Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

ARTICLE TYPE

Distributed Infinite-Horizon Optimal Control of Continuous-Time Linear Systems over Network

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

[1]DIAG, Department of Computer Control and Menagement, Sapienza University of Rome, Rome, Italy
[2]Campus Bio-Medico di Roma, Rome, Italy

Correspondence
*Email: mdangelo@diag.uniroma1.it

Summary

This paper deals with the distributed infinite-horizon Linear-Quadratic-Gaussian optimal control problem for continuous-time systems over networks. In particular, the feedback controller is composed of local control stations, which receive 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 a design parameter is set sufficiently large. Numerical simulation validate the theoretical results.

KEYWORDS:
Distributed Optimal Control, Consensus filters, Networked Control Systems, Cyber-Physical Systems

1 | INTRODUCTION

Large-scale cyber-physical systems play a fundamental role in the current and future era of smart industries, smart cities and smart homes[1]. Large-scale systems grasp structures of interconnected subsystems coupled through their dynamics. Among them, cyber-physical systems feature tightly coupled physical components, as system dynamics, sensors, controllers, environments, and cyber components, as data, communication, and computation[2]. Although a lot of efforts have been made in the last decades, many future generation co-design techniques are still missing in order to successfully cope with disturbances, faults, threats, and attacks. In this respect, optimal control theory finds its significant value in developing such solutions. Most of the time, these complex systems are increasingly spatially dispersed over large areas connected through networks, and thus a key necessary ingredient to control these systems is to built distributed solutions which allow to decentralized physical components, computation and storage[3].A decentralized controller is intended as a feedback controller composed of local control stations, which receive some measurement data from the plant process and regulates a portion of the input signal. 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. Thus, the main peculiar feature between decentralized and distributed schemes is the distribution of information.

Many solutions have been proposed on these problems, and a comprehensive review has been given in[4]. Solutions to the decentralized problem for continuous-time deterministic liner-systems can be tracked back to the 1970[5,6,7]. In[8] the robust decentralized control of a general servomechanism is addressed. In[9,10] previous results are generalized with time-varying feedback laws. A multi-rate distributed solution to state estimation and control has been proposed in[11]. In recent years, distributed solutions to the decentralized control problem of some interesting applications has been given for large population multi-agent systems[12], smart grids[13], satellite formations[14]. A parameter optimization approach of a fixed controller structure has been performed in[15] for a large flexible space structures.

On the other hand, distributed schemes make use of the communication capabilities of a network aiming to improve the performance of the solution. The literature concerning distributed strategies consists of a large number of methods with different approaches based on the nature of the problem. However, most of these approaches divide the control problem and the estimation problem, and tackle only one area. In particular, a lot of efforts have been given to the discrete-time distributed Linear Quadratic Gaussian (LQG) control with communication delays. The LQG problem under one-step delay information sharing pattern has been solved in [16,17,18,19], and some generalizations to other delay structures have been given in [20] and [21], where a computationally efficient solution for the LQG output-feedback problem with communication delays was presented using a state space formulation and covariance constraints, and the controller is obtained from the corresponding semi-definite programming solution. In [22], the authors consider LQG control with communication delays for the three interconnected systems. While they provide an explicit solution, their approach is restricted to state-feedback and assumes independence of disturbances acting on each subsystem. Finally, [23] generalizes the latter results to output-feedback and correlated disturbances. In [24] a limited number of simultaneous connections constraint is assumed, in [25] a holistic approach is adopted for resource-constrained networked systems, whilst in [26] a cooperative scheme for filtering from relative and absolute distance measurements is presented. Finally, an interesting scalable (state-feedback) method is proposed in [27] and [28] presents an explicit solution to a two player distributed LQG problem in which communication between controllers occurs across a communication link with varying delay. More recently, solutions to the distributed control problem of some interesting applications has been given for smart grids [29], high-speed trains [30], unmanned underwater vehicles [31].

To the best of our knowledge, there is still a lack regarding the fundamental problem of the distributed stochastic optimal regulator in continuous-time, namely a distributed solution which solves the classical Linear-Quadratic control problem for Gaussian systems over a general structure of interconnections of nodes. In the spirit of some recent advances in the context of distributed filtering [32,33], this paper tackles the infinite-horizon optimal distributed regulator for Gaussian continuous-time linear systems with partial state information. We shall clarify in the paper what we mean by *distributed* and the information we assume each node has.

The paper is organized as follows. The problem is introduced in Section 2 with specifications on the network structure, assumptions, and other preliminaries. Section 3 solves the distributed infinite-horizon output-feedback control with global information on the system matrices, whilst Section 4 addressed a solution with *local* information on the system matrices. Numerical simulations (Section 5) validate the theoretical results, and conclusion follows.

Notation.

$\mathbb{R}$ denotes the set of real numbers. For a square matrix A, $\mathrm{tr}(A)$ 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 $\mathrm{row}_i()$, $\mathrm{col}_i()$, $\mathrm{diag}_i()$ denote respectively the horizontal, vertical and diagonal compositions of matrices and vectors indexed by i. If A is a squared matrix, we write $A \geq 0$ $(A > 0)$ if A is positive semi-definite (respectively, positive definite). We denote by I_n the identity matrix of size n, by $1_n = \mathrm{col}_{i=1}^n(1)$ the column vector with all entries one, and by $U_n = 1_n 1_n^\top$ the matrix with all entries one. A network, e.g. Fig. 1, is topologically described by a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ which defines the information exchange among N nodes, where $\mathcal{V} = \{1, 2 \dots, N\}$ 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. A graph is undirected if 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 $\mathcal{N}_i$. 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.

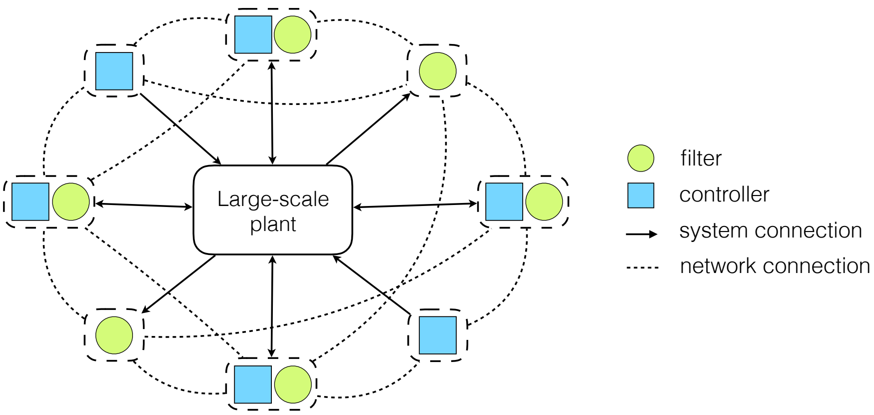


FIGURE 1 Example of a network diagram for the distributed control of a large-scale plant. In particular, it has been reported the connections among the filter and/or controller and the plant (*system connections*), and the connections among neighbors (*network connections*).

