Distributed Infinite-Horizon Optimal Control of Discrete-Time Linear Systems over Networks

Stefano Battilotti[1], Filippo Cacace[2], and Massimiliano d'Angelo[1,*]

***Abstract*— In this paper we consider the distributed infinite-horizon Linear-Quadratic-Gaussian optimal control problem for discrete-time systems over networks. In particular, the feedback controller is composed of local control stations, which receives some measurement data from the plant process and regulates a portion of the input signal. We provide a solution when the nodes have information on the structural data of the whole network but takes local actions, and also when only local information on the network are available to the nodes. The proposed solution is arbitrarily close to the optimal centralized one (in terms of cost index) when the intermediate consensus steps are sufficiently large.**

***Keywords*: discrete-time systems; distributed optimal control; sensor networks.**

I. Introduction

Technological developments in ad hoc networking and the availability of low-cost reliable computing, data storage, and sensing devices have made scenarios possible where the coordination of many subsystems extends the range of human capabilities [33]. In this context, large-scale cyber-physical systems play a fundamental role in the current and future era of smart industries, smart cities, smart homes and smart health care [20]. In these applications, the ability of a network system to fuse information, compute common estimates of unknown quantities, and agree on a common view of the environment in a decentralized/distributed fashion is critical. In particular, sensor networks and multi-agent systems, composed by a set of homogeneous or heterogeneous systems that communicate over a network, represented by a graph, have attracted much attention in recent years due to their potential application in many scenarios, for instance monitoring through wireless sensor networks [39], [22], formation control of mobile vehicles [26], [3], distributed optimization [19], cooperative control [2], [25], power networks [14], and also to systems with time-delay [23], [24], among many others.

In this paper we focus on the distributed optimal control problem of discrete-time linear stochastic systems with Gaussian noise over sensor networks. A distributed controller is composed of local control stations which receive some measurement data from the plant process and regulates a portion of the input signal by exchanging data among its neighbors and without the need of converging some information in specific points of the network ([40], [21], [35], [28], [16]). In particular, we adopt the distributed discrete-time framework where the neighboring nodes are able to communicate with some intermediate steps between two time instant (*e.g.*, [1], [31], [17], [32]). Indeed, a dynamic averaging of some important terms is performed by the local nodes. As for the centralized classical optimal control problem with incomplete information (*i.e.* when the state is not completely available), the distributed optimal control descends from the distributed optimal filter.

Although a lot of work has been done in the last decades, *e.g.* [38], [12], [27], [13], [29], [18], to the best of our knowledge only the recent result of [4] is able to provide a complete analysis on the stability and the optimality of the distributed filter for discrete-time linear system with Gaussian noise over sensor networks (with intermediate consensus steps). For, in the spirit of these recent advances, this paper tackles the distributed infinite-horizon optimal control problem for Gaussian discrete-time linear systems with partial state information. We note that this problem faces a fundamental difficulty. In fact, it is not possible to simply resort to the results of [4] since each node can not implement the full control law (which is composed by the local input signals) received by the plant (see Remark 3 for further details). We exploit in the paper both the cases in which global information of the system parameters and local information of the system parameters only of the network are available.

Finally, for the continuous-time analog we refer to [7] for the optimal filtering problem and to [2], [15] for the optimal control problem.

Notation. $\mathbb{R}$ and $\mathbb{C}$ denote real and complex numbers. For a square matrix A, tr(A) is the trace and $\sigma(A)$ is the spectrum. A is said to be Hurwitz stable if $\sigma(A) \subset \mathbb{C}_{-}$, the set of complex numbers with negative real part. $\mathbb{E}\{\cdot\}$ denotes expectation. $\otimes$ is the Kronecker product between vectors or matrices. The operators row$_i()$, col$_i()$, diag$_i()$ denote respectively the horizontal, vertical and diagonal compositions of matrices and vectors indexed by i. Let $\mathcal{S}(n) \in \mathbb{R}^{n\times n}$ be the set of symmetric matrices of size n, then $\mathcal{P}(n)$ (resp., $\mathcal{P}_+(n)) \subset \mathcal{S}(n)$ denotes the set of positive semi-definite (definite) matrices in $\mathcal{S}(n)$. We denote I_n the identity matrix of size n and by $U_n = 1_n 1_n^\top$, $1_n = \mathrm{col}_{i=1}^n(1)$, the square matrix of size n having 1 in each entry.

II. Problem formulation and preliminaries

We use a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ to describe the information exchange between the N nodes, where $\mathcal{V} = \{1, 2 \ldots, N\}$ is the set of vertices representing the N agents and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges of the graph. An edge of $\mathcal{G}$ is denoted by (i, j), representing that nodes i and j can exchange information between them. The graph is undirected, that is, the edges (i, j) and $(j, i) \in \mathcal{E}$ are considered to be the same. Two nodes i and j, with $i \neq j$, are neighbors to each other if $(i, j) \in \mathcal{E}$. The set of neighbors of node i is denoted by $\mathcal{N}^{(i)} := \{j \in \mathcal{V} : (j, i) \in \mathcal{E}\}$. A path is a sequence of connected edges in a graph. A graph is connected if there is a path between every pair of vertices. The adjacency matrix $\mathcal{A}$ of a graph $\mathcal{G}$ is an $N \times N$ matrix, whose (i,j)-th entry is 1 if (i,j) is an edge of $\mathcal{G}$ and 0 otherwise. The degree matrix $\mathcal{D}$ of $\mathcal{G}$ is a diagonal matrix whose i-th diagonal element is equal to the cardinality of $\mathcal{N}^{(i)}$, denoted $\#\mathcal{N}^{(i)}$. The Laplacian of $\mathcal{G}$ is defined to be a $N \times N$ matrix $\mathcal{L}$ such that $\mathcal{L} = -\mathcal{A} + \mathcal{D}$. $\mathcal{L}$ is symmetric if and only if the graph is undirected. Moreover, $0 = \lambda_1(\mathcal{L}) < \lambda_2(\mathcal{L}) \leq \cdots \leq \lambda_N(\mathcal{L})$, where $\lambda_i(\mathcal{L})$ denotes an eigenvalue of $\mathcal{L}$, if and only if the graph is connected. An eigenvector associated to $\lambda_1(\mathcal{L})$ is 1_N.

