2019 IEEE Conference on Decision and Control (CDC)
Nice, France, Dec. 11-13, 2019

Networked control of coupled subsystems: Spectral decomposition and low-dimensional solutions

Shuang Gao and Aditya Mahajan

***Abstract*— In this paper, we investigate optimal networked control of coupled subsystems where the dynamics and the cost couplings depend on an underlying weighted graph. We use the spectral decomposition of the graph adjacency matrix to decompose the overall system into $(L+1)$ systems with decoupled dynamics and cost, where L is the rank of the adjacency matrix. Consequently, the optimal control input at each subsystem can be computed by solving $(L+1)$ decoupled Riccati equations. A salient feature of the result is that the solution complexity depends on the rank of the adjacency matrix rather than the size of the network (i.e., the number of nodes). Therefore, the proposed solution framework provides a scalable method for synthesizing and implementing optimal control laws for large-scale systems.**

I. Introduction

A. Motivation

The recent proliferation of low cost sensors and actuators has resulted in many applications such as Internet of Things, smart grids, smart buildings, etc., where multiple subsystems are connected over a network. In such systems, the evolution of the state of a subsystem depends on its local state and control input and is also influenced by the states and controls of its neighbors.

Such networks, which are often referred to as large-scale systems or complex networks, have been investigated since the early 1970s [1], [2]. Various aspects of such systems have been investigated including issues such as controllability [3], [4], observability [4], [5], control energy metric [6], decentralized control [7]–[10] and adaptive control [10].

A key theme for investigating large-scale systems is to identify conditions under which the optimal control laws may be synthesized and implemented with low-complexity. These include simplified control objectives (e.g., consensus [11]–[13] or synchronization [14]), simplified control inputs (e.g., pinning control [15]–[17] or ensemble control [18]), simplified coupling between subsystems (e.g., symmetric interconnections [4], [7], [8], [19], [20], exchangeable or anonymous subsystems [21], [22], or patterned systems [23]), or approximate optimality as the number of subsystems become large (e.g., mean-field games [24]–[27] or graphon-based control [28]). In this paper, we present a framework for the control of coupled subsystems where network couplings appear in the dynamics, control and cost. For systems where the adjacency matrix has low rank, our proposed framework provides a low-complexity method to synthesize and implement optimal control laws.

S. Gao and A. Mahajan are with the Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada. Emails: `sgao@cim.mcgill.ca, aditya.mahajan@mcgill.ca`.

B. Contributions of this paper

In this paper, we investigate a control system with multiple subsystems connected over an undirected graph. Each subsystem has a local state and takes a local control action. The evolution of the state of each subsystem depends on its local state and local control as well as a weighed combination (which we call the network field) of the states and controls of its neighbors. Each subsystem is also coupled to its multi-hop neighbors via a quadratic cost. The objective is to choose the control inputs of each subsystem to minimize the total cost over time.

The above model is a linear quadratic regulation problem and a centralized solution can be obtained by solving $nd_x \times nd_x$-dimensional Riccati equation, where n is the number of subsystems and d_x is the dimension of the state of each subsystem. In this paper, we propose an alternative solution that has low complexity and may be implemented in a distributed manner.

Our solution proceeds as follows. Let $(\lambda^1, \dots, \lambda^L)$ be the non-zero eigenvalues of the adjacency matrix M for the underlying network and let $(w^1, \dots, w^L)$ be the corresponding orthonormal eigenvectors. We define eigenstates $\{x_i^\ell\}_{\ell=1}^L$ and eigencontrols $\{u_i^\ell\}_{\ell=1}^L$ for each subsystem i as

$$x_i^\ell(t) = \sum_{j=1}^n x_j(t) w_j^\ell w_i^\ell \quad \text{and} \quad u_i^\ell(t) = \sum_{j=1}^n u_j(t) w_j^\ell w_i^\ell$$

and show that the eigenstates $\{x_i^\ell(t)\}_{\ell=1}^L$ have decoupled dynamics that are identical for all subsystems i. We then define auxiliary states and controls

$$\breve{x}_i(t) = x_i(t) - \sum_{\ell=1}^L x_i^\ell(t) \quad \text{and} \quad \breve{u}_i(t) = u_i(t) - \sum_{\ell=1}^L u_i^\ell(t)$$

and show that the auxiliary states $\{\breve{x}_i(t)\}_{i=1}^n$ have identical and decoupled dynamics that don't depend on the eigenstates $\{x_i^\ell(t)\}_{\ell=1}^L$. Next, we show that the instantaneous cost can be decoupled in terms of $(x_i^\ell(t), u_i^\ell(t))$ and $(\breve{x}_i(t), \breve{u}_i(t))$.

