Distributed Filter Design for Cooperative $\mathscr{H}_\infty$-type Estimation

Jingbo Wu   Li Li   Valery Ugrinovskii   Frank Allgöwer

Abstract— In this paper, we consider the distributed robust filtering problem, where estimator design is based on a set of coupled linear matrix inequalities (LMIs). We separate the problem and show that the method of multipliers can be applied to obtain a solution efficiently and in a decentralized fashion, i.e. all local estimators can compute their filter gains locally, with communications restricted to their neighbors.

I. INTRODUCTION

Estimator design has been an essential part of controller design ever since the development of state-space based controllers. A milestone was laid by the Kalman Filter in 1960 [1].

While in the classical estimator design one estimator is used for the entire system, distributed estimators have gained attention since a distributed Kalman Filter was presented in [2], [3]. In a distributed estimator setup, multiple estimators create an estimate of the system's state, either individually [4] or cooperatively. In the latter case, even when every single estimator may be able to obtain an estimate of the state on its own, cooperation reduces the effects of model and measurement disturbances [5]. Also, the situations are not uncommon where individual estimators are unable to obtain an estimate of the state on their own and cooperation becomes an essential prerequisite [6], [7]. The node estimators may even not have a model of the full system, but only know a part of the system [8].

However, even though the setup consists of distributed estimation units without a central coordinator, in many known approaches the design process itself requires a central coordination unit. In some practical application examples, where the design process can be done offline, this may not be a significant drawback. On the other hand, in many applications especially those involving distributed sensor networks with varying communication topology, a centralized computation of observer parameters represents a severe limitation. Practicality of a distributed system demands that the estimator design process is to be carried out in a distributed

manner as well. If network needs to adapt to some changes, such as a change in the plant or change in the network structure, this allows each node to reconfigure using only local communications and computation only.

In this paper, we provide a complete analysis of one distributed estimation problem where such a distributed design scheme is possible. Specifically, we adopt the setup from [6] concerned with the problem of distributed estimation with $\mathscr{H}_\infty$ consensus performance. As a matter of fact, in [6] a gradient-descent-type algorithm was proposed that can be used to calculate the filter gains in a distributed manner. Although the proposed gradient type algorithm demonstrated a possibility of computing the estimator parameters in principle, a practical application of that algorithm is hindered due to slow convergence observed even in low dimensional examples. Also, implementation of the decentralized design scheme proposed in [6] requires bidirectional communications between the network nodes, which essentially requires the communication graph to be undirected for the purpose of the estimator design. In this paper, we address the problem of designing distributed estimators by using distributed optimization methods presented in [9],[10]. Distributed optimization methods are widely applied in networked systems, see e.g. [11],[12],[13]. The contribution of this paper is to show that the problem of designing distributed estimators is amenable to the methodology of distributed optimization as well. Although the design scheme is proposed for the specific class of algorithms in [6], it illustrates all the steps necessary to devise similar design schemes for other distributed estimation algorithms and distributed optimization subject to LMI-constraints in general.

The rest of the paper is organized as follows: We first introduce the notation and some preliminaries on graph theory. Then, we revisit some essential results published in [6] and discuss there limitations with respect to numerical optimization. Section III is dedicated for introducing the proposed optimization scheme. In Section IV, we give a mathematical example, and Section V concludes the paper.

II. PRELIMINARIES AND BACKGROUND

In this section, we introduce the basic definitions and results which our main results will build on.

A. Notation

Let P be a symmetric matrix. If P is positive definite, it is denoted $P > 0$, and we write $P < 0$, if P is negative definite. 0 denotes a matrix of suitable dimension, with all entries equal 0. Moreover, for vectors $x \in \mathscr{R}^n$ we use the Euclidean vector norm $\|x\| = \sqrt{x^\top x}$ and the weighted vector norm

$\|x\|_P = \sqrt{x^\top P x}$ for symmetric matrices $P > 0$. For matrices $A \in \mathscr{R}^{n \times m}$, we use the Frobenius norm $\|A\| = \sqrt{tr(A^\top A)}$ and the induced norm $\|A\|_2 = \sup_{\|x\| \neq 0}(\|Ax\|/\|x\|)$. $\mathscr{L}_2$ denotes the Lebesgue space of $\mathbb{R}^n$−valued vector-functions $z(\cdot)$, defined on the time interval $[0, \infty)$ with the norm $\|z\|_2 = \sqrt{\int_0^\infty \|z(t)\|^2 dt}$. The vectorization $vec(\cdot)$ maps any matrix $A \in \mathscr{R}^{n \times m}$ to the $n \cdot m$-dimensional vector $vec(A)$ formed by the stacked columns of A.

B. Communication graphs

In this section we summarize some notation from the graph theory. We use directed, unweighted graphs $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ to describe the communication topology between the individual agents. $\mathscr{V} = \{v_1, ..., v_N\}$ is the set of vertices, where $v_k \in \mathscr{V}$ represents the k-th agent. $\mathscr{E} \subseteq \mathscr{V} \times \mathscr{V}$ are the sets of edges, which model the information flow, i.e. the k-th agent receives information from agent j if and only if $(v_j, v_k) \in \mathscr{E}$. The set of vertices that agent k receives information from is called the neighborhood of agent k, which is denoted by $\mathscr{N}_k = \{j : (v_j, v_k) \in \mathscr{E}\}$. The set of vertices that receive information from agent k is called the out-neighborhood of agent k, which is denoted by $\mathscr{M}_k = \{j : (v_k, v_i) \in \mathscr{E}\}$. The in-degree p_k and out-degree q_k of a vertex k is defined as the number of edges in $\mathscr{E}$, which have v_k as their head and tail, respectively.