Consider the process

$$x_{k+1} = Ax_k + \sum_{i=1}^{N} B^{(i)} u_k^{(i)} + f_k, \tag{1}$$

$$y_k^{(i)} = C^{(i)} x_k + g_k^{(i)}, \quad i = 1, \ldots, N, \tag{2}$$

where $x_k \in \mathbb{R}^n$ is the state of the system, $y_k^{(i)} \in \mathbb{R}^{q_i}$, $q_i \geq 0$, is the measurement received by sensor i-th, $u_k^{(i)} \in \mathbb{R}^{p_i}$, $p_i \geq 0$, is the input signal sent by sensor i-th, and f_k and $g_k^{(i)}$, $i = 1, \ldots, N$, are zero-mean white noises, mutually independent with covariance respectively $Q \in \mathcal{P}_+(n)$, $R^{(i)} \in \mathcal{P}_+(q_i)$ $i = 1, \ldots, N$. The matrices Q and $R = \text{diag}_i(R_i)$ are nonsingular. Also $\bar{x}_0$ is random with mean $\bar{x}_0 := \mathbb{E}\{\bar{x}_0\}$ and covariance $\Sigma_{\bar{x}_0} := \mathbb{E}\{(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^\top\}$. Moreover, let us define $B = \text{row}_i(B^{(i)})$, $C = \text{col}_i(C^{(i)})$, $u_k = \text{col}_i(u_k^{(i)})$, $y_k = \text{col}_i(y_k^{(i)})$, $g_k = \text{col}_i(g_k^{(i)})$.

The infinite-horizon cost function is given by

$$J = \lim_{T \to \infty} \frac{1}{T} \mathbb{E} \left\{ \sum_{k=1}^{T} x_k^\top Q_c x_k + u_k^\top R_c u_k \right\}, \tag{3}$$

where Q_c and $R_c = \text{diag}_i\{R_c^{(i)}\}$ are positive semi-definite matrices (strictly positive in the case of R_c).

In order to design a distributed asymptotically optimal output-feedback control, we need the following assumptions.

Assumption 1: The underlying graph $\mathcal{G}$ is undirected and connected.

Assumption 2: The couples (A, B) and (A, Q_c) are controllable, and the couples (C, A), (R_c, A) are observable.

Assumption 3: Matrices A, $B^{(i)}$, $C^{(i)}$, Q, $R^{(i)}$, Q_c, $R_c^{(i)}$ are available data for the node $i \in \{1, \ldots, N\}$ of the network.

We point out that Assumption 2 is a standard global controllability/observability assumption, and we do not assume anything for the single pairs $(A, B^{(i)})$ and $(C^{(i)}, A)$. As regard the knowledge on the system matrices of the single node, Assumption 3 corresponds to local information only for the nodes.

The aim of this paper is the following.

Goal. Find the distributed optimal output-feedback control policy $u_k^{(i)}$ for each sensor $i = 1, \ldots, N$, of system (1)–(2) *such that the cost function J defined in* (3) *is minimized.*

The term *distributed* explicitly refers to the constraints on communication local knowledge of the nodes (as specified by Assumption 3).

We conclude this section by noticing that the model (1)–(2) can be written in a compact form as

$$x_{k+1} = Ax_k + Bu_k + f_k, \tag{4}$$

$$y_k = Cx_k + g_k, \tag{5}$$

and, in order to clarify some notations, we recall the following standard result.

Proposition 1 ([34]):

- Given the system (4) with the cost function (3), the optimal state-feedback control u_k is given by

$$u_k^\star = -L\, x_k, \tag{6}$$

  where

$$L = R_c^{-1} B^\top S \tag{7}$$

  with S the solution to the algebraic Riccati equation

$$S = A^\top \left(S - SB(B^\top SB + R_c)^{-1} B^\top S \right) A + Q_c. \tag{8}$$

- Given the system (4)–(5) with the cost function (3), the optimal output-feedback control u_k is given by

$$u_k^\star = -L\, \check{x}_k \tag{9}$$

  where L is again given by (7) and $\check{x}_k$ is the optimal estimate (in the minimum variance sense) provided by the centralized asymptotic Kalman filter, i.e. the Kalman filter that employs all the measurements $y_k^{(i)}$ for all $i = 1, \ldots, N$, namely

$$\check{x}_{k+1} = \check{x}_{k+1|k} + K(y_{k+1} - C\check{x}_{k+1|k}) \tag{10}$$

$$\check{x}_{k+1|k} = A\check{x}_k + Bu_k. \tag{11}$$

$$K = PC^\top R^{-1}, \tag{12}$$

  where P is the solution to the algebraic Riccati equation

$$P = A \left(P - PC^\top (CPC^\top + R)^{-1} CP \right) A^\top + Q. \tag{13}$$

  Moreover, the (centralized) optimal cost is given by

$$J^\star = \text{tr}\{SQ\} + \text{tr}\{PL^\top B^\top SA\}. \tag{14}$$

We point out that we use the notation $\check{x}$ to denote the ideal centralized optimal Kalman filter.