Based on the above decomposition, we show that the optimal control input $u_i(t)$ may be written as

$$u_i(t) = -\breve{K}(t)\breve{x}_i(t) - \sum_{\ell=1}^L K^\ell(t) x_i^\ell(t)$$

where the control gains $(\breve{K}(t), K^1(t), \dots, K^L(t))$ are the same for all subsystems and are obtained by solving $L+1$ decoupled Riccati equations.

The decoupling method for generating the optimal control is inspired by [22], [29]. The coupling in this paper it takes into account the network weights and hence is more general than (weighted or unweighted) mean-field coupling in [10], [27], and the spectral decomposition of networks is further required.

The solution has the following salient features.

- The optimal control law is obtained by solving $L+1$ Riccati equations of dimension $d_x \times d_x$. In contrast, obtaining the centralized solution requires solving a $nd_x \times nd_x$-dimensional Riccati equation.
- To implement the optimal control input, subsystem i needs to know the $(L+1)d_x$-dimensional vector $(\breve{x}_i(t), x_i^1(t), \dots, x_i^L(t))$. In contrast, to implement the centralized solution, each subsystem needs to know the nd_x dimensional global state $(x_1(t), \dots, x_n(t))$.

In many real-world network applications [30]–[33], $\operatorname{rank}(M) \ll \dim(M)$ (i.e., $L \ll n$). Therefore, the method proposed in paper leads to considerable simplification in synthesizing and implementing the optimal control law.

C. Notation

We use $\mathbb{N}$ and $\mathbb{R}$ to denote the sets of natural and real numbers. For a matrix A, $A^\intercal$ denotes its transpose. Given vectors $v_1, \dots, v_n$, $V = \operatorname{cols}(v_1, \dots, v_n)$ denotes the matrix formed by horizontally stacking the vectors. $\mathbb{1}_{n\times n}$ denotes the $n \times n$ matrix of ones.

II. System model and problem formulation

A. System model

Consider a network consisting of n nodes connected over an undirected weighted graph $\mathcal{G}(\mathcal{N}, \mathcal{E}, M)$, where $\mathcal{N} = \{1, ..., n\}$ is the set of nodes, $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$ is the unordered set of edges, and $M = [m_{ij}] \in \mathbb{R}^{n\times n}$ is the weighted adjacency matrix.[1] For any node $i \in \mathcal{N}$, let $\mathcal{N}_i \coloneqq \{j \in \mathcal{N} : (i,j) \in \mathcal{E}\}$ denote the set of neighbors of node i.

The system operates in continuous time for a finite horizon $[0,T]$. A state $x_i(t) \in \mathbb{R}^{d_x}$ and a control input $u_i(t) \in \mathbb{R}^{d_u}$ are associated with each node $i \in \mathcal{N}$. At time $t = 0$, the system starts from an initial state $(x_i(0))_{i\in\mathcal{N}}$ and for $t > 0$, the state of node i evolves according to

$$\dot{x}_i(t) = Ax_i(t) + Bu_i(t) + Dz_i(t) + Ev_i(t), \tag{1}$$

where A, B, D and E are matrices of appropriate dimensions and

$$z_i(t) = \sum_{j\in\mathcal{N}_i} m_{ij}x_j(t) \quad \text{and} \quad v_i(t) = \sum_{j\in\mathcal{N}_i} m_{ij}u_j(t) \tag{2}$$

are the locally perceived *network field* of states and control actions at node i.

We follow an atypical representation of the "vectorized" dynamics. Define

$$x(t) = \operatorname{cols}(x_1(t), \dots, x_n(t)),$$
$$u(t) = \operatorname{cols}(u_1(t), \dots, u_n(t)),$$

as the global state and control actions of the system. Similarly define

$$z(t) = \operatorname{cols}(z_1(t), \dots, z_n(t)), v(t) = \operatorname{cols}(v_1(t), \dots, v_n(t)),$$

as the global network field of states and actions. Note that $x(t), z(t) \in \mathbb{R}^{d_x\times n}$ and $u(t), v(t) \in \mathbb{R}^{d_u\times n}$ are matrices and not vectors.

