Distributed Composite Control of Clustered Network ⋆

B. Adhikari [a,1], J. Veetaseveera [a,1], V. Varma [a], I.-C. Morarescu [a], E. Panteley [b]

[a] *Université de Lorraine, CNRS, CRAN, F-54000 Nancy, France*

[b] *L2S, CNRS, CentraleSupélec, Université Paris-Saclay, 91192 Gif-sur-Yvette, France.*

Abstract

We consider a clustered network in which connections inside the cluster are dense and between clusters are sparse. This leads us to a classical decoupling into fast (intra-cluster) dynamics and slow (inter-cluster) one. Our objective is to provide a computationally efficient method to design control strategies, that guarantee a certain bound on the cost for each cluster. Basically, we design a composite synchronizing controller with two terms: one responsible for the intra-cluster synchronization, and the other achieving the synchronization between clusters. The first one does not require much computational effort, since it is described by an analytic expression. The second term is designed through a satisfaction equilibrium approach. In other words, the internal (fast) and external (slow) controllers are independently designed, and they ensure a guaranteed satisfactory cost for each cluster. Moreover, we show that the internal control affects the cluster cost only for a short period of time. Finally, numerical simulations illustrate the theoretical results.

Key words: Synchronization; Distributed Control; Networked system; Singularly perturbation; Time-Scale modeling; Clustered network

1 Introduction

Due to its application in various disciplines such as physics, biology, economics, telecommunications, medicine, the study of synchronization of network interconnected systems has received significant attention in the literature. Examples of such networks include small-world networks [21], power systems [7], [17], wireless sensor networks [15], social networks [20], etc.

A major problem related to the synchronization of large-scale networks is the computational load associated with the design of effective controllers. In most of the existing literature, the cost related to the synchronization of the network is either not considered, or considered to be global. In [9], the authors take into account the global cost to design the energy-aware control that reduces the communication and computational load. The control design with optimal global cost in the framework of multi-agent systems is presented in [4], but the problem is NP-hard due to the information structure imposed by the graph. In [2], the authors presented the decentralized control design approach with global cost guarantees however, the assumption of the same gain for all the agents presented in the network is quite restrictive. A decentralized control design strategy that achieves the synchronization of the multi-agent system is proposed in [19] with individual performance guarantees. The proposed strategy works well with small-scale networks, however, for large-scale networks the computational effort required to obtain the gain is huge. Thus, the main goal of this paper is to provide an effective control design strategy for large-scale networks that reduces the computational effort while satisfying the performance guarantees.

One methodology to address the synchronization of the large-scale networks is based on model reduction. The objective is to decrease the size of the system state while approximating its overall dynamic behavior. One effective method to achieve such model reduction uses the *Singular Perturbation Theory* (*SPT*) that exploits the timescale properties of clustered networks. Classical results in [6], [7] develop a simplified model using SPT on the networks of linear interconnected systems with diffusive coupling. These results have been extended in [5] to networks of non-

⋆ This work was partially funded by CEFIPRA through the project "High-performance formation control in the presence of uncertainties and communication constraints" and by ANR under grants HANDY ANR-18-CE40-0010 and NICETWEET ANR-20-CE48-0009.

Email addresses: bikash.adhikari@univ-lorraine.fr (B. Adhikari), jomphop.veetaseveera@univ-lorraine.fr (J. Veetaseveera), vineeth.satheeskumar-varma@univ-lorraine.fr (V. Varma), constantin.morarescu@univ-lorraine.fr (I.-C. Morarescu), elena.panteley@centralesupelec.fr (E. Panteley).

[1] B. Adhikari and J. Veetaseveera contributed equally to this work as first authors.

linear systems interacting over fixed undirected networks. In [17], the authors relax the requirement of some assumptions in [6] and also extend the results for the weighted graphs. Furthermore, in [14], a model reduction using averaging theory is obtained for the time-varying directed graphs.

It is noteworthy that most of the existing works deal with the synchronization analysis of open-loop dynamics. In other words, they do not consider the control design for the synchronization of clustered networks. A particular setup for the synchronization of clustered networks is considered in [3]. However, none of the previously mentioned works consider the synchronization problem under the requirements of costs optimization. These requirements are on one hand timely, and on the other induce a high computational load, preventing the design of (sub-)optimal controllers in a centralized manner.

In this paper, we address the problem of distributed controller design while exploiting the network structure, and bound an associated cost at the same time. Our approach aims to significantly reduce the problem's complexity and the computational effort necessary to obtain the controller. We separate the control design and the cost optimization to the cluster-level.

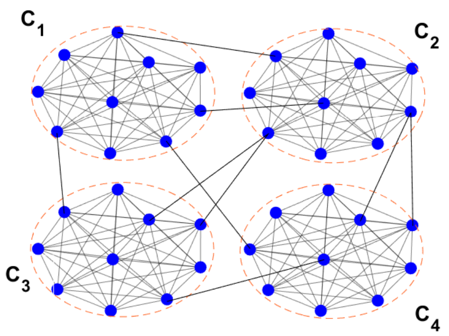

Fig. 1. A network partitioned into 4 clusters.

In a clustered network, the interconnections are dense inside the clusters and sparse between them. This results in a fast convergence inside the cluster towards a local agreement and then slowly towards the global consensus. The approach exploits this network property to divide the control design problem into computationally tractable sub-problems.

Provided that connections are much denser inside clusters than between clusters, we show that the network exhibits slow and fast dynamics that can be decoupled, via *Time-Scale Separation* (*TSS*) techniques. The fast variables represent the synchronization error inside the clusters, whereas the slow variables represent the aggregate behavior of the agent states within each cluster. The long-term behavior of the network depends on these slow dynamics. Decoupling these dynamics, we can approximate the behavior of the original system while independently designing the controllers for the intra and inter-clusters synchronization. For each agent present in the network, we define a composite control as a sum of the internal and external control. The two controls are designed independently at the cluster-level and employ a simplified model that significantly reduces the computational load and the control design complexity. The internal control, related to the fast dynamics, achieves the consensus inside the clusters while minimizing a local cost. As the connections are dense inside the clusters, the control design is described analytically assuming that clusters are characterized by all-to-all (complete topology) connections. This assumption is only made for the design purpose, and it is not required to be satisfied in practice. As for the external control, it synchronizes all the clusters employing a satisfaction equilibrium technique [19]. The only requirement for synchronization is the connectivity of the graph representing the inter-cluster network. Finally, applying the results from SPT [11], we show that the closed-loop response due to the composite control is close to that of the approximate models. In addition to the distributed control design, we also provide an approximation of the cluster cost after a short period of time required to synchronize the agents inside clusters.

The main contributions of this paper can be outlined as follows, (1) the control design is broken into independent designs of slow and fast controllers that aims at minimizing the intra-cluster and inter-cluster synchronization effort, (2) only one controller is designed per cluster for all the agents inside, irrespective of the number of agents in the cluster. This significantly reduces the computational effort and the complexity of the control design, (3) we propose sub-optimal control strategies as the internal cost is minimized while the external cost is satisfactory for each cluster, (4) a cost approximation is induced by the time-scale decoupling approach.

The remainder of the paper is organized as follows. The model with the objectives is stated in Section II. The time-scale modeling is described in detail in Section III. Then, the internal and external design procedures are developed in Section IV. In Section V, we provide an approximation of the cluster cost. Finally, numerical results are presented in Section VI before concluding in Section VII. To make the paper easily readable, some proofs are included in the Appendix.

1.1 Notation and Preliminaries

The symbol $\otimes$ represents the Kronecker product. Let $(x, y) \in \mathbb{R}^{n+m}$ stand for $[x^\top y^\top]^\top$. The identity matrix of size n is denoted by I_n and by $\mathbb{1}_n \in \mathbb{R}^n$, the column vector whose components are all 1. For a matrix $A \in \mathbb{R}^{m\times n}$, $A^\top$ denotes the transpose of A. For a vector $x \in \mathbb{R}^n$, we denote by $\|x\|_2 := \sqrt{x^\top x}$ its Euclidean norm and, for a matrix A, $\|A\|_2 := \sqrt{\lambda_{max}(A^\top A)}$. For a square matrix $M \in \mathbb{R}^{n\times n}$, let $\lambda_{min}(M)$ and $\lambda_{max}(M)$ be the minimum and the maximum eigenvalue, respectively. The *measure of the square matrix* M is defined as $\nu(M) = \frac{1}{2}\lambda_{max}(M + M^\top)$. We said that a matrix

$M \in \mathbb{R}^{n\times n}$ is orthonormal if $M^\top M = MM^\top = I_n$. We denote by $M_{-k} \in \mathbb{R}^{(n-1)\times(n-1)}$ the matrix M with its k-th row and column removed. By $B = \mathrm{diag}(B_1, ..., B_N)$, we denote a block-diagonal matrix with the entries B_1,...,B_N on the diagonal and $B_{-k} := \mathrm{diag}(B_1, ..., B_{k-1}, B_{k+1}, ..., B_N)$ the block-matrix with the k-th block removed. A function $f(t) : (0, \infty) \to \mathbb{R}^n$ is $\mathcal{O}(\epsilon)$ iff there exists constant $c > 0$ such that $\|f(t)\| \le c\epsilon$. Consider a connected, undirected graph $\mathcal{G} := (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} := \{1, 2, \ldots, n\}$ is the agent set and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the edge set. The adjacency matrix $\mathcal{A} = (a_{ij})_{n\times n}$ is defined as: $a_{ij} \neq 0$ if $(j, i) \in \mathcal{E}$ and $a_{ij} = 0$, otherwise. The Laplacian of the graph $\mathcal{G}$ is defined as $\mathcal{L}$, has $-a_{ij}$ off-diagonal elements and $\sum_{j=1}^{n} a_{ij}$ diagonal ones. Let $\mathbf{G} = (\mathcal{V}, \{\mathcal{K}_i\}_{i\in\mathcal{V}}, \{u_i\}_{i\in\mathcal{V}})$, be a strategic form game, where $\mathcal{V} = \{1, 2, \ldots, n\}$ is the set of players (agents), $\mathcal{K}_i$ is the set of strategies of player i , and u_i is an utility function of the player i and $\{f_1, \ldots, f_n\}$ be n set-valued satisfaction functions. Then the strategy profile $\mathcal{K}^* = (\mathcal{K}_1^*, \ldots, \mathcal{K}_n^*)$ is a *Satisfaction Equilibrium (SE)* if and only if for all $i \in \mathcal{V}$, we have, $\mathcal{K}_i^* \in f_i(\mathcal{K}_{-i}^*)$, where $\mathcal{K}_{-i}^* := (\mathcal{K}_1^*, \ldots, \mathcal{K}_{i-1}^*, \mathcal{K}_{i+1}^*, \ldots, \mathcal{K}_n^*)$ denotes the reduced profile with the component $\mathcal{K}_i^*$ removed.

2 Problem Statement

2.1 System Model

Consider a network of n agents partitioned into m non-empty clusters $\mathcal{C}_1, \ldots, \mathcal{C}_m \subset \mathcal{V}$. Clustered network refers to a network which is divided into distinct groups of agents having dense connection structure, whereas the connections between the clusters are sparse. Let us denote by $\mathcal{M} := \{1, 2, \ldots, m\}$, the set of clusters while n_k represents the cardinality of the cluster $\mathcal{C}_k$ such that $n = \sum_{k=1}^{m} n_k$. Each agent in the network is identified by a couple $(k, i) \in \mathcal{C}_k$, where k refers to the cluster $\mathcal{C}_k$ and i the index of the agent. The notation $(k, j) \in \mathcal{N}_{k,i}$ represents the neighbors of the agent (k, i) in the same cluster $\mathcal{C}_k$. To each agent $(k, i) \in \mathcal{C}_k$, $k \in \mathcal{M}$, one assigns a state $x_{k,i} \in \mathbb{R}^{n_x}$ whose dynamics is

$$\dot{x}_{k,i} = Ax_{k,i} + Bu_{k,i}, \tag{1}$$

where $u_{k,i} \in \mathbb{R}^{n_u}$, $A \in \mathbb{R}^{n_x\times n_x}$ and $B \in \mathbb{R}^{n_x\times n_u}$.

For each cluster $\mathcal{C}_k$, let $x_k := (x_{k,1}, \ldots, x_{k,n_k}) \in \mathbb{R}^{n_k\cdot n_x}$ be the cluster state and $u_k := (u_{k,1}, \ldots, u_{k,n_k}) \in \mathbb{R}^{n_k\cdot n_u}$ the cluster control. Thus, the cluster dynamics takes the following form

$$\dot{x}_k = (I_{n_k} \otimes A)x_k + (I_{n_k} \otimes B)u_k, \quad \forall k \in \mathcal{M}. \tag{2}$$

Moreover, in the presence of clusters, the Laplacian of the network can be written as $\mathcal{L} := \mathcal{L}^{int} + \mathcal{L}^{ext}$. The internal Laplacian of the network $\mathcal{L}^{int} := \mathrm{diag}(\mathcal{L}_1^{int}, ..., \mathcal{L}_m^{int})$ and the external Laplacian of the network $\mathcal{L}^{ext}$ represents the connections between agents from different clusters.