2 | PROBLEM FORMULATION AND PRELIMINARIES

We consider a network described by a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ together with the system model

$$\dot{x}(t) = Ax(t) + \sum_{i=1}^{N} B_i u_i(t) + w(t), \tag{1}$$

$$y_i(t) = C_i x(t) + v_i(t), \quad i = 1, \dots, N, \tag{2}$$

where $x(t) \in \mathbb{R}^n$ is the state of the system, $y_i(t) \in \mathbb{R}^{q_i}$, $q_i \geq 0$, is the measurement received by sensor i-th, $u_i(t) \in \mathbb{R}^{p_i}$, $p_i \geq 0$, is the input signal sent by sensor i-th, and $w(t)$ and $v_i(t)$, $i = 1, \dots, N$, are zero-mean white Gaussian noises, mutually independent with covariances $W \geq 0$, $V_i > 0$ $i = 1, \dots, N$, respectively. Also, at the initial time t_0, the initial condition $x(t_0) =: x_0$ is a Gaussian random vector with mean $\bar{x}_0 := \mathbb{E}\{x_0\}$ and intensity $\Sigma_{\bar{x}_0}$.

The infinite-horizon cost function $\bar{J}$ is

$$\bar{J} = \lim_{\substack{t_f \to +\infty \\ t_0 \to -\infty}} \frac{1}{t_f - t_0} J, \tag{3}$$

with

$$J = \mathbb{E}\left\{ \int_{t_0}^{t_f} x^\top(t) Q x(t) + \sum_{i=1}^{N} u_i^\top(t) R_i u_i(t) \mathrm{d}t \right\}, \tag{4}$$

where $Q \geq 0$ and $R_i > 0$ for all $i \in \mathcal{V}$.

By defining $B = \mathrm{row}_i(B_i)$, $C = \mathrm{col}_i(C_i)$, $R = \mathrm{diag}_i(R_i)$ $u(t) = \mathrm{col}_i(u_i(t))$, $y(t) = \mathrm{col}_i(y_i(t))$, $v(t) = \mathrm{col}_i(v_i(t))$, the model (1)–(2) can be written in a compact form as

$$\dot{x}(t) = Ax(t) + Bu(t) + w(t), \tag{5}$$

$$y(t) = Cx(t) + v(t), \tag{6}$$

with the function (4) given by

$$J = \mathbb{E}\left\{ \int_{t_0}^{t_f} x^\top(t) Q x(t) + u^\top(t) R u(t) \mathrm{d}t \right\}. \tag{7}$$

We also denote by $\bar{J}(u_t)$ the cost $\bar{J}$ when the control sequence $\{u(t)\}_{t \geq 0}$ is applied to the plant. In order to design a distributed optimal output-feedback control, we consider the following assumptions.

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

Assumption 2. The couples (A, B), $(A, W^{\frac{1}{2}})$ are stabilizable, and the couples (C, A), $(Q^{\frac{1}{2}}, A)$ are detectable.

Assumption 3. Each node $i \in \mathcal{V}$ of the network has information on the matrices A, B, C, W, V, Q, R.

Assumption $3'$**.** Each node $i \in \mathcal{V}$ of the network has information on the matrices A, B_i, C_i, W, V_i, Q, R_i.

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). Instead, Assumption 3 reflects the paradigm "*known-global-act-local*" (e.g.[34] Ch. 11,[35]), in which each node has information on the structural data of the whole network but takes local actions. We shall see in Section 4, how to relax Assumption 3, in particular how to relax the knowledge on the matrices B, C, V, R (Assumption $3'$), and to solve the problem when only local information are available to the nodes of the network.

The aim of this paper is the following.

Goal: *Find the distributed optimal output-feedback control policy* $\{u_i(t)\}_{t\geq 0}$ *for each sensor* $i \in \mathcal{V}$*, of system* (1)–(2) *such that the cost function* J *defined in* (3) *is minimized.*

We notice that the term *distributed* refers to the fact that the local controllers make use of a control policy that is implementable with the local information of the node and the neighbors information (Section 4), where the underlying graph is only connected, and there is not the need of converging some information in specific points of the network. In particular, the algorithm satisfies the following linear complexity online constraints on communication and computation[1]:

1. *Local information exchange*: for any $t \geq 0$ each node $i \in \mathcal{V}$ can exchange information with each $j \in \mathcal{N}_i$.
2. *Local computation*: for any $t \geq 0$ the computational load of each node $i \in \mathcal{V}$ should be at most $O\left(|\mathcal{N}_i|\right)$.

In other words, this means that for any $t \geq 0$, each $u_i(t)$ can be generated by using the local measurement $y_i(t)$ and other data received from its neighbors $j \in \mathcal{N}_i$ satisfying the constraints above.

Finally, we recall the following standard result (e.g.[36]).

Proposition 1. Given a system in the form (5)–(6) with the cost function (3) and J as in (7), under the Assumption 2, the optimal output-feedback control $\{u^\star(t)\}_{t\geq 0}$ is given by

$$u^\star(t) = -L\,\check{x}(t) \tag{8}$$

where

$$L = R^{-1}B^\top S, \tag{9}$$

with S solution to the algebraic Riccati equation

$$0 = A^\top S + SA - SBR^{-1}B^\top S + Q, \tag{10}$$

and $\check{x}(t)$ is the optimal estimate (in the minimum variance sense) provided by stationary centralized asymptotic Kalman-Bucy filter (CKBF), namely

$$\dot{\check{x}}(t) = A\check{x}(t) + Bu(t) + K(y(t) - C\check{x}(t)) \tag{11}$$

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

where P is the solution to the algebraic Riccati equation

$$0 = AP + PA^\top - PC^\top V^{-1}CP + W. \tag{13}$$

Finally, the optimal cost $\bar{J}(u_t^\star)$ is given by

$$\bar{J}(u_t^\star) = \mathrm{tr}(SW + PL^\top RL). \tag{14}$$

We also recall that the closed-loop dynamic matrix of the stationary CKBF (11), namely $A_C := A - KC$, is Hurwitz by construction under the hypothesis (C, A) detectable (Assumption 2).

Remark 1. We note that the stationary CKBF (11) employs all the measurements $y_i(t)$ and all the input signal $u_i(t)$ of all the nodes of the network $i \in \mathcal{V}$.

With reference to the distributed optimal control problem, because of Remark 1, we refer to (8) and (11) as the *centralized* output-feedback LQG regulator.

[1] excluding *off-line* computation.

3 | IDEAL DISTRIBUTED LQG REGULATOR

In this section we shall design a distributed counterpart of Proposition 1. We start with the trivial case of state-feedback and then, we shall see how to extend this solution to the more interesting case when the state is not available.

Over the underlying network described by the graph $\mathcal{G}$ and under the assumption of *full state information* (namely the state $x(t)$ is available to each node $i \in \mathcal{V}$) we can trivially rewrite the control law (8) in a distributed fashion as indicated in the following lemma.

Lemma 1. Given the system (1) over the graph $\mathcal{G}$ with the cost function (3), under the Assumption 1–2–3, the optimal distributed state-feedback control $\{u_i(t)\}_{t\geq 0}$ for each node $i \in \mathcal{V}$, is given by

$$u_i(t) = -L_i x(t), \tag{15}$$

where $L_i = R_i^{-1} B_i S$ is the i-th $p_i \times n$ row block of the gain L given by (9).

Remark 2. We note that because of Assumption 3, each node $i \in \mathcal{V}$ can compute off-line the gain L_i since it can compute off-line the matrix S solution to (10).

The following theorem solves the infinite-horizon distributed optimal output-feedback control problem.

