(k) \triangleq \{p_{ij}(k)\}_{i,j=1}^N$ are time invariant, we can also define the infinite horizon filtering cost function

$$J_\infty(\mathbf{L}) = \lim_{K\to\infty} \frac{1}{K} J_K(\mathbf{L}) = \lim_{k\to\infty} \sum_{i=1}^{N} E[\|\epsilon_i(k)\|^2], \tag{12}$$

where $\mathbf{L} \triangleq \{L_i\}_{i=1}^N$ is the set of steady state filtering gains. By solving the optimization problem

$$\mathbf{L}^* = \arg\min_{\mathbf{L}} J^\infty(\mathbf{L}), \tag{13}$$

we obtain the optimal steady-state filter gains.

In the next sections we will address the following problems:

Problem 2.1: (*Detectability conditions*) Under the above setup, we want to find conditions under which the system (2) is detectable in the sense of Definition 2.1.

Problem 2.2: (*Sub-optimal scheme for consensus based distributed filtering*) Ideally, we would like to obtain the optimal filter gains by solving the optimization problems (11) and (13), respectively. Due to the complexity of these problems, we will not provide the optimal filtering gains but rather focus on providing a sub-optimal scheme with quantifiable performances.

Problem 2.3: (*Connection with the linear filtering of a Markovian jump linear system*) We make a parallel between the consensus-based distributed linear filtering scheme and the linear filtering of a particular Markovian jump linear system.

III. Distributed detectability

We start with a toy example motivating our interest in the distributed detectability problem under the CBDLF scheme. Let us assume that no single pair (A, C_i) is detectable in the classical sense, but the pair (A, C) is detectable, where $C' = (C'_1, \ldots, C'_N)$. In this case, we can design a stable (centralized) Luenberger observer filter. The question is, can we obtain a stable consensus-based distributed filter? As the following example will show, in general this is not true. That is why is important to find conditions under which the CBDLF can produce stable estimates.

Example 3.1: (*Centralized detectable but not distributed detectable*) Consider a linear dynamics as in (2-3), with two sensors, where

$$A = \begin{pmatrix} 10 & 0 \\ 0 & 10 \end{pmatrix},\ C_1 = \begin{pmatrix} 1 & 0 \end{pmatrix} \text{ and } C_2 = \begin{pmatrix} 0 & 1 \end{pmatrix}.$$

Obviously, the pairs (A, C_1) and (A, C_2) are not detectable while the pair (A, C) it is, where $C' = (C'_1\ C'_2)$ is. Let $L'_1 = (l_1\ l_2)$ and $L'_2 = (l_3\ l_4)$. For this example, the matrix that dictates the stability property of (9) is given by

$$\mathcal{A} = \begin{pmatrix} p_{11}(10-l_1) & 0 & 10p_{12} & -p_{12}l_3 \\ -p_{11}l_2 & 10p_{11} & 0 & p_{12}(10-l_4) \\ p_{21}(10-l_1) & 0 & 10p_{22} & -p_{22}l_3 \\ -p_{21}l_2 & 10p_{21} & 0 & p_{22}(10-l_4) \end{pmatrix}$$

For $p_{11} = 0.9$, $p_{12} = 0.1$, $p_{21} = 0.7$ and $p_{22} = 0.3$, the characteristic polynomial of the above matrix is given by

$$\mathbf{q}(s) = s^4 + q_3(l_1, l_3)s^3 + q_2(l_1, l_4, l_2l_3)s^2 + q_1(l_1, l_4) + q_0(l_1, l_4),$$

where

$$q_3(l_1, l_3) = -24 + 0.9l_1 + 0.3l_4,$$

$$q_2(l_1, l_4, l_2l_3) = -0.07l_2l_3 - 5.6l_4 + 184 - 12.8l_1 + 0.27l_1l_4,$$

$$q_1(l_1, l_4) = 30l_4 - 480 - 2.4l_1l_4 + 42l_1,$$

$$q_0(l_1, l_4) = -40l_1 - 40l_4 + 4l_1l_4 + 400,$$

and let $\lambda_i(l_1, l_4, l_2l_3)$ denote the eigenvalues of $\mathcal{A}$. We define $\lambda_{max}(l_1, l_4, l_2l_3) = \max_i |\lambda_i(l_1, l_4, l_2l_3)|$. The system (2-3) is not detectable in the sense of Definition 2.1 if $\lambda_{max}(l_1, l_4, l_2l_3) > 1$ for all values of l_1, l_2 and of the product $l_2 l_3$. We introduce also the quantity $\lambda^{23}_{max}(l_2 l_3) = \min_{l_1,l_4} \lambda_{max}(l_1, l_4, l_2 l_3)$.

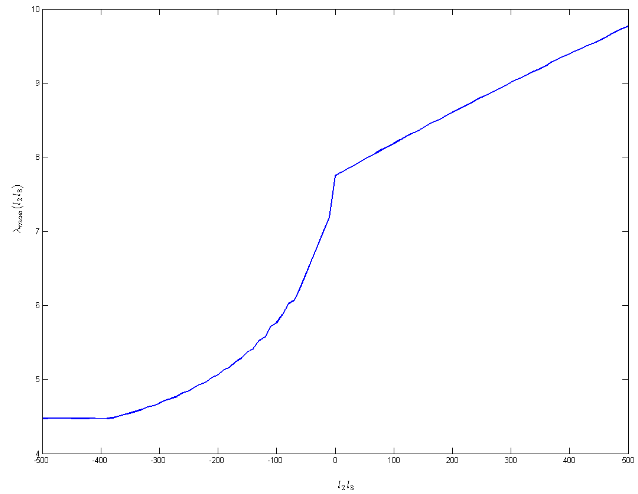

Fig. 1. The evolution of $\lambda^{23}_{max}(l_2 l_3)$

From Figure 1, we note that $\min_{l2l3} \lambda^{23}_{max}(l_2 l_3) = 4.498$, which shows that, for the given consensus weights, and matrices A, C_1 and C_2, there are no values for l_1, l_2, l_3 and l_4, such that (9) can be made asymptotically stable.

The CBDLF (8) uses only one consensus step and we have seen, through Example 3.1, that in general this does not guarantee stable estimates, even in the case where the pair (A, C) is detectable. However, as the next proposition suggests, stable estimates might be achieved if a large enough number of consensus steps is used, i.e. we set $\mathcal{P}(k) = P(k)^{\eta} \otimes I$, for some positive integer value η, large enough.

Proposition 3.1: Consider the linear dynamics (2)-(3). Assume that in the CBDLF scheme (6), we have $p_{ij} = \frac{1}{N}$ and that $\hat{x}_i(0) = x_0$, for all $i, j = 1 \dots N$. If the pair (A, C) is detectable, then the system (2) is detectable as well, in the sense of Definition 2.1.