2.2 Objective

We consider the problem of network synchronization, and the network is said to be asymptotically synchronized when for all $(k, i) \in \mathcal{C}_k$, $(l, j) \in \mathcal{C}_l$ and $k, l \in \mathcal{M}$, $\lim\limits_{t\to+\infty} \|x_{k,i}(t) - x_{l,j}(t)\| = 0$. If the network dimension is large, the problem of designing a control under certain cost constraints is difficult to tackle. Hence, for each cluster $k \in \mathcal{C}_k$, we propose a composite control of the form:

$$u_k := u_k^{int} + u_k^{ext}, \quad \forall k \in \mathcal{M}, \tag{3}$$

where $u_k^{int} := (u_{k,1}^{int}, ..., u_{k,n_k}^{int})$, $u_k^{ext} := (u_{k,1}^{ext}, ..., u_{k,n_k}^{ext})$ and

$$\begin{cases} u_{k,i}^{int} := -K_k^{int} \sum_{(k,j)\in\mathcal{N}_{k,i}} (x_{k,i} - x_{k,j}), \\ u_{k,i}^{ext} := -K_k^{ext} \sum_{(l,p)\in\mathcal{N}_{k,i}} (x_{k,i} - x_{l,p}). \end{cases} \tag{4}$$

The notation $(l, p) \in \mathcal{N}_{k,i}$ indicates the neighbors belonging to a different cluster, with $(l, p) \neq (k, i)$.

The main objective is to design a distributed control synchronizing the network while minimizing the cost. As the network size increases, the problem becomes too complex. Thus, we split the total cost into the sum of cost of clusters and minimize cost for each cluster. The cluster cost, J_k associated with each cluster $\mathcal{C}_k$, $k \in \mathcal{M}$, is defined as

$$\begin{aligned} J_k = \int_0^{+\infty} & x_k^\top(t)(\mathcal{L}_k^{int} \otimes I_{n_x})x_k(t) + x^\top(t)(\mathcal{L}_k^{ext} \otimes I_{n_x})x(t) \\ & + u_k^\top(t)(I_{n_k} \otimes R_k)u_k(t)\, \mathrm{dt}, \end{aligned} \tag{5}$$

where the internal Laplacian $\mathcal{L}_k^{int} \in \mathbb{R}^{n_k\times n_k}$ captures the connections inside $\mathcal{C}_k$, and the external Laplacian $\mathcal{L}_k^{ext} \in \mathbb{R}^{n\times n}$ expresses the external connections between $\mathcal{C}_k$ and the neighboring clusters.

By substituting the individual control (4), we recast the cost (5) as the sum of internal, external and a cross term as

$$\begin{aligned} J_k = & \underbrace{\int_0^{+\infty} x_k^\top(t)(\mathcal{L}_k^{int} \otimes I_{n_x})x_k(t) + {u_k^{int}}^\top(t)(I_{n_k} \otimes R_k)u_k^{int}(t)\, \mathrm{dt}}_{J_k^{int}} \\ & + \underbrace{\int_0^{+\infty} x^\top(t)(\mathcal{L}_k^{ext} \otimes I_{n_x})x(t) + {u_k^{ext}}^\top(t)(I_{n_k} \otimes R_k)u_k^{ext}(t)\, \mathrm{dt}}_{J_k^{ext}} \\ & \underbrace{+2 \int_0^{+\infty} {u_k^{ext}}^\top(t)(I_{n_k} \otimes R_k)u_k^{int}(t)\, \mathrm{dt}\,.}_{J_k^{cross}} \end{aligned} \tag{6}$$

The internal control u_k^{int} is the effort required to achieve local agreement, whereas the external control u_k^{ext} is the energy necessary to synchronize the agents between the clusters.

The second objective is to bound the total cluster cost with the sum of internal (J_k^{int}) and external (J_k^{ext}) cost and a constant term. In this way, we replace the original problem of optimizing the cost function with the problem of optimization of the internal and the external cost. So, the problem is to find the parameters of equation (4) i.e., K_k^{int} and K_k^{ext}, such that the internal and the external cost are minimized or bounded by a desired value. The cross term in the equation (6) can be bounded by a constant term multiplied by the norm of the initial conditions (See Theorem 2). To solve this problem, we propose an approach based on timescale separation that we describe in the following section.

3 Time Scale Modeling

In this section, we provide a procedure to decouple the closed-loop network dynamics into two subsystems, evolving on different time scales. First, we perform a coordinate transformation to exhibit the collective dynamics of the network: the average and the synchronization error dynamics. Then, we apply the TSS to decouple the collective dynamics into slow and fast subsystems. In two timescale, the slow variable corresponds to the average while the fast variable to the synchronization error.

3.1 Coordinate transformation

Let us consider the clusters without external connections. Then based on the internal Laplacian $\mathcal{L}_k^{int}$, following from [16], [1], we introduce the coordinate transformation for the cluster $\mathcal{C}_k$. For a connected graph, the Jordan decomposition of the symmetric Laplacian matrix is,

$$\mathcal{L}_k^{int} = T_k \begin{bmatrix} 0 & 0 \\ 0 & \Lambda_k^{int} \end{bmatrix} T_k^\top, \quad \forall k \in \mathcal{M}, \tag{7}$$

where T_k is a orthonormal matrix and $\Lambda_k^{int} = \text{diag}(\lambda_{k,2}^{int}, \dots, \lambda_{k,n_k}^{int})$ collects the $n_k - 1$ eigenvalues of $\mathcal{L}_k^{int}$. Moreover, the matrix T_k can be expressed as

$$T_k = \begin{bmatrix} v_{k,1} & V_k \end{bmatrix}, \quad \forall k \in \mathcal{M}, \tag{8}$$

where $v_{k,1}^\top = \frac{1}{\sqrt{n_k}} \mathbb{1}_{n_k}^\top$ is the eigenvector associated with the 0 eigenvalue and the matrix $V_k \in \mathbb{R}^{n_k \times (n_k - 1)}$ is constituted of the eigenvectors corresponding to the nonzero eigenvalues of $\mathcal{L}_k^{int}$. Furthermore, it can be verified that, $v_{k,1}^\top V_k = 0$ and $V_k^\top V_k = I_{n_k - 1}$.

Now, we define the coordinate transformation as

$$\bar{x}_k := \begin{bmatrix} y_k \\ \xi_k \end{bmatrix} = (\frac{1}{\sqrt{n_k}} T_k^\top \otimes I_{n_x}) x_k, \quad \forall k \in \mathcal{M}. \tag{9}$$

Then, from (8) and (9), the change of variables yields, for all $k \in \mathcal{M}$,

$$y_k := \left(\frac{\mathbb{1}_{n_k}^\top}{n_k} \otimes I_{n_x} \right) x_k =: H_k x_k \in \mathbb{R}^{n_x} \tag{10}$$

$$\xi_k := \left(\frac{V_k^\top}{\sqrt{n_k}} \otimes I_{n_x} \right) x_k =: Z_k x_k \in \mathbb{R}^{(n_k - 1).n_x}. \tag{11}$$

The first component y_k corresponds to the average of the respective agents' states in the cluster $\mathcal{C}_k$. The second component, ξ_k corresponds to the synchronization error. Since the matrix T_k orthonormal i.e., $T_k^\top = T_k^{-1}$, the inverse of the transformation (9) yields

$$\begin{aligned} x_k &= (\sqrt{n_k} T_k \otimes I_{n_x}) \bar{x}_k (\mathbb{1}_{n_k} \otimes I_{n_x}) y_k + (\sqrt{n_k} V_k \otimes I_{n_x}) \xi_k \\ &=: \widetilde{H}_k y_k + \widetilde{Z}_k \xi_k, \quad \forall k \in \mathcal{M}. \end{aligned} \tag{12}$$

Let us define the vectors collecting the states, the average and the error as $x: = (x_1, \dots, x_m) \in \mathbb{R}^{n.n_x}$, $y: = (y_1, \dots, y_m) \in \mathbb{R}^{m.n_x}$ and $\xi: = (\xi_1, \dots, \xi_m) \in \mathbb{R}^{(n-m).n_x}$, respectively. Then for overall network, we obtain

$$y = Hx, \quad \xi = Zx \quad \text{and} \quad x = \widetilde{H} y + \widetilde{Z} \xi, \tag{13}$$

where $H = \text{diag}(H_1, ..., H_m)$ and $Z = \text{diag}(Z_1, ..., Z_m)$ and $\widetilde{H} = \text{diag}(\widetilde{H}_1, ..., \widetilde{H}_m)$ and $\widetilde{Z} = \text{diag}(\widetilde{Z}_1, ..., \widetilde{Z}_m)$. Now, we recast the overall network dynamics in terms of the new coordinate variables. The overall network dynamics in the presence of the control (3) is

$$\begin{aligned} \dot{x} = ((I_n \otimes A) &- (I_n \otimes B) K^{int} (\mathcal{L}^{int} \otimes I_{n_x}) \\ &- (I_n \otimes B) K^{ext} (\mathcal{L}^{ext} \otimes I_{n_x})) x, \end{aligned} \tag{14}$$

where $K^{int} = \text{diag}((I_{n_1} \otimes K_1^{int}), \dots, (I_{n_m} \otimes K_m^{int}))$ and $K^{ext} = \text{diag}((I_{n_1} \otimes K_1^{ext}), \dots, (I_{n_m} \otimes K_m^{ext}))$. Then, using (13) and (14), the overall dynamics is recast in new coordinates as follows,

$$\begin{cases} \dot{y} &= \bar{A}_{11} y + \bar{A}_{12} \xi, \\ \dot{\xi} &= \bar{A}_{21} y + (\bar{A}_{22}^1 + \bar{A}_{22}^2) \xi, \end{cases} \tag{15}$$

where

$$\begin{cases} \bar{A}_{11} &= ((I_m \otimes A) - H(I_n \otimes B)K^{ext}(\mathcal{L}^{ext} \otimes I_{n_x})\widetilde{H}), \\ \bar{A}_{12} &= -H(I_n \otimes B)K^{ext}(\mathcal{L}^{ext} \otimes I_{n_x})\widetilde{Z}, \\ \bar{A}_{21} &= -Z(I_n \otimes B)K^{ext}(\mathcal{L}^{ext} \otimes I_{n_x})\widetilde{H}, \\ \bar{A}_{22}^1 &= -Z(I_n \otimes B)K^{ext}(\mathcal{L}^{ext} \otimes I_{n_x})\widetilde{Z}, \\ \bar{A}_{22}^2 &= ((I_{n-m} \otimes A) - (I_n \otimes B)K_{n-m}^{int}(\Lambda^{int} \otimes I_{n_x})), \end{cases} \tag{16}$$

and $K_{n-m}^{int} = \text{diag}((I_{n_1-1} \otimes K_1^{int}), \dots, (I_{n_m-1} \otimes K_m^{int}))$ and $\Lambda^{int} = \text{diag}(\Lambda_1^{int}, \dots, \Lambda_m^{int})$.

3.2 Two Time-Scaled Network Dynamics

To study the time-scale behavior and analyze the synchronizing behavior, we define the network parameters as follows

$$\begin{cases} \mu^{ext}: &= \|(I_n \otimes B)K^{ext}(\mathcal{L}^{ext} \otimes I_{n_x})\|, \\ \mu^{int}: &= \min_{k\in\mathcal{M}} \|(\Lambda_k^{int} \otimes BK_k^{int})\|, \\ \epsilon: &= \frac{\mu^{ext}}{\mu^{int}}. \end{cases} \tag{17}$$

Remark 1 *For the rest of this section, we assume that ϵ is small such that time-scale separation occurs and this will be enforced during the control design in the next section.*

The network parameter ϵ is the ratio of the strength of the controls between and within the clusters. It's worth noting, in our case, the network parameter ϵ can be tuned by the choice of the control gains.

In the absence of agents' internal dynamics, the authors in [6], [14], are able to express the closed-loop dynamics (14) into a *Standard Singular Perturbation Form* (*SSPF*). Upon analyzing the orders of the state matrices of the collective dynamics (15), it appears that the closed-loop network dynamics cannot be expressed in SSPF without the knowledge of the state matrix A. Thus we make the following assumption in the order of the matrix A.

Assumption 1 *The state matrix A satisfy the following*

$$\|A\| \leq \mathcal{O}(\mu^{ext}).$$

However, we note that since μ^{ext} depends on K^{ext}, we can always choose K^{ext} sufficiently large such that the assumption 1 is satisfied. The assumption on the order of the matrix A is necessary for representing the dynamics (15) in SSPF. In the following lemma, we analyze the order of the matrices in equation (16) under assumption on the matrix A.

Lemma 1 *Under Assumption 1, the matrices in* (16) *satisfy the following conditions,*

- $\|\bar{A}_{11}\|$, $\|\bar{A}_{12}\|$, $\|\bar{A}_{21}\|$, $\|\bar{A}_{22}^1\|$ *are of order* $\mathcal{O}(\epsilon\mu^{int})$
- $\|\bar{A}_{22}^2\|$ *is of order* $\mathcal{O}(\mu^{int})$

PROOF. See appendix A. ■

As a consequence of the Assumption 1, all the matrices in (16) are of order $\mathcal{O}(\epsilon\mu^{int})$ except A_{22}^2, which is of order $\mathcal{O}(\mu^{int})$. This implies that the matrix A_{22}^2 has larger eigenvalues compared to the remaining matrices in the equation (15) which can also be observed with respect to the definition of the ϵ in equation (17) since $\mu^{int} >> \mu^{ext}$. Since the dynamics of the variable y and ξ are dominated by the matrix $\bar{A}_{11}$ and $\bar{A}_{22}$, the y and ξ behaves as a slow and fast variable, respectively.

Now, to reveal the TSS, we rescale the time with μ^{int} to obtain a fast time-scale as $t_f = \mu^{int}t$, and a slow time-scale $t_s = \epsilon t_f$. This allows us to represent the overall dynamics (15) in SSPF. Then using the tools from singular perturbation theory we can decouple the overall dynamics in SSPF into reduce-ordered slow and fast subsystems. The matrices $\bar{A}_{ij}$ in (16) are re-scaled as follows,

$$A_{11} = \frac{\bar{A}_{11}}{\epsilon\mu^{int}}, \qquad A_{12} = \frac{\bar{A}_{12}}{\epsilon\mu^{int}}, \qquad A_{21} = \frac{\bar{A}_{21}}{\epsilon\mu^{int}},$$
$$A_{22}^1 = \frac{\bar{A}_{22}^1}{\epsilon\mu^{int}}, \qquad A_{22}^2 = \frac{\bar{A}_{22}^2}{\mu^{int}}. \tag{18}$$

In the following, we decouple the original dynamics in into slow and fast subsystems using the TSS.