The system dynamics may be written as

$$\dot{x}(t) = Ax(t) + Bu(t) + Dz(t) + Ev(t). \tag{3}$$

Furthermore, we may write

$$z(t) = x(t)M^\intercal = x(t)M \quad \text{and} \quad v(t) = u(t)M^\intercal = u(t)M.$$

B. System performance and control objective

At any time $t \in [0,T)$, the system incurs an instantaneous cost

$$c(x(t), u(t)) = \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}} \big[g_{ij}x_i(t)^\intercal Qx_j(t) + h_{ij}u_i(t)^\intercal Ru_j(t)\big] \tag{4}$$

and at the terminal time T, the system incurs a terminal cost

$$c_T(x(T)) = \sum_{i\in\mathcal{N}}\sum_{j\in\mathcal{N}} g_{ij}x_i(T)^\intercal Q_T x_j(T), \tag{5}$$

where Q, Q_T, and R are matrices of appropriate dimensions and g_{ij} and h_{ij} are real-valued weights. Let $G = [g_{ij}]$ and $H = [h_{ij}]$. We assume that the weight matrices G and H are polynomials of M, i.e., $G = \sum_{k=0}^{K_G} q_k M^k$ and $H = \sum_{k=0}^{K_H} r_k M^k$ where K_G and K_H denote the degree of the polynomials and $\{q_k\}_{k=0}^{K_G}$ and $\{r_k\}_{k=0}^{K_H}$ are real-valued coefficients.

Since M is real and symmetric, it has real eigenvalues. Let L denote the rank of M and $\lambda^1, \dots, \lambda^L$ denote the non-zero eigenvalues. For ease of notation, for $\ell \in \{1, \dots, L\}$, define

$$q^\ell = \sum_{k=0}^{K_G} q_k(\lambda^\ell)^k \quad \text{and} \quad r^\ell = \sum_{k=0}^{K_H} r_k(\lambda^\ell)^k.$$

We impose the following assumptions on the cost function.

(A1) The matrices Q and Q_T are symmetric and positive semi-definite and R is symmetric and positive definite.

(A2) For $\ell \in \{1, \dots, L\}$, q^ℓ is non-negative and r^ℓ is strictly positive; $q_0 \geq 0$ and $r_0 > 0$.

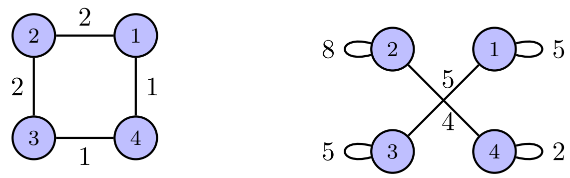
We are interested in the following optimization problem.

Problem 1 *Choose a control trajectory* $u\colon [0,T) \to \mathbb{R}^{d_u\times n}$ *to minimize*

$$J(u) = \int_0^T c(x(t), u(t))dt + c_T(x(T)). \tag{6}$$

Remark 1 The assumption that the system is time-homogeneous is made only for notational simplicity. It will be evident from the solution approach that results extend to systems with time-varying A, B, D, E, Q and R matrices. See e.g. [34]. The assumption (A2) ensures that, for any $y \in \mathbb{R}^n$, $y^\intercal Gy \geq 0$ and $y^\intercal Hy > 0$. □

[1] Since the graph is undirected, M is symmetric.



(a) A graph $\mathcal{G}$ (b) 2-hop neighborhood of $\mathcal{G}$

Fig. 1: A graph and its 2-hop neighborhood.

C. Salient features of the model

We highlight salient features of the model via an example. Consider a system with 4 nodes connected via a network shown in Fig. 1(a), with

$$G = q_0 I + q_1 M + q_2 M^2 \text{ and } H = r_0 I + r_1 M + r_2 M^2,$$

where M and M^2 are the weighted adjacency matrix of the graph $\mathcal{G}$ and that of the 2-hop neighborhood of $\mathcal{G}$, respectively, given by

$$M = \begin{bmatrix} 0 & 2 & 0 & 1 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \quad \text{and} \quad M^2 = \begin{bmatrix} 5 & 0 & 5 & 0 \\ 0 & 8 & 0 & 4 \\ 5 & 0 & 5 & 0 \\ 0 & 4 & 0 & 2 \end{bmatrix}.$$

1) Salient features of the dynamics: For this example,

$$z_1(t) = 2x_1(t) + x_4(t), \quad z_2(t) = 2x_1(t) + 2x_3(t),$$
$$z_3(t) = 2x_2(t) + x_4(t), \quad z_4(t) = x_1(t) + x_3(t).$$

Thus, each subsystem is affected by its neighbors. The influence of each neighbor is not homogeneous but depends on the weight associated with the corresponding edge in the graph. Furthermore, the "network field" $z(t)$ is not homogeneous and varies from subsystem to subsystem.