Proof: Rewrite the matrix C as

$$C = \sum_{i=1}^{N} \bar{C}_i,$$

where $\bar{C}_i' = (O_{n\times r_1} \dots O_{n\times r_{i-1}}\ C_i'\ O_{n\times r_{i+1}} \dots O_{n\times r_N})$. Ignoring the noise, we define the measurements

$$\bar{y}_i(k) = \bar{C}_i x(k),$$

which are equivalent to the ones in (3). Under the assumption that $p_{ij} = \frac{1}{N}$ and $\hat{x}_i = x_0$ for all $i, j = 1 \ldots N$, it follows that the estimation errors respect the dynamics

$$\epsilon(k+1) = \frac{1}{N}\sum_{i=1}^{N}(A - L_i\bar{C}_i)\epsilon(k). \tag{14}$$

Setting $L_i = NL$ for $i = 1 \ldots N$, it follows that

$$\epsilon(k+1) = (A - LC)\epsilon(k).$$

Since the pair (A, C) is detectable, there exists a matrix L such that $A - LC$ has all eigenvalues within the unit circle and therefore the dynamics (14) is asymptotically stable, which implies that (2) is detectable in the sense of Definition 2.1. ■

The previous proposition tells us that if we achieve (average) consensus between the state estimates at each time instant, and if the pair (A, C) is detectable (in the classical sense), then the system (2) is detectable in the sense of Definition 2.1. However, achieving consensus at each time instant can be time and numerically costly and that is why is important to find (testable) conditions under which the CBDLF produces stable estimates.

Lemma 3.1: (*sufficient conditions for distributed detectability*) If there exists a set of symmetric, positive definite matrices $\{Q_i\}_{i=1}^N$ and a set of matrices $\{L_i\}_{i=1}^N$ such that

$$Q_i = \sum_{j=1}^{N} p_{ji}(A - L_jC_j)'Q_j(A - L_jC_j) + S_i,\ i = 1 \ldots N, \tag{15}$$

for some positive definite matrices $\{S_i\}_{i=1}^N$, then the system (2) is detectable in the sense of Definition 2.1.

Proof: The dynamics of the estimation error without noise is given by

$$\epsilon_i(k+1) = \sum_{j=1}^{N} p_{ij}(A - L_jC_j)\epsilon_j(k),\ i = 1 \ldots N. \tag{16}$$

In order to prove the stated result we have to show that (16) is asymptotically stable. We define the Lyapunov function

$$V(k) = \sum_{i}^{N} x_i(k)'Q_ix_i(k),$$

and our goal is to show that $V(k+1) - V(k) < 0$ for all $k \geq 0$. The Lyapunov difference is given by

$$V(k+1) - V(k) = \sum_{i=1}^{N}\left(\sum_{j=1}^{N} p_{ij}(A - L_jC_j)\epsilon_j(k)\right)' Q_i \left(\sum_{j=1}^{N} p_{ij}(A - L_jC_j)\epsilon_j(k)\right) - \epsilon_i(k)'TQ_i\epsilon_i(k) \leq$$

$$\leq \sum_{i=1}^{N}\left(\sum_{j=1}^{N} p_{ij}\epsilon_j(k)'(A-L_jC_j)'Q_i(A-L_jC_j)\epsilon_j(k)\right)-\epsilon_i(k)'Q_i\epsilon_i(k), \tag{17}$$

where the inequality followed from Remark 1.1. By changing the summation order we can farther write

$$V(k+1)-V(k)\leq \sum_{i=1}^{N}\epsilon_i(k)'\left(\sum_{j=1}^{N} p_{ji}(A-L_jC_j)'Q_j(A-L_jC_j)-Q_i\right)\epsilon_i(k).$$

Using (15) yields

$$V(k+1)-V(k)\leq -\sum_{i=1}^{N}\epsilon_i(k)'S_i\epsilon_i(k)$$

From the fact that $\{S_j\}_{j=1}^{N}$ are positive definite matrices, we get

$$V(k+1)-V(k)<0,$$

which implies that (16) is asymptotically stable. ■

The following result relates the existence of the sets of matrices $\{Q_i\}_{i=1}^{N}$ and $\{L_i\}_{i=1}^{N}$ such that (15) is satisfied, with the feasibility of a set of linear matrix inequalities (LMI).

Proposition 3.2: (*distributed detectability test*) The linear system (2) is *detectable* in the sense of Definition 2.1 if the following linear matrix inequalities, in the variables $\{X_i\}_{i=1}^{N}$ and $\{Y_i\}_{i=1}^{N}$, are feasible

$$\begin{pmatrix} X_i & \sqrt{p_{1i}}(A'X_1-C_1'Y_1') & \sqrt{p_{2i}}(A'X_2-C_2'Y_2') & \cdots & \sqrt{p_{Ni}}(A'X_N-C_N'Y_N') \\ \sqrt{p_{1i}}(X_1A-Y_1C1) & X_1 & 0 & \cdots & 0 \\ \sqrt{p_{2i}}(X_2A-Y_2C2) & 0 & X_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sqrt{p_{Ni}}(X_NA-Y_NC_N) & 0 & 0 & \cdots & X_N \end{pmatrix}>0, \tag{18}$$

for $i=1\dots N$ and where $\{X_i\}_{i=1}^{N}$ are symmetric. Moreover, a stable CBDLF is obtained by choosing the filter gains as $L_i=X_i^{-1}Y_i$ for $i=1\dots N$.

Proof: First we note that, by the Schur complements Lemma, the linear matrix inequalities (18) are feasible if and only if there exist a set a symmetric matrices $\{X_i\}_{i=1}^{N}$ and a set of matrices $\{Y_i\}_{i=1}^{N}$, such that

$$X_i-\sum_{j=1}^{N}(X_jA-Y_jC_j)'X_j^{-1}(X_jA-Y_jC_j)>0,\ X_i>0$$

for all $i=1\dots N$. We further have that,

$$X_i-\sum_{j=1}^{N}(A-X_j^{-1}Y_jC_j)'X_j(X_jA-X_j^{-1}Y_jC_j)>0,\ X_i>0$$

By defining $L_i \triangleq X_i^{-1} Y_i$, it follows that

$$X_i - \sum_{j=1}^{N} (A - L_j C_j)' X_j (A - L_j C_j) > 0, \; X_i > 0.$$

Therefore, if thr matrix inequalities (18) are feasible, there exists a set of positive definite matrices $\{X_i\}_{i=1}^N$ and a set of positive matrices $\{S_i\}_{i=1}^N$, such that

$$X_i = \sum_{j=1}^{N} (A - L_j C_j)' X_j (A - L_j C_j) + S_i.$$

By Lemma 3.1, it follows that the linear dynamics (7), without noise, is asymptotically stable, and therefore the system (2 is detectable in the sense of Definition 2.1. ■

IV. Sub-Optimal Consensus-Based Distributed linear Filtering

Obtaining the closed form solution of the optimization problem (11) is a challenging problem, which is in the same spirit as the decentralized optimal control problem. In this section we provide a sub-optimal algorithm for computing the filter gains of the CBDLF, with quantifiable performance in the sense that we compute a set of filtering gains which guarantee a certain level of performance with respect the quadratic cost (10).