C. Distributed $\mathscr{H}_\infty$ estimation and LMI conditions

We now present the underlying distributed estimation problem from [6]. It involves LMI design conditions, which are the main object of interest in the paper. Our main objective is to show that these LMI design conditions are amendable to a distributed solution by using the multiplier method (c.f. [9]) .

The distributed estimation problem with $\mathscr{H}_\infty$ consensus of estimates posed in [6] involves estimation of the state of the uncertain LTI system described by the differential equation

$$\dot{x} = Ax + B\xi(t), \tag{1}$$

where $x \in \mathbb{R}^n$ is the system state variable to be estimated and $\xi(t) \in \mathscr{L}_2$ is a disturbance function. N estimators are applied, each estimator receives a r_k-dimensional measurement

$$y_k = C_k x + D_k \xi + \overline{D}_k \eta_k(t). \tag{2}$$

In (2), $\eta_k(t) \in \mathscr{L}_2$ represents the measurement uncertainty of the local estimator k. In particular, it is assumed that $E_k = D_k D_k^\top + \overline{D}_k \overline{D}_k^\top > 0$.

Remark 1: The assumption that $E_k > 0$ is a standard technical assumption made in nonsingular $\mathscr{H}_\infty$ control problems [14]. It is obviously satisfied in the case when all measurements are affected by disturbances, which is evidently satisfied in practical applications. This assumption is later used to guarantee boundedness of the solution set.

The estimators form a network of interconnected $\mathscr{H}_\infty$ filters of the form

$$\dot{\hat{x}}_k = A\hat{x}_k + L_k(y_k - C_k \hat{x}_k) + K_k \sum_{j \in \mathscr{N}_k} (\hat{x}_j - \hat{x}_k) \tag{3}$$

with initial condition $\hat{x}_k(0) = 0$. Here the matrices $L_k \in \mathbb{R}^{n \times r_k}$ and $K_k \in \mathbb{R}^{n \times n}$ are the filter gains to be designed.

As it can be seen in (3), the estimators are distributed, i.e. the local estimators create an estimation of the system's state x, solely based on the local output y_k and communication with neighbouring estimators. The problem in [6] was to determine estimator gains L_k, K_k in (3) to satisfy natural internal stability and $\mathscr{H}_\infty$ gain conditions. To introduce these conditions, define the local estimator error as $e_k = x - \hat{x}_k$, and the estimator disagreement function is defined as

$$\Psi(\hat{x}) = \frac{1}{N} \sum_{k=1}^{N} \sum_{j \in \mathscr{N}_k} \|\hat{x}_j - \hat{x}_k\|^2, \tag{4}$$

where $\hat{x} = [\hat{x}_1^\top, ..., \hat{x}_N^\top]^\top$ and $e = [e_1^\top, ..., e_N^\top]^\top$. The estimator design problem is concerned with achieving the following properties:

(i) In the absence of model and measurement disturbances (i.e., when ξ, $\eta_k = 0$), the estimation errors decay so that $e_k \to 0$ asymptotically for all $k = 1, ..., N$.

(ii) The estimators (3) provide guaranteed $\mathscr{H}_\infty$ performance

$$\sup_{x_0, (\xi, \eta_k) \neq 0} \frac{\int_0^\infty \Psi(\hat{x}(t)) dt}{\|x_0\|_P^2 + (1/N) \sum_{k=1}^N \|\eta_k\|_2^2 + \|\xi\|_2^2} \leq \gamma$$

$$\frac{1}{N} \sum_{k=1}^{N} \|e_k\|_2^2 \leq \overline{\gamma} \left(\|x_0\|_P^2 + \frac{1}{N} \sum_{k=1}^{N} \|\eta_k\|_2^2 + \|\xi\|_2^2 \right), \tag{5}$$

for some positive definite matrix P, some $\overline{\gamma} > 0$, and performance index $\gamma > 0$.

Property (ii) requires both the local estimation errors and the estimator disagreement to be bounded with respect to the disturbances in an $\mathscr{H}_\infty$−sense. As shown in [6], LMI-conditions can be found, where the solution delivers estimator gains sufficient for solving the above problem. To present these LMI conditions, define the matrices

$$\widetilde{A}_k = A + \alpha_k I - BD_k^\top E_k^{-1} C_k,$$
$$Q_k = X_k \widetilde{A}_k + \widetilde{A}_k^\top X_k - C_k^\top E_k^{-1} C_k + \beta(p_k + q_k)I,$$
$$\widetilde{B}_k = [B(I - D_k^\top E_k^{-1} D_k) \quad -BD_k^\top E_k^{-1} \overline{D}_k],$$

where $X_k \in \mathbb{R}^{n \times n}$ is a symmetric, positive definite matrix and α_k, β are positive parameters. For the remainder of this paper, we will make two assumptions on the system class.

Assumption 1 The communication graph $\mathscr{G}$ is connected and balanced, i.e. $q_k = p_k$ for all $k = 1, ..., N$.

Assumption 2 For all $k = 1, ..., N$, the tuple $(\widetilde{A}_k, \widetilde{B}_k)$ is controllable.

The LMIs used for designing the estimator gains are

proposed as

$$\begin{bmatrix} Q_k - p_k F_k - p_k F_k^\top & X_k \widetilde{B}_k \\ * & -I \end{bmatrix} < 0 \qquad (6)$$

$$\begin{bmatrix} -\frac{2\alpha_k}{q_k+1}X_k & -\beta I + F_k & \cdots & -\beta I + F_k \\ * & -\frac{2\alpha_{j_1^k}}{q_{j_1^k}+1}X_{j_1^k} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & 0 & \cdots & -\frac{2\alpha_{j_{p_k}^k}}{q_{j_{p_k}^k}+1}X_{j_{p_k}^k} \end{bmatrix} < 0, \qquad (7)$$

$$\begin{bmatrix} -\rho X_k & -F_k^\top \\ -F_k & -X_k \end{bmatrix} < 0, \qquad (8)$$

where X_k, F_k, and β are the variables, $\rho > 0$ is a constant parameter, and $\mathcal{N}_k = \{j_1^k, ..., j_{p_k}^k\}$. We can now formulate a variation of the main result from [6].