Theorem 1. Given the system (1)–(2) over the graph $\mathcal{G}$, with the cost function (3), under the Assumption 1–2–3, for $i \in \mathcal{V}$, the distributed output-feedback control

$$\hat{u}_i(t) = -L_i \hat{x}_i(t), \tag{16}$$

where $L_i = R_i^{-1} B_i S$ and $\hat{x}_i(t)$ given by

$$\dot{\hat{x}}_i(t) = A\hat{x}_i(t) - BL\hat{x}_i(t) + K_i(y_i(t) - C_i\hat{x}_i(t)) + \gamma P \sum_{j\in\mathcal{N}_i} (\hat{x}_j(t) - \hat{x}_i(t)), \tag{17}$$

with $K_i = NPC_i^\top V_i^{-1}$, P solution of (13), is such that

$$\lim_{\gamma\to+\infty} \bar{J}(\hat{u}_t) = \bar{J}(u_t^\star), \tag{18}$$

where $\hat{u}_t = \{\hat{u}(t)\}_{t\geq 0}$ and $\hat{u}(t) = \mathrm{col}_i(\hat{u}_1(t), \dots, \hat{u}_N(t))$, and $u_t^\star = \{u^\star(t)\}_{t\geq 0}$ is the control process of the centralized LQG regulator (8). In particular, there exists $\gamma_0 > 0$ such that for any $\gamma > \gamma_0$ the cost $\bar{J}(\hat{u}_t)$ is finite.

In other words, Theorem 1 states that, as the parameter γ of the filter (21) tends to $+\infty$, the distributed control law (20) is equivalent to the centralized optimal solution provided by the control law (8).

Remark 3. It is worth noticing that each individual node of the network have no information of the whole control signal applied to the plant. This is reflected in the term $-BL\hat{x}_i(t) = -\sum_{j=1}^N B_j L_j \hat{x}_i$ in the filter equation (21) (which employs the local estimate $\hat{x}_i(t)$ only) that is mimicking the full control term $B\hat{u}(t) = -\sum_{j=1}^N B_j L_j \hat{x}_j$ applied to the plant.

Remark 4. In the papers[32,33] the analysis of the filtering algorithm is carried on for autonomous systems, *i.e.* systems without control inputs. We notice that if the term $-BL\hat{x}_i(t)$ in the filter equation (21) were set to $B\hat{u}(t)$, then the estimation error dynamics of the filter (21) would coincide with the one of the filter of[32,33], and thus it would follow that the estimate $\hat{x}_i(t)$ of (21) would be asymptotically optimal in the minimum variance sense (as $t \to +\infty$ and $\gamma \to +\infty$), and the optimality of the control policy (20) would follow by Lemma 1. However, the latter choice is clearly not feasible since each node $i \in \mathcal{V}$ would require the full control signal $\hat{u}(t)$ as pointed out in Remark 3. Thus, further analysis is required.

In what follows, we denote the estimation error of the filter of the node $i \in \mathcal{V}$ with $e_i(t) := x(t) - \hat{x}_i(t)$, the total estimation error with $e(t) := \mathrm{col}_i(e_i(t))$, and the total estimation error intensity matrix $X(t) := \mathbb{E}[e(t)e^\top(t)]$. Clearly, $X(t)$ depends on γ (since each $\hat{x}_i(t)$ of (21) depends on γ), but we omit this dependence for notational simplicity. Moreover, let $X_\infty^C := U_N \otimes P$, with U_N defined in the notation section and P solution of (13), be the asymptotic error covariance of N identical CBKFs implemented at each node using the whole output $y(t)$.

The proof of Theorem 1 is exploited through the following fundamental lemma proved in the Appendix.

Lemma 2. Given the system (1)–(2) over the graph $\mathcal{G}$, under the Assumption 1–2–3, with the cost function (3) and the control law (20), there exists $\gamma_0 > 0$ such that for any $\gamma > \gamma_0$ the state estimation (21) for each node $i \in \mathcal{V}$ is unbiased and the total covariance of the estimation error $X(t)$ is uniformly bounded. Moreover, it holds

$$\lim_{\gamma \to +\infty} \lim_{t \to +\infty} X(t) = X_\infty^C := U_N \otimes P. \tag{19}$$

In other words, Lemma 2 states that, as $\gamma \to +\infty$, the covariance matrix of the estimation error of the filter (21) of each node $i \in \mathcal{V}$ tends asymptotically (*i.e.* as $t \to +\infty$) to the covariance matrix P of the estimation error of the centralized Kalman-Bucy filter (11).

Proof of Theorem 1. The thesis follows directly from Lemma 1 and Lemma 2. In fact, Lemma 2 guarantees that the filter (21) converges as $t \to +\infty$ and $\gamma \to +\infty$ to the CKBF, while Lemma 1 guarantees the optimality of the control law (20) which minimizes the cost function (3). □

We conclude this section with two important remarks.

Remark 5. It is worth remarking that in any discrete-time implementation of the control law (20), and thus of the filter (21), the value of γ cannot be chosen arbitrarily large due to discretization issues. Roughly speaking, a larger γ requires a smaller integration step that constraints the communication lag among nodes. Consequently, any implementable version of the proposed solution will suffer a certain performance degradation with respect to the centralized LQG regulator.

Remark 6. The control law is extremely simple to implement and the information exchange among nodes is reduced to a minimum. The matrices L and P can be computed by solving the algebraic Riccati equations (10) and (13), which are of dimension $n \times n$ and do not depend on the graph structure. Clearly, with many sensor and actuator nodes, the size of the pairs (B, R) and (C, V) can be large, but $BR^{-1}B^\top$ and $C^\top V^{-1}C$ is again an $n \times n$ matrix. Finally, L and P can be computed off-line.

4 | 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 3 by assuming that each node $i \in \mathcal{V}$ has only local information as clarified by Assumption 2′ we recall here.

Assumption 3′**.** Each node $i \in \mathcal{V}$ of the network has information on the matrices A, B_i, C_i, W, V_i, Q, R_i.

In the first place it is worth remarking that the computation of S and P by solving (10) and (13) is trivial, since S and P are the solution to algebraic equations in $\mathbb{R}^{n\times n}$ that does not depend on the size of the graph. Thus, nodes with limited computational power can easily solve (10) and (13) provided that the values of $F := BR^{-1}B^\top$ and $G := C^\top V^{-1}C$ are available, respectively. We note that F can be written as $F = \sum_{i=1}^N B_i R_i^{-1} B_i^\top$ and, when measurement noises are independent (which is our case), G can be expressed similarly with the sum $G = \sum_{i=1}^N C_i^\top V_i^{-1} C_i$.

A distributed computation of the terms F and G can thus be achieved by resorting to distributed algorithms to compute aggregate functions over graphs[37]. In Figure 2 we report an algorithm derived from the *Protocol Push-Sum* of[37] to compute F and G in a distributed way. Thus, the knowledge of F and G allows to compute the solutions S and P of the algebraic Riccati equations (10) and (13) exploiting local information only.

Also, there is another critical issue in the implementation of the term $BL\hat{x}_i(t)$ in the filter (21) of each node $i \in \mathcal{V}$. In fact, it seems that the latter term can be only implemented if each node $i \in \mathcal{V}$ has knowledge on the full matrices B and L. However, as before, the computation of the product BL can be exploited with local information only. In fact, it is worth remarking that the equality $BL = FS$ holds true, and both F and S can be computed in a fully distributed way as described above.