3.3 Fast Dynamics

Performing the time rescale $t_f = \mu^{int}t$, and from equations (15), (17) and (18) we obtain the dynamics in fast time-scale as follows,

$$\begin{cases} \frac{d\hat{y}}{dt_f}(t_f) &= \epsilon A_{11}\hat{y}(t_f) + \epsilon A_{12}\hat{\xi}(t_f), \\ \frac{d\hat{\xi}}{dt_f}(t_f) &= \epsilon A_{21}\hat{y}(t_f) + (\epsilon A_{22}^1 + A_{22}^2)\hat{\xi}(t_f). \end{cases} \tag{19}$$

Then, setting $\epsilon = 0$, we have $\frac{d\hat{y}}{dt_f} = 0$ implying $\hat{y}$ is constant and the fast dynamics is

$$\frac{d\hat{\xi}_f}{dt_f}(t_f) = A_{22}^2\hat{\xi}_f(t_f). \tag{20}$$

The fast dynamics (20) in original time-scale t is

$$\dot{\xi}_f(t) = (I_{n-m} \otimes A)\xi_f(t) + (I_{n-m} \otimes B)u_f(t), \tag{21}$$

where $u_f(t) = -K_{n-m}^{int}(\Lambda^{int} \otimes I_{n_x})\xi_f(t)$. The fast dynamics (21) corresponds to the intra-cluster dynamics and hence the dynamics is dominated by the internal gain and the eigenvalues of the intra-cluster Laplacian. With the suitable choice of the internal gain K_k^{int}, the system (21) is exponentially stable.

3.4 Slow Dynamics

The overall dynamics in slow time-scale $t_s = \epsilon t_f$ is

$$\begin{cases} \frac{d\tilde{y}}{dt_s}(t_s) &= A_{11}\tilde{y}(t_s) + A_{12}\tilde{\xi}(t_s), \\ \epsilon\frac{d\tilde{\xi}}{dt_s}(t_s) &= \epsilon A_{21}\tilde{y}(t_s) + (\epsilon A_{22}^1 + A_{22}^2)\tilde{\xi}(t_s). \end{cases} \tag{22}$$

Setting $\epsilon = 0$ in (22) yields $\tilde{\xi}_s(t_s) = 0$ and the slow dynamics is,

$$\frac{d\tilde{y}_s}{dt_s}(t_s) = A_{11}\tilde{y}_s(t_s). \tag{23}$$

Since we have $t_s = \epsilon t_f = \epsilon\mu^{int}t$, the slow dynamics is

$$\dot{y}_s(t) = (I_m \otimes A)y_s(t) + (I_m \otimes B)u_s(t), \tag{24}$$

where $u_s(t) = -HK^{ext}(\mathcal{L}^{ext} \otimes I_{n_x})\widetilde{H}y_s(t)$. The slow dynamics (24) represents the collective behavior of the cluster which may or may not be stable.

Note that for a singular perturbation based analysis of the overall network dynamics, we re-scale the dynamics into slow time-scale (t_s) in (22), and fast time-scale (t_f) in (19), to obtain the slow (23) and fast dynamics (20). However, since the control design is done in the original time-scale, we change the slow and fast dynamics to original time-scale as in equation (24) and (21), respectively in the next section. This is possible because the transformations are invertible and it can be verified by the definition of t_f and t_s.

3.5 Singular Perturbation Approximation

Under Assumption 1, the closed-loop network dynamics (**??**) is reformulated into the network dynamics in new coordinates (15) namely, the average and error dynamics. Afterwards, the average and error dynamics (15) are decoupled into the slow (24) and fast dynamics (21) using TSS and SPT.

Now, we show that the fast (20) and slow (23) dynamics are an approximation of the synchronization error and average dynamics (15), respectively. But before stating the result we make the following assumption on the existence of the control gains.

Assumption 2 *There exists an internal gain K^{int} and an external gain K^{ext} such that the slow dynamics (24) is synchronized and the fast dynamics (21) is stabilized.*

Remark 2 *Although, we assume the existence of the synchronizing internal and external gain, it will be ensured by design in the next section that such gains exist. The internal and external gains are designed independently, and the obtained internal gain is optimal while the external gain is sub-optimal.*

Then, under Assumption 1 and 2 , we provide an approximation of the original system by the reduced-order subsystems in the following theorem. The proof follows from Theorem 5.1, Chapter 2, [11].

Theorem 1 *Under the Assumption 1, if Re $\lambda(A_{22}^2) < 0$, there exists an $\epsilon^* > 0$ such that, for all $\epsilon \in (0, \epsilon^*]$, the original system (15) starting from any bounded initial conditions $y(t_0)$ and $\xi(t_0)$, is approximated for all finite time $t \geq t_0$ by*

$$\begin{cases} y(t) &= y_s(t) + \epsilon\Psi(\epsilon)\xi_f(\mu^{int}t), \\ \xi(t) &= \xi_f(\mu^{int}t) - \Omega(\epsilon)y_s(t) - \epsilon\Omega(\epsilon)\Psi(\epsilon)\xi_f(\mu^{int}t), \end{cases} \tag{25}$$

where $y_s(t) \in \mathbb{R}^{m.n_x}$ and $\xi_f(\mu^{int}t) \in \mathbb{R}^{(n-m).n_x}$ are the respective the slow variable (24) and the fast variable (20). The terms $\epsilon\Psi(\epsilon)\xi_f(\mu^{int}t)$ and $\Omega(\epsilon)y_s(t) + \epsilon\Omega(\epsilon)\Psi(\epsilon)\xi_f(\mu^{int}t)$ are of order $\mathcal{O}(\epsilon)$. The approximation of the functions Ω and Ψ are,

$$\begin{aligned} \Omega(\epsilon) &= \epsilon(A_{22}^2)^{-1}A_{21} + \mathcal{O}(\epsilon^2), \\ \Psi(\epsilon) &= A_{12}(A_{22}^2)^{-1} + \epsilon((A_{22}^2)^{-1}A_{11}A_{12}(A_{22}^2)^{-1} - A_{12}) \\ &\quad + \mathcal{O}(\epsilon^2). \end{aligned} \tag{26}$$

PROOF. See Appendix. ■

In the next section, we discuss in detail the procedure to design the gains to stabilize the fast subsystems and synchronize the slow subsystems.

4 Design procedure

In this section, we give a controller design strategy for system (15), that provides internal and external control independent of each other. We use the idea of time scale separation and split the design procedure into two independent parts. First, based on the fast dynamics (21) we design an internal control using the local information that ensures the asymptotic synchronization inside the cluster. Roughly speaking, after the phase of fast dynamics, all agents in the cluster synchronize quickly and the behavior of any node can be approximated by a single node. On the other hand, the interactions between agents in different clusters can be approximated by the slow dynamics. We thus use the slow dynamics in (24) for the design of external control to achieve the synchronization between the clusters. While the internal controller is optimal, the external control is designed by bounding the cost under a given threshold. Finally, Theorem 1 is used to justify such a separation of the system analysis in two steps.

The two next assumptions aim to model the graph structure that we use to choose gain. To ensure the synchronization of the entire network, we assume that no cluster is isolated.

Assumption 3 *The graph of clusters is connected.*

Due to the dense communications between the agents inside the clusters, we make the following assumption on the communications inside the cluster.

Assumption 4 *The internal graphs are complete for all clusters.*

This assumption is imposed only for the control design purpose, and the obtained controller can be implemented even if the Assumption 4 does not hold. The main advantage of this assumption is that it greatly reduces the computational effort required to obtain the control; it allows us to decouple the control design into agents' level and to obtain an analytical expression.

Remark 3 *Under Assumption 4, the zero eigenvalue of the internal Laplacian $\mathcal{L}_k^{int}$ is simple and all the non-zero eigenvalues are n_k for $k \in \mathcal{M}$.*

In the following, we first address the internal control by giving an analytical gain expression under Assumption 4. The key idea is to break down the control design from the cluster's level to the agent's level.

4.1 Internal (Fast) Control Design

As the fast variable ξ_f is an approximation of the synchronization error ξ inside the clusters, it is still relevant to consider the fast subsystems (21) for the internal control design. We denote by $\xi_{f,k} \in \mathbb{R}^{(n_k-1).n_x}$ the component of $\xi_f := (\xi_{f,1}, ..., \xi_{f,m})$ corresponding to the k-th cluster. For each cluster $\mathcal{C}_k$, for $k \in \mathcal{M}$, we have the following dynamics

$$\begin{cases} \dot{\xi}_{f,k}(t) &= (I_{n_k-1} \otimes A)\xi_{f,k}(t) + (I_{n_k-1} \otimes B)u_{f,k}(t), \\ u_{f,k}(t) &= -(\Lambda_k^{int} \otimes K_k^{int})\xi_{f,k}(t). \end{cases} \tag{27}$$

The cluster cost associated with the cluster $\mathcal{C}_k$ takes the form

$$J_{f,k} = \int_0^{+\infty} \xi_{f,k}^\top (\Lambda_k^{int} \otimes I_{n_x})\xi_{f,k} + u_{f,k}^\top (I_{n_k-1} \otimes R_k)u_{f,k}\, \mathrm{d}t\,. \tag{28}$$

Furthermore, the matrices in equations (27) and (28) have block-diagonal form, thus they can decoupled into $n_k - 1$ independent subsystems. For each cluster $\mathcal{C}_k$, similarly to the ξ_k defined in equation (11), let us denote the fast subsystems and the associated control by $\xi_{f,k} := (\xi_{f,k,1}, ..., \xi_{f,k,n_k-1})$ and $u_{f,k} := (u_{f,k,1}, ..., u_{f,k,n_k-1})$, respectively. Then, for $i = 1, ..., n_k - 1$ and for all $k \in \mathcal{M}$, the dynamics are

$$\begin{cases} \dot{\xi}_{f,k,i}(t) &= A\xi_{f,k,i}(t) + n_k B u_{f,k,i}(t), \\ u_{f,k,i}(t) &= -K_k^{int}\xi_{f,k,i}(t), \end{cases} \tag{29}$$

and the associated individual cost is

$$J_{f,k,i} = \int_0^{+\infty} n_k \xi_{f,k,i}^\top \xi_{f,k,i} + n_k^2 u_{f,k,i}^\top R_k u_{f,k,i}\, \mathrm{d}t\,. \tag{30}$$

Thus, the cost (28) can be expressed as the sum of individual costs (30) as follows, $J_{f,k} = \sum_{i=1}^{n_k-1} J_{f,k,i}, \quad \forall k \in \mathcal{M}$.

Remark 4 *The decoupling of dynamics* (27) *into $n_k - 1$ subsystems* (29) *is not only limited to all-to-all connections. In the case, where we know the eigenvalues of the Laplacian or the Laplacian eigenvalues can be characterized in terms of n_k (for example, star graph), similar decoupling can be achieved.*

Remark 5 *It is noteworthy that the gain K_k^{int} is same for all the agents belonging to the same cluster $\mathcal{C}_k$. As a result, the rewriting of* (28) *as a sum of individual cost* (30) *reduces the computational effort for the control design. Indeed, one can solve only one optimization problem* (29)-(30) *for each cluster, it is equivalent to optimizing the cluster cost* (28).

Next we show that the system (29) is stabilizable with a simple linear controller while we recall that the system (29) corresponds to fast dynamics of our original system. Finally, we apply the LQ-control [10] to stabilize (29) while minimizing the cost (30).

Lemma 2 *Consider the system* (29), *under assumption 3, if the pair (A, B) is stabilizable and $(A, (R_k)^{1/2})$ is detectable, then for every $k \in \mathcal{M}$, the system* (29) *is stabilizable while minimizing the cost* (30) *by a controller $u_{f,k,i}(t) = -K_k^{int}\xi_{f,k,i}(t)$ with the gain*

$$K_k^{int} = \frac{R_k^{-1}}{n_k} B^\top P_k^{int}, \quad k \in \mathcal{M}, \tag{31}$$

where P_k^{int} is the solution of the Algebraic Riccati Equation (ARE)

$$P_k^{int} A + A^\top P_k^{int} - P_k^{int} B R_k^{-1} B^\top P_k^{int} + n_k I_{n_x} = 0. \tag{32}$$

It follows from Lemma 2 , the fast dynamics (29) is exponentially stable i.e., $\xi_f(t) \to 0$ as $t \to \infty$ and we pass to the design of the external controller.

4.2 External (Slow) Control Design

In this sub-section we present the external controller design based on the slow dynamics (24). In general, the results present in the literature consider the leader-follower architecture where all the agents present in a cluster communicate with the leader, and the communication between the clusters is done through leaders only. However, in our case, we consider a more general form of communication architecture where any agent in one cluster can communicate with any other agent in a different cluster, as shown in Figure 1.

To achieve the synchronization between the clusters, we propose a method based on [19]. First, we define the average slow variable. Then, the synchronization problem is transformed to a stabilization problem using a change of variable. Finally, we design the control to stabilize the system while upper bounding an associated cost.

Recall that after the clusters are synchronized, the agents in each cluster behave like a single node, and the number of nodes representing the external network equals the number of clusters. Thus, the external graph of agents that is connecting the clusters is only connected and hence the standard optimization or the optimal control approaches cannot be applied directly. In this context, inspired by the notion in game theory, we use the satisfaction equilibrium approach and satisfaction games [18]. A set of actions are said to be in satisfaction equilibrium when the individual cost for each agent is upper-bounded by a given threshold.