With reference to the distributed optimal control problem, we refer to the control (9) also as the *centralized* output-feedback LQG Regulator, since it make use of all the measurements of the network, that means it is not a distributed solution.

III. Distributed Infinite-Horizon LQG Regulator

In this section we shall design a distributed counterpart of Proposition 1. We start with the simple case when the nodes have global information on the system matrices and then, we shall see how to extend this solution to the more interesting case when local information only are available (Assumption 3). For, let us consider the following assumption.

Assumption 4: Matrices A, B, C, Q, R, Q_c, R_c are available data for the node $i \in \{1, \ldots, N\}$ of the network.

Also, let us consider the state-feedback case as indicated in the following lemma for which we omit the trivial proof.

Lemma 1: Given the system (1), the cost function (3) with (A, B) and (A, Q_c) controllable pairs and the couple (R_c, A)

observable pair, a network with the topology specified by a graph $\mathcal{G}$, let Assumption 4 holds true. The optimal distributed state-feedback control $u_k^{(i)}$ for each $i = 1, \ldots, N$, is given by

$$u_k^{(i)} = -L^{(i)} x_k, \tag{15}$$

where $L^{(i)}$ is the i-th row of the gain L given by (7).

Remark 1: We note that because of Assumption 4, each node i can compute off-line the gain $L^{(i)}$ since it can compute (7) and (8).

A. Distributed optimal control with global information

We are now able to prove the main theorem of the paper which solves the infinite-horizon distributed optimal output-feedback control problem with global information on the system parameters (*i.e.* Assumption 4).

Theorem 1: Given the system (1)–(2) with the cost function (3), a network with the topology specified by a graph $\mathcal{G}$ with N nodes, let Assumptions 1, 2 and 4 hold true. Consider the distributed output-feedback control for $i = 1, \ldots, N$ given by

$$u_k^{(i)} = -L^{(i)} \hat{x}_k^{(i)}, \tag{16}$$

where $L^{(i)}$ is the i-th row of the gain L of (7) and $\hat{x}_k^{(i)}$ given by

$$v_k^{(i)} = -\sum_{j=1}^{N} B^{(j)} L^{(j)} \hat{x}_k^{(i)} \tag{17}$$

$$\hat{x}_{k+1|k}^{(i)} = A\hat{x}_k^{(i)} + v_k^{(i)} \tag{18}$$

$$\begin{cases} z_{k+1,0}^{(i)} = & \hat{x}_{k+1|k}^{(i)} + K^{(i)}(y_{k+1}^{(i)} - C_i \hat{x}_{k+1|k}^{(i)}), \\ z_{k+1,h+1}^{(i)} = & z_{k+1,h}^{(i)} + \frac{1}{\delta} \sum_{j \in \mathcal{N}^{(i)}} (z_{k+1,h}^{(j)} - z_{k+1,h}^{(i)}) \end{cases} \tag{19}$$

$$\hat{x}_{k+1}^{(i)} = z_{k+1,\gamma}^{(i)}, \tag{20}$$

where $h = 0, \ldots, \gamma - 1$, $K^{(i)} := NPC^{(i)^\top} R^{(i)^{-1}}$, P is the solution of (13), $\delta > \lambda_N(\mathcal{L})$. Then the infinite-horizon cost function J of (3) depending on the parameter $\gamma \in \mathbb{N}$ satisfies

$$\lim_{\gamma \to +\infty} J = J^\star, \tag{21}$$

where $J^\star$ is the optimal cost (14) of the centralized LQG Regulator.

A sketch of the proof is given in the Appendix.

In other words, Theorem 1 states that, as the parameter γ of the filter tends to $+\infty$, the distributed control law (16) is equivalent to the centralized optimal solution provided by the control law (9).

Remark 2: We notice that for the node i-th, the computation of the matrices $B^{(j)}$ and $L^{(j)}$ for $j = 1, \ldots, N$ in (17), and $K^{(i)}$ in (28) is possible because of Assumption 4. In the next section we shall substitute the latter assumption with Assumption 3, *i.e.* local knowledge instead of global knowledge.

Remark 3: The role of $v_k^{(i)}$ in (17) is to mimic the full control input $Bu_k = \sum_{i=1}^{N} B^{(i)} u_k^{(i)} = -\sum_{j=1}^{N} B^{(j)} L^{(j)} \hat{x}_k^{(j)}$. In fact, it is easy to see that if the input term $v_k^{(i)}$ were set to Bu_k, then the estimation error of the filter would coincide with the one of [4] and also the optimality of the control would trivially follows. However, the latter choice of $v_k^{(i)}$ is not feasible since each node i would require the control $u_k^{(j)}$ for *all* $j = 1, \ldots, N$, which is not the case.

B. Distributed LQG regulator with local information

In order to implement the control law of Theorem 1 each node $i \in \mathcal{V}$ needs to compute (or to know) the value of the matrices S and P, that depend on all the nodes of the graph. Although this solution tackles the problem with the paradigm "*known-global-act-local*", it may seem to impair a truly distributed computation in which each node has local information only. Thus the aim of this section is to show how the solution of Theorem 1 can be implemented in a completely distributed manner in order to solve the problem with the paradigm "*known-local-act-local*". In particular, we relax Assumption 4 by assuming that each node i has only local information as clarified by Assumption 3.