We remark that the main difference of the algorithm of Figure 2 with respect to[37] is that the communication is supposed to be point-to-point, whereas the algorithm in Figure 2 assumes that a node can broadcast messages to all its neighbors. The speed of convergence of the local estimate to the true value of F and G can be analyzed in the light of the results of[37]. This estimation step can be executed before the controlling phase for static graphs, or it can be kept running during the operating phase in order to adjust the value of F and G in presence of a dynamical graph where nodes appear or disconnect. Finally, the value of N can be also computed by the same distributed algorithm when it is not known at the nodes.

Algorithm Broadcast Push-Sum

1: In all nodes set $s_{0,i} = B_i R_i^{-1} B_i^\top$ 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 G at step t (if $w_{t,i} = 0$ the estimate is not specified or 0).

FIGURE 2 A modified version of the *Push-Sum* algorithm of[37] that makes possible the distributed computation of F (and G if $s_{0,i} = C_i^\top V_i^{-1} C_i$).

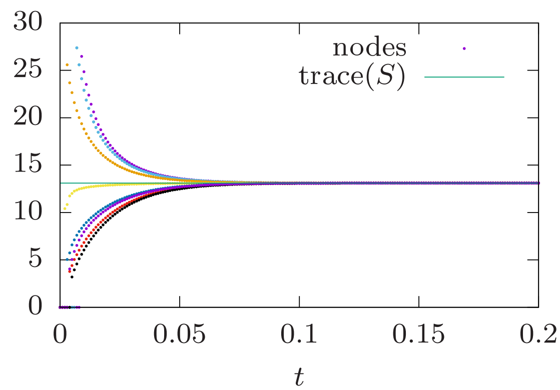


FIGURE 3 Distributed computation of the trace S solution to the algebraic Riccati equation (10), computed by each node of the network, when the algorithm Push-Sum of Fig. 2 is employed by the nodes to compute the product $BR^{-1}B^\top$.

By considering the system model, the graph structure (Fig. 5) and the integration step of Section 5.1, we show in Figure 3 the evolution of the matrix S (for simplicity we only show the trace), solution to the algebraic Riccati equation (10), computed by each node of the network, when the algorithm Push-Sum of Fig. 2 is employed by the nodes to compute the product $BR^{-1}B^\top$ in a distributed way. In particular, Figure 3 shows the convergence towards the trace of the matrix S computed by using global information.

We conclude this section with a corollary that complete the analysis of the problem.

Corollary 1. Consider the system (1)–(2) over the graph $\mathcal{G}$, with the cost function (3), under the Assumption 1–2–3′, for $i \in \mathcal{V}$, and let the distributed output-feedback control be

$$\hat{u}_i(t) = -L_i \hat{x}_i(t), \tag{20}$$

with $L_i = R_i^{-1} B_i S$, where S is the solution to (10) when the term $F = BR^{-1}B^\top$ is computed by each node $i \in \mathcal{V}$ through the Push-Sum algorithm of Figure 2, and $\hat{x}_i(t)$ given by

$$\dot{\hat{x}}_i(t) = A\hat{x}_i(t) - FS\hat{x}_i(t) + K_i(y_i(t) - C_i\hat{x}_i(t)) + \gamma P \sum_{j \in \mathcal{N}_i} (\hat{x}_j(t) - \hat{x}_i(t)), \tag{21}$$

with $K_i = NPC_i^\top V_i^{-1}$, where P is the solution to (13) when the term $G = C^\top V^{-1} C$ is computed by each node $i \in \mathcal{V}$ through the Push-Sum algorithm of Figure 2. Then, it holds

$$\lim_{\gamma \to +\infty} \bar{J}(\hat{u}_t) = \bar{J}(u_t^\star), \tag{22}$$

where $\hat{u}_t = \{\hat{u}(t)\}_{t\geq 0}$ and $\hat{u}(t) = \mathrm{col}_i(\hat{u}_1(t), \dots, \hat{u}_N(t))$, and $u_t^\star = \{u^\star(t)\}_{t\geq 0}$ is the control process of the centralized LQG regulator (8). In particular, there exists $\gamma_0 > 0$ such that for any $\gamma > \gamma_0$ the cost $\bar{J}(\hat{u}_t)$ is finite.

The proof of this corollary is identical to the one of Theorem 1 by simply taking into account that the limit points of the approximated F and G through the Push-Sum algorithm converge to the real values of F and G, respectively, and thus the same happens for the solutions S and P to the Riccati equations (10) and (13), respectively.

Also, Corollary 1 implicitly requires that the computations through the Push-Sum algorithm of F and G are made offline. However, this is not strictly necessary and online computations of F and G with the corresponding solutions S and P to the Riccati equations (10) and (13), respectively, can be done. We remark that this latter case does not cause, in general, a degradation of the performance since we are considering the infinite-horizon case. In fact, even in the ideal case, i.e. when F and G are available, the steady-state optimal gains L (of the controller) and K (of the filter), computed through the matrices S and P respectively, are not optimal in the transient phase.

To sum-up:

- *offline computation of F and G*: through the Push-Sum algorithm each node can compute in a distributed way the values of F and G.
  Pros: once each node has computed F and G, it is possible to compute matrices P and S solutions to the Riccati equations (10) and (13) (just once) and to implement the distributed regulator that *coincides* with the ideal regulator of Section 3 (but with the exception that only local information has been exploited).
  Cons: a quick pre-processing in order to compute F and G.

- *online computation of F and G*: the Push-Sum algorithm is kept running in order to compute adaptively in a distributed way the value of F and G.
  Pros: no need of the pre-processing phase; capability of accommodate link insertion or link deactivation (or failure).
  Cons: the Push-Sum algorithm needs to be kept running during all the operating phase toughener with the computation of the matrices P and S.

5 | SIMULATION EXAMPLES

In this section we consider two example in order to show the validity of the theoretical results and the effectiveness of the proposed method: in Section 5.1 we report an academic example in which each node together with its neighbors does not constitute an observable system, and only the collective observability assumption holds true. In the second example (Section 5.2) we consider a real-plant example. The examples show the generality of application and the actual effectiveness of the proposed control design.

5.1 | Academic example

In this academic example we consider the graph $\mathcal{G}$ of Figure 5 consisting of nine nodes. The plant (1) is characterized by the unstable dynamical matrix A[2]

$$A = \begin{bmatrix} -0.1 & 0 & 0 & -1 & 1 & 0 & 0 & 1 \\ 2.5 & -0.5 & -1.6 & -1.5 & 2 & 0 & 1.6 & 1.5 \\ 2.6 & -0.5 & -0.7 & -1.5 & 1.5 & 0.5 & 0.5 & 1.5 \\ -2 & 0 & 1 & 0 & -0.1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.1 & 0 & 0 & 0 \\ -0.5 & 0 & 1 & 0 & 0.5 & -0.5 & -1 & 0 \\ 3.8 & -0.5 & -1.8 & -0.5 & 2 & 0.5 & 1.6 & 0.5 \\ -1 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \end{bmatrix}, \tag{23}$$

and the state and output noises have intensity $W = 0.09 \cdot I_n$ and $V_i = 0.36$ with $i = 1, \dots, 9$. The output and input matrices of the nodes are

$$C_1 = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad C_2 = \begin{bmatrix} -2 & 1 & 1 & 1 & -1 & 0 & -1 & -1 \end{bmatrix} C_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad C_4 = \begin{bmatrix} -3 & 1 & 2 & 1 & -1 & 0 & -1 & -1 \end{bmatrix}$$
$$C_5 = \begin{bmatrix} 1 & 0 & -1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}, \quad C_6 = \begin{bmatrix} 2 & -1 & -1 & -1 & 1 & 1 & 1 & 1 \end{bmatrix}, \quad C_7 = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}, C_8 = \begin{bmatrix} 1 & 0 & -1 & -1 & 2 & 0 & 1 & 1 \end{bmatrix},$$