4.2.1 Average slow variables

The slow dynamics obtained after time-scale separation in equation (24) defines the dynamics of the average of each cluster. Following from equation (24), the average dynamics can be written as

$$\dot{y}_s(t) = ((I_m \otimes A) - (I_m \otimes B)\overline{K}^{ext}(\overline{\mathcal{L}}^{ext} \otimes I_{n_x}))y_s(t), \tag{33}$$

where $\overline{K}^{ext} = \text{diag}\,(K_1^{ext}, ..., K_m^{ext})$ is the external gain and $(\overline{\mathcal{L}}^{ext} \otimes I_{n_x}) = H(\mathcal{L}^{ext} \otimes I_{n_x})\widetilde{H}$ with the following form

$$\overline{\mathcal{L}}^{ext} = \begin{pmatrix} \sum_{l=2}^{m} \frac{a_{1l}^{ext}}{n_1} & -\frac{a_{12}^{ext}}{n_1} & \cdots & -\frac{a_{1m}^{ext}}{n_1} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{a_{m1}^{ext}}{n_m} & -\frac{a_{m2}^{ext}}{n_m} & \cdots & \sum_{l=1}^{m-1} \frac{a_{ml}^{ext}}{n_m} \end{pmatrix} \in \mathbb{R}^{m\times m},$$

is the average Laplacian matrix related to (33). In average Laplacian $\overline{\mathcal{L}}^{ext}$ the diagonal elements represents the total number of external connections from a cluster $k \in \mathcal{M}$ to the rest of the network and the non-diagonal entries a_{kl}^{ext} represents the total number of connections between cluster $\mathcal{C}_k$ and $\mathcal{C}_l$.

Let us denote by $y_{s,k} \in \mathbb{R}^{n_x}$ the k-th component of the variable y_s. Then, the average dynamics of each cluster $\mathcal{C}_k$, for $k \in \mathcal{M}$, based on equation (33) is

$$\begin{cases} \dot{y}_{s,k} &= Ay_{s,k} + Bu_{s,k}^{ext}, \\ u_{s,k}^{ext} &= -K_k^{ext} \sum\limits_{l\in\mathcal{N}_{C_k}} \frac{a_{kl}^{ext}}{n_k}(y_{s,k} - y_{s,l}) \end{cases} \tag{34}$$

where $u_{s,k}^{ext}$ can be viewed as the control on the cluster level, since it represents the sum of the individual controllers. For system (34) we define the average cost for each cluster $\mathcal{C}_k, k \in \mathcal{M}$, as

$$\overline{J}_k^{ext} = \int_0^{+\infty} \sum_{l\in\mathcal{N}_{C_k}} \frac{a_{kl}^{ext}}{n_k}(y_{s,k} - y_{s,l})^2 + n_k \sum_{i=1}^{n_k} \widehat{u}_{k,i}^{ext\top} R_k \widehat{u}_{k,i}^{ext}\, \mathrm{d}t \tag{35}$$

where

$$\widehat{u}_{k,i}^{ext} := -K_k^{ext} \sum_{l\in\mathcal{N}_{C_k}} \frac{a_{(k,i)\leftrightarrow C_l}^{ext}}{n_k}(y_{s,k} - y_{s,l}) \quad \forall i \in C_k,$$

and $a_{(k,i)\leftrightarrow\mathcal{C}_l}^{ext}$ is the total number of connections between the i-th agent belonging to $\mathcal{C}_k$ and the cluster $\mathcal{C}_l$ and clearly $a_{(k,i)\leftrightarrow\mathcal{C}_l}^{ext} \le n_k$. The control $\widehat{u}_{k,i}^{ext}$ is the external control (4) expressed in the average variable y_s. In addition, we have the relation $u_{s,k}^{ext} = \sum_{i=1}^{n_k} \widehat{u}_{k,i}^{ext}$ and $a_{kl}^{ext} = \sum_{i=1}^{n_k} a_{(k,i)\leftrightarrow\mathcal{C}_l}^{ext}$.

Notice that the average cost (35) is different from the external cost function that appears in equation (6) in several ways:

- the average variable $y_{s,k}$ is used instead of the original state variables x_k for each cluster, and
- the cost due to the control effort is expressed in terms of the variable $\widehat{u}_{k,i}^{ext}$, for all $i \in \mathcal{C}_k$, $k \in \mathcal{M}$, which is the individual control (4) expressed in terms of the average variables $y_{s,k}$ for all $k \in \mathcal{M}$, instead of the average control $u_{s,k}^{ext}$. It would be possible to define the cost function as

$$\overline{J}_k^{ext} = \int_0^{+\infty} \sum_{l\in\mathcal{N}_{C_k}} \frac{a_{kl}^{ext}}{n_k}(y_{s,k} - y_{s,l})^2 + u_{s,k}^{ext\top} R_k u_{s,k}^{ext}\, \mathrm{d}t, \tag{36}$$

however, we remark that optimization of the average control $u_{s,k}^{ext}$ does not necessarily imply optimization of individual control $u_{k,i}^{ext}$. Although the clusters have merged into single node, the agents still apply the individual control (4) rather than the average control (34). Thus we express the individual external control (4) in average variables y_s and define the average cost (35) in term of the original control.

In the following, we perform the change of variables to design an external gain synchronizing the network of clusters.

4.2.2 Change of Variables

To study the consensus between the clusters, we define the external error variable for each cluster $\mathcal{C}_k$. Let us define $Y_{k,l} = (y_{s,l} - y_{s,k}), l \neq k$. Then the external error variable for each cluster $\mathcal{C}_k$, $k \in \mathcal{M}$ is defined as

$$Y_k := \Big(Y_{k,1}, \ldots, Y_{k,k-1}, Y_{k,k+1}, \ldots, Y_{k,m}\Big) \in \mathbb{R}^{(m-1).n_x}. \tag{37}$$

Then, based on equation (37), the corresponding external error dynamics is

$$\dot{Y}_k = \mathbf{A}_k Y_k + \mathbf{B}_k u^{ext}_{s,k}, \quad \forall k \in \mathcal{M},$$

where,

$$\begin{aligned}\mathbf{A}_k &= (I_{m-1} \otimes A) - (I_{m-1} \otimes B)\overline{K}^{ext}_{-k}(\overline{\mathcal{L}}^{ext}_{-k} \otimes I_{n_x}),\\ \mathbf{B}_k &= -(\mathbb{1}_{m-1} \otimes B).\end{aligned} \tag{38}$$

Here, $\overline{K}^{ext}_{-k} = \mathrm{diag}(K^{ext}_1, \ldots, K_{k-1}, K_{k+1}, \ldots, K_m)$ is not a control action, but it represents the behavior of the network.

Now to represent the average cost function (35) in terms of the error variable $Y_k, k \in \mathcal{M}$, we study the structure of the external Laplacian $\mathcal{L}^{ext}$. Upon close inspection, we see that it has the following form,

$$\mathcal{L}^{ext} = \begin{pmatrix} \mathcal{L}^{ext}_{1,1} & \mathcal{L}^{ext}_{1,2} & \ldots & \mathcal{L}^{ext}_{1,m} \\ \mathcal{L}^{ext}_{2,1} & \mathcal{L}^{ext}_{2,2} & \ldots & \mathcal{L}^{ext}_{2,m} \\ \vdots & \vdots & & \vdots \\ \mathcal{L}^{ext}_{m,1} & \mathcal{L}^{ext}_{m,2} & \ldots & \mathcal{L}^{ext}_{m,m} \end{pmatrix} \in \mathbb{R}^{n \times n}, \tag{39}$$

where $\mathcal{L}^{ext}_{p,q} \in \mathbb{R}^{n_p \times n_q}$ for $p, q \in \mathcal{M}$. We denote by $\mathcal{L}^{ext}_{k,row} \in \mathbb{R}^{n_k \times n}$ the k-th row of the block-matrix (39) for all $k \in \mathcal{M}$. It describes the connections of the cluster $\mathcal{C}_k$ with the rest of the agents in the network. The matrix $\mathcal{L}^{ext}_{k,red} \in \mathbb{R}^{n_k \times (n-n_k)}$ is obtained by removing the $\mathcal{L}^{ext}_{k,k}$ block from the $\mathcal{L}^{ext}_{k,row}$. For example, $\mathcal{L}^{ext}_{2,row} = [\mathcal{L}^{ext}_{2,1}\ \mathcal{L}^{ext}_{2,2}\ \ldots\ \mathcal{L}^{ext}_{2,m}]$ and $\mathcal{L}^{ext}_{2,red} = [\mathcal{L}^{ext}_{2,1}\ \mathcal{L}^{ext}_{2,3}\ \ldots\ \mathcal{L}^{ext}_{2,m}]$. Then, we express the external cost (35) in terms of new variables as

$$\overline{J}^{ext}_k = \int_0^{+\infty} Y_k^\top Q^{ext}_{k,1} Y_k + Y_k^\top \frac{Q^{ext}_{k,2}}{n_k} Y_k \,\mathrm{d}t \tag{40}$$

where

$$\begin{aligned} Q^{ext}_{k,1} &= \left(\mathrm{diag}\left(\frac{a^{ext}_{k,1}}{n_k}, \ldots, \frac{a^{ext}_{k,k-1}}{n_k}, \frac{a^{ext}_{k,k+1}}{n_k}, \ldots, \frac{a^{ext}_{k,m}}{n_k}\right) \otimes I_{n_x}\right),\\ Q^{ext}_{k,2} &= U^\top_{-k}(\mathcal{L}^{ext\top}_{k,red}\mathcal{L}^{ext}_{k,red} \otimes K^{ext\top}_k R_k K^{ext}_k) U_{-k},\\ U &= (\mathrm{diag}(\mathbb{1}_{n_1}, \ldots, \mathbb{1}_{n_m}) \otimes I_{n_x}),\\ R_k &> 0. \end{aligned} \tag{41}$$

The matrices $Q^{ext}_{k,1}$ and $Q^{ext}_{k,2}$ simplify the expressions in (35) such that $Y_k^\top Q^{ext}_{k,1} Y_k = \sum_{l \in \mathcal{N}_{C_k}} \frac{a^{ext}_{kl}}{n_k}(y_{s,k} - y_{s,l})^2$ and $Y_k^\top \frac{Q^{ext}_{k,2}}{n_k} Y_k = n_k \sum_{i=1}^{n_k} \widehat{u}^{ext\top}_{k,i} R_k \widehat{u}^{ext}_{k,i}$.

4.2.3 Control Design

To design the external control, we use the satisfaction equilibrium approach proposed in [19] for a network of homogeneous systems. Given the external error dynamics (38), it characterizes the external gain profile synchronizing the network in such a way that each cost (35) is bounded, i.e.,

$$\overline{J}^{ext}_k \leq \gamma^{ext} \|Y_k(0)\|^2, \quad \text{for} \quad k \in \mathcal{M}. \tag{42}$$

The term $\|Y_k(0)\|$ represents the initial condition of the cluster $\mathcal{C}_k$ while γ^{ext} is a given threshold. In particular, the following proposition is valid.

Proposition 1 (Prop 1, [19]) *Let a gain profile $\overline{K}^{ext} = \mathrm{diag}\left(K^{ext}_1, \ldots, K^{ext}_m\right)$ be given. The following statements are equivalent,*

(1) The gain profile $\overline{K}^{ext}$ is an SE of the satisfaction game (38) *for all $k \in \mathcal{M}$.*
(2) For all $k \in \mathcal{M}$, there exists a positive-definite matrix $P^{ext}_k > 0$ such that

$$\begin{cases} P^{ext}_k \mathbf{A}_{k,cl}(K^{ext}_k) + \mathbf{A}^\top_{k,cl}(K^{ext}_k) P^{ext}_k + \mathbf{Q}^{ext}_k(K^{ext}_k) < 0,\\ P^{ext}_k - \gamma^{ext} I_{(m-1).n_x} < 0, \end{cases} \tag{43}$$

where

$$\begin{cases} \mathbf{A}_{k,cl}(K^{ext}_k) = \mathbf{A}_k + \mathbf{B}_k K^{ext}_k (F_k \otimes I_{n_x}),\\ F_k = \left(\frac{a^{ext}_{k,1}}{n_k}, \ldots, \frac{a^{ext}_{k,k-1}}{n_k}, \frac{a^{ext}_{k,k+1}}{n_k}, \ldots, \frac{a^{ext}_{k,m}}{n_k}\right),\\ \mathbf{Q}^{ext}_k = \left(Q^{ext}_{k,1} + \frac{Q^{ext}_{k,2}}{n_k}\right). \end{cases} \tag{44}$$

Next, we present the algorithm that allows us to obtain the gain (K^{ext}) in satisfaction equilibrium. This algorithm greatly reduces the computational effort of obtaining the gain while synchronization the large-scale network in comparison to the algorithm proposed in [19].

4.3 Algorithm

Consider a network of m clusters (which is the number of cluster in our case) with their respective dynamics. We aim to design a synchronizing gain profile $K^{ext} = (K^{ext}_1, \ldots, K^{ext}_m)$ satisfying the cost constraints.

In the following algorithm, we first calculate the internal gain by solving the algebraic riccati equation (32). To design the external gain, we start with the initial gain profile that satisfies the LMI (43). Then to obtain the suboptimal gain, we multiply the gain from the previous iteration with a scalar $\alpha^{ext} \in \mathbb{R}_+ \setminus \{0\}$ and check if satisfies the LMI (43). One approach could be to start with high gain and decrease α^{ext} until the condition (43) is not satisfies, and use the smallest gain that satisfied the condition.