Proposition 1: Suppose the interconnection graph $\mathcal{G}$ and the parameters $\alpha_k > 0, \ k = 1, ..., N$, are such that the set

$$\Gamma = \{\beta > 0 : (6)\text{-}(8) \text{ are feasible for } k = 1, ..., N\} \qquad (9)$$

is not empty. For any $\beta \in \Gamma$, one solution to the distributed estimation problem under consideration, with $\gamma = \frac{1}{\beta}$, is given by the network of estimators (3) in which

$$K_k = X_k^{-1}F_k \quad \text{and} \quad L_k = (X_k^{-1}C_k^\top + BD_k^\top)E_k^{-1}, \qquad (10)$$

where X_k and $F_k, k = 1, ..., N$, belong to the feasibility set of (6) - (8), corresponding to this particular value of β. The weighting matrix P in (5) is given by $P = (1/N)\sum_{k=1}^N X_k$.

Remark 2: Assumption 1 is a restriction toward the class of communication graphs, which is made in order to ensure that the well-known average consensus algorithm is applicable. Assumption 2 is used to ensure boundedness of the feasible sets. It is not restrictive, as it represents the worst case of disturbance, and if not satisfied, small hypothetical disturbances can be added to the system description, i.e. additional columns to B, D_k and $\overline{D}_k$. Furthermore, note that the tuple $(\widetilde{A}_k, C_k)$ is not required to be detectable.

Since the LMIs (6), (7), (8) are coupled, they may be solved in a centralized manner as the optimization problem

$$\min \ (-\beta)$$
$$\text{subject to } (6), (7), (8), k = 1, ..., N, \qquad (11)$$

where the resulting matrices X_k, F_k deliver the estimator gains L_k, K_k according to (10). In the next section we will explore the separation of the problem and parallel computation in order to solve the problem in a distributed manner.

III. DISTRIBUTED CALCULATION OF FILTER GAINS

Parallel and distributed computation is thoroughly discussed e.g. in [9], and in this section, we use some of the methods presented in Section 3 in [9] to calculate our estimator gains in a distributed fashion. Solving the optimization problem (11) can be formulated as a separable problem by defining local representations of the solution

variables, X_j^k, and β^k for all $k = 1, ..., N$ and $j = k, j_1^k, ..., j_{p_k}^k$. The tuple of local variables is denoted by

$$Y_k = (F_k, \beta^k, X_k^k, X_{j_1^k}^k, ..., X_{j_{p_k}^k}^k), \qquad (12)$$

where the upper index k denotes the representation of a variable used by estimator k and all X_j^k are symmetric, positive definite matrices, and $\beta^k \geq 0$.

Problem 1: Find an iterative algorithm, which creates a sequence $Y_k(t), t \in \mathbb{N}$, such that local representations of the variables converge in the sense that

$$\lim_{t\to\infty} \left(\beta^{k_1}(t) - \beta^{k_2}(t)\right) = 0, \qquad (13)$$

for all $k_1, k_2 = 1, ..., N$ and

$$\lim_{t\to\infty} \left(X_j^{k_1}(t) - X_j^{k_2}(t)\right) = 0, \qquad (14)$$

for all $j = 1, ..., N$ and $k_1, k_2 \in \mathcal{M}_j \cup j$. All iterations $Y_k(t)$ shall satisfy the LMIs (6)-(8) when setting $\beta = \beta^k, X_k = X_k^k, X_{j_1^k} = X_{j_1^k}^k, ..., X_{j_{p_k}^k} = X_{j_{p_k}^k}^k$. Furthermore, the iteration steps of the local variables $Y_k(t+1)$ shall be calculated in a distributed fashion, i.e. interaction with the neighbors $j \in \mathcal{N}_k$ only.

As a first step, in order to ensure that both (11) and Problem 1 are well-posed, we establish a statement about the boundedness of the feasible set of the LMIs (6)-(8). The proof of this theorem will later be used in order to ensure that solutions of local optimizations are always attainable.

Theorem 1: Suppose the pairs $(\widetilde{A}_k, \widetilde{B}_k)$ are controllable. Then, for any $\rho > 0$, the feasible set

$$\Omega = \{(\beta, X_k, F_k, k = 1, ..., N)| \\ (6), (7), (8) \text{ hold for } k = 1, ..., N\} \qquad (15)$$

is bounded.

Proof: Suppose $(\beta, X_k, F_k, k = 1, ..., N) \in \Omega$. Using the Schur complement, it follows from (6), (8) that for an arbitrary $\tau_k > 0$,

$$X_k\widetilde{A}_k + \widetilde{A}_k^\top X_k - C_k^\top E_k^{-1}C_k + \beta(p_k + q_k)I \\ -p_k F_k - p_k F_k^\top + X_k\widetilde{B}_k\widetilde{B}_k^\top X_k \\ +\tau_k(F_k^\top X_k^{-1}F_k - \rho X_k) < 0. \qquad (16)$$

Completing the squares on the left-hand side yields

$$X_k\widetilde{A}_k + \widetilde{A}_k^\top X_k - C_k^\top E_k^{-1}C_k + \beta(p_k + q_k)I \\ +\tau_k(F_k - \frac{p_k}{\tau_k}X_k)^\top X_k^{-1}(F_k - \frac{p_k}{\tau_k}X_k) \\ -\tau_k(\frac{p_k^2}{\tau_k^2} + \rho)X_k + X_k\widetilde{B}_k\widetilde{B}_k^\top X_k < 0. \qquad (17)$$