A. Finite Horizon Sub-Optimal Consensus-Based Distributed Linear Filtering

The sub-optimal scheme for computing the CBDLF gains results from minimizing an upper bound of the quadratic filtering cost (10). The following proposition gives upper-bounds for the covariance matrices of the estimation errors.

Proposition 4.1: Consider the following coupled difference equations

$$Q_i(k+1) = \sum_{i=1}^{N} p_{ij}(k) \left[\left(A(k) - L_j(k) C_j(k)\right) Q_j(k) \left(A(k) - L_j(k) C_j(k)\right)' + L_j(k) \Sigma_{v_j}(k) L_j(k) \right] + \Sigma_w(k), \tag{19}$$

with $Q_i(0) = \Sigma_i(0)$, for $i = 1 \ldots N$. The following inequality holds

$$\Sigma_i(k) \leq Q_i(k), \tag{20}$$

for $i = 1 \ldots N$ and for all $k \geq 0$.

Proof: The matrix $\Sigma_i(k+1)$ can be explicitly written as

$$\Sigma_i(k+1) = E[\epsilon_i(k+1)' \epsilon_i(k+1)] =$$

$$= E\left[\left(\sum_{j=1}^{N} p_{ij}(k)\Big(A(k)-L_j(k)C_j(k)\Big)\epsilon_j(k)+w(k)-\sum_{j=1}^{N} p_{ij}(k)L_j(k)v_j(k)\right)'\right.$$
$$\left.\left(\sum_{j=1}^{N} p_{ij}(k)\Big(A(k)-L_j(k)C_j(k)\Big)\epsilon_j(k)+w(k)-\sum_{j=1}^{N} p_{ij}(k)L_j(k)v_j(k)\right)\right].$$

Using the fact that the noises $w(k)$ and $v_i(k)$ have zero mean, and they are independent with respect to themselves and the initial state, for every time instant, we can further write

$$\Sigma_i(k+1) = E\left[\left(\sum_{j=1}^{N} p_{ij}(k)\Big(A(k)-L_j(k)C_j(k)\Big)\epsilon_j(k)\right)'\left(\sum_{j=1}^{N} p_{ij}(k)\Big(A(k)-L_j(k)C_j(k)\Big)\epsilon_j(k)\right)\right]+$$
$$+E\left[\left(\sum_{j=1}^{N} p_{ij}(k)L_j(k)v_j(k)\right)'\left(\sum_{j=1}^{N} p_{ij}(k)L_j(k)v_j(k)\right)\right]+\Sigma_w(k).$$

By Remark 1.2, it follows that

$$E\left[\left(\sum_{j=1}^{N} p_{ij}(k)\Big(A(k)-L_j(k)C_j(k)\Big)\epsilon_j(k)\right)'\left(\sum_{j=1}^{N} p_{ij}(k)\Big(A(k)-L_j(k)C_j(k)\Big)\epsilon_j(k)\right)\right] \preceq$$
$$\preceq \sum_{j=1}^{N} p_{ij}(k)\Big(A(k)-L_j(k)C_j(k)\Big)\Sigma_j(k)\Big(A(k)-L_j(k)C_j(k)\Big)'$$

and

$$E\left[\left(\sum_{j=1}^{N} p_{ij}(k)L_j(k)v_j(k)\right)'\left(\sum_{j=1}^{N} p_{ij}(k)L_j(k)v_j(k)\right)\right] \preceq \sum_{j=1}^{N} p_{ij}(k)L_j(k)\Sigma_{v_j}(k)L_j(k)',\ i=1\ldots N.$$

From the previous two expressions, we obtain that

$$\Sigma_i(k+1) \preceq \sum_{j=1}^{N} p_{ij}(k)\Big(A(k)-L_j(k)C_j(k)\Big)\Sigma_j(k)\Big(A(k)-L_j(k)C_j(k)\Big)'+$$
$$+\sum_{j=1}^{N} p_{ij}(k)L_j(k)\Sigma_{v_j}(k)L_j(k)+\Sigma_w(k)$$

We prove (20) by induction. Assume that $\Sigma_i(k) \preceq Q_i(k)$ for all $i=1\ldots N$. Then

$$(A(k)-L_i(k)C_i(k))\,\Sigma_i(k)\,(A(k)-L_i(k)C_i(k))' \preceq (A(k)-L_i(k)C_i(k))\,Q_i(k)\,(A(k)-L_i(k)C_i(k))',$$

and

$$L_i(k)\Sigma_i(k)L_i(k)' \preceq L_i(k)Q_i(k)L_i(k)',\ i=1\ldots N.$$

and therefore

$$\Sigma_i(k+1) \le Q_i(k+1),\ i = 1\ldots N.$$

■

Defining the finite horizon quadratic cost function

$$\bar{J}_K(\mathbf{L}(\cdot)) = \sum_{k=1}^{K}\sum_{i=1}^{N} tr(Q_i(k)), \tag{21}$$

the next Corollary follows immediately.

Corollary 4.1: The following inequalities hold

$$J^K(\mathbf{L}(\cdot)) \le \bar{J}^K(\mathbf{L}(\cdot)), \tag{22}$$

and

$$\limsup_{K\to\infty} \frac{1}{K} J^K(\mathbf{L}) \le \limsup_{K\to\infty} \frac{1}{K} \bar{J}^K(\mathbf{L}) \tag{23}$$

Proof: Follows immediately from Proposition 4.1. ■

In the previous corollary we obtained an upper bound on the filtering cost function. Our sub-optimal consensus based distributed filtering scheme will result from minimizing this upper bound in terms of the filtering gains $\{L_i(k)\}_{i=1}^{N}$:

$$\min_{\mathbf{L}(\cdot)} \bar{J}_K(\mathbf{L}(\cdot)). \tag{24}$$

Proposition 4.2: The optimal solution for the optimization problem (24) is

$$L_i^*(k) = A(k)Q_i^*(k)C_i(k)'\left[\Sigma_{v_i}(k) + C_i(k)Q_i^*(k)C_i(k)'\right]^{-1}, \tag{25}$$

and the optimal value is given by

$$\bar{J}_K^*(\mathbf{L}^*(\cdot)) = \sum_{k=1}^{K}\sum_{i=1}^{N} tr(Q_i^*(k)),$$

where $Q_i^*(k)$ is computed using

$$\begin{aligned} Q_i^*(k+1) = \sum_{j=1}^{N} p_{ij}(k)\Big[&A(k)Q_j^*(k)A(k)' + \Sigma_w(k) - A(k)Q_j^*(k)C_j(k)'\cdot \\ &\cdot\left(\Sigma_{v_j}(k) + C_j(k)Q_j^*(k)C_j(k)'\right)^{-1} C_j(k)Q_j^*(k)A(k)'\Big], \end{aligned} \tag{26}$$

with $Q_i^*(0) = \Sigma_i(0)$ and for $i = 1\ldots N$.