Furthermore, we should also make sure the network parameter ϵ is smaller so that control design using time-scale separation holds. Thus, to ensure this, we multiply the internal

Algorithm 1 Sequential Satisfaction Algorithm

Data: A, B and $n_k, k \in \mathcal{M}$;
Set: iterations $itr = 1$, maximum number of iterations itr_{max}, $0 < \epsilon^* \ll 1$ and $K^{ext}(0) = (K_1^{ext}(0), ..., K_m^{ext}(0))$ initial gain profile synchronizing the system ;
Calculate: P_k^{int} and K_k^{int} using equation (32) and (31) for all $k \in \mathcal{M}$, respectively;
while LMIs (43) not satisfied **OR** $itr \leq itr_{max}$ **do**
  $K^{ext}(itr+1) \leftarrow \alpha^{ext} K^{ext}(itr), \quad \alpha^{ext} \in \mathbb{R}_+ \setminus \{0\}$;
  Calculate: ϵ;
  if $\epsilon > \epsilon^*$ **then**
    $K_k^{int}(itr+1) \leftarrow \frac{\epsilon}{\epsilon^*} K_k^{int}(itr)$;
  else
    $K_k^{int}(itr+1) \leftarrow K_k^{int}(itr)$;
  end if
end while

gain K_k^{int} with ϵ/ϵ^* to obtain the new internal gain such that $\epsilon \leq \epsilon^*$.

Remark 6 *Notice that with such an approach we only scale the whole matrix K_k^{int} and K_k^{ext} on each step, while keeping the structure of the matrix intact.*

In the algorithm 1, to obtain the initial gain profile $K^{ext}(0)$ we use the algorithm in [2] which has the computational complexity of $\mathcal{O}(m)$ for m clusters. Then, the computation complexity to obtain the internal gain is of order $\mathcal{O}(m)$. Notice that the dimension of the matrix P_k^{int} in equation (32) does not depends on the number of agent (n_k) in the cluster thus the problem of finding the internal control K_k^{int} is independent of the number of agents in the cluster. To obtain the external gain K^{ext} we use the SeDuMi [12]. The computational complexity of verifying, if the gain profile satisfies the LMI condition (50) using SeDuMi is $\mathcal{O}(m^{5.5})$. Thus the overall computational complexity of the algorithm 1 is $\mathcal{O}(m)+\mathcal{O}(m)+\mathcal{O}(m^{5.5})$. Moreover, from Lemma 2 we obtain the stabilizing internal gain K^{int} and if the algorithm successfully converge to synchronizing external gain (K^{ext}) that satisfies LMI conditions (43), then they will satisfy the Assumption 2.

5 Global System Analysis

In this section, we analyze the overall networked system with the controller gains K^{int} and K^{ext} defined by the Algorithm 1 and designed for reduced slow and fast subsystems. First, we present the proposition which ensures that the slow and fast controllers, designed independent of each other, synchronize the overall network. And finally, we prove that the cluster cost $J_k(T, +\infty)$ is approximated only by the external cost $J_k^{ext}(T, +\infty)$, where $T > 0$ is a finite time after which the agents are synchronized.

5.1 Overall network behavior under reduced order control design

Based on the controller design procedure presented in the section 4, we ensure that the Assumption 2 is satisfied i.e., the internal gain stabilizing the fast dynamics and the external gain synchronizing the slow dynamics exists. Note that the presented design strategy optimize the internal cost function (28) while upper bound the corresponding external cost function (35) associated with the fast and the slow dynamics, respectively. The obtained internal control gain is optimal while the external control gain is sub-optimal. Now we apply these gains to achieve the synchronization in the overall network and the synchronization is ensured by the following proposition.

Proposition 2 *Consider the closed-loop network dynamics (14) and let the internal and external control gains are chosen based on Lemma 2 and Proposition 1, then the overall network synchronizes and satisfy the following bounds,*

$$\begin{aligned} y(t) &= y_s(t) + \mathcal{O}(\epsilon) \\ \xi(t) &= \xi_f(\mu^{int} t) + \mathcal{O}(\epsilon). \end{aligned} \tag{45}$$

PROOF. The proof follows from Theorem 2. ■

5.2 Cost Approximation

In this subsection, we prove that the cluster cost can be approximated by the average cost after finite time T. The motivation is derived from the fact that the internal dynamics converges rapidly to the consensus and the dominating network behavior is illustrated by the external dynamics. We prove that for the time $t \in [T, +\infty)$, the cluster cost J_k is approximated by n_k times the average external cost, i.e., $n_k \overline{J}_k^{ext}$.

To provide this approximation result, firstly we define the internal error bound which helps us to characterize the time $T > 0$. And secondly, we ensure that the exponential stability of the fast dynamics (21) implies the exponential stability of the error dynamics (15).

The necessity of the internal error bound arises in the approximation of the cluster cost. During the control design, we remind that the internal consensus is considered to be achieved before designing the external control. Thus, we need to characterize an error bound in finite time T, at which the cluster is very close to the internal consensus. More precisely, the bound at the time $T > 0$ such that $|\xi_{f,k}(T)| \leq \epsilon$ for all $k \in \mathcal{M}$.

The closed-loop fast dynamics is

$$\dot{\xi}_{f,k}(t) = \big((I_{n_k-1} \otimes A) - (\Lambda_k^{int} \otimes BK_k^{int})\big)\xi_{f,k}(t),$$

and

$$\xi_{f,k}(t) = e^{Cl_{f,k}t}\xi_{f,k}(0),$$

where $Cl_{f,k} := \big((I_{n_k-1} \otimes A) - (\Lambda_k^{int} \otimes BK_k^{int})\big)$ and $Cl_{f,k} < 0$ due to Lemma 2. Now, taking norm on both sides and from the definition of the measure of the matrix (see notations and preliminaries for definition), we obtain,

$$\|\xi_{f,k}(t)\| = e^{\nu(Cl_{f,k})t}\|\xi_{f,k}(0)\| \leq e^{\nu(Cl_f)t}\|\xi_{f,k}(0)\|$$

where $\nu(Cl_f) = \max_{k\in\mathcal{M}} \nu(Cl_{f,k})$. Then, as an internal error bound, we choose smallest $T \geq 0$ such that

$$\|\xi_{f,k}(T)\| \leq e^{\nu(Cl_f)T} \max_{k\in\mathcal{M}} \|\xi_{f,k}(0)\| \leq \epsilon.$$

This bound characterizes the local consensus inside each cluster in the finite time T. And hence, it yields

$$\|\xi_{f,k}(t)\| \leq \epsilon e^{\nu(Cl_f)(t-T)} \qquad \forall k \in \mathcal{M},$$

and

$$\|\xi_f(t)\| \leq \epsilon\sqrt{n-m}.e^{\nu(Cl_f)(t-T)}. \tag{46}$$

Next, in equation (25), we notice that the approximation of ξ defined in equation (15) depends on the fast variable ξ_f and the slow variable y_s, but the slow variable may or may not be stable. For the network to achieve synchronization, ξ should be stable. Thus, we prove the following lemma which ensures the exponential stability of ξ provided that ξ_f is exponentially stable.

Lemma 3 *The exponential stability of the fast dynamics* (21) *and the external error dynamics* (38) *implies the exponential stability of the error dynamics in* (15).

PROOF. See Appendix. ■

Next, with error bound for a finite time T we present the cluster cost approximation for $t \in [T, +\infty)$. The proposition is stated as follows:

Proposition 3 *During the time interval* $[T, +\infty)$*, the following approximation holds,*

$$J_k(T, +\infty) = n_k\bar{J}_k^{ext}(T, +\infty) + \mathcal{O}(\epsilon), \quad \forall k \in \mathcal{M}. \tag{47}$$

PROOF. See Appendix. ■

Finally, we present the following theorem that bounds the total cluster cost with the sum of internal, external and the constant term.

Theorem 2 *The total cluster cost for all clusters* $\mathcal{C}_k$*,* $k \in \mathcal{M}$ *satisfy the following bound:*

$$J_k \leq (\|P_k^{int}\| + n_k\gamma_k + \boldsymbol{C}_k)\|x(0)\|^2 + \mathcal{O}(\epsilon) \tag{48}$$

where P_k^{int} *is the solution of the Riccati equation* (32) *and* $\boldsymbol{C}_k$ *is a constant.*

PROOF. See Appendix. ■

6 Simulation

In this section, we provide numerical results to illustrate the effectiveness of the control procedure and the cost approximation using three scenarios. The agent's dynamics are given by (1), where

$$A = \begin{pmatrix} 0.15 & 0.98 \\ -0.98 & 0.15 \end{pmatrix}, B = \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{49}$$

The external graph between the agents in different clusters is generated using Erdos-Renyi [8] random graph generator. Then the internal graph with all-to-all connections for each cluster are generated and added to the external graph to obtain the graph of the network. For the numerical illustration, we consider the multiple scenarios.

- **Scenario 1:** Graph $\mathcal{G}_1$ with four clusters $m = 4$ with 630 agents in total. Each cluster has **all-to-all** internal connections and 299 external connections between the clusters in total. The threshold for the external cost is $\gamma^{ext} = 0.8$.
- **Scenario 2:** Same as Scenario 1 with **dense** internal connections instead of all-to-all internal connections.
- **Scenario 3:** Comparison of control design presented in this paper with the satisfactory control approach in [19] and guaranteed cost approach proposed in [2].

The details of the simulations are present in Tables 2 - 5. In the tables, n_k represent the number of agents in cluster $\mathcal{C}_k$, $error(k) = \frac{|J_k - n_k\bar{J}_k^{ext}|}{J_k} \times 100$, is the error percentage between the total cost and the external cost after time T, and K^{ext} and K^{int} are the respective external and internal gains.

6.1 Scenario 1: All-to-all connections in Clusters

The Figure 2 represents the synchronization of the agents in a network with graph $\mathcal{G}_1$. For the graph $\mathcal{G}_1$, the network parameter is $\epsilon_1 = 0.06$. In the figure, we can observe the four branches appearing and merging into one. Each branch represents the local agreement within the clusters. Next, Figure 3 illustrates the cost approximation for the cluster $\mathcal{C}_4$ by comparing the total cluster cost J_4 and the external cost $n_4\bar{J}_4^{ext}$, after finite time $T = 2s$. More details of the simulations are presented in the table.

Clusters	K^{int}	K^{ext}
$\mathcal{C}_1$	$[1.5352, -0.1102]$	$[0.85, 0.16]$
$\mathcal{C}_2$	$[1.5349, -0.1114]$	$[1.17, 0.22]$
$\mathcal{C}_3$	$[1.5346, -0.1128]$	$[0.59, 0.11]$
$\mathcal{C}_4$	$[1.5344, -0.1137]$	$[1.05, 0.2]$

Table 1
Internal and External gains

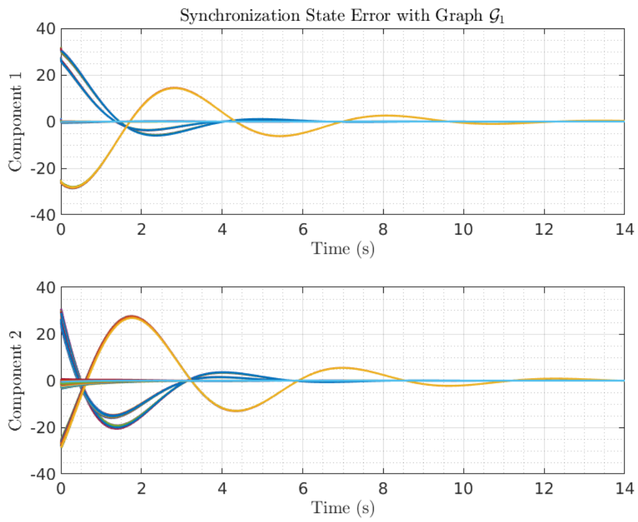

Fig. 2. Evolution of the error between the agents' state in graph $\mathcal{G}_1$ with all-to-all connections inside clusters.

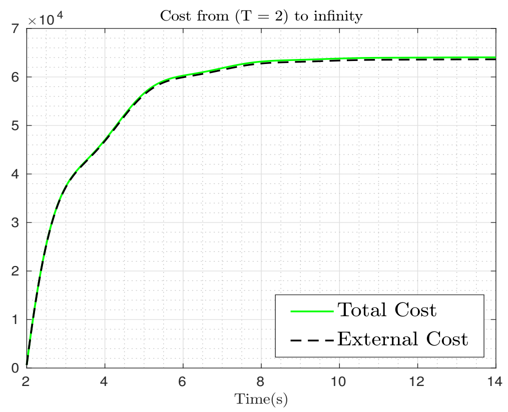

Fig. 3. Evolution of the costs J_4 and $n_4 \bar{J}_4^{ext}$ with all-to-all connections inside clusters.

$\epsilon = 0.06$, $\gamma = 0.8$			
	n_k	$J_k(\times 10^5)$	$error(k)$
$\mathcal{C}_1$	120	0.8966	0.45%
$\mathcal{C}_2$	140	0.5768	0.86%
$\mathcal{C}_3$	170	1.8950	0.24%
$\mathcal{C}_4$	200	0.6405	0.65%

Table 2
Network with 630 agents and 299 external connections with dense connections in clusters.

6.2 *Scenario 2: Connected Clusters*

In this scenario, we consider the graph where the clusters have dense interconnections instead of the all-to-all connections compared to Scenario 1. However, the number of agents and number of external connections remains the same as in the graph $\mathcal{G}_1$. Let us denote this graph as $\mathcal{G}_2$. The same gains from the Scenario 1 (Table 1) are applied to network system with the graph $\mathcal{G}_2$. The details of the simulation are presented in table 3.

$\epsilon = 0.06$, $\gamma = 0.8$			
	n_k	$J_k(\times 10^5)$	$error(k)$
$\mathcal{C}_1$	120	0.8983	0.64%
$\mathcal{C}_2$	140	0.5780	1.07%
$\mathcal{C}_3$	170	1.8975	0.37%
$\mathcal{C}_4$	200	0.6415	0.81%

Table 3
Network with 630 agents and 299 external connections with all-to-all connections inside every cluster .