2) Salient features of the cost: Since M is the weighted adjacency matrix of the graph $\mathcal{G}$, the matrix M^k, $k \in \mathbb{N}$, represents the weighted adjacency matrix of the k-hop neighborhood of $\mathcal{G}$. Thus, $G = q_0 I + q_1 M + q_2 M^2$ means that the each node has a coupling of q_0 with its own state, a coupling of q_1 with it's 1-hop neighborhood and a coupling of q_2 with its 2-hop neighborhood. Similar interpretation holds for H. Note that

$$G = q_0 I + q_1 M + q_2 M^2 = \begin{bmatrix} q_0+5q_2 & 2q_1 & 5q_2 & q_0+q_1 \\ 2q_1 & q_0+8q_2 & 2q_1 & 4q_2 \\ 5q_2 & 2q_1 & q_0+5q_2 & q_1 \\ q_1 & 4q_2 & q_1 & q_0+2q_2 \end{bmatrix}.$$

Thus, the agents are not interchangeable, i.e., in general, $G_{ii} \neq G_{jj}$ and $G_{ki} \neq G_{kj}$.

III. Spectral decomposition of the system

Since the weight matrix M is real and symmetric, it admits a spectral factorization. In particular, there exist non-zero eigenvalues $(\lambda^1, \dots, \lambda^L)$ and orthonormal eigenvectors $(w^1, \dots, w^L)$ such that

$$M = \sum_{\ell=1}^{L} \lambda^\ell w^\ell w^{\ell\mathsf{T}}. \tag{7}$$

In the rest of this section, we decompose the dynamics and the cost based on the above spectral decomposition.

A. Spectral decomposition of the dynamics

For $\ell \in \{1, \dots, L\}$, define eigenstates and eigencontrol actions as

$$x^\ell(t) = x(t) w^\ell w^{\ell\mathsf{T}}, \tag{8}$$
$$u^\ell(t) = u(t) w^\ell w^{\ell\mathsf{T}}, \tag{9}$$

respectively. Multiplying both sides of (3) by $w^\ell w^{\ell\mathsf{T}}$, we get

$$\dot{x}^\ell(t) = (A + \lambda^\ell D) x^\ell(t) + (B + \lambda^\ell E) u^\ell(t), \tag{10}$$

where we have used the fact that $M w^\ell w^{\ell\mathsf{T}} = \lambda^\ell w^\ell w^{\ell\mathsf{T}}$. Let $x_i^\ell(t)$ and $u_i^\ell(t)$ denote the i-th column of these matrices, i.e.,

$$x^\ell(t) = \operatorname{cols}(x_1^\ell(t), \dots, x_n^\ell(t)),$$
$$u^\ell(t) = \operatorname{cols}(u_1^\ell(t), \dots, u_n^\ell(t)).$$

Therefore, the dynamics (10) can be written as a collection of decoupled "local" dynamics: for $i \in \mathcal{N}$,

$$\dot{x}_i^\ell(t) = (A + \lambda^\ell D) x_i^\ell(t) + (B + \lambda^\ell E) u_i^\ell(t). \tag{11}$$

Using the spectral factorization (7), we may write:

$$z(t) = x(t) M = \sum_{\ell=1}^{L} \lambda^\ell x^\ell(t), \tag{12}$$
$$v(t) = u(t) M = \sum_{\ell=1}^{L} \lambda^\ell u^\ell(t). \tag{13}$$

Now, define auxiliary state and control actions as

$$\breve{x}(t) = x(t) - \sum_{\ell=1}^{L} x^\ell(t) \quad \text{and} \quad \breve{u}(t) = u(t) - \sum_{\ell=1}^{L} u^\ell(t).$$

Then, by subtracting (10) from (3) and substituting (12) and (13), we get

$$\dot{\breve{x}}(t) = A \breve{x}(t) + B \breve{u}(t). \tag{14}$$

Note that $\breve{x}(t) \in \mathbb{R}^{d_x \times n}$ and $\breve{u}(t) \in \mathbb{R}^{d_u \times n}$. Let $\breve{x}_i(t)$ and $\breve{u}_i(t)$ denote the i-th column of these matrices, i.e.,

$$\breve{x}(t) = \operatorname{cols}(\breve{x}_1(t), \dots, \breve{x}_n(t)),$$
$$\breve{u}(t) = \operatorname{cols}(\breve{u}_1(t), \dots, \breve{u}_n(t)).$$

Therefore, the dynamics (14) of the auxiliary state can be written as a collection of decoupled "local" dynamics:

$$\dot{\breve{x}}_i(t) = A \breve{x}_i(t) + B \breve{u}_i(t), \quad i \in \mathcal{N}. \tag{15}$$

The above decomposition may be summarized as follows.