Proof:

Let $\bar{J}_K(\mathbf{L}(\cdot))$ be the cost function when an arbitrary set of filtering gains $\mathbf{L}(\cdot) \triangleq \{L_i(k), k = 0\ldots K-1\}_{i=1}^{N}$ is used in (19). We will show that $\bar{J}_K^*(\mathbf{L}^*(\cdot)) \le \bar{J}_K(\mathbf{L}(\cdot))$, which in turn will show that $\mathbf{L}^*(\cdot) \triangleq \{L_i(k)^*, k = 0\dots K-1\}_{i=1}^N$ is the optimal solution of the optimization problem (24). Let $\{Q_i^*(k)\}_{i=1}^N$ and $\{Q_i(k)\}_{i=1}^N$ be the matrices obtained when $\mathbf{L}^*(\cdot)$ and $\mathbf{L}(\cdot)$, respectively are substituted in (19). In what follows we will show by induction that $Q_i^*(k) \le Q_i(k)$ for $k \ge 0$ and $i = 1\dots N$, which basically prove that $\bar{J}_K^*(\mathbf{L}^*(\cdot)) \le \bar{J}_K(\mathbf{L}(\cdot))$, for any $\mathbf{L}(\cdot)$. For simplifying the proof, we will omit in what follows the time index for some matrices and for the consensus weights.

Substituting $\{L_i^*(k), k \ge 0\}_{i=1}^N$ in (19), after some matrix manipulations we get

$$Q_i^*(k+1) = \sum_{j=1}^N p_{ij}\Big[AQ_j^*(k)A' + \Sigma_w - AQ_j^*(k)C_j'(\Sigma_{v_j} +$$

$$+C_jQ_j^*(k)C_j')^{-1}C_jQ_j^*(k)A'\Big],\ Q_i^*(0) = \Sigma_i(0),\ i = 1\dots N.$$

We can derive the following matrix identity (for simplicity we will give up the time index):

$$(A + L_iC_i)Q_i(A_i + L_iC_i)' + L_i\Sigma_{v_i}L_i' = (A + L_i^*C_i)Q_i(A_i + L_i^*C_i)' + L_i^*\Sigma_{v_i}L_i^{*\prime} +$$

$$+(L_i - L_i^*)(\Sigma_{v_i} + C_iQ_iC_i')(L_i - L_i^*). \tag{27}$$

Assume that $Q_i^*(k) \le Q_i(k)$ for $i = 1\dots N$. Using identity (27), the dynamics of $Q_i(k)^*$ becomes

$$Q_i^*(k+1) = \sum_{j=1}^N p_{ij}\Big((A + L_j(k)C_j)Q_j(k)(A + L_j(k)C_j)' + L_j(k)\Sigma_{v_j}L_j(k)' -$$

$$-(L_j(k) - L_j^*(k))(\Sigma_{v_j} + C_jQ_j(k)C_j')(L_j(k) - L_j^*(k))' + \Sigma_w\Big).$$

The difference $Q_i^*(k+1) - Q_i(k+1)$ can be written as

$$Q_i(k+1)^* - Q_i(k+1) = \sum_{j=1}^N p_{ij}\Big((A + L_j(k)C_j)(Q_j^*(k) - Q_j(k))(A + L_j(k)C_j)' -$$

$$-(L_j(k) - L_j^*(k))(\Sigma_{v_j} + C_jQ_j(k)C_j')(L_j(k) - L_j^*(k))'\Big).$$

Since $\Sigma_{v_i} + C_iQ_i(k)C_i'$ is positive definite for all $k \ge 0$ and $i = 1\dots N$ and since we assumed that $Q_i^*(k) \le Q_i(k)$, it follows that $Q_i^*(k+1) \le Q_i(k+1)$. Hence we obtain that

$$\bar{J}_K^*(\mathbf{L}^*(\cdot)) \le \bar{J}_K(\mathbf{L}(\cdot)),$$

for any set of filtering gains $\mathbf{L}(\cdot) = \{L_i(k), k = 0\dots K-1\}_{i=1}^N$, which concludes the proof. ■

We summarize in the following algorithm the sub-optimal CBDLF scheme resulted from Proposition 4.2.

Algorithm 1: Consensus Based Distributed Linear Filtering Algorithm

Input: $\mu_0,\ P_0$

Initialization: $\hat{x}_i(0) = \mu_0,\ Y_i(0) = \Sigma_0$

while new data exists

Compute the filter gains:

$$L_i \leftarrow AY_iC_i'(\Sigma_{v_i} + C_iY_iC_i')^{-1}$$

Update the state estimates:

$$\varphi_i \leftarrow A\hat{x}_i + L_i(y_i - C - i\hat{x}_i)$$

$$\hat{x}_i \leftarrow \sum_j p_{ij}\varphi_j$$

Update the matrices Y_i:

$$Y_i \leftarrow \sum_{j=1}^{N} p_{ij}\Big((A - L_jC_j)Y_j(A - L_jC_j)' + L_j\Sigma_{v_j}L_j'\Big) + \Sigma_w$$

B. Infinite Horizon Consensus Based Distributed Filtering

We now assume that the matrices $A(k)$, $\{C_i(k)\}_{i=1}^N$, $\{\Sigma_{v_i}(k)\}_{i=1}^N$ and $\Sigma_w(k)$ and the weights $\{p_{ij}(k)_{i,j=1}^N\}$ are time invariant. We are interested in finding out under what conditions Algorithm 1 converges and if the filtering gains produce stable estimates. From the previous section we note that the optimal infinite horizon cost can be written

$$\bar{J}_\infty^* = \lim_{k\to\infty}\sum_{i=1}^{N} tr(Q_i^*(k)),$$

where the dynamics of $Q_i(k)^*$ is given by

$$Q_i^*(k+1) = \sum_{j=1}^{N} p_{ij}\Big[AQ_j^*(k)A' + \Sigma_w - AQ_j^*(k)C_j'\Big(\Sigma_{v_j} + C_jQ_j^*(k)C_j'\Big)^{-1}C_jQ_j^*(k)A'\Big], \tag{28}$$

and the optimal filtering gains are given by

$$L_i^*(k) = AQ_i^*(k)C_i'\left[\Sigma_{v_i} + C_iQ_i^*(k)C_i'\right]^{-1},$$

for $i = 1\ldots N$. Assuming that (28), converges, the optimal value of the cost $\bar{J}_\infty^*$ is given by

$$\bar{J}_\infty^* = \sum_{i=1}^{N} tr(\bar{Q}_i),$$

where $\{\bar{Q}_i\}_{i=1}^N$ satisfy

$$\bar{Q}_i = \sum_{j=1}^{N} p_{ij}\left[A\bar{Q}_jA' + \Sigma_w - A\bar{Q}_jC_j'(\Sigma_{v_j} + C_j\bar{Q}_jC_j')^{-1}C_j\bar{Q}_jA'\right]. \quad (29)$$

Sufficient conditions under which there exists a unique solution of (29) are provided by Proposition A.7, which says that if $(\mathbf{p}, \mathbf{L}, \mathbf{A})$ is detectable and $(\mathbf{A}, \Sigma_v^{1/2}, \mathbf{p})$ is stabilizable in the sens of Definitions A.1 and A.2, respectively, then there is a unique solution of (29) and $\lim_{k\to\infty} Q_i^*(k) = \bar{Q}_i$.