6.3 *Scenario 3*

In the last scenario, we consider a network of $m = 4$ clusters with $n_k = 10$ agents in each. We recall that $\gamma^{ext} = 1$ is chosen for both controls. A comparison is done between the composite control proposed in this paper and the satisfactory control approach proposed in [19]. The design procedure in [19] needs 13752 seconds (3.8 hours) to compute the gains for $n = 40$ agents, while the composite design in this paper requires 13 seconds. However, we can observe an incontestable difference in performance on the cluster costs due to satisfactory control, as shown in table 4. This emphasizes the trade-off between the computing time/resources to obtain the required controller. Despite being less effective, we have to keep in mind that the composite control suits betters for large-scale networks and present an important benefit in terms of computation loads and time.

	$\mathcal{C}_1$	$\mathcal{C}_2$	$\mathcal{C}_3$	$\mathcal{C}_4$
n_k	10	10	10	10
J_k	17204	5452	6943	16949
J_k^*	10164	3303	3080	9714

Table 4
Comparison of cost.

Next, we compare the strategy in [2] with the composite control. In [2], every agents apply the same gain independently of their neighborhoods and aim to bound a global cost. Applying the control [2] on the graph $\mathcal{G}_1$, it results in a cluster cost which we label by $J_k^{\dagger}$. From Table 5, we observe that our strategy significantly outperforms the strategy in [2], the first cluster cost obtained via the composite control is 20 times smaller. One may observe the same for the other clusters.

7 Conclusion

In this paper, we propose a distributed composite control design strategy for the clustered network. Using coordinate transformation, the network dynamics is transformed into

	$\mathcal{C}_1$	$\mathcal{C}_2$	$\mathcal{C}_3$	$\mathcal{C}_4$
n_k	120	140	170	200
$J_k(\times 10^6)$	0.385	0.269	0.689	0.262
$J_k^\dagger(\times 10^6)$	6.7	8.1	16.7	20.5

Table 5
Comparison of the cost.

standard singular perturbation form and decoupled into slow and fast dynamics using time-scale separation. This decoupling of the network dynamics also decouple the control into fast (internal) and slow (external). The internal control is responsible for intra-cluster synchronization, while the external synchronize the network while satisfying the imposed cost criterion. This independent design greatly reduces the computational effort required to obtain the control. Finally, we show that the cluster cost is approximated only by the external cost after a short period of time.

A Proofs

Proof of Lemma 1
We know from [13] $\|(A\otimes B)\| = \|A\|\|B\|$ for any matrix $A\in\mathbb{R}^{n\times n}$, $B\in\mathbb{R}^{m\times m}$. Let us define $\bar{n}=\max_k n_k$ and $\underline{n}=\min_k n_k$. In addition, $\|H\|=\frac{1}{\sqrt{\underline{n}}}$, $\|\widetilde{H}\|=\sqrt{\bar{n}}$ and $\|Z\|=\frac{1}{\sqrt{\underline{n}}}$, $\|\widetilde{Z}\|=\sqrt{\bar{n}}$. From the Assumption 1, there exists a strictly positive constant $c_1\in\mathbb{R}$ such that $\|A\|=c_1\mu^{ext}$. It follows that,

$$\begin{aligned}\|\bar{A}_{11}\| &= \|(I_m\otimes A) - H(I_n\otimes B)K^{ext}(\mathcal{L}^{ext}\otimes I_{n_x})\widetilde{H}\|\\ &\le \|A\| + \|H\|.\|(I_n\otimes B)K^{ext}(\mathcal{L}^{ext}\otimes I_{n_x})\|.\|\widetilde{H}\|\\ &= (c_1+\sqrt{\tfrac{\bar{n}}{\underline{n}}})\mu^{ext} = (c_1+\sqrt{\tfrac{\bar{n}}{\underline{n}}})\epsilon\mu^{int}.\end{aligned} \tag{A.1}$$

The bounds of $\bar{A}_{12}$, $\bar{A}_{21}$ and $\bar{A}_{22}^1$ are derived similarly, that's why we only prove for $\bar{A}_{12}$,

$$\begin{aligned}\|\bar{A}_{12}\| &= \|H(I_n\otimes B)K^{ext}(\mathcal{L}^{ext}\otimes I_{n_x})\widetilde{Z}\|\\ &\le \sqrt{\tfrac{\bar{n}}{\underline{n}}}\mu^{ext} = \epsilon\sqrt{\tfrac{\bar{n}}{\underline{n}}}\mu^{int}.\end{aligned} \tag{A.2}$$

Then, we lower-bound the matrix $\bar{A}_{22}^2$ such that

$$\begin{aligned}\|\bar{A}_{22}^2\| &= \|(I_{n-m}\otimes A) - (I_{n-m}\otimes B)K_{n-m}^{int}(\Lambda^{int}\otimes I_{n_x})\|\\ &\ge |\|A\| - \|(I_{n-m}\otimes B)K_{n-m}^{int}(\Lambda^{int}\otimes I_{n_x})\||.\end{aligned} \tag{A.3}$$

From (17), we understand that the second term in (A.3) is much larger than the first one. Thus, by taking the difference between the largest value of the first term and the smallest value of the second term it yields a lower-bound as

$$\|\bar{A}_{22}^2\| \ge |c_1\epsilon\mu^{int} - \mu^{int}| = |1-c_1\epsilon|\mu^{int}, \tag{A.4}$$

where $\mu^{int} = \min_{k\in\mathcal{M}} \|(\Lambda_k^{int}\otimes BK_k^{int})\|$.

Proof of Lemma 3
Integrating the error dynamics in (15), we obtain

$$\begin{aligned}\xi(t) &= e^{\bar{A}_{22}t}\xi(0) + \textstyle\int_0^t e^{\bar{A}_{22}(t-\tau)}\bar{A}_{21}y(\tau)\,\mathrm{d}\tau\\ &= e^{\bar{A}_{22}t}\xi(0) + \textstyle\int_0^t e^{\bar{A}_{22}(t-\tau)}\bar{A}_{21}(y_s(\tau)+\epsilon\Psi(\epsilon)\xi_f(\tau))\,\mathrm{d}\tau\\ &= e^{\bar{A}_{22}t}\xi(0) + \textstyle\int_0^t e^{\bar{A}_{22}(t-\tau)}Z^T MY(\tau)\,\mathrm{d}\tau\\ &\quad +\epsilon\textstyle\int_0^t e^{\bar{A}_{22}(t-\tau)}\bar{A}_{21}\Psi(\epsilon)\xi_f(\tau)\,\mathrm{d}\tau\end{aligned}$$

where $M = diag(M_1,...,M_m)$ and $M_k = (\mathcal{L}_{k,red}^{ext}\otimes BK_k^{ext})U_{-k}$. By taking norm on both sides, we have

$$\begin{aligned}\|\xi(t)\| \le \|e^{\bar{A}_{22}t}\|\|\xi(0)\| + \|Z^T M\|\textstyle\int_0^t \|e^{\bar{A}_{22}(t-\tau)}\|\|Y(\tau)\|\,\mathrm{d}\tau\\ +\epsilon\|\bar{A}_{21}\Psi(\epsilon)\|\textstyle\int_0^t \|e^{\bar{A}_{22}(t-\tau)}\|\|\xi_f(\tau)\|\,\mathrm{d}\tau\end{aligned} \tag{A.5}$$

Also, from the design of internal and external control we know that, for all $t\ge 0$,

$$\begin{cases} Y(t) = e^{\mathbf{A}_{cl}t}Y(0)\\ \xi_f(t) = e^{\bar{A}_{22}^2 t}\xi_f(0)\end{cases} \Rightarrow \begin{cases}\|Y(t)\| \le e^{\nu(\mathbf{A}_{cl})t}\|Y(0)\|\\ \|\xi_f(t)\| \le e^{\nu(\bar{A}_{22}^2)t}\|\xi_f(0)\|\end{cases} \tag{A.6}$$

where $\mathbf{A}_{cl} = diag(\mathbf{A}_{1,cl},...,\mathbf{A}_{m,cl})$ is the closed-loop dynamics of the external error (38). Then, it follows that

$$\begin{aligned}&\|\xi(t)\| \le e^{\nu(\bar{A}_{22})t}\|\xi(0)\|\\ &\quad +\|Z^T M\|\|Y(0)\|\textstyle\int_0^t e^{\nu(\bar{A}_{22})(t-\tau)}e^{\nu(\mathbf{A}_{cl})\tau}\,\mathrm{d}\tau\\ &\quad +\epsilon\|\bar{A}_{21}\Psi(\epsilon)\|\|\xi_f(0)\|\textstyle\int_0^t e^{\nu(\bar{A}_{22})(t-\tau)}e^{\nu(\bar{A}_{22}^2)\tau}\,\mathrm{d}\tau\,.\end{aligned}$$

By integrating the second term in (A.5), we have

$$\begin{aligned}&\|Z^T M\|\|Y(0)\|\textstyle\int_0^t e^{\nu(\bar{A}_{22})(t-\tau)}e^{\nu(\mathbf{A}_{cl})\tau}\,\mathrm{d}\tau\\ &\quad = \|Z^T M\|\|Y(0)\|e^{\nu(\bar{A}_{22})t}\textstyle\int_0^t e^{(\nu(\mathbf{A}_{cl})-\nu(\bar{A}_{22}))\tau}\,\mathrm{d}\tau\\ &= \tfrac{\|Z^T M\|\|Y(0)\|}{\nu(\mathbf{A}_{cl})-\nu(\bar{A}_{22})}\left[e^{\nu(\mathbf{A}_{cl})t} - e^{\nu(\bar{A}_{22})t}\right].\end{aligned}$$

In the same manner, the third term is

$$\begin{aligned}&\epsilon\|\bar{A}_{21}\Psi(\epsilon)\|\|\xi_f(0)\|\int_0^t e^{\nu(\bar{A}_{22})(t-\tau)}e^{\nu(\bar{A}_{22}^2)\tau}\,\mathrm{d}\tau\\ &\quad = \frac{\epsilon\|\bar{A}_{21}\Psi(\epsilon)\|\|\xi_f(0)\|}{\nu(\bar{A}_{22}^2)-\nu(\bar{A}_{22})}\left[e^{\nu(\bar{A}_{22}^2)t} - e^{\nu(\bar{A}_{22})t}\right].\end{aligned} \tag{A.7}$$

Finally, we have

$$\begin{aligned}\|\xi(t)\| &\le \mathbf{C}_1 e^{\nu(\mathbf{A}_{cl})t} + \epsilon\mathbf{C}_2 e^{\nu(\bar{A}_{22}^2)t}\\ &\quad + (\|\xi(0)\| - \mathbf{C}_1 - \epsilon\mathbf{C}_2)\,e^{\nu(\bar{A}_{22})t},\end{aligned} \tag{A.8}$$

where $\mathbf{C}_1 = \frac{\|Z^T M\|\|Y(0)\|}{\nu(\mathbf{A}_{cl})-\nu(\bar{A}_{22})}$ and $\mathbf{C}_2 = \frac{\|\bar{A}_{21}\Psi(\epsilon)\|\|\xi_f(0)\|}{\nu(\bar{A}_{22}^2)-\nu(\bar{A}_{22})}$. Moreover, we know that $\nu(\bar{A}_{22}^2) < \nu(\bar{A}_{22}) < \nu(\mathbf{A}_{cl}) < 0$.

Thus, we conclude that ξ converges exponentially to zero and the rate of convergence can be bounded as

$$\|\xi(t)\| \leq \|\xi(0)\| e^{\nu(\mathbf{A}_{cl})t}. \tag{A.9}$$

Proof of Theorem 1
The proof follows the reasoning in Theorem 5.1, Chapter 2, [11]. In [11], via similarity transformation, the authors express and decouple the original slow and fast variables into the approximated variables. The singularly perturbed system dynamics (22) is slightly different from the one in the [11]. Thus, we adapt the result from [11] to our system model to obtain the approximation results. The similarity transformations [11] for the decoupling of the dynamics (15) are

$$\begin{aligned}
\begin{bmatrix} y \\ \xi \end{bmatrix} &= \begin{bmatrix} I_{m.n_x} & \epsilon\Psi(\epsilon) \\ -\Omega(\epsilon) & I_{n_x(n-m)} - \epsilon\Omega(\epsilon)\Psi(\epsilon) \end{bmatrix} \begin{bmatrix} y_s \\ \xi_f \end{bmatrix} \\
\begin{bmatrix} y_s \\ \xi_f \end{bmatrix} &= \begin{bmatrix} I_{m.n_x} - \epsilon\Psi(\epsilon)\Omega(\epsilon) & -\epsilon\Psi(\epsilon) \\ \Omega(\epsilon) & I_{n_x(n-m)} \end{bmatrix} \begin{bmatrix} y \\ \xi \end{bmatrix},
\end{aligned} \tag{A.10}$$

where the functions Ω and Ψ should satisfy the following,

$$\begin{aligned}
R(\Omega(\epsilon), \epsilon) &= \epsilon A_{21} - \epsilon A_{22}^1 \Omega(\epsilon) - A_{22}^2 \Omega(\epsilon) \\
&\quad + \epsilon\Omega(\epsilon) A_{11} - \epsilon\Omega(\epsilon) A_{12}\Omega(\epsilon) = 0, \\
S(\Psi(\epsilon), \epsilon) &= \epsilon A_{11}\Psi(\epsilon) + A_{12} - \epsilon A_{12}\Omega(\epsilon)\Psi(\epsilon) \\
&\quad - \epsilon\Psi(\epsilon) A_{22}^1 - \Psi(\epsilon) A_{22}^2 - \epsilon\Psi(\epsilon)\Omega(\epsilon) A_{12} = 0.
\end{aligned}$$

The approximation of Ω and Ψ, obtained with the Taylor development w.r.t. ϵ, are

$$\begin{aligned}
&\Omega(\epsilon) = \epsilon (A_{22}^2)^{-1} A_{21} + \mathcal{O}(\epsilon^2), \\
&\Psi(\epsilon) = A_{12}(A_{22}^2)^{-1} + \epsilon((A_{22}^2)^{-1} A_{11} A_{12} (A_{22}^2)^{-1} - A_{12}) + \mathcal{O}(\epsilon^2).
\end{aligned}$$

From Lemma (3), we know that $\xi(t)$ and $\xi_f(t_f)$ converge to zero exponentially as t and t_f tend to $+\infty$, respectively. Thus, we can claim that $\Omega(\epsilon) y_s(t)$ has an exponential decrease to zero w.r.t. t. Finally, from the above transformation (A.10) and (A.11), we obtain the approximations (25).