Proposition 1 *The local state and control at each node $i \in \mathcal{N}$ may be decomposed as*

$$x_i(t) = \breve{x}_i(t) + \sum_{\ell=1}^{L} x_i^\ell(t), \tag{16}$$
$$u_i(t) = \breve{u}_i(t) + \sum_{\ell=1}^{L} u_i^\ell(t), \tag{17}$$

where the dynamics of $\breve{x}_i(t)$ depend on only $\breve{u}_i(t)$ and are given by (15) *and the dynamics of $x_i^\ell(t)$ depends on only $u_i^\ell(t)$ and are given by* (11). □

B. Spectral decomposition of the cost

For any $n \times n$ matrix $P = [p_{ij}]$, any $d \times n$ matrices $x = \operatorname{cols}(x_1, \dots, x_n)$ and $y = \operatorname{cols}(y_1, \dots, y_n)$, we use the following short hand notation:

$$\langle x, y \rangle_P = \sum_{i \in \mathcal{N}} \sum_{j \in \mathcal{N}} p_{ij} x_i^\intercal y_j. \tag{18}$$

Proposition 2 *The instantaneous cost may be written as*

$$c(x(t), u(t)) = \langle x(t), Qx(t) \rangle_G + \langle u(t), Ru(t) \rangle_H,$$

which can be simplified as follows:

$$\begin{aligned}
&\langle x(t), Qx(t) \rangle_G \\
&\qquad = \sum_{i \in \mathcal{N}} \Big[q_0 \breve{x}_i(t)^\intercal Q \breve{x}_i(t) + \sum_{\ell=1}^{L} q^\ell x_i^\ell(t)^\intercal Q x_i^\ell(t) \Big], \\
&\langle u(t), Ru(t) \rangle_H \\
&\qquad = \sum_{i \in \mathcal{N}} \Big[r_0 \breve{u}_i(t)^\intercal R \breve{u}_i(t) + \sum_{\ell=1}^{L} r^\ell u_i^\ell(t)^\intercal R u_i^\ell(t) \Big]. \quad \Box
\end{aligned}$$

See Appendix for the proof.

IV. Main result: structure and synthesis of optimal control strategies

The main result of the paper is the following.

Theorem 1 *For $\ell \in \{1, \dots, L\}$, let $P^\ell \colon [0,T] \to \mathbb{R}^{d_x \times d_x}$ be the solution to the backward Riccati differential equation*

$$\begin{aligned}
-\dot{P}^\ell(t) &= (A + \lambda^\ell D)^\intercal P^\ell(t) + P^\ell(t)(A + \lambda^\ell D) \\
&\quad - P^\ell(t)(B + \lambda^\ell E)(r^\ell R)^{-1}(B + \lambda^\ell E)^\intercal P^\ell(t) + q^\ell Q
\end{aligned} \tag{19}$$

with the final condition $P^\ell(T) = q^\ell Q_T$. Similarly, let $\breve{P} \colon [0,T] \to \mathbb{R}^{d_x \times d_x}$ be the solution to the backward Riccati differential equation

$$-\dot{\breve{P}}(t) = A^\intercal \breve{P}(t) + \breve{P}(t) A - \breve{P}(t) B (r_0 R)^{-1} B^\intercal \breve{P}(t) + q_0 Q \tag{20}$$

with the final condition $\breve{P}(T) = q_0 Q_T$.

Then, under assumptions (A1) and (A2), the optimal control strategy for Problem 1 is given by

$$u_i(t) = -\breve{K}(t)\breve{x}_i(t) - \sum_{\ell=1}^{L} K^\ell(t) x_i^\ell(t), \tag{21}$$

where

$$\begin{aligned}
\breve{K}(t) &= (r_0 R)^{-1} B^\intercal \breve{P}(t), \\
K^\ell(t) &= (r^\ell R)^{-1} (B + \lambda^\ell E)^\intercal P^\ell(t).
\end{aligned} \qquad \Box$$

PROOF Consider the following collections of dynamical systems:

- Eigensystem (ℓ, i), $\ell \in \{1, \dots, L\}$, $i \in \mathcal{N}$, with state $x_i^\ell(t)$, control inputs $u_i^\ell(t)$, dynamics

$$\dot{x}_i^\ell(t) = (A + \lambda^\ell D) x_i^\ell(t) + (B + \lambda^\ell E) u_i^\ell(t),$$

and cost

$$\begin{aligned}
J_i^\ell(u_i^\ell) = \int_0^T &\big[q^\ell x_i^\ell(t)^\intercal Q x_i^\ell(t) + r^\ell u_i^\ell(t)^\intercal R u_i^\ell(t) \big] dt \\
&+ q^\ell x_i^\ell(T)^\intercal Q x_i^\ell(T).
\end{aligned}$$