Mimicking Theorem A.12 of [6], it can be shown that a numerical approach to solve (29) (if it has a solution) consists in solving the following convex programming optimization problem

$$\max \quad tr\left(\textstyle\sum_{i=1}^N Q_i\right) \quad (30)$$

$$\begin{pmatrix} -Q_i + \sum_{j=1}^N p_{ij}AQ_jA' + \Sigma_w & \sqrt{p_{i1}}C_1Q_1A' & \dots & \sqrt{p_{i1}}C_NQ_NA' \\ \sqrt{p_{i1}}AQ_1C_1' & \Sigma_{v_1} + C_1Q_1C_1' & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{p_{iN}}AQ_NC_N' & 0 & \dots & \Sigma_{v_N} + C_NQ_NC_N' \end{pmatrix} \geq 0$$

$$Q_i \succ 0,\ i = 1 \dots N.$$

V. Connection with the Markovian Jump Linear Systems state estimation

In this section we present a connection between the detectability of (2) in the sense of Definition 2.1 and the detectability property of a MJLS, which is going to be defined in what follows. We also show that the optimal gains of a linear filter for the state estimation of the aforementioned MJLS can be used to approximate the solution of the optimization problem (11), which gives the optimal CBDLF. We assume that the matrix $P(k)$ describing the communication topology of the sensors is *irreducible* and *doubly stochastic* and we assume, without loss of generality, that the matrices $\{C_i(k), k \geq 0\}_{i=1}^N$ in the sensing model (3), have the same dimension. We define the following Markovian jump linear system

$$\begin{aligned} \xi(k+1) &= \tilde{A}_{\theta(k)}(k)\xi(k) + \tilde{B}_{\theta(k)}(k)\tilde{w}(k) \\ z(k) &= \tilde{C}_{\theta(k)}(k)\xi(k) + \tilde{D}_{\theta(k)}(k)\tilde{v}(k),\ \xi(0) = \xi_0, \end{aligned} \quad (31)$$

where $\xi(k)$ is the state, $z(k)$ is the output, $\theta(k) \in \{1, \dots, N\}$ is a Markov chain with probability transition matrix $P(k)'$, $\tilde{w}(k)$ and $\tilde{v}(k)$ are independent Gaussian noises with zero mean and identity covariance matrices. Also, ξ_0 is a Gaussian noise with mean μ_0 and covariance matrix Σ_0. We denote by $\pi_i(k)$ the probability distribution of $\theta(k)$ ($Pr(\theta(k) = i) = \pi_i(k)$) and we assume that $\pi_i(0) > 0$. We have that $\tilde{A}_{\theta(k)}(k) \in \{\tilde{A}_i(k)\}_{i=1}^N$, $\tilde{B}_{\theta(k)}(k) \in \{\tilde{B}_i(k)\}_{i=1}^N$, $\tilde{C}_{\theta(k)}(k) \in \{\tilde{C}_i(k)\}_{i=1}^N$ and $\tilde{D}_{\theta(k)}(k) \in \{\tilde{D}_i(k)\}_{i=1}^N$, where the index i refers to the state i of $\theta(k)$. We set

$$\begin{aligned} \tilde{A}_i(k) = A(k), \qquad & \tilde{B}_i(k) = \tfrac{\sqrt{\pi_i(0)}}{\sqrt{\pi_i(k)}}\Sigma_w^{1/2}(k), \\ \tilde{C}_i(k) = \tfrac{1}{\sqrt{\pi_i(0)}}C_i(k), \quad & \tilde{D}_i(k) = \tfrac{1}{\sqrt{\pi_i(k)}}\Sigma_{v_i}^{1/2}(k), \end{aligned} \tag{32}$$

for all $i, k \geq 0$ (note that since $P(k)$ is assumed doubly stochastic and irreducible and $\pi_i(0) > 0$, we have that $\pi_i(k) > 0$ for all $i, k \geq 0$). In addition, ξ_0, $\theta(k)$, $\tilde{w}(k)$ and $\tilde{v}(k)$ are assumed independent for all $k \geq 0$. The random process $\theta(k)$ is also called *mode*. Assuming that the mode is directly observed, a linear filter for the state estimation is given by

$$\hat{\xi}(k+1) = \tilde{A}_{\theta(k)}(k)\hat{\xi}(k) + M_{\theta(k)}(k)(z(k) - \tilde{C}_{\theta(k)}(k)\hat{\xi}(k)), \tag{33}$$

where we assume that the filter gain $M_{\theta(k)}$ depends only on the current mode. The dynamics of the estimation error $e(k) \triangleq \xi(k) - \hat{\xi}(k)$ is given by

$$\begin{aligned} e(k+1) = &\left(\tilde{A}_{\theta k}(k) - M_{\theta(k)}(k)\tilde{C}_{\theta(k)}(k)\right)e(k) + \\ &+ \tilde{B}_{\theta(k)}(k)w(k) - M_{\theta(k)}(k)\tilde{D}_{\theta(k)}(k)v(k). \end{aligned} \tag{34}$$

Let $\mu(k)$ and $Y(k)$ denote the mean and the covariance matrix of $e(k)$, i.e. $\mu(k) \triangleq E[e(k)]$ and $Y(k) \triangleq E[e(k)e(k)']$, respectively. We define also the mean and the covariance matrix of $e(k)$, when the system is in mode i, i.e. $\mu_i(k) \triangleq E[e(k)\mathbb{1}_{\{\theta(k)=i\}}]$ and $Y_i(k) \triangleq E[e(k)e(k)'\mathbb{1}_{\{\theta(k)=i\}}]$, where $\mathbb{1}_{\{\theta(k)=i\}}$ is the indicator function. It follows immediately that $\mu(k) = \sum_{i=1}^N \mu_i(k)$ and $Y(k) = \sum_{i=1}^N Y_i(k)$.

Definition 5.1: The optimal linear filter (33) is obtain by minimizing the following quadratic finite horizon cost function

$$\tilde{J}_K(\mathbf{M}(\cdot)) = \sum_{k=1}^{K} tr(Y(k)) = \sum_{k=1}^{K}\sum_{i=1}^{N} tr(Y_i(k)), \tag{35}$$

where $\mathbf{M}(\cdot) \triangleq \{M_i(k), k = 0 \ldots K-1\}_{i=1}^N$ are the filter gains and where $M_i(k)$ corresponds to $M_{\theta(k)}(k)$ when $\theta(k)$ is in mode i. We can give a similar definition for an optimal steady state filter using the infinite horizon quadratic cost function.

Definition 5.2: Assume that the matrices $\tilde{A}_i(k)$, $\tilde{C}_i(k)$ and $P(k)$ are constant for all $k \geq 0$. We say that the Markovian jump linear system (31) is *mean square detectable* if there exits $\{M_i\}_{i=1}^N$ such that $\lim_{k\to\infty} E[\|e(k)\|^2] = 0$, when the noises $\tilde{w}(k)$ and $\tilde{v}(k)$ are set to zero.
The next result makes the connection between the detectability of the MJLS defined above the distributed detectability of the process (2).