Proof of Proposition 3
The cost J_k is split into the sum of the internal and external costs and composite term as shown in equation (6). Then, we bound the internal and external costs from time T to infinity. We proceed similarly with the composite term.
Internal Cost: Substituting $x_k = \widetilde{H}_k y_k + \widetilde{Z}_k \xi_k$ from equation (12) into J_k^{int} in equation (6) and with $\widetilde{H}_k^\top (\mathcal{L}_k^{int} \otimes I_{n_x}) = 0$, it yields

$$\begin{aligned}
J_k^{int}(T, +\infty) &= \textstyle\int_T^{+\infty} \xi_k^\top \widetilde{Z}_k ((\mathcal{L}_k^{int} \otimes I_{n_x}) \\
&\quad + (\mathcal{L}_k^{int\top} \mathcal{L}_k^{int} \otimes K_k^{int\top} R_k K_k^{int})) \widetilde{Z}_k \xi_k \,\mathrm{dt}, \\
&= \textstyle\int_T^{+\infty} n_k \xi_k^\top \Big((\Lambda_k^{int} \otimes I_{n_x}) \\
&\quad + \Big({\Lambda_k^{int}}^2 \otimes P_k^{int\top} B \frac{R_k^{-1}}{n_k^2} B^\top P_k^{int} \Big) \Big) \xi_k \,\mathrm{dt}, \\
&= \int_T^{+\infty} n_k \xi_k^\top \big((\Lambda_k^{int} \otimes I_{n_x}) \\
&\quad + \big(I_{n_k - 1} \otimes P_k^{int\top} B R_k^{-1} B^\top P_k^{int} \big) \big) \xi_k \,\mathrm{dt}, \\
&\leq \mathbf{C}_{3,k} \textstyle\int_T^{+\infty} \|\xi_k\|^2 \,\mathrm{dt} \leq \mathbf{C}_{3,k} \int_T^{+\infty} \|\xi(t)\|^2 \,\mathrm{dt}\,.
\end{aligned}$$

where,
$\mathbf{C}_{3,k} = \|n_k \big((\Lambda_k^{int} \otimes I_{n_x}) + \big(I_{n_k-1} \otimes P_k^{int\top} B R_k^{-1} B^\top P_k^{int} \big) \big)\|$.
From Lemma 3 and equation (A.9), we have $\|\xi(t)\| \leq \|\xi(T)\| e^{\nu(\mathbf{A}_{cl})(t-T)}$, for all $t \in [T, +\infty)$. Thus, with $\nu(\mathbf{A}_{cl}) < 0$, we have,

$$\int_T^{+\infty} \|\xi(t)\|^2 \,\mathrm{dt} \leq -\frac{\|\xi(T)\|^2}{2\nu(\mathbf{A}_{cl})} = \mathbf{C}_4 \|\xi(T)\|^2 \tag{A.11}$$

where $\mathbf{C}_4 := \big(-\frac{1}{2\nu(\mathbf{A}_{cl})} \big)$. Thus, from (A.11)-(A.11) and the approximation of ξ in equation (25),

$$\begin{aligned}
&J_k^{int}(T, +\infty) \leq \mathbf{C}_{3,k} \mathbf{C}_4 \|\xi_f(T) + \mathcal{O}(\epsilon)\|^2 \\
&\leq \mathbf{C}_{3,k} \mathbf{C}_4 \left(\|\xi_f(T)\|^2 + 2\mathcal{O}(\epsilon) \|\xi_f(T)\| + \mathcal{O}(\epsilon^2) \right).
\end{aligned}$$

Finally, replacing $\|\xi_f(T)\| \leq \epsilon\sqrt{n-m}$ from (46) we have

$$J_k^{int}(T, +\infty) \leq \mathcal{O}(\epsilon^2). \tag{A.12}$$

External cost: First, we recast the collective external control (4) in the external error variable Y_k, as follows

$$\begin{aligned}
u_k^{ext}(t) &= -(I_{n_k} \otimes K_k^{ext})(\mathcal{L}_{k,row}^{ext} \otimes I_{n_x}) x(t) \\
&= -(\mathcal{L}_{k,row}^{ext} \otimes K_k^{ext})(\widetilde{H} y(t) + \widetilde{Z}\xi(t)) \\
&= -(\mathcal{L}_{k,row}^{ext} \otimes K_k^{ext})(\widetilde{H} y_s(t) + \epsilon \widetilde{H} \Psi(\epsilon) \xi_f(t_f) + \widetilde{Z}\xi(t)) \\
&= (\mathcal{L}_{k,red}^{ext} \otimes K_k^{ext}) U_{-k} Y_k(t) \\
&\quad - (\mathcal{L}_{k,row}^{ext} \otimes K_k^{ext})(\epsilon \widetilde{H} \Psi(\epsilon) \xi_f(t_f) + \widetilde{Z}\xi(t)),
\end{aligned} \tag{A.13}$$

where $\mathcal{L}_{k,row}^{ext}$ is the k-th block-row of $\mathcal{L}^{ext}$ and $\mathcal{L}_{k,red}^{ext}$ is obtained by removing the $\mathcal{L}_{k,k}^{ext}$ block from $\mathcal{L}_{k,row}^{ext}$. Then, it yields

$$\begin{aligned}
&u_k^{ext\top}(t)(I_{n_k} \otimes R_k) u_k^{ext}(t) \\
&= Y_k^\top(t) Q_{k,2}^{ext} Y_k(t) + \epsilon^2 \xi_f^\top(t_f) D_{1,k} \xi_f(t_f) + \xi^\top(t) D_{2,k} \xi(t) \\
&\quad - \epsilon Y_k^\top(t) D_{3,k} \xi_f(t_f) - Y_k^\top(t) D_{4,k} \xi(t) + \epsilon \xi^\top(t) D_{5,k} \xi_f(t_f),
\end{aligned} \tag{A.14}$$

where

$$\begin{cases} Q_{k,2}^{ext} = U_{-k}^\top(\mathcal{L}_{k,red}^{ext\top}\mathcal{L}_{k,red}^{ext}\otimes K_k^{ext\top}R_kK_k^{ext})U_{-k}, \\ D_{1,k} = \Psi(\epsilon)^\top\widetilde{H}^\top(\mathcal{L}_{k,row}^{ext^\top}\mathcal{L}_{k,row}^{ext}\otimes K_k^{ext^\top}R_kK_k^{ext})\widetilde{H}\Psi(\epsilon), \\ D_{2,k} = \widetilde{Z}^\top(\mathcal{L}_{k,row}^{ext^\top}\mathcal{L}_{k,row}^{ext}\otimes K_k^{ext^\top}R_kK_k^{ext})\widetilde{Z}, \\ D_{3,k} = 2U_{-k}^\top(\mathcal{L}_{k,red}^{ext^\top}\mathcal{L}_{k,row}^{ext}\otimes K_k^{ext^\top}R_kK_k^{ext})\widetilde{H}\Psi(\epsilon), \\ D_{4,k} = 2U_{-k}^\top(\mathcal{L}_{k,red}^{ext\top}\mathcal{L}_{k,row}^{ext}\otimes K_k^{ext\top}R_kK_k^{ext})\widetilde{Z}, \\ D_{5,k} = 2\widetilde{Z}^\top(\mathcal{L}_{k,row}^{ext\top}\mathcal{L}_{k,row}^{ext}\otimes K_k^{ext\top}R_kK_k^{ext})\widetilde{H}\Psi(\epsilon). \end{cases}$$

Secondly, let consider the state part in the external cost. To simplify the expression, we use $(\mathcal{L}_k^{ext}\otimes I_{n_x})\widetilde{H}y_s(t) = -(\mathcal{L}_{k,col}^{ext}\otimes I_{n_x})U_{-k}Y_k(t)$ where $\mathcal{L}_{k,col}^{ext}$ is the matrix $\mathcal{L}_k^{ext}$ with its k-th block-column removed. Then, we obtain

$$\begin{aligned} & x^\top(t)(\mathcal{L}_k^{ext}\otimes I_{n_x})x(t) \\ & = \star^\top(\mathcal{L}_k^{ext}\otimes I_{n_x})(\widetilde{H}y_s(t)+\epsilon\widetilde{H}\Psi(\epsilon)\xi_f(t_f)+\widetilde{Z}\xi(t)) \\ & = n_kY_k^\top(t)Q_{k,1}^{ext}Y_k(t)+\epsilon^2\xi_f(t_f)^\top M_{1,k}\xi_f(t_f) \\ & +\xi^\top(t)M_{2,k}\xi(t)-\epsilon Y_k^\top(t)M_{3,k}\xi_f(t_f)-Y_k^\top(t)M_{4,k}\xi(t) \\ & +\epsilon\xi^\top(t)M_{5,k}\xi_f(t_f) \end{aligned} \tag{A.15}$$

where

$$\begin{cases} M_{1,k} = \Psi(\epsilon)^\top\widetilde{H}^\top(\mathcal{L}_k^{ext}\otimes I_{n_x})\widetilde{H}\Psi(\epsilon) \\ M_{2,k} = \widetilde{Z}^\top(\mathcal{L}_k^{ext}\otimes I_{n_x})\widetilde{Z} \\ M_{3,k} = 2U_{-k}^\top(\mathcal{L}_{k,col}^{ext\top}\otimes I_{n_x})\widetilde{H}\Psi(\epsilon) \\ M_{4,k} = 2U_{-k}^\top(\mathcal{L}_{k,col}^{ext\top}\otimes I_{n_x})\widetilde{Z} \\ M_{5,k} = 2\widetilde{Z}^\top(\mathcal{L}_k^{ext}\otimes I_{n_x})\widetilde{H}\Psi(\epsilon). \end{cases}$$

Then, replacing (A.14) and (A.15) into the external cost (J_k^{ext}) in equation (6), we get

$$\begin{aligned} & J_k^{ext}(T,+\infty) \\ & = n_k\int_T^{+\infty} Y_k^\top(t)Q_{k,1}^{ext}Y_k(t)+Y_k^\top(t)\frac{Q_{k,2}^{ext}}{n_k}Y_k(t)\,\mathrm{d}t+\Delta_1 \\ & = n_k\bar{J}_k^{ext}(T,+\infty)+\Delta_1, \end{aligned} \tag{A.16}$$

where $\Delta_1 = \Delta_1^1+\Delta_1^2+\Delta_1^3+\Delta_1^4+\Delta_1^5$ and

$$\begin{cases} \Delta_1^1 = \epsilon^2\int_T^{+\infty}\xi_f(t_f)^\top(M_{1,k}+D_{1,k})\,\xi_f(t_f)\,\mathrm{d}t, \\ \Delta_1^2 = \int_T^{+\infty}\xi^\top(t)\,(M_{2,k}+D_{2,k})\,\xi(t)\,\mathrm{d}t, \\ \Delta_1^3 = -\epsilon\int_T^{+\infty}Y_k^\top(t)\,(M_{3,k}+D_{3,k})\,\xi_f(t_f)\,\mathrm{d}t, \\ \Delta_1^4 = -\int_T^{+\infty}Y_k^\top(t)\,(M_{4,k}+D_{4,k})\,\xi(t)\,\mathrm{d}t, \\ \Delta_1^5 = \epsilon\int_T^{+\infty}\xi^\top(t)\,(M_{5,k}+D_{5,k})\,\xi_f(t_f)\,\mathrm{d}t\,. \end{cases} \tag{A.17}$$

$$\begin{aligned} \Delta_1^1 & \le \epsilon^2\|M_{1,k}+D_{1,k}\|\int_T^{+\infty}\|\xi_f(t_f)\|^2\,\mathrm{d}t \\ & \le -\epsilon^2\frac{\|M_{1,k}+D_{1,k}\|\|\xi_f(0)\|^2}{2\nu(\bar{A}_{22}^2)}e^{2\nu(\bar{A}_{22}^2)T} = \mathcal{O}(\epsilon^2). \end{aligned}$$

$$\begin{aligned} \Delta_1^2 & \le \mathbf{C}_{3,k}\|M_{2,k}+D_{2,k}\|\textstyle\int_T^{+\infty}\|\xi(t)\|^2\|\xi(T)\|^2\,\mathrm{d}t \\ & \le \mathbf{C}_{3,k}\|M_{2,k}+D_{2,k}\|\left(\|\xi_f(T)\|^2+2\mathcal{O}(\epsilon)\|\xi_f(T)\|+\mathcal{O}(\epsilon^2)\right) \\ & \le \mathcal{O}(\epsilon^2). \end{aligned}$$

$$\begin{aligned} \Delta_1^3 & \le \epsilon\|M_{3,k}+D_{3,k}\|\|Y(0)\|\|\xi_f(0)\|\textstyle\int_T^{+\infty}e^{\nu(\mathbf{A}_{cl})t}e^{\nu(\bar{A}_{22}^2)t}\,\mathrm{d}t \\ & = -\epsilon\frac{\|M_{3,k}+D_{3,k}\|\|Y(0)\|\|\xi_f(0)\|}{\nu(\mathbf{A}_{cl})+\nu(\bar{A}_{22}^2)}e^{\left(\nu(\mathbf{A}_{cl})+\nu(\bar{A}_{22}^2)\right)T} = \mathcal{O}(\epsilon). \end{aligned}$$