In the first place it is worth remarking that the computation of S, solution to the algebraic Riccati equation (8), and P, solution to the algebraic Riccati equation (13) does not depend on the size of the graph. Moreover, the algebraic Riccati equation (8) can be written as

$$\begin{aligned} S = \tilde{A_C} S \tilde{A}_C^\top + (I - SBR_c^{-1}B^\top) Q_c (I - SBR_c^{-1}B^\top)^\top \\ + SBR_c^{-1}B^\top S \end{aligned} \tag{22}$$

with $\tilde{A_C} := (I - SBR_c^{-1}B^\top)A^\top$, and the algebraic Riccati equation (13) can be written as

$$\begin{aligned} P = A_C P A_C^\top + (I - PC^\top R^{-1} C) Q (I - PC^\top R^{-1} C)^\top \\ + PC^\top R^{-1} CP \end{aligned} \tag{23}$$

with $A_C := (I - PC^\top R^{-1} C)A$. Thus, it is clear from (22) and (23) that all the nodes can compute S and P provided that the values of $F := BR_c^{-1}B^\top$ (for (22)) and $G := C^\top R^{-1} C$ (for (23)) are available. We note that F can be written as $F = \sum_{i=1}^{N} B^{(i)} R_c^{(i)^{-1}} B^{(i)^\top}$ and, when measurement noises are independent (which is our case), G can be expressed similarly with the sum $G = \sum_{i=1}^{N} C^{(i)^\top} R^{(i)^{-1}} C^{(i)}$.

A distributed computation of F and G can thus be achieved by resorting to distributed algorithms to compute aggregate functions over graphs [30], [11]. In Fig. 1 we report an algorithm derived from the *Protocol Push-Sum* of [30] to compute G in a distributed way. The main difference is that [30] is a gossip algorithm with peer-to-peer communication, whereas the algorithm in Fig. 1 is a diffusion protocol with the node that broadcasts messages to all its neighbors. The speed of convergence of the local estimate to the true value of G can be analyzed in the light of the results of [30]. This estimation phase can be executed before the filtering phase for static graphs, or it can be kept running during the execution of the filter in order to adjust the value of S and G in presence of a dynamical graph where nodes appear or disconnect. Finally, the value of N can be computed by the same distributed algorithm when it is not known at the nodes. More details can be found in [11]. We conclude with the summarizing corollary.

Algorithm Broadcast Push-Sum

1: In all nodes set $s_{0,i} = B^{(i)} {R_c^{(i)}}^{-1} {B^{(i)}}^\top$ and $w_{0,i} = 0$, except for $w_{0,1} = 1$.
2: At time 0 each nodes sends $(s_{0,i}, w_{0,i})$ to itself.
3: At time t each node executes:
   1. Let $\{s_r, w_r\}$ be the pairs sent to i in round $t-1$.
   2. Let $s_{t,i} = \sum_r s_r$, $w_{t,i} = \sum_r w_r$.
   3. Send to all neighbors and to i (yourself):
   $$\left(\frac{1}{|\mathcal{N}^{(i)}|+1} s_{t,i},\ \frac{1}{|\mathcal{N}^{(i)}|+1} w_{t,i}\right)$$
   4. $s_{t,i}/w_{t,i}$ is the estimate of F at step t (if $w_{t,i} = 0$ the estimate is not specified or 0).

Fig. 1. A modified version of the *Push-Sum* algorithm of [30] that makes possible the distributed computation of F (and G if $s_{0,i} = {C^{(i)}}^\top {R^{(i)}}^{-1} C^{(i)}$).

Corollary 2: Given the system (1)–(2) with the cost function (3), a network with the topology specified by a graph $\mathcal{G}$ with N nodes, let Assumptions 1, 2 and 3 hold true. Consider the distributed output-feedback control for $i = 1, \dots, N$ given by

$$u_k^{(i)} = -L^{(i)} \hat{x}_k^{(i)}, \tag{24}$$

where $L^{(i)} = {R_c^{(i)}}^{-1} {B^{(i)}}^\top S$ is the i-th row of the gain L of (7) when the term $F := BR_c^{-1}B^\top$ in (22) is computed by each node $i \in \mathcal{V}$ through the Push-Sum algorithm of Figure 1, and $\hat{x}_k^{(i)}$ given by

$$v_k^{(i)} = -FS\hat{x}_k^{(i)} \tag{25}$$
$$\hat{x}_{k+1|k}^{(i)} = A\hat{x}_k^{(i)} + v_k^{(i)} \tag{26}$$
$$\begin{cases} z_{k+1,0}^{(i)} = & \hat{x}_{k+1|k}^{(i)} + K^{(i)}(y_{k+1}^{(i)} - C_i\hat{x}_{k+1|k}^{(i)}), \\ z_{k+1,h+1}^{(i)} = & z_{k+1,h}^{(i)} + \frac{1}{\delta}\sum_{j\in\mathcal{N}^{(i)}}(z_{k+1,h}^{(j)} - z_{k+1,h}^{(i)}) \end{cases} \tag{27}$$
$$\hat{x}_{k+1}^{(i)} = z_{k+1,\gamma}^{(i)}, \tag{28}$$

where $h = 0, \dots, \gamma-1$, F (and thus S) is computed again through the algorithm of Fig. 1, $K^{(i)} := NP{C^{(i)}}^\top {R^{(i)}}^{-1}$, P is the solution of (23) where the term $G := C^\top R^{-1} C$ is computed by each node $i \in \mathcal{V}$ through the algorithm of Figure 1, $\delta > \lambda_N(\mathcal{L})$. Then the infinite-horizon cost function J of (3) depending on the parameter $\gamma \in \mathbb{N}$ satisfies

$$\lim_{\gamma\to+\infty} J = J^\star, \tag{29}$$

where $J^\star$ is the optimal cost (14) of the centralized LQG Regulator.