- Auxiliary system i, $i \in \mathcal{N}$, with state $\breve{x}_i(t)$, control inputs $\breve{u}_i(t)$, dynamics

$$\dot{\breve{x}}_i(t) = A \breve{x}_i(t) + B \breve{u}_i(t),$$

and cost

$$\begin{aligned}
\breve{J}_i(\breve{u}_i) = \int_0^T &\big[q_0 \breve{x}_i(t)^\intercal Q \breve{x}_i(t) + r_0 \breve{u}_i(t)^\intercal R \breve{u}_i(t) \big] dt \\
&+ q_0 \breve{x}_i(T)^\intercal Q \breve{x}_i(T).
\end{aligned}$$

Note that all systems have decoupled dynamics and decoupled nonnegative cost. By Proposition 2, we have

$$J(u) = \sum_{i \in \mathcal{N}} \Big[\breve{J}_i(\breve{u}_i) + \sum_{\ell=1}^{L} J_i^\ell(u_i^\ell) \Big].$$

Thus, instead of solving:

- **(P1)** choose control trajectory $u \colon [0,T) \to \mathbb{R}^{d_u \times n}$ to minimize $J(u)$,

we can equivalently solve the following optimization problems:

- **(P2)** choose control trajectory $u_i^\ell \colon [0,T) \to \mathbb{R}^{d_u}$ to minimize $J_i^\ell(u_i^\ell)$ for $i \in \mathcal{N}$, $\ell \in \{1, \dots, L\}$,
- **(P3)** choose control trajectory $\breve{u}_i \colon [0,T) \to \mathbb{R}^{d_u}$ to minimize $\breve{J}_i(\breve{u}_i)$ for $i \in \mathcal{N}$.

Given the solutions of Problems (P2) and (P3), we can use Proposition 1 and choose $u_i(t)$ according to (17).

Problems (P2) and (P3) are standard optimal control problems and their solution are given as follows. Let $P^\ell \colon [0,T] \to \mathbb{R}^{d_x \times d_x}$ and $\breve{P} \colon [0,T] \to \mathbb{R}^{d_x \times d_x}$ be as given by (19) and (20). Then, for all $i \in \mathcal{N}$, the optimal solution of (P2) is given by $u_i^\ell(t) = K^\ell(t) x_i^\ell(t), \ell \in \{1, \dots, L\}$, and the solution of (P3) is given by $\breve{u}_i(t) = \breve{K}(t)\breve{x}_i(t)$. The result follows by combining the above two equations using (17). ■

A. Remarks on the implementation of the optimal strategy

Since we are interested in regulating a deterministic system, we may implement the optimal control law either using open-loop (i.e. pre-computed) control inputs or using closed-loop (i.e. state feedback) control inputs. For the both implementations, the eigenvalues $\{\lambda^\ell\}_{\ell=1}^{L}$ need to be known at all subsystems.

For the open-loop implementation, one can write

$$u_i(t) = -\breve{K}(t)\breve{\Phi}(t,0)\breve{x}_i(0) - \sum_{\ell=1}^{L} K^\ell(t)\Phi^\ell(t,0)x_i^\ell(0), \tag{22}$$

where the state transition matrices $\breve{\Phi}(t,0)$ and $\Phi^\ell(t,0)$ are given by

$$\begin{aligned}
\breve{\Phi}(t,0) &= \exp\Big(\textstyle\int_0^t \big(A - B\breve{K}(s)\big) ds \Big), \\
\Phi^\ell(t,0) &= \exp\Big(\textstyle\int_0^t \big(A + \lambda^\ell D - (B + \lambda^\ell E) K^\ell(s)\big) ds \Big).
\end{aligned}$$

Thus, to implement the control action, subsystem i needs to know $\breve{x}_i(0)$ and $\{x_i^\ell(0)\}_{\ell=1}^L$, which can be obtained using one of the following three information structures:

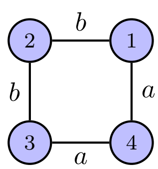
1) All subsystems know the initial condition $x(0)$ and the eigendirections $\{w^\ell\}_{\ell=1}^L$. Using these, subsystem i can compute $\{x_i^\ell(0)\}_{\ell=1}^L$ and $\breve{x}_i(0)$, and implement (22).
2) Subsystem $i, i \in \mathcal{N}$, knows its local initial state $x_i(0)$ and its local initial eigensystem states $\{x_i^\ell(0)\}_{\ell=1}^L$. Then subsystem i can compute $\breve{x}_i(0)$ and implement (22).
3) All subsystems knows the initial state $\{x(0)w^\ell\}_{\ell=1}^L$. In addition, subsystem i knows $w_i := (w_i^1, \cdots, w_i^L)$ and its local initial state $x_i(0)$. Then subsystem i can compute $\{x_i^\ell(0)\}_{\ell=1}^L$ and $\breve{x}_i(0)$, and implement (22).