Hence, we conclude that (β, X_k) satisfy the Riccati inequality

$$X_k(\widetilde{A}_k - \frac{p_k^2 + \tau_k^2\rho}{2\tau_k}I) + (\widetilde{A}_k - \frac{p_k^2 + \tau_k^2\rho}{2\tau_k}I)^\top X_k \\ -C_k^\top E_k^{-1}C_k + \beta(p_k + q_k)I + X_k\widetilde{B}_k\widetilde{B}_k^\top X_k < 0. \quad (18)$$

After pre- and post-multiplying (18) by X_k^{-1}, (18) reduces to

$$(\widetilde{A}_k - \frac{p_k^2 + \tau_k^2\rho}{2\tau_k}I)X_k^{-1} + X_k^{-1}(\widetilde{A}_k - \frac{p_k^2 + \tau_k^2\rho}{2\tau_k}I)^\top$$
$$-X_k^{-1}(C_k^\top E_k^{-1}C_k - \beta(p_k+q_k)I)X_k^{-1} + \widetilde{B}_k\widetilde{B}_k^\top \quad < 0 \quad (19)$$

Associated with this Riccati inequality, consider the Riccati equation

$$(\widetilde{A}_k - \frac{p_k^2 + \tau_k^2\rho}{2\tau_k}I)Z_k + Z_k(\widetilde{A}_k - \frac{p_k^2 + \tau_k^2\rho}{2\tau_k}I)^\top$$
$$-Z_k(C_k^\top E_k^{-1}C_k - \frac{1}{\gamma}(p_k+q_k)I)Z_k + \widetilde{B}_k\widetilde{B}_k^\top \quad = 0 \quad (20)$$

and define

$$\gamma^\circ = \inf\left\{ \begin{array}{c} \gamma > 0: \text{ equation (20) has a nonnegative-}\\ \text{definite solution} \end{array}\right\}(21)$$

From the $\mathscr{H}_\infty$ control theory [14, Theorems 4.8 and 9.7], it is known that the set whose infimum determines γ° is nonempty if the pair $(\widetilde{A}_k - \frac{p_k^2+\tau_k^2\rho}{2\tau_k}I, C_k)$ is detectable and the pair $(\widetilde{A}_k - \frac{p_k^2+\tau_k^2\rho}{2\tau_k}I, \widetilde{B}_k)$ is stabilizable. Note that by the condition of the theorem, the pair $(\widetilde{A}_k, \widetilde{B}_k)$ is controllable; this implies the stabilizability of $(\widetilde{A}_k - \frac{p_k^2+\tau_k^2\rho}{2\tau_k}I, \widetilde{B}_k)$. Now, let us choose $\tau_k > 0$ such that all unstable unobservable modes of the matrix pair $(\widetilde{A}_k, C_k)$ lie in the region $\text{Re } s < \frac{p_k^2+\tau_k^2\rho}{2\tau_k}$. This will guarantee that the pair $(\widetilde{A} - \frac{p_k^2+\tau_k^2\rho}{2\tau_k}I, C_k)$ is detectable. Thus, we conclude that $\gamma^\circ < \infty$.

The feasibility of the Riccati inequality (19) also implies that the following state-feedback $\mathscr{H}_\infty$ control problem involving the system

$$\dot{x} = (\widetilde{A}_k - \frac{p_k^2+\tau_k^2\rho}{2\tau_k}I)^\top x + C_k^\top u + (p_k+q_k)^{1/2}w, \quad (22)$$

$$z_k = \begin{bmatrix} \widetilde{B}_k^\top \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ E_k^{-1/2} \end{bmatrix} u$$

and the $\mathscr{H}_\infty$ performance criterion

$$\int_0^\infty \|z_k\|^2 dt < \frac{1}{\beta}\int_0^\infty \|w\|^2 dt \quad \forall w \in L_2, \quad (x(0)=0), (23)$$

has a solution. Indeed, it follows from (19) that

$$(\widetilde{A}_k - \frac{p_k^2+\tau_k^2\rho}{2\tau_k}I - X_k^{-1}C_k^\top E_k^{-1}C_k)X_k^{-1}$$
$$+X_k^{-1}(\widetilde{A}_k - \frac{p_k^2+\tau_k^2\rho}{2\tau_k}I - X_k^{-1}C_k^\top E_k^{-1}C_k)^\top$$
$$+X_k^{-1}(C_k^\top E_k^{-1}C_k + \beta(p_k+q_k)I)X_k^{-1} + \widetilde{B}_k\widetilde{B}_k^\top < 0. \quad (24)$$

Since $X_k^{-1} > 0$ and (24) is a strict inequality, the matrix

$$\widetilde{A} - \frac{p_k^2+\tau_k^2\rho}{2\tau_k}I - C_k^\top E_k^{-1}C_k X_k^{-1} \quad (25)$$

is Hurwitz. Thus, the closed loop system consisting of the system (22) with $w=0$ and the state-feedback controller

$$u = -E_k^{-1}C_k X_k^{-1} x$$

is exponentially stable. Also, using the completion of squares, it is easy to show from (19) that the above controller

guarantees the $\mathscr{H}_\infty$ attenuation property (23). Since the pair $(\widehat{A}_k, \widehat{B}_k)$ is controllable, these observations guarantee that the Riccati equation (20) with $\gamma = \frac{1}{\beta}$ has a unique nonnegative definite stabilizing solution $\widehat{Z}_k$ (e.g., see [15, Theorem 3.2.2]). Thus, $\frac{1}{\beta} > \gamma^\circ$. Furthermore, since $(\widehat{A}_k, \widehat{B}_k)$ is assumed to be controllable, $Z_k > 0$ and is invertible.