$C_9 = [0, 0, 1, 1, -1, 0, 0, 0]$, and $B_i = C_i^\top$ for all $i = 1, \dots, 9$. Also, the cost index (3) has $Q = I_7$ and $R = I_9$. 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 $\left(A, \text{row}_{j \in \mathcal{N}_i \cup \{i\}}(B_j)\right)$ is not controllable for any $i \in \mathcal{V}$ and the couple $\left(\text{col}_{j \in \mathcal{N}_i \cup \{i\}}(C_j), A\right)$ 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 (20) are applied to the plant, in accordance with the expected cost (3) of the optimal centralized solution, we define the pseudo-cost function

$$\widehat{J}_\gamma(\hat{u}_t) = \text{tr}\left(SW + \widehat{P}(\gamma) L^\top R L\right), \tag{24}$$

where $\widehat{P}(\gamma)$ is the arithmetic mean of the covariance matrices of the estimation errors of the nodes $i \in \mathcal{V}$, namely $\widehat{P}(\gamma) = \frac{1}{N} \sum_{i=1}^N X_i(\gamma)$ with $X_i(\gamma)$ covariance of the estimation error $e_i(t)$. Clearly, by mean of the result of Theorem 1, it follows that $\lim_{\gamma \to +\infty} \widehat{J}_\gamma(\hat{u}_t) = \bar{J}(u_t^\star)$, where $u_t^\star$ is the optimal centralized control (8). Figure 4 (right) shows the convergence of $\widehat{J}_\gamma(\hat{u}_t)$ of (24) to the optimal cost when γ tends to infinity, and similarly Figure 4 (left) shows the convergence of the traces of the covariance of the estimation error of all the nodes towards the optimal value of the trace of the covariance of the centralized Kalman-Bucy filter.

A simulation with 100 independent realizations of initial condition and noise processes has been also performed with final time $t_f = 200$ and an Euler-Maruyama approximation scheme with samples $\Delta = 10^{-3}$. The consensus gain parameter is set $\gamma = 200$. Table 1 summarizes the results: it shows the *a priori* optimal values of estimation error and cost, namely the trace of the covariance of the estimation error of the centralized Kalman-Bucy filter, *i.e.* $\text{tr}(P)$, and the optimal cost of the centralized LQG control, *i.e.* $\bar{J}$ computed as (14). Moreover, by averaging over the 100 realizations, Table 1 shows the Average Mean Square Errors of the centralized Kalman-Bucy filter ($\text{AMSE}_{\text{CKBF}}$) and the mean of the AMSE among the nodes of the network ($\text{AMSE}_{\text{nodes}}$) with its standard deviation (σ_{AMSE}). Also, it shows the actual cost of the centralized LQG control $\widetilde{J}_C$ and of the proposed distributed control $\widetilde{J}_D$.

Table 1 summarizes the results: it shows the *a priori* optimal value of estimation error, namely the trace of the covariance of the estimation error of the centralized Kalman-Bucy filter, *i.e.* $\text{tr}(P)$, and, by averaging over the 100 realizations, the Average Mean Square Errors of the centralized Kalman-Bucy filter ($\text{AMSE}_{\text{CKBF}}$) and the mean of the AMSE among the nodes of the network ($\text{AMSE}_{\text{nodes}}$) with its standard deviation (σ_{AMSE}). Also, Table 1 shows the *a priori* optimal cost of the centralized LQG control, *i.e.* $\bar{J}$ computed as (14) and, by averaging over the 100 realizations, shows the actual cost of the centralized LQG control $\widetilde{J}_C$ and of the proposed distributed control $\widetilde{J}_D$.

[2] A has a positive eigenvalue in 1.1.

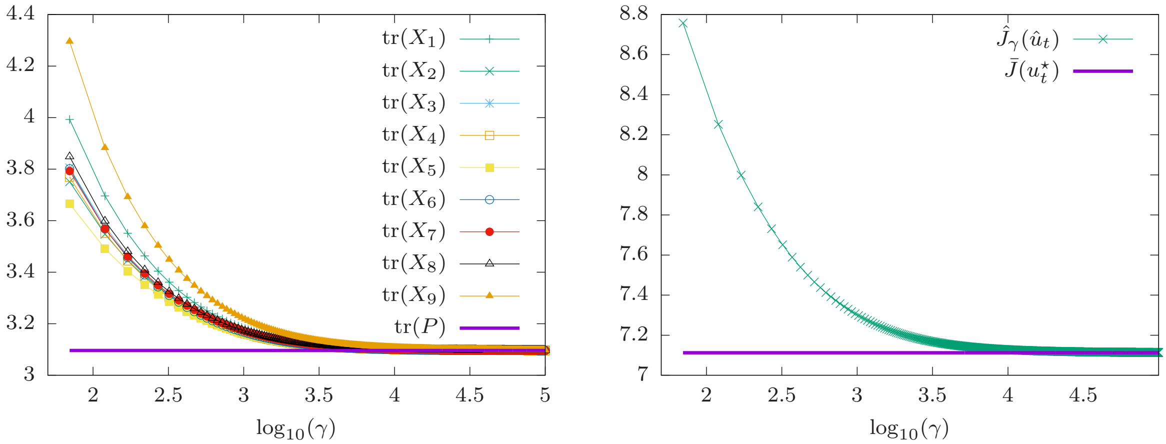
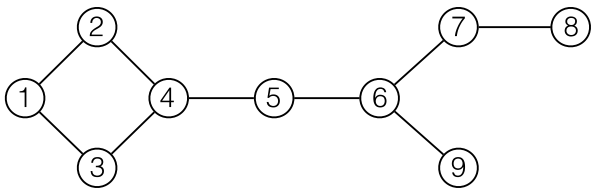


FIGURE 4 Traces of the covariance matrices of the estimation errors tr(X_i) of the nodes and trace of the covariance of the CKBF tr(P) (left). Convergence of the pseudo-cost function (24) towards the centralized optimal one. (right).



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

FIGURE 5 Communication graph $\mathcal{G}$ and its Laplacian matrix $\mathcal{L}$.

	γ	tr(P)	AMSE$_{\text{CKBF}}$	AMSE$_{\text{nodes}}$	σ_{AMSE}	$\bar{J}$	$\widetilde{J}_C$	$\widetilde{J}_D$
Example 1	200	3.096	3.142	3.503	0.006	7.113	7.635	9.435
Example 2	50	0.0243	0.0371	0.0596	$< 10^{-5}$	0.0512	0.0721	0.1658

TABLE 1 *A priori* optimal values of estimation error and cost, namely the trace of the covariance of the estimation error of the CKBF, *i.e.* tr(P), and the optimal cost of the centralized LQG control, *i.e.* $\bar{J}$ computed as (14). Empirical Average Mean Square Errors of the CKBF (AMSE$_{\text{CKBF}}$) and the mean of the AMSE (AMSE$_{\text{nodes}}$) with its standard deviation (σ_{AMSE}) of the nodes of the network. Empirical cost of the centralized LQG control $\widetilde{J}_C$ and of the proposed distributed control $\widetilde{J}_D$.