Similarly, Δ_1^4 and Δ_1^5 are of order $\mathcal{O}(\epsilon)$. Finally, from (A.16) and bounds in (A.17) for Δ_1, we obtain

$$J_k^{ext}(T,+\infty) = n_k\bar{J}_k^{ext}(T,+\infty)+\mathcal{O}(\epsilon). \tag{A.18}$$

Composite term: We rewrite the external control (A.13) and the internal control (4) as

$$\begin{aligned} u_k^{ext}(t) & = -\mathbf{C}_{5,k}Y_k(t)-\epsilon\mathbf{C}_{6,k}\xi_f(t_f)-\mathbf{C}_{7,k}\xi(t) \\ u_k^{int}(t) & = (\mathcal{L}_k^{int}\otimes K_k^{int})\widetilde{Z}_k\xi_k(t) =: \mathbf{C}_{8,k}\xi_k(t). \end{aligned} \tag{A.19}$$

where $\mathbf{C}_{5,k} = (\mathcal{L}_{k,red}^{ext}\otimes K_k^{ext})U_{-k}$, $\mathbf{C}_{6,k} = (\mathcal{L}_{k,row}^{ext}\otimes K_k^{ext})\widetilde{H}\Psi(\epsilon)$ and $\mathbf{C}_{7,k} = (\mathcal{L}_{k,row}^{ext}\otimes K_k^{ext})\widetilde{Z}$. Then, taking the norm and substituting from equations (A.19) into the J_k^{cross} term in equation (6), we get,

$$\begin{aligned} & J_k^{cross}(T,+\infty) \le 2\|R_k\|\textstyle\int_T^{+\infty}\|u_k^{ext^\top}(t)\|\|u_k^{int}(t)\|\,\mathrm{d}t \\ & \le 2\|R_k\|\int_T^{+\infty}\|\mathbf{C}_{5,k}Y_k(t)+\epsilon\mathbf{C}_{6,k}\xi_f(t_f)+\mathbf{C}_{7,k}\xi(t)\|. \\ & \qquad \|\mathbf{C}_{8,k}\xi_k(t)\|\,\mathrm{d}t \end{aligned} \tag{A.20}$$

With simple calculation it can be shown that the first integral in the above equation is of order $\mathcal{O}(\epsilon)$ and the second and the third integrals are of order $\mathcal{O}(\epsilon^2)$. Thus, we have,

$$J_k^{cross}(T,+\infty) \le \mathcal{O}(\epsilon). \tag{A.21}$$

Finally, from (6), (A.12), (A.18) and (A.21), we conclude the proof.

Proof of Theorem 2

Internal Cost: Following the similar approximation as the approximation of the internal cost in Proposition 3, we obtain the following approximation for the internal cost for

$$J_k^{int} = n_kJ_{f,k}+\mathcal{O}(\epsilon) \tag{A.22}$$

Moreover, due to LQ-control design, the optimal fast cost $J_{f,k} = \xi_{f,k}(0)^\top(I_{n_{k-1}}\otimes P_k^{int})\xi_{f,k}(0)$. The substituting the approximation $\xi_k = \xi_{f,k}+\mathcal{O}(\epsilon)$, we get, $J_{f,k} = \xi_k(0)^\top(I_{n_{k-1}}\otimes P_k^{int})\xi_k(0)+\mathcal{O}(\epsilon)$. Then, from the transformation (11), it yields,

$$\begin{aligned} J_k^{int} & = n_k.x_k(0)^\top Z_k^\top(I_{n_{k-1}}\otimes P_k^{int})Z_kx_k(0)+\mathcal{O}(\epsilon) \\ & = x_k(0)^\top(I_{n_k}\otimes P_k^{int})x_k(0)+\mathcal{O}(\epsilon) \\ & \le \|P_k^{int}\|\|x_k(0)\|^2+\mathcal{O}(\epsilon). \end{aligned} \tag{A.23}$$

External Cost: Substituting $x_k = \widetilde{H}y + \widetilde{Z}\xi$ in the external cost J_k^{ext} in equation (6), and performing the similar operation as in the approximation of the external cost in Proposition 3, we obtain,

$$J_k^{ext} \leq n_k \overline{J}_k^{ext} + \Pi_1 + \mathcal{O}(\epsilon) \tag{A.24}$$

with

$$\begin{aligned}\Pi_1 = 2 \int_0^{+\infty} \xi^\top \widetilde{Z}^\top \big((\mathcal{L}_k^{ext} \otimes I_{n_x}) + \\ ({\mathcal{L}_{k,row}^{ext}}^\top \mathcal{L}_{k,row}^{ext} \otimes {K_k^{ext}}^\top R_k K_k^{ext})\big)\widetilde{H}y \,\mathrm{dt} \\ + \int_0^{+\infty} \xi^\top \big(\widetilde{Z}^\top (\mathcal{L}_k^{ext} \otimes I_{n_x}) \\ + ({\mathcal{L}_{k,row}^{ext}}^\top \mathcal{L}_{k,row}^{ext} \otimes {K_k^{ext}}^\top R_k K_k^{ext})\widetilde{Z}\big)\xi \,\mathrm{dt}\,.\end{aligned} \tag{A.25}$$

Furthermore, substituting $(\mathcal{L}_k^{ext} \otimes I_{n_x})\widetilde{H}y_s = -(\mathcal{L}_{k,col}^{ext} \otimes I_{n_x})U_{-k}Y_k$ and $(\mathcal{L}_{k,row}^{ext} \otimes I_{n_x})\widetilde{H}y_s = -(\mathcal{L}_{k,red}^{ext} \otimes I_{n_x})U_{-k}Y_k$ in equation (A.25) and taking the norm we have,

$$\Pi_1 \leq \mathbf{C}_{9,k}\|x(0)\|^2 + \mathcal{O}(\epsilon) \tag{A.26}$$

where $\mathbf{C}_{9,k} := \mathbf{C}_4\|Z\|\big(\|Y_k(0)\|\|2\widetilde{Z}^\top\big((\mathcal{L}_{k,col}^{ext} \otimes I_{n_x}) + ({\mathcal{L}_{k,row}^{ext}}^\top \mathcal{L}_{k,red}^{ext} \otimes {K_k^{ext}}^\top R_k K_k^{ext})\big)U_{-k}\| + \|\big(\widetilde{Z}^\top(\mathcal{L}_k^{ext} \otimes I_{n_x}) + ({\mathcal{L}_{k,row}^{ext}}^\top \mathcal{L}_{k,row}^{ext} \otimes {K_k^{ext}}^\top R_k K_k^{ext})\widetilde{Z}\big)\mathbf{C}_4\big)$.

Cross Term: Substituting from equation (A.19) and from Theorem 1 into the cross term in equation (6) and after further calculation, we get,

$$\begin{aligned}J_k^{cross} &\leq 2\|R_k\|\|\mathbf{C}_{5,k}\|\|\mathbf{C}_{8,k}\|\|\mathbf{C}_4\|\|Y_k(0)\|\|Z\|\|x(0)\| \\ &\quad +2\|R_k\|\|\mathbf{C}_{7,k}\|\|\mathbf{C}_{8,k}\|\|\mathbf{C}_4\|\|Z\|^2\|x(0)\|^2 + \mathcal{O}(\epsilon)\end{aligned}$$

By definition of the variable Y_k in equation (37), it satisfies $\|Y_k\| \leq \sqrt{n_k}\|H\|\|x(0)\| + \mathcal{O}(\epsilon)$ and substituting it in the above equation leads to

$$J_k^{cross} \leq \boldsymbol{C}_{10,k}\|x(0)\|^2 + \mathcal{O}(\epsilon), \tag{A.27}$$

where $\boldsymbol{C}_{10,k} := 2\|R_k\|\|\mathbf{C}_{8,k}\|\|\mathbf{C}_4\|\big(\sqrt{n_k}\|H\|\|\mathbf{C}_{5,k}\| + \|\mathbf{C}_{7,k}\|\|Z\|\big)\|Z\|$. Then from equation (6), (A.23), (A.24), (A.26) and (A.27), we have,

$$J_k \leq \|P_k^{int}\|\|x_k(0)\|^2 + n_k\overline{J}_k^{ext} + \boldsymbol{C}_k\|x(0)\|^2 + \mathcal{O}(\epsilon)$$

where $\boldsymbol{C}_k := (\boldsymbol{C}_{9,k} + \boldsymbol{C}_{10,k})$. Moreover, we have $\|Y_k(0)\| \leq \|x(0)\|^2 + \mathcal{O}(\epsilon)$ and substituting from equation (42),

$$\begin{aligned}J_k &\leq \|P_k^{int}\|\|x_k(0)\|^2 + n_k\gamma_k\|x(0)\|^2 + \boldsymbol{C}_k\|x(0)\|^2 + \mathcal{O}(\epsilon) \\ &\leq (\|P_k^{int}\| + n_k\gamma_k + \boldsymbol{C}_k)\|x(0)\|^2 + \mathcal{O}(\epsilon)\end{aligned} \tag{A.28}$$

References

[1] B. Adhikari, I-C. Morărescu, and E. Panteley. An emerging dynamics approach for synchronization of linear heterogeneous agents interconnected over switching topologies. *IEEE Control Systems Letters*, 5(1):43–48, 2021.

[2] J. Ben Rejeb, I-C. Morărescu, and J. Daafouz. Guaranteed cost control design for synchronization in networks of linear singularly perturbed systems. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pages 1602–1607, 2017.

[3] A. M. Boker, T. R. Nudell, and A. Chakrabortty. On aggregate control of clustered consensus networks. In *2015 American Control Conference (ACC)*, pages 5527–5532, 2015.

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

[5] E. Bıyık and M. Arcak. Area aggregation and time-scale modeling for sparse nonlinear networks. *Systems & Control Letters*, 57(2):142–149, 2008.

[6] J. Chow and P. Kokotovic. Time scale modeling of sparse dynamic networks. *IEEE Transactions on Automatic Control*, 30(8):714–722, Aug 1985.

[7] J. H. Chow, Ed. Time-scale modeling of dynamic networks with applications to power systems. *Lecture Notes in Control and Information Sciences*, 1982.

[8] P. Erdos, A. Rényi, et al. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

[9] H. Jaleel and J. S. Shamma. Decentralized energy aware co-optimization of mobility and communication in multiagent systems. In *2016 IEEE 55th Conference on Decision and Control (CDC)*, pages 2665–2670, 2016.

[10] Rudolf Emil Kalman et al. Contributions to the theory of optimal control. *Bol. Soc. Mat. Mexicana*, 5(2):102–119, 1960.

[11] P. Kokotović, H. K. Khalil, and J. O'Reilly. *Singular Perturbation Methods in Control: Analysis and Design*. Society for Industrial and Applied Mathematics, 1999.

[12] Y. Labit, D. Peaucelle, and D. Henrion. Sedumi interface 1.02: a tool for solving lmi problems with sedumi. In *Proceedings. IEEE International Symposium on Computer Aided Control System Design*, pages 272–277. IEEE, 2002.

[13] A. J. Laub. *Matrix analysis for scientists and engineers*, volume 91. Siam, 2005.

[14] S. Martin, I-C. Morărescu, and D. Nesic. Time scale modeling for consensus in sparse directed networks with time-varying topologies. *2016 IEEE 55th Conference on Decision and Control (CDC)*, Dec 2016.

[15] J. Mytum-Smithson. Wireless sensor networks: An information processing approach. *Sensor Review*, 25(2), Jun 2005.

[16] E. Panteley and A. Loría. Synchronization and dynamic consensus of heterogeneous networked systems.

IEEE Transactions on Automatic Control, 62(8):3758–3773, 2017.

[17] D. Romeres, F. Dörfler, and F. Bullo. Novel results on slow coherency in consensus and power networks. In *2013 European Control Conference (ECC)*, pages 742–747, 2013.

[18] Stéphane Ross and Brahim Chaib-draa. Satisfaction equilibrium: Achieving cooperation in incomplete information games. In Luc Lamontagne and Mario Marchand, editors, *Advances in Artificial Intelligence*, pages 61–72, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[19] J. Veetaseveera, V. S. Varma, I-C. Morărescu, and J. Daafouz. Decentralized control for guaranteed individual costs in a linear multi-agent system: A satisfaction equilibrium approach. *IEEE Control Systems Letters*, 3(4):918–923, 2019.

[20] S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications*. Structural Analysis in the Social Sciences. Cambridge University Press, 1994.

[21] D. J. Watts and S. H. Strogatz. Collective dynamics of "small-world" networks. *Nature*, 393(6684):440–442, Jun 1998.

bikas.jpeg

Bikash Adhikari obtained his Bachelor's degree in Mechanical Engineering from Motilal Nehru National Institute of Tehcnology (MNNIT), Allahabd, India in 2014 and Master's in Erasmus Mundus Dual Masters in Mathematical Modeling (MATHMODS) from University of L'Aquila, Italy and University of Hamburg, Germany in 2018. Currently, he is doing his PhD in Automatic Control Centre de Recherche en Automatique de Nancy (CRAN) in France. His areas of interest are multi-agent systems, singularly perturbed systems and distributed control.

varma.jpeg

Vineeth S Varma obtained his Bachelor's in Physics with Honors from Chennai Mathematical Institute, India in 2008, his dual Masters in Science and Technology from Friedrich-Schiller-University of Jena in 2009 and Warsaw University of Technology in 2010. He was awarded his PhD from LSS/Supélec on energy efficient wireless telecommunications. He did one year of post doctoral research at Singapore University of Technology and Design from 2014-2015. Since 2016, he is CNRS researcher at the Centre de Recherche en Automatique de Nancy (CRAN) in France since 2016. His areas of interest are analysis, control and games over various networks.