From Theorem 4.8 in [14], we know that $\gamma^\circ > 0$. These observations imply that $\beta < \gamma^{\circ-1}$. Also, using the relationship between solutions to the Riccati equation (20) and the corresponding Riccati inequality (19) [16, Lemma 8.1], it follows that $X_k < Z_k^{-1}$.

This discussion leads us to conclude that there exist upper bounds on feasible β and $\|X_k\|$. Indeed, γ° and Z_k are defined using the conditions involving the properties of the matrices $\widetilde{A}_k$, C_k and B_k and the constants ρ, p_k. Hence, these constant and the matrix are not dependent on the choice of the feasible β and X_k.

It remains to show that there is an upper bound on the feasible F_k as well. Using the Schur complement, (8) is equivalent to $F_k^\top X_k^{-1} F_k < \rho X_k$. This further implies

$$X_k^{-1/2}F_k^\top X_k^{-1}F_k X_k^{-1/2} < \rho I$$
$$tr\left(X_k^{-1/2}F_k^\top X_k^{-1}F_k X_k^{-1/2}\right) < n\rho$$
$$\|X_k^{-1/2}F_k X_k^{-1/2}\| < \sqrt{n\rho}.$$

For the Frobenius-norm of F_k, we can now conclude

$$\|F_k\| = \|X_k^{1/2}X_k^{-1/2}F_k X_k^{-1/2}X_k^{1/2}\|$$
$$\leq \|X_k^{1/2}\|\|X_k^{-1/2}F_k X_k^{-1/2}\|\|X_k^{1/2}\|$$
$$< \sqrt{n\rho}\|X_k^{1/2}\|^2,$$

which is bounded due to boundedness of X_k. ∎

Now, the decoupled version of the LMI conditions (6), (7), (8) is proposed as

$$\begin{bmatrix} Q_k^k - p_kF_k - p_kF_k^\top & X_k^k\widetilde{B}_k \\ * & -I \end{bmatrix} \leq -\begin{bmatrix} \delta X_k^k & 0 \\ 0 & 0 \end{bmatrix} \quad (26)$$

$$\begin{bmatrix} -\frac{2\alpha_k}{q_k+1}X_k^k & -\beta^kI+F_k & \cdots & -\beta^kI+F_k \\ * & -\frac{2\alpha_{j_1^k}}{q_{j_1^k}+1}X_{j_1^k}^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & 0 & \cdots & -\frac{2\alpha_{j_{p_k}}}{q_{j_{p_k}}+1}X_{j_{p_k}}^k \end{bmatrix} \leq -\delta I \quad (27)$$

$$\begin{bmatrix} -\rho X_k^k & -F_k^\top \\ -F_k & -X_k^k \end{bmatrix} \leq 0, \quad (28)$$

with $Q_k^k = X_k^k\widetilde{A}_k + \widetilde{A}_k^\top X_k^k - C_k^\top E_k^{-1}C_k + \beta^k(p_k+q_k)I$. Note that the LMI-conditions are formulated as non-strict inequalities, but with additional parameter $\delta > 0$. However, as δ can be chosen arbitrarily small, it introduces no conservativeness.

We denote the feasible set of the k-th group of the LMIs as $\Omega_k = \{Y_k \mid (26),(27),(28) \text{ hold true}\}$. Then, the separable

convex program can be written as

$$\text{minimize}\left(-\sum_{k=1}^{N}\beta^k\right)$$

$$\text{subject to } Y_k \in \Omega_k, \quad \beta^k = \widetilde{\beta} \tag{29}$$

$$X_k^k = \widetilde{X}_k, \quad X_{j_1^k}^k = \widetilde{X}_{j_1^k}, \quad \dots \quad, X_{j_{p_k}^k}^k = \widetilde{X}_{j_{p_k}^k}$$

for every $k = 1,...,N$. Here, $\widetilde{\beta}, \widetilde{X}_j, j = 1,...,N$ are additional variables that are needed to make the problem separable.

Remark 3: The optimization problem (29) can be varied in the way that for a given performance parameter $\beta > 0$, filter gains for (3) are to be found. Then, (29) turns to a pure feasibility problem without optimization objective, and therefore, the variables $\beta^k, \widetilde{\beta}$ and their iterations in the following algorithm can be omitted.

The dual problem has the form

$$\text{maximize } \mathbf{q}(\widetilde{\Lambda}^1,...,\widetilde{\Lambda}^N) \tag{30}$$

where $\widetilde{\Lambda}^k = (\lambda^k, \Lambda_k^k, \Lambda_{j_1^k}^k, ..., \Lambda_{j_{p_k}^k}^k)$ for $k = 1,...,N$ is the suitable tuple of Lagrange multipliers and the dual function $\mathbf{q}(\cdot)$ is defined as

$$\mathbf{q}(\widetilde{\Lambda}^1,...,\widetilde{\Lambda}^N) = \inf_{Y_k \in \Omega_k, k=1,...,N} L(Y_1,...,Y_N,\widetilde{\Lambda}^1,...,\widetilde{\Lambda}^N). \tag{31}$$

$L(\cdot)$ is the augmented Lagrangian function (cf. [9])

$$L(Y_k, \widetilde{\Lambda}^k) = \sum_{k=1}^{N}\left(-\beta^k + \lambda^k(\widetilde{\beta} - \beta^k) + \frac{c}{2}|\widetilde{\beta} - \beta^k|^2\right)$$

$$+ \sum_{k=1}^{N}\sum_{j \in \mathcal{N}_k \cup k}\left(tr\left(\Lambda_j^{k\top}(\widetilde{X}_j - X_j^k)\right) + \frac{c}{2}\|\widetilde{X}_j - X_j^k\|^2\right) \tag{32}$$

with design parameter $c > 0$. The optimization problem (30) can now be solved iteratively with Algorithm 1, which is initialized with $Y_k(0) \in \Omega_k$, $\lambda > 0$ and symmetric $\Lambda_j^k > 0$.