5.2 | Oscillator

In the second example we consider an oscillator example where the state vector consists of two position and two velocity components. The plant in the form (1)–(2) is characterized by the marginally stable system

$$A = \begin{bmatrix} 0 & -0.5 & 0 & 0 \\ 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 \\ 0 & 0 & -0.5 & 0 \end{bmatrix}, \qquad C_i = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}, \; i = 1, 6. \qquad C_i = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}, \; i = 2, 7.$$

$$C_i = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix}, \; i = 4, 8. \qquad C_i = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}, \; i = 3, 5, 9,$$

$B_i = C_i^\top$ for all $i = 1, \dots, 9$, the same graph of Figure 5, where the state and output noises have intensity $W = 0.1 \cdot I_n$ and $V_i = 0.3$ with $i = 1, \dots, 9$. Also, the cost index (3) has $Q = I_4$ and $R = I_9$. This setting is such that the couple (A, B_i) (respectively (C_i, A)) is not controllable (respectively observable) for all $i \in \mathcal{V}$. However, the hypotheses of Assumption 2 are satisfied, in particular (A, B) is controllable and (C, A) is observable.

A simulation with 100 independent realizations of initial condition and noise processes has been performed with final time $t_f = 100$ and an Euler-Maruyama approximation scheme with samples $\Delta = 10^{-3}$. The consensus gain parameter is set to $\gamma = 50$. The second row of Table 1 summarizes the results: it shows the *a priori* optimal values of estimation error and cost, namely the trace of the covariance of the estimation error of the centralized Kalman-Bucy filter, *i.e.* $\mathrm{tr}(P)$, and the optimal cost of the centralized LQG control, *i.e.* $\bar{J}$ computed as (14). Moreover, by averaging over the 100 realizations, Table 1 shows the Average Mean Square Errors of the centralized Kalman-Bucy filter ($\mathrm{AMSE}_{\mathrm{CKBF}}$) and the mean of the AMSE among the nodes of the network ($\mathrm{AMSE}_{\mathrm{nodes}}$) with its standard deviation (σ_{AMSE}). Also, it shows the actual cost of the centralized LQG control $\widetilde{J}_C$ and of the proposed distributed control $\widetilde{J}_D$.

6 | CONCLUSIONS

Further extensions deserve additional investigation, for example the introduction of communications delays or packet dropouts[38,39], link failure[40], or polynomial techniques for non-Gaussian noises[41,42,43]. Also, other future directions are secure and resilient solutions that account for disturbances, faults, threats, and attacks and/or disturbances. Finally, the proposed control scheme does not extend trivially to the discrete-time case, as mentioned in Remark 5. Thus, it is of interest to derive a discrete-time implementation of the proposed solution and to characterize the loss of accuracy and consensus of this discrete-time counterpart with respect to the optimal case as a function of the size of the discretization interval.

□

APPENDIX

For what follows, we need to consider the orthonormal transformation $T = \begin{pmatrix} T_1 \\ T_2 \end{pmatrix}$, with $T_1 = \frac{1}{\sqrt{N}} 1_N^\top$ and $T_2 \in \mathbb{R}^{(N-1)\times N}$, such that

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

where $\Lambda = \mathrm{diag}(\lambda_2(\mathcal{L}), \dots, \lambda_N(\mathcal{L}))$, and $\mathcal{L} T_1^\top = 0$ and $\mathcal{L} T_2^\top = T_2^\top \Lambda$ hold true.

Proof of Lemma 2

Let the control (20) be applied to the plant (1). Omitting time dependencies, the state estimation error $e_i = x - \hat{x}_i$ of each node $i \in \mathcal{V}$ is given by

$$\dot{e}_i = A_i e_i + \xi_i - \gamma P \sum_{j \in \mathcal{N}_i} (e_i - e_j) + h_i, \tag{1}$$

where $A_i = A - K_i C_i$, $h_i = w - K_i v_i$ and the control mismatch ξ_i given by

$$\xi_i = Bu + BL\hat{x}_i = -\sum_{j=1}^{N} B_j L_j e_i + \sum_{j=1}^{N} B_j L_j e_j. \tag{2}$$

By noticing that $\xi = \mathrm{col}(\xi_i)$ can be written as

$$\xi = -\left[\left(I_N \otimes \sum_{j=1}^{N} B_j L_j \right) - \left(1_N \otimes \mathrm{row}_j^N (B_j L_j) \right) \right] e = -\Xi\, e, \tag{3}$$

with

$$\Xi = \begin{bmatrix} \sum_{\substack{j=1\\ j\neq 1}}^{N} B_j L_j & -B_2 L_2 & \cdots & -B_N L_N \\ -B_1 L_1 & \sum_{\substack{j=1\\ j\neq 2}}^{N} B_j L_j & \cdots & -B_N L_N \\ \vdots & \vdots & \ddots & \vdots \\ -B_1 L_1 & -B_2 L_2 & \cdots & \sum_{\substack{j=1\\ j\neq N}}^{N} B_j L_j \end{bmatrix}, \tag{4}$$

we can write the total estimation error $e(t) = \mathrm{col}_i(e_i(t))$ as

$$\dot{e} = \big(\mathrm{diag}_i\{A_i\} - \gamma(\mathcal{L} \otimes P) - \Xi\big)e + h \tag{5}$$

where $h(t) = \mathrm{col}_i(h_i(t))$ has intensity matrix $H = \mathbb{E}\{h(t)h^\top(t)\} = U_N \otimes W + \mathrm{diag}_i(K_i)V\mathrm{diag}_i(K_i^\top)$.

Moreover, by noticing that $(\mathcal{L} \otimes P)X_\infty^C = (\mathcal{L} \otimes P)(U_N \otimes P) = 0$, and $\Xi X_\infty^C = \Xi(U_N \otimes P) = \frac{1}{N}\Xi(U_N \otimes I_n)(U_N \otimes P) = 0$[3], we obtain that X_∞^C satisfies

$$\begin{aligned} 0 = &\big(\mathrm{diag}_{i=1}^N(A_C) - \gamma(\mathcal{L} \otimes P_\infty) - \Xi\big) X_\infty^C + X_\infty^C \big(\mathrm{diag}_{i=1}^N(A_C) - \gamma(\mathcal{L} \otimes P_\infty) - \Xi\big)^\top + U_N \otimes Q \\ &+ \mathrm{diag}_{i=1}^N(K_\infty)(U_N \otimes V)\mathrm{diag}_{i=1}^N(K_\infty^T), \end{aligned} \tag{6}$$

and, by introducing the covariance mismatch $E := X - X_\infty^C$ we obtain after some manipulations

$$\dot{E} = A_D(\gamma)E + EA_D^\top(\gamma) + \Sigma, \tag{7}$$

where $A_D = \big(\mathrm{diag}_i\{A_i\} - \gamma(\mathcal{L} \otimes P) - \Xi\big)$ and