The closed-loop implementation, which is given by (21), can be obtained by using one of the three information structures described above where $x(0)$, $x_i(0)$ and $x_i^\ell(0)$ are replaced by $x(t)$, $x_i(t)$ and $x_i^\ell(t)$, respectively.

Furthermore, for the information structures in 2) and 3), a mixed implementation which combines open-loop and close-loop implementations can also be obtained via only replacing $x_i(0)$ by $x_i(t)$ in 2) and 3).

B. An illustrative example

Consider a network with $n = 4$ subsystems connected over a graph $\mathcal{G}$, as shown in Fig. 2, with the adjacency matrix M. Note that $L = \operatorname{rank}(M) = 2$. Consider the following



$$\begin{bmatrix} 0 & a & 0 & b \\ a & 0 & a & 0 \\ 0 & a & 0 & b \\ b & 0 & b & 0 \end{bmatrix}$$

Fig. 2: Graph $\mathcal{G}$ with $n = 4$ nodes and its adjacency matrix

couplings in the cost

$$G = I - 2M + M^2 \quad \text{and} \quad H = I. \tag{23}$$

For the ease of notation define $\rho = \sqrt{(a^2 + b^2)/2}$ and $\theta = \tan^{-1}(b/a)$. Then it is easy to verify that the non-zero eigenvalues of M are $\lambda^1 = -\rho$ and $\lambda^2 = \rho$. The corresponding eigenvectors are $\omega^1 = \left[-\frac{1}{2} \; \frac{\sin(\theta)}{\sqrt{2}} \; -\frac{1}{2} \; \frac{\cos(\theta)}{\sqrt{2}} \right]^\intercal$ and $\omega^2 = \left[\frac{1}{2} \; \frac{\sin(\theta)}{\sqrt{2}} \; \frac{1}{2} \; \frac{\cos(\theta)}{\sqrt{2}} \right]^\intercal$. Observe that $q^\ell = (1 - \lambda^\ell)^2$ is non-negative and $r^\ell = 1$ is strictly positive, $\ell \in \{1, 2\}$. Thus the model satisfies assumption (A2).

To illustrate how to use the result of Theorem 1, let's pick a subsystem say subsystem 1, and consider the calculations that need to be carried out at that subsystem. Recall that for all $i \in \mathcal{N}$, $x_i^\ell(0) = x(0)\omega^\ell w_i^\ell$. Thus

$$x_1^1(0) = \frac{1}{4}x_1(0) - \frac{\sin(\theta)}{2\sqrt{2}}x_2(0) + \frac{1}{4}x_3(0) - \frac{\cos(\theta)}{2\sqrt{2}}x_4(0),$$
$$x_1^2(0) = \frac{1}{4}x_1(0) + \frac{\sin(\theta)}{2\sqrt{2}}x_2(0) + \frac{1}{4}x_3(0) + \frac{\cos(\theta)}{2\sqrt{2}}x_4(0).$$

Following the mixed implementation with information structure 3) described in Section IV-A, subsystem 1 can calculate the trajectory for $x_1^1(t)$, $x_1^2(t)$, $t \in (0, T]$ based on the initial conditions. This together with real time local observation $x_1(t)$ yields $\breve{x}_1(t)$.

Subsystem 1 solves three Riccati equations to compute $P^1(t)$, $P^2(t)$, and $\breve{P}(t)$ for $t \in [0, T]$, and then applies the optimal control action given by

$$u_1(t) = -R^{-1}\bigg(B^\intercal \breve{P}(t)\breve{x}_1(t) + (B - \rho E)^\intercal P^1(t)x_1^1(t) + (B + \rho E)^\intercal P^2(t)x_1^2(t)\bigg)$$

according to Theorem 1. Similar implementations hold for other subsystems.

Note that if each $x_i(t) \in \mathbb{R}^{d_x}$ then $x(t) \in \mathbb{R}^{4d_x}$. A naive centralized optimal solution of the above system would involve solving a $4d_x \times 4d_x$-dimensional Riccati equation. In contrast, the above solution involves solving three $d_x \times d_x$-dimensional Riccati equations. These computational savings increase with the size of the networks. For example, consider the graph $\mathcal{G}_{4n} = \mathcal{G} \otimes \mathcal{K}_n$ where $\mathcal{G}$ is the 4-node graph shown in Fig. 2 and $\mathcal{K}_n$ is the complete graph with n-nodes and each edge weight is $\frac{1}{n}$. The graph $\mathcal{G}_{4n}$ has $4n$ nodes and its adjacency matrix is given by $M_{4n} = M \otimes K_n$, where M and $K_n = \frac{1}{n}\mathbb{1}_{n \times n}$ are the adjacency matrices of graph $\mathcal{G}$ and $\mathcal{K}_n$ respectively. The only non-zero eigenvalue of K_n is 1. Thus, the eigenvalues of M_{4n} are the same as eigenvalues of M. The corresponding eigenvectors are different. Note also that the Riccati equations in Theorem 1 only depends on the eigenvalues. So the Riccati equations for all graphs $\mathcal{G}_{4n}, n \in \mathbb{N}$ are the same.