Remark 4: Out of the three steps in Algorithm 1, clearly 2) and 3) can be run in parallel by the individual estimators separately. Calculation of Step 1 of Algorithm 1 requires the evaluation of the mean value, which can be done in a distributed manner by applying a consensus algorithm. Under Assumption 1, average consensus algorithms can be used to calculate $\widetilde{\beta}(t+1)$. In particular, discrete time algorithms are preferable to keep the concept of an iterative algorithm [17] and algorithm which converge in finite-time are useful to ensure exact convergence [18], [19].

For the calculation of $\widetilde{X}_k(t+1)$ in the case of undirected graphs, only two steps are needed: All neighbors $j \in \mathcal{M}_k$ pass their $X_k^j(t)$ and $\Lambda_k^j(t)$ to estimator k. Then, estimator k calculates $\widetilde{X}_k(t+1)$ and passes it back to its neighbors. The calculation of $\widetilde{X}_k(t+1)$ in the case of directed graphs is more demanding with respect to the graph topology: Usual average consensus algorithms can be applied when for every $k = 1,...,N$, the subgraph $\mathcal{G}_k$ induced by node k and its out-neighborhood $\mathcal{M}_k$, is a balanced graph. This however can be relaxed by adding additional variables $X_j^k, j \notin \mathcal{N}_k$, to Y_k and

Algorithm 1: Calculation of iteration step t+1

1) Set the fusion variables for $j = 1,...,N$

$$\widetilde{\beta}(t+1) = \frac{1}{N}\sum_{k=1}^{N}\beta^k(t) - \frac{1}{Nc}\sum_{k=1}^{N}\lambda^k(t)$$

$$\widetilde{X}_j(t+1) = \frac{1}{q_j}\sum_{k \in \mathcal{M}_j}X_j^k(t) - \frac{1}{q_jc}\sum_{k \in \mathcal{M}_j}\Lambda_j^k(t)$$

2) Calculate the new variables for $k = 1,...,N$

$$Y_k(t+1)$$
$$= \underset{Y_k \in \Omega_k}{\text{argmin}}\left(-\beta^k - \lambda^k(t)\beta^k + \frac{c}{2}|\widetilde{\beta}(t+1) - \beta^k|^2\right.$$
$$\left. + \sum_{j \in \mathcal{N}_k \cup k}\left(-tr\left(\Lambda_j^{k\top}(t)X_j^k\right) + \frac{c}{2}\|\widetilde{X}_j(t+1) - X_j^k\|^2\right)\right)$$

3) Set the Lagrange variables for $k = 1,...,N$

$$\lambda^k(t+1) = \lambda^k(t) + c(\widetilde{\beta}(t+1) - \beta^k(t+1))$$

and for all $k = 1,...,N, j \in \mathcal{N}_k$

$$\Lambda_j^k(t+1) = \Lambda_j^k(t) + c(\widetilde{X}_j(t+1) - X_j^k(t+1))$$

adding $X_j^k = \widetilde{X}_j$ as equality constraint. For instance, if for all $k = 1,...,N$ $Y_k = (F_k, \beta^k, X_1^k, ..., X_N^k)$, then $\widetilde{X}_k(t+1), k = 1,...,N$ can be calculated under Assumption 1 using average consensus. This will later be demonstrated in the numerical example.

In order to show the convergence of Algorithm 1, two lemmas need to be introduced.

Lemma 1: The Lagrangian (32) can be written in terms of the vectorized variables, i.e.

$$L(Y_k, \widetilde{\Lambda}^k) = \sum_{k=1}^{N}\left(-\beta^k + \lambda^k(\widetilde{\beta} - \beta^k) + \frac{c}{2}|\widetilde{\beta} - \beta^k|^2\right)$$

$$+ \sum_{k=1}^{N}\sum_{j \in \mathcal{N}_k \cup k}\left(vec(\Lambda_j^k)^\top vec(\widetilde{X}_j - X_j^k) + \frac{c}{2}\|vec(\widetilde{X}_j - X_j^k)\|^2\right) \tag{33}$$

Proof: We have the equalities

$$tr(A^\top B) = \sum_i \sum_j A_{ji}B_{ji} = vec(A)^\top vec(B)$$

$$\|A\|^2 = tr(A^\top A) = \sum_i \sum_j A_{ji}^2 = \|vec(A)\|^2.$$

∎

This Lemma shows, that we can recast the problem into a problem of a standard form defined on a finite dimensional vector space.

Lemma 2: For fixed $\widetilde{\beta}, \widetilde{X}_k, \widetilde{X}_{j_1^k}, ..., \widetilde{X}_{j_{p_k}^k}, \lambda^k, \Lambda_{j_1^k}^k, ..., \Lambda_{p_k}^k$, the minimization

$$\underset{Y_k \in \Omega_k}{\text{arg min}}\left(-\beta^k - \lambda^k\beta^k + \frac{c}{2}|\widetilde{\beta} - \beta^k|^2\right.$$

$$\left. + \sum_{j \in \mathcal{N}_k \cup k}\left(-tr\left(\Lambda_j^{k\top}X_j^k\right) + \frac{c}{2}\|\widetilde{X}_j - X_j^k\|^2\right)\right) \tag{34}$$

is always attainable.

Proof: First, note that the LMI conditions (26)-(28) are non-strict inequalities. The definition range of the solution matrices $X_k^k > 0, X_j^k > 0, j \in \mathcal{N}_k$ are strict inequalities, but (26)-(28) imply that there exists a $\overline{\delta} > 0$ such that $X_k^k \geq \overline{\delta} I$ and $X_j^k \geq \overline{\delta} I$ for $j \in \mathcal{N}_k$. Thus, the feasible set Ω_k is closed and convex.