Proposition 5.1: If the Markovian jump linear system (31) is mean square detectable, then the linear stochastic system (2) is detectable in the sense of Definition 2.1.

Proof: In the context of this proposition, the dynamics of the estimation error for the MJLS (31) becomes

$$e(k+1) = (A - M_{\theta(k)}\tilde{C}_{\theta(k)})e(k),\ e(0) = e_0,$$

where $\tilde{C}_i = C_i$. It is not difficult to check that the dynamic equations for the covariance matrices $\{Y_i(k)\}_{i=1}^N$ and the mean vectors $\{\mu_i(k)\}_{i=1}^N$ are given by

$$Y_i(k+1) = \sum_{j=1}^{N} p_{ij}(A - M_j \frac{1}{\sqrt{\pi_i(0)}} C_j) Y_j(k)(A - M_j \frac{1}{\sqrt{\pi_i(0)}} C_j)', \tag{36}$$

with $Y_i(0) = Y_i^0$ and

$$\mu_i(k+1) = \sum_{j=1}^{N} p_{ij}(A - M_j \frac{1}{\sqrt{\pi_i(0)}} C_j)\mu_j(k), \mu_i(0) = \mu_i^0, \tag{37}$$

for $i = 1 \ldots N$. Since the MJLS is assumed mean square detectable it follows that there exists a set of matrices $\{M_i\}_{i=1}^N$ such that (36) is asymptotically stable. But this also implies (see for instance Proposition 3.6 of [6]) that (37) is asymptotically stable as well. Setting $L_i = \pi_i(0)M_i$, we see that (37) is identical to equation (7) and therefore (7) is asymptotically stable (when ignoring the noise). Hence, (2) is detectable in the sense of Definition 2.1. ■

The next result establishes that the optimal gains of the filter (33) can be used to approximate the solution of the optimzation problem (11).

Proposition 5.2: Let $\mathbf{M}^*(\cdot) \triangleq \{M_i^*(k), k = 0, \ldots, K-1\}_{i=1}^N$ be the optimal gains of the linear filter (33). If we set $L_i(k) = \frac{1}{\sqrt{\pi_i(0)}} M_i^*(k)$ as filtering gains in the CBDLF scheme, then the filter cost function (10) is guaranteed to be upper bounded by

$$J_K(\mathbf{L}(\cdot)) \leq \sum_{k=0}^{K} \sum_{i=1}^{N} \frac{1}{\pi_i(0)} tr(Y_i^*(k)), \tag{38}$$

where $Y_i^*(k)$ are the covariance matrices resulting from minimizing (35).

Proof:

By Theorem 5.5 of [6], the filtering gains that minimize (35) are given by

$$M_i^*(k) = \tilde{A}_i(k) Y_i^*(k) \tilde{C}_i(k)' \left[\pi_i(k)\tilde{D}_j(k)\tilde{D}_j(k)' + \tilde{C}_i(k) Y_i^*(k) \tilde{C}_i(k)'\right]^{-1}, \tag{39}$$

for $i = 1 \ldots N$, where $Y_i^*(k)$ satisfies

$$\begin{aligned} Y_i^*(k+1) = \textstyle\sum_{j=1}^{N} p_{ij}(k) \Big[& \tilde{A}_j(k) Y_j^*(k) \tilde{A}_j(k)' + \pi_j(k)\tilde{B}_j(k)\tilde{B}_j(k)' - \\ & -\tilde{A}_j(k) Y_j^*(k) \tilde{C}_j(k)' \left(\pi_j(k)\tilde{D}_j(k)\tilde{D}_j(k)' + \tilde{C}_j(k) Y_j^*(k) \tilde{C}_j(k)'\right)^{-1} \tilde{C}_j(k) Y_j^*(k) \tilde{A}_j(k)' \Big]. \end{aligned} \tag{40}$$

In what follows we will show by induction that $Y_i^*(k) = \pi_i(0)Q_i^*(k)$ for all $i,k \geq 0$, where $Q_i^*(k)$ satisfies (26). For $k=0$ we have $Y_i^*(0) = \pi_i(0)Y^*(0) = \pi_i(0)\Sigma_0 = \pi_i(0)Q_i^*(0)$. Let us assume that $Y_i^*(k) = \pi_i(0)Q_i^*(k)$. Then, from (32) we have

$$\begin{aligned} &\pi_j(k)\tilde{B}_j(k)\tilde{B}_j(k)' = \pi_i(0)\Sigma_w(k), \quad \pi_j(k)\tilde{D}_j(k)\tilde{D}_j(k)' = \Sigma_{v_i}(k), \\ &\pi_j(k)\tilde{D}_j(k)\tilde{D}_j(k)' + \tilde{C}_j(k)Y_j^*(k)\tilde{C}_j(k)' = \Sigma_{v_j}(k) + C_j(k)Q_j^*(k)C_j(k)'. \end{aligned} \tag{41}$$

Also,

$$M_i^*(k) = \pi_i(0)A(k)Q_i^*(k)C_i(k)'\left[\Sigma_{v_j}(k) + C_j(k)Q_j^*(k)C_j(k)'\right]^{-1}, \tag{42}$$

and from (25) we get that $M_i^*(k) = \sqrt{\pi_i(0)}L_i^*(k)$. From (40) and (41) it can be easily argued that $Y_i^*(k+1) = \pi_i(0)Q_i^*(k+1)$. By Corollary 4.1 we have that

$$J_K(\mathbf{L}(\cdot)) \leq \bar{J}_K(\mathbf{L}(\cdot)),$$

for any set of filtering gains $\mathbf{L}(\cdot)$ and in particular for $L_i(k) = \frac{1}{\pi_i(0)}M_i^*(k) = L_i^*(k)$, for all i and k. But since

$$\bar{J}_K(\mathbf{L}^*(\cdot)) = \sum_{k=0}^{K}\sum_{i=1}^{N}\frac{1}{\pi_i(0)}Y_i^*(k),$$

the result follows.

■

VI. Conclusions

In this note we addressed three problems. First we provided (testable) sufficient conditions under which stable consensus-based distributed linear filters can be obtained. Second, we gave a sub-optimal, linear filtering scheme, which can be implemented in a distributed manner and is valid for time varying communication topologies as well, and which guarantees a certain level of performance. Third, under the assumption that the stochastic matrix used in the consensus step is doubly stochastic we showed that if an appropriately defined Markovian jump linear system is detectable, then the stochastic process of our interest is detectable as well. We also showed that the optimal gains of the consensus-based distributed linear filter scheme can be approximated by using the optimal linear filter for the state estimation of a particular Markovian jump linear system.

References

[1] M. Alighanbari and J.P. How, ”An Unbiased Kalman Consensus Algorithm”, *Proceedings of the 2006 American Control Conference*, pages 3519-3524, June 14-16, 2006.