Remark 4: We note that (25) is the same as (17) since $\sum_{j=1}^N B^{(j)}L^{(j)} = BL = FS$, but we stress the fact that F and consequently S can be computed through the distributed algorithm of Fig. 1.

IV. Example

We consider the academic, but challenging, example of [2] in which the network topology is characterized in Figure 2, the system matrices are given by

$$A = \begin{bmatrix} -0.1 & 0 & 0 & -1 & 1 & 0 & 0 & 1 \\ 2.5 & -0.5 & -1.6 & -1.5 & 2 & 0 & 1.6 & 1.5 \\ 2.6 & -0.5 & -0.7 & -1.5 & 1.5 & 0.5 & 0.5 & 1.5 \\ -2 & 0 & 1 & 0 & -0.1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.1 & 0 & 0 & 0 \\ -0.5 & 0 & 1 & 0 & 0.5 & -0.5 & -1 & 0 \\ 3.8 & -0.5 & -1.8 & -0.5 & 2 & 0.5 & 1.6 & 0.5 \\ -1 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \end{bmatrix},$$

and the state and output noises have intensity $Q = 0.09 \cdot I_n$ and $R^{(i)} = 0.36$ with $i = 1, \dots, 9$. The output and input matrices of the nodes are $C_1 = [0,0,1,0,0,0,0,0]$, $C_2 = [-2,1,1,1,-1,0,-1,-1]$, $C_3 = [1,0,0,0,0,0,0,0]$, $C_4 = [-3,1,2,1,-1,0,-1,-1]$, $C_5 = [1,0,-1,0,1,0,1,0]$, $C_6 = 2,-1,-1,-1,1,1,1,1]$, $C_7 = [0,0,0,0,1,0,0,0]$, $C_8 = [1,0,-1,-1,2,0,1,1]$, $C_9 = [0,0,1,1,-1,0,0,0]$, and $B_i = C_i^\top$ for all $i = 1, \dots, 9$. Also, the cost index (3) has $Q_c = I_7$ and $R_c = I_9$. This setting is very general since the couple $(A, B^{(i)})$ (respectively $(C^{(i)}, A)$) is not controllable (respectively observable) for all $i \in \mathcal{V}$. Also, controllability and observability property are not satisfied even *locally*, namely the couple $\left(A, \mathrm{row}_{j\in\mathcal{N}_i\cup\{i\}}(B_j)\right)$ is not controllable for any $i \in \mathcal{V}$ and the couple $\left(\mathrm{col}_{j\in\mathcal{N}_i\cup\{i\}}(C^{(j)}), A\right)$ is not observable for any $i \in \mathcal{V}$. However, the hypotheses of Assumption 2 are satisfied, in particular (A, B) is controllable and (C, A) is observable.

When the control laws (16) are applied to the plant, in accordance with the expected optimal cost (14) of the optimal centralized solution, we define the pseudo-cost function

$$\bar{J}_\gamma = \mathrm{tr}\{SQ\} + \mathrm{tr}\{\bar{P}(\gamma)L^\top B^\top SA\}. \tag{30}$$

where $\bar{P}(\gamma)$ is the arithmetic mean of the covariance matrices of the estimation errors of the nodes $i \in \mathcal{V}$, namely $\bar{P}(\gamma) = \frac{1}{N}\sum_{i=1}^N X^{(i)}(\gamma)$ with $X_i(\gamma)$ covariance of the estimation error $e_i(t)$. By mean of the result of Theorem 1, it follows that $\lim_{\gamma\to+\infty} \bar{J}_\gamma = J^*$, where J^* is the optimal cost (14) of the ideal centralized regulator. Figure 3 (right) shows the convergence of $\bar{J}_\gamma$ of cost when γ tends to infinity, and similarly Figure 3 (left) shows the convergence of the traces of the covariance of the estimation error of all the nodes, namely $\mathrm{tr}(X^{(i)}(\gamma))$ towards the optimal value of the trace of the covariance of the centralized Kalman filter.

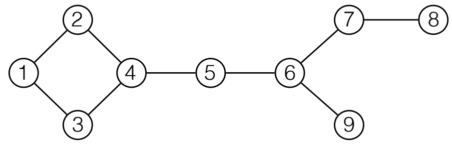
The interested reader could find (or request) the journal version of this paper ([5]) which includes a detailed proof, some extensions and a real-plant example.

V. Conclusions

Further extensions deserve additional investigation, for example the introduction of communications delays or packet dropouts [41], [10], link failure [6], nonlinear systems under sampling [37], multi-agents systems [36], or methods for non-Gaussian noises [9], [8]. Also, other future directions are secure and resilient solutions that account for threats and attacks.

Appendix

Sketch of the proof. of Theorem 1. Because of Lemma 1, we need to prove that the local estimates $\hat{x}_k(i)$ for all



$$\mathcal{L} = \begin{bmatrix} 2 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 3 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 3 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \end{bmatrix}$$

Fig. 2. Communication graph $\mathcal{G}$ and its Laplacian matrix $\mathcal{L}$.