$$\Sigma := N^2(I_N \otimes P_\infty)G_d(I_N \otimes P_\infty) + U_N \otimes (P_\infty G P_\infty) - N(I_N \otimes P_\infty)G_d(U_N \otimes P_\infty) - N(U_N \otimes P_\infty)G_d(I_N \otimes P_\infty), \tag{8}$$

with $G_i := C_i^\top R_i^{-1} C_i$, $G_d := \mathrm{diag}_i(G_i)$, and $G := C^\top R^{-1} C$. Define $S := T \otimes I_n$, where T is defined at the beginning of the Appendix, and let $\widetilde{E}(t) := S\,E(t)S^\top$. By noticing that

$$G = \sum_{i=1}^N G_i = \sum_{i=1}^N C_i^\top R_i^{-1} C_i, \tag{9}$$

we have after some manipulations

$$\dot{\widetilde{E}} = \widetilde{A}_D(\gamma)\widetilde{E} + \widetilde{E}\widetilde{A}_D^\top(\gamma) + N^2\widetilde{\Sigma}, \tag{10}$$

with[4]

$$\widetilde{A}_D(\gamma) = \begin{pmatrix} A_C & \Pi_{1,2} & \cdots & \Pi_{1,N} \\ \Pi_{2,1} & \Pi_{2,2} - \gamma\lambda_2 P & \cdots & \Pi_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \Pi_{N,1} & \Pi_{N,2} & \cdots & \Pi_{N,N} - \gamma\lambda_N P \end{pmatrix}, \tag{11}$$

where, for $i = 2, \dots, N$ and $j = 2, \dots, N$, we define[5]

$$\Pi_{1,j} := -\sqrt{N}\sum_{\ell=1}^N t_{j,\ell} P G_\ell - B_\ell L_\ell, \tag{12}$$

$$\Pi_{j,1} := -\sqrt{N}\sum_{\ell=1}^N t_{j,\ell} P G_\ell, \tag{13}$$

$$\Pi_{i,j} = \Pi_{j,i} := \delta_{i,j}(A - BL) - N\sum_{\ell=1}^N t_{i,\ell} t_{j,\ell} P G_\ell, \tag{14}$$

[3] since $\Xi(U_N \otimes I_n) = 0$ because of the structure of Ξ in (4).
[4] we recall $A_C = A - KC$.
[5] we use $t_{i,j}$ to denote the (i,j)-th element of the matrix T.

and

$$\widetilde{\Sigma} = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & \widetilde{\Sigma}_{2,2} & \cdots & \widetilde{\Sigma}_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \widetilde{\Sigma}_{N,2} & \cdots & \widetilde{\Sigma}_{N,N} \end{pmatrix}, \qquad \widetilde{\Sigma}_{i,j} = \widetilde{\Sigma}_{j,i} := \sum_{\ell=1}^{N} t_{i,\ell} t_{j,\ell} P G_\ell P. \tag{15}$$

By (11) and the fact that A_C is Hurwitz by construction, it is not difficult to see that there exists a $\gamma_0 > 0$ such that the matrix $\widetilde{A}_D(\gamma)$ is Hurwitz for all $\gamma > \gamma_0$, and thus the estimate are unbiased. Notice also that since $\sigma(\widetilde{A}_D(\gamma)) \cap \sigma(-\widetilde{A}_D^\top(\gamma)) = \emptyset$ for all $\gamma > \gamma_0$, for each $\gamma > \gamma_0$ there exists a unique symmetric solution $\widetilde{E}_\infty(\gamma)$ to

$$0 = \widetilde{A}_D(\gamma)\widetilde{E}_\infty(\gamma) + \widetilde{E}_\infty(\gamma)\widetilde{A}_D^\top(\gamma) + N^2\widetilde{\Sigma}. \tag{16}$$

The proof is concluded if we show that the asymptotic estimation error intensity $\widetilde{E}_\infty$ satisfies

$$\lim_{\gamma\to+\infty} \widetilde{E}_\infty(\gamma) = 0. \tag{17}$$

In fact, the solution $\widetilde{E}_\infty(\gamma)$ to (16) can be parametrized in γ as follows. Let

$$W_1 := \mathrm{row}_{i=2}^{N}\Pi_{1i} \quad W_2 := \mathrm{col}_{i=2}^{N}\Pi_{1i} \tag{18}$$

$$W_0 := \begin{pmatrix} \Pi_{2,2} & \cdots & \Pi_{2,N} \\ \vdots & \ddots & \vdots \\ \Pi_{N,2} & \cdots & \Pi_{N,N} \end{pmatrix}, \quad \Lambda := \begin{pmatrix} \widetilde{\Sigma}_{2,2} & \cdots & \widetilde{\Sigma}_{2,N} \\ \vdots & \ddots & \vdots \\ \widetilde{\Sigma}_{N,2} & \cdots & \widetilde{\Sigma}_{N,N} \end{pmatrix}. \tag{19}$$

With this definitions the equation (16) reads out as

$$\begin{pmatrix} 0 & 0 \\ 0 & \Lambda \end{pmatrix} = \begin{pmatrix} A_C & W_1 \\ W_2 & W_0 - \gamma D \otimes P_\infty \end{pmatrix} \widetilde{E}_\infty(\gamma) + \widetilde{E}_\infty(\gamma) \begin{pmatrix} A_C & W_1 \\ W_2 & W_0 - \gamma D \otimes P \end{pmatrix}^\top, \tag{20}$$

where $D = \mathrm{diag}_{i=2}^{N}(\lambda_i)$. The solution $\widetilde{E}_\infty(\gamma)$ is analytic in $\gamma > 0$ and can be written (using a Taylor expansion) as

$$\widetilde{E}_\infty(\gamma) = \frac{1}{\gamma} \begin{pmatrix} Y_{1,1} + O\left(\frac{1}{\gamma^2}\right) & \frac{1}{\gamma} Y_{2,1} + O\left(\frac{1}{\gamma^2}\right) \\ \frac{1}{\gamma} Y_{2,1}^\top + O\left(\frac{1}{\gamma^2}\right) & Y_{3,1} + \frac{1}{\gamma} Y_{3,2} + O\left(\frac{1}{\gamma^2}\right) \end{pmatrix} \tag{21}$$

where $Y_{3,1}$ is the unique (since $\sigma(D \otimes P) \cap \sigma(-D \otimes P) = \emptyset$) solution of

$$Y_{3,1}(D \otimes P) + (D \otimes P)Y_{3,1} = N^2\Lambda,$$

$Y_{3,2}$ is the unique (since $\sigma(D \otimes P) \cap \sigma(-D \otimes P) = \emptyset$) solution of

$$(D \otimes P)Y_{3,2} + Y_{3,2}(D \otimes P) = W_0 Y_{3,1} + Y_{3,1} W_0^\top,$$

$Y_{2,1}$ is defined as

$$Y_{2,1} := W_1 Y_{3,1} (D \otimes P)^{-1}$$

and $Y_{1,1}$ is the unique (since $\sigma(A_C) \cap \sigma(-A_C^\top) = \emptyset$) solution of

$$A_C Y_{1,1} + Y_{1,1} A_C^\top = -(W_1 Y_{2,1}^\top + Y_{2,1} W_1^\top).$$

From (21) it follows that $\lim_{\gamma\to+\infty} \widetilde{E}_\infty(\gamma) = 0$, and thus we can finally conclude that

$$\lim_{\gamma\to+\infty} \lim_{t\to+\infty} \widetilde{E}(t) = \lim_{\gamma\to+\infty} \lim_{t\to+\infty} E(t) = 0, \tag{22}$$

that implies the thesis of the lemma.