[2] V. Borkar and P. Varaiya, ”Asymptotic agreement in distributed estimation”, *IEEE Transaction on Automatic Control*, vol. ac-27, no. 3, 1982.

[3] R. Carli, A. Chiuso, L. Schenato and S. Zampieri, ”Distributed Kalman Filtering Based on Consensus Strategies”, *IEEE Journal on selected area in communication*, vol. 26, no. 4, pages 622-633, May 2008.

[4] O.L.V. Costa and M.D. Fragoso, ”Stability results for discrete-time linear systems with Markovian jumping parameters,“ *Journal of Mathematical Analysis and Application*, no. 179(1993), pp. 154-178.

[5] O.L.V. Costa, ” Discrete-Time Coupled Riccati Equations for Systems with Markov Switching Parameters,“ *Journal of Mathematical Analysis and Applications*, no. 194(1995), pp. 197-216.

[6] O.L.V. Costa, M.D, Fragoso and R.P. Marques, ”Discrete-Time Markov Jump Linear Systems,“ Springer 2005.

[7] D. Teneketzis and P. Varaiya, ”Consensus in Distributed Estimation,” *Advances in Statistical Signal Processing*, JAI Press, H.V. Poor Ed., pp. 361-386, January 1988.

[8] R. Olfati-Saber, ”Distributed Kalman Filtering for Sensor Networks”, *Proceedings of the* 46^{th} *IEEE Conference on Decision and Control*, pages 5492-5498, 2007

[9] R. Olfati-Saber, ”Distributed Kalman Filter with Embedded Consensus Filters,” *Proc. of the* 44^{th} *IEEE Conference on Decision and Control*, Dec. 2005.

[10] R. Olfati-Saber, ”Distributed Tracking for Mobile Sensor Networks with Information-Driven Mobility”, *Proc. of the 2007 American Control Conference*, pages 4606-4612, July 11-13, 2007.

[11] A. Speranzon, C. Fischione, K. H. Johansson and A. Sangiovanni-Vincentelli, ”A Distributed Minimum Variance Estimator for Sensor Networks”, *IEEE Journal on Select Areas in Communication*, vol. 26, no. 4, pages 609-621, May 2008.

[12] Y. Zhu, Z. You, J. Zhao, K. Zhang, X.R. Li, ”The optimality for the distributed Kalman ”ltering fusion with feedback”, *Automatica* no. 37, pages 1489-1493, 2001.

Appendix

A. Properties of a special class of difference matrix equations

Given two positive integers N and n, a sequence of positive numbers $\mathbf{p} = \{p_{ij}\}_{i,j=1}^{N}$ and a set of matrices $\mathbf{F} = \{F_i\}_{i=1}^{N}$, we consider the following matrix difference equations

$$W_i(k+1) = \sum_{j=1}^{N} p_{ij} F_j W_j(k) F_j',\ W_i(0) = W_i^0,\ i = 1 \dots N. \tag{43}$$

Additionally, consider a similar set of matrix difference equations

$$W_i(k+1) = \sum_{j=1}^{N} p_{ji} F_j' W_j(k) F_j,\ W_i(0) = W_i^0,\ i = 1 \dots N. \tag{44}$$

Proposition A.1 ([4]): The dynamics (43) is asymptotically stable if and anly if the dynamics (44) is asymptotically stable.

Related to the above dynamic equations, we introduce the following stabilizability and detectability definitions.

Definition A.1 ([5]): Given a set of matrices $\mathbf{C} = \{C_i\}_{i=1}^N$, we say that $(\mathbf{p}, \mathbf{L}, \mathbf{A})$ is *detectable* if there exists a set of matrices $\mathbf{L} = \{L_i\}_{i=1}^N$ such that the dynamics (43) is asymptotically stable, where $F_i = A_i - L_i C_i$, for $i = 1 \dots N$.

Definition A.2 ([5]): Given a set of matrices $\mathbf{C} = \{C_i\}_{i=1}^N$, we say that $(\mathbf{A}, \mathbf{L}, \mathbf{p})$ is *stabilizable*, if there exists a set of matrices $\mathbf{L} = \{L_i\}_{i=1}^N$ such that the dynamics (43) is asymptotically stable, where $F_i = A_i - C_i L_i$, for $i = 1 \dots N$.

Remark A.1: Given a semipositive definite matrix X and a positive definite matrix Y, the following holds:

$$\min_{i=1\dots n} \lambda_i(Y) tr(X) \leq tr(YX) \leq \max_{i=1\dots n} \lambda_i(Y) tr(X)$$

Proposition A.2: If there exists a set of symmetric positive definite matrices $\{V_i\}_{i=1}^N$ such that

$$V_i = \sum_{j=1}^N p_{ji} F_i' V_j F_i + S_i, \tag{45}$$

for some set of symmetric positive set of matrices $\{S_i\}_{i=1}^N$, then the dynamics (43) is asymptotically stable.

Proof: We use the same idea as in the proof of Theorem 3.19 of [6] and define the following Lyapunov function

$$\Phi(k) = \sum_{i=1}^N tr(W_i(k) V_i).$$

In the following we show that the difference $\Phi(k+1) - \Phi(k)$ is negative for all $k \geq 0$, from which we infer the asymptotic stability of (43). We get that

$$\Phi(k+1) - \Phi(k) = tr\left(\sum_{i=1}^N \left(\sum_{j=1}^N p_{ij} F_j W_j(k) F_j'\right) V_i - W_i(k) V_i\right) =$$

$$= tr\left(\sum_{i=1}^N W_i(k) \left(\sum_{j=1}^N p_{ji} F_i V_j(k) F_i' - V_i\right)\right) = \sum_{i=1}^N tr(W_i(k) S_i).$$

Since $\{W_i(k)\}_{i=1}^N$ are positive semi-definite matrices for $k \geq 0$ and $\{S_i\}_{i=1}^N$ are positive definite, by Remark A.1, it follows that

$$\Phi(k+1) - \Phi(k) < 0, \ k \geq 0.$$

■

Proposition A.3: If there exists a set of symmetric positive definite matrices $\{V_i\}_{i=1}^N$ such that

$$V_i = \sum_{j=1}^{N} p_{ij} F_i' V_j F_i + S_i, \tag{46}$$

for some set of symmetric positive set of matrices $\{S_i\}_{i=1}^N$, then the dynamics (43) is asymptotically stable.

Proof: Using the same approach as in the previous proposition, we prove the asymptotic stability of dynamics (44). Using Proposition A.1, the result follows. ■

Proposition A.4: If the following linear matrix inequalities are feasible

$$\begin{pmatrix} X_i & \sqrt{p_{1i}} X_i F_i & \sqrt{p_{2i}} F_i' X_i & \cdots & \sqrt{p_{Ni}} F_i' X_i \\ \sqrt{p_{1i}} X_i F_i & X_1 & 0 & \cdots & 0 \\ \sqrt{p_{2i}} X_i F_i & 0 & X_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sqrt{p_{Ni}} X_i F_i & 0 & 0 & \cdots & X_N \end{pmatrix} \succ 0, \; i = 1 \ldots N, \tag{47}$$

where $\{X_i\}_{i=1}^N$ are the unknown variable, then the dynamics (43) is asymptotically stable.