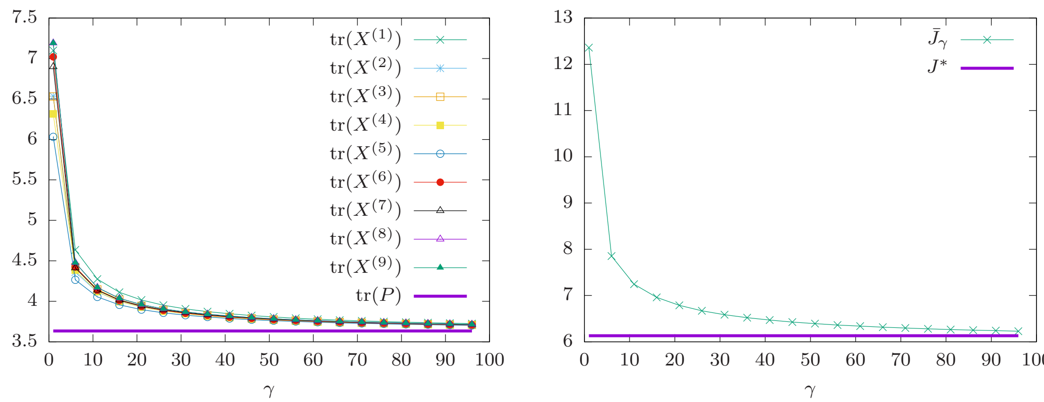


Fig. 3. Traces of the covariance matrices of the estimation errors tr$(X^{(i)})$ of the nodes and trace of the covariance of the ideal centralized KF tr(P) (left). Convergence of the pseudo-cost function (30) towards the centralized optimal one. (right).

$i = 1, \ldots, N$ provided by (17)–(28) tend to the centralized estimate of the Kalman filter, *i.e.* the estimate given by (10)–(12).

The first step is the definition of the matrix $\Upsilon = (I_N - \frac{1}{\delta}\mathcal{L}) \otimes I_n$ and the aggregate vectors $z_{k,h} = \mathrm{col}_i(z_{k,h}^{(i)})$, $v_{k,h} = \mathrm{col}_i(v_{k,h}^{(i)})$. From (19) we get $z_{k,h+1} = \Upsilon z_{k,h}$. Let $e_k^{(i)} := x_k - \hat{x}_k^{(i)}$ be the estimation error of the node i. After some manipulations the dynamics of the aggregate vector of the estimation error, namely $e_k = \mathrm{col}_i\{e_k^{(i)}\}$, can be written as

$$e_{k+1} = \mathcal{A}(\gamma) e_k + h_k, \tag{31}$$

where $\mathcal{A}(\gamma) := \Upsilon^\gamma(\mathrm{diag}_i\{M^{(i)}A\} - \mathrm{diag}_i\{M^{(i)}\}\Xi)$, with $M^{(i)} := I - K^{(i)}C^{(i)}$, Ξ is a matrix such that $\Xi(1_N \otimes I_n) = 0$, and $h_k := \mathrm{col}_i\{h_k^{(i)}\}$, with $h_k^{(i)} := M^{(i)} f_k - K^{(i)} g_{k+1}^{(i)}$. Then, by considering the orthonormal transformation $T = \begin{pmatrix} V \\ W \end{pmatrix}$, with $V = \frac{1}{\sqrt{N}} 1_N^\top$ and $W \in \mathbb{R}^{(N-1)\times N}$, such that

$$T\mathcal{L}T^\top = \begin{pmatrix} 0 & 0 \\ 0 & \Lambda \end{pmatrix},$$

where $\Lambda = \mathrm{diag}(\lambda_2(\mathcal{L}), \ldots, \lambda_N(\mathcal{L}))$ it is possible to see that

$$(T \otimes I_n)\mathcal{A}(\gamma)(T^\top \otimes I_n) = \begin{pmatrix} A_C & H_{12} \\ H_{21}(\gamma) & H_{22}(\gamma) \end{pmatrix}, \tag{32}$$

where $H_{12} = (V \otimes I_n)\Gamma(W^\top \otimes I_n))$, $H_{21}(\gamma) = (S^\gamma W \otimes I_n)\Gamma(V^\top \otimes I_n)$, $H_{22}(\gamma) = (S^\gamma W \otimes I_n)\Gamma(W^\top \otimes I_n))$, with $\Gamma := (\mathrm{diag}_i\{A^{(i)}\} - \mathrm{diag}_i\{M^{(i)}\}\Xi)$ and $S := I_{N-1} - \frac{\Lambda}{\delta}$. It is not difficult to see that

$$H_{21}(\gamma) \to 0 \text{ and } H_{22}(\gamma) \to 0 \quad \text{as } \gamma \to +\infty. \tag{33}$$

Thus, since A_C is Schur stable by construction, it follows that there exists $\gamma_0 \in \mathbb{N}$ such that for all integer $\gamma > \gamma_0$, $\mathcal{A}(\gamma)$ is Schur stable. Furthermore, we prove that, for each node, as γ increases, the covariance of the estimation error tends to the covariance of the ideal CKF. In other words, if we denote with U_N the $N \times N$ matrix with all ones, we shall prove that $X(\gamma) := \lim_{k\to+\infty} \mathbb{E}[e_k e_k^\top] \to X_C := U_N \otimes P$, with P solution to (13), as $\gamma \to +\infty$. Firstly, note that X_C satisfies

$$\begin{aligned} X_C = {} & \mathrm{diag}_i\{A_C\} X_C \,\mathrm{diag}_i\{A_C^\top\} \\ & + \mathrm{diag}_i\{I - KC\}(U_N \otimes Q)\mathrm{diag}_i\{(I - KC)^\top\} \\ & + \mathrm{diag}_i\{K\}(U_N \otimes R)\mathrm{diag}_i\{K^\top\}. \end{aligned} \tag{34}$$