Following again the proof of Theorem 1, (26) and (28) imply that F_k, β^k, X_k^k are bounded for all $k = 1,...,N$. In constrast, the variables X_j^k for $j \in \mathcal{N}_k$ are not restricted to a bounded set by the LMIs (26)-(28). However, note that the cost function of (34) is quadratic in the variables $X_j^k, j \in \mathcal{N}_k$. Thus, due to the boundedness of F_k, β^k, X_k^k, we conclude that the sub-level sets of (34)

$$\left\{ Y_k \in \Omega_k | -\beta^k - \lambda^k \beta^k + \frac{c}{2}|\widetilde{\beta} - \beta^k|^2 \right.$$
$$\left. + \sum_{j \in \mathcal{N}_k \cup k} \left(-tr\left(\Lambda_j^{k\top} X_j^k\right) + \frac{c}{2}\|\widetilde{X}_j - X_j^k\|^2 \right) < \overline{c} \right\} \tag{35}$$

for $\overline{c} \in \mathbb{R}$ are bounded. Following the argument in Proposition 4.1 in [9], Chapter 3, we can conclude that we can equivalently search for the minimum of the cost function over a non-empty sub-level set (35) instead of Ω_k. Therefore, we can conclude that (34) is always attainable. ∎

Theorem 2: Algorithm 1 is a solution to Problem 1. In particular, the iteration steps $Y_k(t), k = 1,...,N$, can be calculated in parallel, and satisfy the convergence conditions (13), (14).

Proof: Using Lemma 1 and 2, we can follow the steps from [9], Section 3.3 and 3.4, in order to prove convergence of the iterations. ∎

IV. NUMERICAL EXAMPLE

Like in [6], we consider a system of the form (1), with

$$A = \begin{bmatrix} 0.3775 & 0 & 0 & 0 & 0 & 0 \\ 0.2959 & 0.3510 & 0 & 0 & 0 & 0 \\ 1.4751 & 0.6232 & 1.0078 & 0 & 0 & 0 \\ 0.2340 & 0 & 0 & 0.5596 & 0 & 0 \\ 0 & 0 & 0 & 0.4437 & 1.1878 & 0.0215 \\ 0 & 0 & 0 & 0 & 2.2023 & 1.0039 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.1\, I_6 & 0 \end{bmatrix} \qquad \overline{D}_k = 0.01\, I_2 \text{ for all } k = 1,...,N$$

which is observed by six sensor nodes, sensing two coordinates each. For every sensor an estimator is implemented, where none of the estimators is able to estimate the complete state vector without communication. The communication topology is assumed to be a directed circulant graph and we use Algorithm 1 to calculate the filter gains. For the numerical calculations we use YALMIP [20]. Since we are dealing with a directed but balanced graph, we apply the method described in Remark 4 and use complete local representations of all variables X_j at every estimator k. The algorithm is run with both fixed performance parameter $\beta^{const} = 100$ as discussed in Remark 3, and also using optimization over
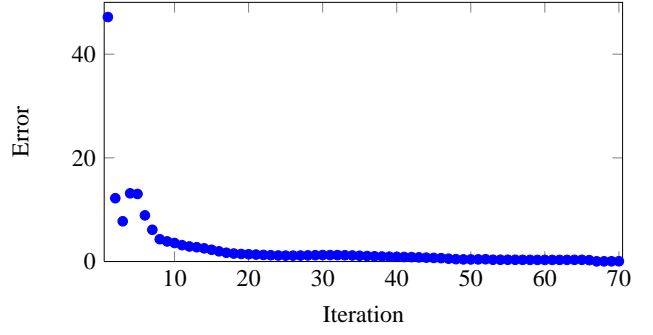

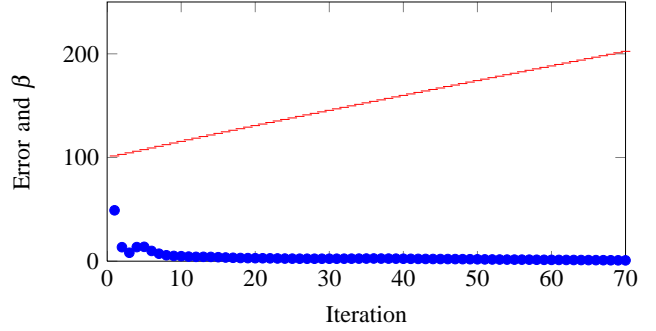Fig. 1. Evolution of the error during iteration for fixed β.

Fig. 2. Evolution of the error (blue dots) and performance index β^{ave} (red line) in the algorithm involving optimisation over variables β^k.

the variable performance parameter $\beta^k(t), k = 1,...,N$, (29), where $\beta^k(0) = 100$ is set as initialization.

In the first case, where β is fixed, we evaluate the matrix convergence condition (14) by calculating the average value $X_j^{ave} = \frac{1}{q_j} \sum_{k \in \mathcal{M}_j \cup j} X_j^k$ and subsequently Error $= \sum_{j=1}^N \sum_{k=1}^N \|X_j^k - X_j^{ave}\|^2$.

In the second case, involving optimisation over β_k, we additionally calculate $\beta^{ave} = \frac{1}{N} \sum_{k=1}^N \beta^k$ and subsequently we have Error $= \sum_{j=1}^N \sum_{k=1}^N \|X_j^k - X_j^{ave}\|^2 + \sum_{k=1}^N |\beta^k - \beta^{ave}|^2$.