Proof: By the Schur complement lemma, (47) are feasible if and only if

$$X_i - \sum_{j=1}^{N} p_{ji} X_i F_i X_j^{-1} F_i' X_i \succ 0, \; X_i \succ 0, \; i = 1 \ldots N. \tag{48}$$

By defining $V_i \triangleq X_i^{-1}$, $i = 1 \ldots N$, (48), becomes

$$V_i - \sum_{j=1}^{N} p_{ji} F_i V_j F_i' \succ 0, \; V_i \succ 0, \; i = 1 \ldots N.$$

By Proposition A.2, (43) is asymptotically stable. ■

Inspired by Proposition A.4, detectability and stabilizability tests, in the sense of Definitions A.1 and A.2, respectively, can be formulated in terms of the feasibility of a set of linear matrix inequalities.

Proposition A.5 (detectability test): If the following matrix inequalitie are feasible

$$\begin{pmatrix} X_i & \sqrt{p_{i1}}(X_i A_i - Y_i C_i) & \sqrt{p_{i2}}(X_i A_i - Y_i C_i) & \cdots & \sqrt{p_{iN}}(X_i A_i - Y_i C_i) \\ \sqrt{p_{i1}}(X_i A_i - Y_i C_i)' & X_1 & 0 & \cdots & 0 \\ \sqrt{p_{i2}}(X_i A_i - Y_i C_i)' & 0 & X_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sqrt{p_{iN}}(X_i A_i - Y_i C_i)' & 0 & 0 & \cdots & X_N \end{pmatrix} \succ 0, \; i = 1 \ldots N, \tag{49}$$

where $\{X_i\}_{i=1}^N$ and $\{Y_i\}_{i=1}^N$ are the unknown variable, then $(\mathbf{p},\mathbf{L},\mathbf{A})$ is detectable in the sense of Definition A.1. Moreover chosing $L_i = X_i^{-1}Y_i$, for $i = 1\dots N$, the dynamics (43) is asymptotically stable.

Proof: By the Schur complement lemma, (49) are feasible if and only if

$$X_i - \sum_{j=1}^{N} p_{ij}(X_iA_i - Y_iC_i)X_j^{-1}(X_iA_i - Y_iC_i)' \succ 0,\ X_i \succ 0,\ i = 1\dots N. \tag{50}$$

By defining $L_i \triangleq X_i^{-1}Y_i$ and $V_i \triangleq X_i^{-1}$, $i = 1\dots N$, (48), becomes

$$V_i - \sum_{j=1}^{N} p_{ij}F_iV_jF_i' \succ 0,\ V_i \succ 0,\ i = 1\dots N.$$

By Proposition A.3, $(\mathbf{p},\mathbf{L},\mathbf{A})$ is detectable in the sense of Definition A.1. ■

Proposition A.6 (stabilizability test): If the following matrix inequalities are feasible

$$\begin{pmatrix} X_i & \sqrt{p_{1i}}(X_iA_i - C_iY_i)' & \sqrt{p_{2i}}(X_iA_i - C_iY_i)' & \cdots & \sqrt{p_{Ni}}((X_iA_i - C_iY_i)' \\ \sqrt{p_{1i}}(X_iA_i - C_iY_i) & X_1 & 0 & \cdots & 0 \\ \sqrt{p_{2i}}(X_iA_i - C_iY_i) & 0 & X_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sqrt{p_{Ni}}(X_iA_i - C_iY_i) & 0 & 0 & \cdots & X_N \end{pmatrix} \succ 0,\ i = 1\dots N, \tag{51}$$

where $\{X_i\}_{i=1}^N$ and $\{Y_i\}_{i=1}^N$ are the unknown variable, then $(\mathbf{A},\mathbf{L},\mathbf{p})$ is stabilizable in the sense of Definition A.2. Moreover chosing $L_i = Y_iX_i^{-1}$, for $i = 1\dots N$, the dynamics (43) is asymptotically stable.

Proof: By the Schur complement lemma, (51) are feasible if and only if

$$X_i - \sum_{j=1}^{N} p_{ji}(X_iA_i - Y_iC_i)'X_j^{-1}(X_iA_i - Y_iC_i) \succ 0,\ X_i \succ 0,\ i = 1\dots N. \tag{52}$$

By defining $L_i \triangleq X_i^{-1}Y_i$ and $V_i \triangleq X_i^{-1}$, $i = 1\dots N$, (48), becomes

$$V_i - \sum_{j=1}^{N} p_{ji}F_i'V_jF_i \succ 0,\ V_i \succ 0,\ i = 1\dots N.$$

By Proposition A.2, $(\mathbf{p},\mathbf{L},\mathbf{A})$ is stabilizable in the sense of Definition A.2. ■

B. Discrete-time coupled Riccati equations

Consider the following coupled Riccati difference equations

$$Q_i(k+1) = \sum_{i=1}^{N} p_{ij}\Big(A_j Q_j(k) A_j' - A_j Q_j(k) C_j' (C_j Q_j(k) C_j' + \Sigma_{v_j})^{-1} C_j Q_j(k) A_j' + \Sigma_w\Big), \tag{53}$$

$Q_i(0) = Q_i^0 \succ 0,\ i = 1\dots N$, where $\{\Sigma_{v_i}\}_{i=1}^N$ and Σ_w are symmetric positive definite matrices.

Proposition A.7: Let $\Sigma_v^{1/2} = \{\Sigma_{v_i}^{1/2}\}_{i=1}^N$, where $\Sigma_{v_i} = \Sigma_{v_i}^{1/2\prime}\Sigma_{v_i}^{1/2}$. Suppose that $(\mathbf{p}, \mathbf{C}, \mathbf{A})$ is detectable and that $(\mathbf{A}, \Sigma_v^{1/2}, \mathbf{p})$ is stabilizable in the sense of Definitions A.1 and A.2, respectively. Then there exists a unique set of symmetric positive definite matrices $\bar{\mathbf{Q}} = \{\bar{Q}_i\}_{i=1}^N$ satisfying

$$\bar{Q}_i = \sum_{i=1}^{N} p_{ij}\Big(A_j \bar{Q}_j A_j' - A_j \bar{Q}_j C_j' (C_j \bar{Q}_j C_j' + \Sigma_{v_j})^{-1} C_j \bar{Q}_j A_j' + \Sigma_w\Big),\ i = 1\dots N. \tag{54}$$

Moreover, for any initial conditions $Q_i^0 \succ 0$, we have that $\lim_{k\to\infty} Q_i(k) = \bar{Q}_i$.

Proof: The proof can be mimicked after the proof of Theorem 1 of [5]. Compared to our case, in Theorem 1 of [5] a scalar term, taking values between zero and one, multiplies the matrix Σ_{v_j} in (54). However it is not difficult to note that the result holds even in the case where this scalar term takes value one, which corresponds to out setup. ■