Moreover, by introducing the covariance mismatch $E(\gamma) := X(\gamma) - X_C$, we can obtain after some manipulations

$$E(\gamma) = \mathcal{A}(\gamma)E(\gamma)\mathcal{A}^\top(\gamma) + \Sigma(\gamma), \tag{35}$$

where the matrix $\Sigma(\gamma)$ is such that

$$(T \otimes I_n)\Sigma(\gamma)(T^\top \otimes I_n) = \begin{pmatrix} 0 & D_{12}(\gamma) \\ D_{21}(\gamma) & D_{22}(\gamma) \end{pmatrix}, \tag{36}$$

with the sub-matrices $D_{12}(\gamma)$, $D_{21}(\gamma)$, and $D_{22}(\gamma)$ such that

$$D_{12}(\gamma), D_{21}(\gamma), D_{22}(\gamma) \to 0 \text{ as } \gamma \to +\infty. \tag{37}$$

From (32), (35) and (36)

$$\begin{aligned} \widetilde{E}(\gamma) & := (T \otimes I_n)E(\gamma)(T^\top \otimes I_n) \\ & = \begin{pmatrix} A_C & H_{12} \\ H_{21}(\gamma) & H_{22}(\gamma) \end{pmatrix} \widetilde{E}(\gamma) \begin{pmatrix} A_C^\top & H_{21}^\top(\gamma) \\ H_{12}^\top & H_{22}^\top(\gamma) \end{pmatrix} \\ & + \begin{pmatrix} 0 & D_{12}(\gamma) \\ D_{21}(\gamma) & D_{22}(\gamma) \end{pmatrix}, \end{aligned} \tag{38}$$

where H_{12}, $H_{21}(\gamma)$ and $H_{22}(\gamma)$ are defined below (32). On account of (33) and (37) the unique solution $\widetilde{E}(\gamma)$ of (38) tends as $\gamma \to +\infty$ to the unique solution $\widetilde{E}_\infty$ of the equation

$$\widetilde{E}_\infty = \begin{pmatrix} A_C & H_{12} \\ 0 & 0 \end{pmatrix} \widetilde{E}_\infty \begin{pmatrix} A_C^\top & 0 \\ H_{12}^\top & 0 \end{pmatrix}$$

and since A_C is Schur stable, it follows that $\widetilde{E}_\infty = 0$, which implies $\lim_{k\to+\infty} \mathbb{E}[e_k e_k^\top] \to U_N \otimes P$ as $\gamma \to +\infty$. $\square$

References

[1] B. Açıkmeşe, M. Mandić, and J. L. Speyer. Decentralized observers with consensus filters for distributed discrete-time linear systems. *Automatica*, 50(4):1037–1052, 2014.

[2] S. Battilotti, F. Cacace, and M. d'Angelo. Distributed infinite-horizon optimal control of continuous-time linear systems over network. *International Journal of Robust and Nonlinear Control*, 31(6):2082–2096, 2021.

[3] S. Battilotti, F. Cacace, M. d'Angelo, and A. Germani. Cooperative filtering with absolute and relative measurements. In *2018 IEEE Conference on Decision and Control (CDC)*, pages 7182–7187. IEEE, 2018.

[4] S. Battilotti, F. Cacace, and M. d'Angelo. A stability with optimality analysis of consensus-based distributed filters for discrete-time linear systems. *Automatica*, 129:109589, 2021.

[5] S. Battilotti, F. Cacace, and M. d'Angelo. Distributed optimal control of discrete-time linear systems over networks. *IEEE Transactions on Control of Network Systems*, under review.

[6] S. Battilotti, F. Cacace, M. d'Angelo, and A. Germani. Distributed kalman filtering over sensor networks with unknown random link failures. *IEEE control systems letters*, 2(4):587–592, 2018.

[7] S. Battilotti, F. Cacace, M. d'Angelo, and A. Germani. Asymptotically optimal consensus-based distributed filtering of continuous-time linear systems. *Automatica*, 122:109189, 2020.

[8] S. Battilotti, F. Cacace, M. d'Angelo, A. Germani, and B. Sinopoli. Kalman-like filtering with intermittent observations and non-Gaussian noise. *IFAC-PapersOnLine*, 52(20):61–66, 2019.

[9] S. Battilotti, F. Cacace, M. d'Angelo, A. Germani, and B. Sinopoli. LQ non-Gaussian regulator with Markovian control. *IEEE Control Systems Letters*, 3(3):679–684, 2019.

[10] S. Battilotti and M. d'Angelo. Stochastic output delay identification of discrete-time gaussian systems. *Automatica*, 109:108499, 2019.

[11] S. Battilotti, Cacace F., and d'Angelo M. Consensus-based robust computation of aggregate functions on networks. In *61th IEEE Conference on Decision and Control*. IEEE, submitted.

[12] G. Battistelli and L. Chisci. Kullback–leibler average, consensus on probability densities, and distributed state estimation with guaranteed stability. *Automatica*, 50(3):707–718, 2014.

[13] G. Battistelli, L. Chisci, G. Mugnai, A. Farina, and A. Graziano. Consensus-based linear and nonlinear filtering. *IEEE Transactions on Automatic Control*, 60(5):1410–1415, 2014.