References

1. Cheng B, Zhang J, Hancke GP, Karnouskos S, Colombo AW. Industrial cyberphysical systems: Realizing cloud-based big data infrastructures. *IEEE Industrial Electronics Magazine* 2018; 12(1): 25–35.

2. Pasqualetti F, Dorfler F, Bullo F. Control-theoretic methods for cyberphysical security: Geometric principles for optimal cross-layer resilient control systems. *IEEE Control Systems Magazine* 2015; 35(1): 110–127.
3. Bakule L. Decentralized control: An overview. *Annual Reviews in Control* 2008; 32(1): 87–98.
4. Bakule L, Papik M. Decentralized control and communication. *Annual Reviews in Control* 2012; 36(1): 1–10.
5. Aoki M. On feedback stabilizability of decentralized dynamic systems. *Automatica* 1972; 8(2): 163–173.
6. Wang SH, Davison EJ. On the stabilization of decentralized control systems. *IEEE Transactions on Automatic Control* 1973; 18(5): 473–478.
7. Corfmat J, Morse AS. Decentralized control of linear multivariable systems. *Automatica* 1976; 12(5): 479–495.
8. Davison E. The robust decentralized control of a general servomechanism problem. *IEEE Transactions on Automatic Control* 1976; 21(1): 14–24.
9. Anderson B, Moore JB. Time-varying feedback laws for decentralized control. *IEEE Transactions on Automatic Control* 1981; 26(5): 1133–1139.
10. Wang S. Stabilization of decentralized control systems via time-varying controllers. *IEEE Transactions on Automatic Control* 1982; 27(3): 741–744.
11. Vadigepalli R, Doyle FJ. A distributed state estimation and control algorithm for plantwide processes. *IEEE Transactions on Control Systems Technology* 2003; 11(1): 119–127.
12. Li T, Zhang J. Asymptotically optimal decentralized control for large population stochastic multiagent systems. *IEEE Transactions on Automatic Control* 2008; 53(7): 1643–1660.
13. Guerrero JM, Chandorkar M, Lee T, Loh PC. Advanced control architectures for intelligent microgrids—Part I: Decentralized and hierarchical control. *IEEE Transactions on Industrial Electronics* 2012; 60(4): 1254–1262.
14. Russell Carpenter J. Decentralized control of satellite formations. *International Journal of Robust and Nonlinear Control* 2002; 12(2-3): 141–161.
15. Shi XQ, Davison DE, Kwong R, Davison EJ. Optimal decentralized control of large scale systems. *Control Theory and Technology* 2016; 14(4): 287–313.
16. Sandell N, Athans M. Solution of some nonclassical LQG stochastic decision problems. *IEEE Transactions on Automatic Control* 1974; 19(2): 108–116.
17. Kurtaran BZ, Sivan R. Linear-quadratic-gaussian control with one-step-delay sharing pattern. *IEEE Transactions on Automatic Control* 1974; 19(5): 571–574.
18. Toda M, Aoki M. Second-guessing technique for stochastic linear regulator problems with delayed information sharing. *IEEE Transactions on Automatic Control* 1975; 20(2): 260–262.
19. Yoshikawa T. Dynamic programming approach to decentralized stochastic control problems. *IEEE Transactions on Automatic Control* 1975; 20(6): 796–797.
20. Rantzer A. A separation principle for distributed control. In: IEEE. ; 2006: 3609–3613.
21. Gattami A. Generalized linear quadratic control. *IEEE Transactions on Automatic Control* 2009; 55(1): 131–136.
22. Lamperski A, Doyle JC. On the structure of state-feedback LQG controllers for distributed systems with communication delays. In: IEEE. ; 2011: 6901–6906.
23. Feyzmahdavian HR, Gattami A, Johansson M. Distributed output-feedback LQG control with delayed information sharing. *IFAC Proceedings Volumes* 2012; 45(26): 192–197.

24. Hristu-Varsakelis D, Zhang L. LQG control of networked control systems with access constraints and delays. *International Journal of Control* 2008; 81(8): 1266–1280.

25. Yılmaz Y, Kurt MN, Wang X. Distributed dynamic state estimation and LQG control in resource-constrained networks. *IEEE Transactions on Signal and Information Processing over Networks* 2018; 4(3): 599–612.

26. Battilotti S, Cacace F, d'Angelo M, Germani A. Cooperative filtering with absolute and relative measurements. In: IEEE. ; 2018: 7182–7187.

27. Mårtensson K, Rantzer A. A scalable method for continuous-time distributed control synthesis. In: IEEE. ; 2012: 6308–6313.

28. Matni N, Doyle JC. Optimal distributed LQG state feedback with varying communication delay. In: IEEE. ; 2013: 5890–5896.

29. Bidram A, Nasirian V, Davoudi A, Lewis FL. Distributed Control of AC Microgrids. In: Springer. 2017 (pp. 67–98).

30. Wang Y, Song Y, Gao H, Lewis FL. Distributed fault-tolerant control of virtually and physically interconnected systems with application to high-speed trains under traction/braking failures. *IEEE Transactions on Intelligent Transportation Systems* 2015; 17(2): 535–545.

31. Liu H, Wang Y, Lewis FL. Robust distributed formation controller design for a group of unmanned underwater vehicles. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 2019.

32. Battilotti S, Cacace F, d'Angelo M, Germani A. Asymptotically Optimal Consensus-based Distributed Filtering of Continuous-Time Linear Systems. *IFAC World Congress 2020* to appear.

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

34. Mesbahi M, Egerstedt M. *Graph theoretic methods in multiagent networks*. Princeton University Press . 2010.

35. Orihuela L, Millán P, Vivas C, Rubio FR. Distributed control and estimation scheme with applications to process control. *IEEE Transactions on Control Systems Technology* 2014; 23(4): 1563–1570.

36. Kwakernaak H, Sivan R. *Linear optimal control systems*. 1. Wiley-interscience New York . 1972.

37. Kempe D, Dobra A, Gehrke J. Gossip-based computation of aggregate information. In: IEEE. ; 2003: 482–491.

38. Battilotti S, d'Angelo M. Stochastic output delay identification of discrete-time Gaussian systems. *Automatica* 2019; 109: 108499.

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

40. Battilotti S, Cacace F, d'Angelo M, Germani A. Distributed Kalman filtering over sensor networks with unknown random link failures. *IEEE Control Systems Letters* 2018; 2(4): 587–592.

41. Battilotti S, Cacace F, d'Angelo M, Germani A. The polynomial approach to the LQ non-Gaussian regulator problem through output injection. *IEEE Transactions on Automatic Control* 2018; 64(2): 538–552.

42. Cacace F, Conte F, d'Angelo M, Germani A. Feedback polynomial filtering and control of non-Gaussian linear time-varying systems. *Systems & Control Letters* 2019; 123: 108–115.

43. Cacace F, Conte F, d'Angelo M, Germani A. Filtering of systems with nonlinear measurements with an application to target tracking. *International Journal of Robust and Nonlinear Control* 2019; 29(14): 4956–4970.