The plots of the error evolution are shown in Figure 1 and 2. Figure 2 additionally shows the evolution of β^{ave}. The graph demonstrates that β^{ave} is monotonically increasing, and since it is bounded from above according to Theorem 1, it must eventually converge to a limit. In fact, it eventually converges to $2.3 \cdot 10^3$.

Better performance β however is achieved at the expense of higher filter gains. For instance, after 70 iterations, the consensus gain K_1 is

$$\begin{bmatrix} 21.1005 & -0.0256 & 0.0196 & -0.6018 & 0.0418 & 0.0117 \\ -0.0215 & 73.3369 & 0.5599 & 0.0073 & 0.0025 & 0.0021 \\ -0.0423 & -0.8806 & 99.8791 & 0.0617 & 0.0981 & 0.0536 \\ -0.6033 & -0.0178 & 0.0618 & 70.6692 & 1.3701 & 2.8005 \\ 0.0415 & -0.0054 & 0.0972 & 1.7726 & 20.7775 & 5.0466 \\ 0.0117 & -0.0003 & 0.0554 & 2.7740 & 3.2740 & 17.7281 \end{bmatrix}$$

in the fixed-β case and

$$\begin{bmatrix} 28.8328 & -0.0572 & 0.0291 & -0.0089 & 0.0333 & 0.0655 \\ -0.0397 & 99.9887 & 0.8962 & 0.0198 & 0.0050 & 0.0044 \\ -0.0222 & -0.8149 & 100.0003 & 0.2984 & 0.1414 & 0.0747 \\ -1.7921 & -0.0044 & 0.3121 & 71.0412 & 1.4502 & 2.5260 \\ 0.0751 & -0.0035 & 0.1384 & 1.8242 & 27.2461 & 6.2623 \\ -0.0177 & 0.0016 & 0.0711 & 2.4613 & 4.4543 & 23.0606 \end{bmatrix}$$

in the variable-β case.

V. Conclusion

We have developed a method for distributed filter design for cooperative $\mathscr{H}_\infty$-type estimation. In order to achieve this we separated the centralized problem by introducing additional variables and then applied an algorithm that works locally and only needs communication for average consensus.

References

[1] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Transactions of the ASME–Journal of Basic Engineering*, 82(Series D):35–45, 1960.

[2] Reza Olfati-Saber, J. Alexander Fax, and Richard M. Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1):215–233, 2007.

[3] Ruggero Carli, Alessandro Chiuso, Luca Schenato, and Sandro Zampieri. Distributed kalman filtering based on consensus strategies. *IEEE J. on Selected Areas in Comm.*, 26(4):622–633, 2008.

[4] Derui Ding, Zidong Wang, Hongli Dong, and Huisheng Shu. Distributed $\mathscr{H}_\infty$ state estimation with stochastic parameters and nonlinearities through sensor networks: The finite-horizon case. *Automatica*, 48(8):1575–1585, 2012.

[5] Maxim V. Subbotin and Roy S. Smith. Design of distributed decentralized estimators for formations with fixed and stochastic communication topologies. *Automatica*, 45(11):2491 – 2501, 2009.

[6] Valery Ugrinovskii. Distributed robust filtering with H_∞ consensus of estimates. *Automatica*, 47(1):1–13, 2011.

[7] Valery Ugrinovskii and Cédric Langbort. Distributed $\mathscr{H}_\infty$ consensus-based estimation of uncertain systems via dissipativity theory. *IET Control Theory & Applications*, 5(12):1458–1469, 2011.

[8] Srdjan S. Stankovic, Milos S. Stankovic, and Dusan M. Stipanovic. Consensus based overlapping decentralized estimator. *IEEE Transactions on Automatic Control*, 54(2):410–415, 2009.

[9] Dimitri Bertsekas and John N. Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods*. Prentice-Hall, 1989.

[10] Stephen Boyd. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends in Machine Learning*, 3(1):1–122, 2010.

[11] Robin L. Raffard, Claire J. Tomlin, and Stephen P. Boyd. Distributed optimization for cooperative agents: application to formation flight. In *43rd IEEE Conf. on Decision and Control (CDC)*, volume 3, pages 2453–2459, 2004.

[12] Michael Rabbat and Robert Nowak. Distributed optimization in sensor networks. In *Proceedings of the 3rd int. symp. on Information processing in sensor networks*, pages 20–27, 2004.

[13] Ion Necoara, Valentin Nedelcu, and Ioan Dumitrache. Parallel and distributed optimization methods for estimation and control in networks. *Journal of Process Control*, 21(5):756–766, 2011.

[14] Tamer Başar and Pierre Bernhard. $\mathscr{H}^\infty$-optimal Control and Related Minimax Design Problems: A Dynamic Game Approach. Birkhäuser, Boston, 2nd edition, 1995.

[15] Ian R. Petersen, Valery Ugrinovskii, and Andrey V. Savkin. *Robust Control Design Using $\mathscr{H}_\infty$ Methods*. Springer Science & Business Media, 2000.

[16] Pascal Gahinet and Pierre Apkarian. A linear matrix inequality approach to H_∞ control. *Int. J. of Robust and Nonlinear Control*, 4(4):421–448, 1994.

[17] Minghui Zhu and Sonia Martínez. Discrete-time dynamic average consensus. *Automatica*, 46(2):322–329, 2010.

[18] Shreyas Sundaram and Christoforos N. Hadjicostis. Finite-time distributed consensus in graphs with time-invariant topologies. In *Proc. American Control Conf.*, pages 711–716, 2007.

[19] Gang Chen, Frank L. Lewis, and Lihua Xie. Finite-time distributed consensus via binary control protocols. *Automatica*, 47(9):1962–1968, 2011.

[20] J. Löfberg. YALMIP : a toolbox for modeling and optimization in MATLAB. In *IEEE Int. Symposium on Computer Aided Control Systems Design*, pages 284 – 289, 2004.