[14] A. Bidram, F. L. Lewis, and A. Davoudi. Distributed control systems for small-scale power networks: Using multiagent cooperative control theory. *IEEE Control systems magazine*, 34(6):56–77, 2014.

[15] F. Borrelli and T. Keviczky. Distributed LQR design for identical dynamically decoupled systems. *IEEE Transactions on Automatic Control*, 53(8):1901–1912, 2008.

[16] K. Cai and W. M. Wonham. Supervisor localization: a top-down approach to distributed control of discrete-event systems. *IEEE Transactions on Automatic Control*, 55(3):605–618, 2010.

[17] R. Carli, A. Chiuso, L. Schenato, and S. Zampieri. Distributed kalman filtering using consensus strategies. In *2007 46th IEEE Conference on Decision and Control*, pages 5486–5491. IEEE, 2007.

[18] T. Chang and S. Shahrampour. Distributed online linear quadratic control for linear time-invariant systems. In *2021 American Control Conference (ACC)*, pages 923–928. IEEE, 2021.

[19] J. Chen and A. H. Sayed. Diffusion adaptation strategies for distributed optimization and learning over networks. *IEEE Transactions on Signal Processing*, 60(8):4289–4305, 2012.

[20] B. Cheng, J. Zhang, G. P. Hancke, S. Karnouskos, and A. W. Colombo. Industrial cyberphysical systems: Realizing cloud-based big data infrastructures. *IEEE Industrial Electronics Magazine*, 12(1):25–35, 2018.

[21] S. S. Chughtai and H. Werner. Distributed control for a class of spatially interconnected discrete-time systems. *IFAC Proceedings Volumes*, 41(2):7761–7766, 2008.

[22] V. De Iuliis, A. D'Innocenzo, A. Germani, and C. Manes. Stability conditions for linear discrete-time switched systems in block companion form. *IET Control Theory & Applications*, 14(19):3107–3115, 2020.

[23] V. De Iuliis, A. D'Innocenzo, A. Germani, and C. Manes. Stability analysis of coupled differential-difference systems with multiple time-varying delays: A positivity-based approach. *IEEE Transactions on Automatic Control*, 66(12):6085–6092, 2021.

[24] V. De Iuliis, A. Germani, and C. Manes. Identification of forward and feedback transfer functions in closed-loop systems with feedback delay. *IFAC-PapersOnLine*, 50(1):12847–12852, 2017.

[25] Z. Ding and Z. Li. Distributed adaptive consensus control of nonlinear output-feedback systems on directed graphs. *Automatica*, 72:46–52, 2016.

[26] J. A. Fax and R. M. Murray. Information flow and cooperative control of vehicle formations. *IEEE transactions on automatic control*, 49(9):1465–1476, 2004.

[27] X. He, W. Xue, and H. Fang. Consistent distributed state estimation with global observability over sensor network. *Automatica*, 92:162–172, 2018.

[28] S. Heinke and H. Werner. A gradient descent method for finite horizon distributed control of discrete time systems. In *2021 60th IEEE Conference on Decision and Control (CDC)*, pages 1905–1910. IEEE, 2021.

[29] A.T. Kamal, J.A. Farrell, and A.K. Roy-Chowdhury. Information weighted consensus filters and their application in distributed camera networks. *IEEE Transactions on Automatic Control*, 58(12):3112–3125, 2013.

[30] D. Kempe, A. Dobra, and J. Gehrke. Gossip-based computation of aggregate information. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science, 2003. Proceedings.*, pages 482–491. IEEE, 2003.

[31] U. A Khan and A. Jadbabaie. On the stability and optimality of distributed kalman filters with finite-time data fusion. In *Proceedings of the 2011 American Control Conference*, pages 3405–3410. IEEE, 2011.

[32] U. A Khan and J. MF Moura. Distributing the kalman filter for large-scale systems. *IEEE Transactions on Signal Processing*, 56(10):4919–4935, 2008.

[33] S. S. Kia, B. Van Scoy, J. Cortes, R. A Freeman, K. M. Lynch, and S. Martinez. Tutorial on dynamic average consensus: The problem, its applications, and the algorithms. *IEEE Control Systems Magazine*, 39(3):40–72, 2019.

[34] H. Kwakernaak, R. Sivan, and B. N. D. Tyreus. Linear optimal control systems. 1974.

[35] A. Marino and F. Pierri. Discrete-time distributed control and fault diagnosis for a class of linear systems. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2974–2979. IEEE, 2015.

[36] M. Mattioni. On multiconsensus of multi-agent systems under aperiodic and asynchronous sampling. *IEEE Control Systems Letters*, 4(4):839–844, 2020.

[37] S. Monaco, D. Normand-Cyrot, M. Mattioni, and A. Moreschini. Nonlinear hamiltonian systems under sampling. *IEEE Transactions on Automatic Control*, 2022.

[38] R. Olfati-Saber. Kalman-consensus filter: Optimality, stability, and performance. In *Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference*, pages 7036–7042. IEEE, 2009.

[39] M. F. Othman and K. Shazali. Wireless sensor network applications: A study in environment monitoring system. *Procedia Engineering*, 41:1204–1210, 2012.

[40] V. Sundarapandian. Distributed control schemes for large-scale interconnected discrete-time linear systems. *Mathematical and Computer Modelling*, 41(2-3):313–319, 2005.

[41] Hongjiu Yang, Hui Li, Yuanqing Xia, and Li Li. Distributed kalman filtering over sensor networks with transmission delays. *IEEE Transactions on Cybernetics*, 51(11):5511–5521, 2020.