Thus, a naive solution requires solving a $4nd_x \times 4nd_x$-dimensional Riccati equation. In contrast, the method proposed in Theorem 1 would require solving the same three $d_x \times d_x$-dimensional Riccati equations as above.

As an illustration, we consider graph $\mathcal{G}$ with the weights $a = 2$ and $b = 1$. Let $d_x = 1$ and $d_u = 1$. Consider the dynamics with parameters $A = 2, B = 1, D = 3, E = 0.5$ and the cost with parameters $Q = 5, Q_T = 6, R = 2$. Recall that G and H are given by (23). As argued above, the matrix M_{4n} has two non-zero eigenvalues and the optimal control at each subsystem can be obtained by solving 3 Riccati equations. Let us set $n = 5$. Then $M_{20} = M \otimes \frac{1}{5}\mathbb{1}_{5\times5}$. The evolutions of the corresponding eigenstates and the auxiliary states along with the eigencontrols and the auxiliary controls are shown in Fig. 3.

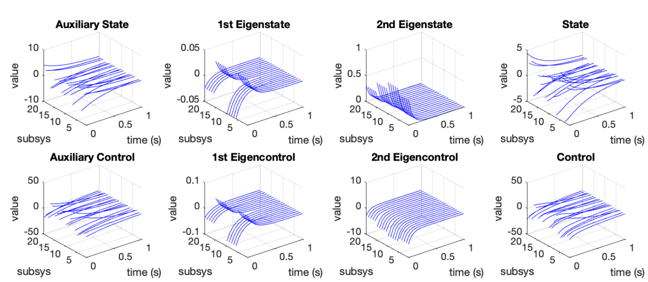


Fig. 3: Numerical example under the proposed optimal control on a network of size 20

C. A special case: mean-field coupling

Suppose the graph $\mathcal{G}$ is a complete graph with all edge weights equal to $\frac{1}{n}$. Then, the adjacency matrix $M = \frac{1}{n}\mathbb{1}_{n\times n}$ has rank 1 and $\lambda_1 = 1$ is the only non-zero eigenvalue with the normalized eigenvector $w^1 = \frac{1}{\sqrt{n}}[1,\dots,1]^\intercal$. Then $x^1(t) = x(t)w^1{w^1}^\intercal = x(t)M$. Thus, the eigenstate $x_i^1(t) = \frac{1}{n}\sum_{j=1}^n x_j(t), i \in \mathcal{N}$, is the same for all subsystems and we denote it by $\bar{x}(t)$. Moreover, $q^1 = \sum_{k=0}^{K_G} q_k := \bar{q}$ and $r^1 = \sum_{k=0}^{K_H} r_k := \bar{r}$. According to Theorem 1, the Riccati equation of eigensystem is given by

$$\begin{aligned} -\dot{\bar{P}}(t) &= (A+D)^\intercal \bar{P}(t) + \bar{P}(t)(A+D) \\ &\quad - \bar{P}(t)(B+E)(\bar{r}R)^{-1}(B+E)^\intercal \bar{P}(t) + \bar{q}Q, \quad (24) \end{aligned}$$

where $\bar{P}(t) := P^1(t)$ and the final condition $\bar{P}(T) = \bar{q}Q_T$. The Riccati equation for the auxiliary system is given by

$$-\dot{\breve{P}}(t) = A^\intercal \breve{P}(t) + \breve{P}(t)A - \breve{P}(t)B(r_0R)^{-1}B^\intercal \breve{P}(t) + q_0 Q$$

with the final condition $\breve{P}(T) = q_0 Q_T$. The optimal control strategy is given by $u_i(t) = -\breve{K}(t)(x_i(t) - \bar{x}(t)) - \bar{K}(t)\bar{x}(t)$, where $\breve{K}(t) = (r_0R)^{-1}B^\intercal \breve{P}(t)$ and $\bar{K}(t) = (\bar{r}R)^{-1}(B+E)^\intercal \bar{P}(t)$.

The above result is similar in spirit to [22, Theorem 1].

V. Conclusion

We consider the optimal networked control of coupled subsystems where the dynamics and the cost couplings depend on an underlying weighted graph. The main idea of a low-dimensional decomposition is to project the state $x(t)$ into L orthogonal eigendirections which generates $\{x^\ell(t)\}_{\ell=1}^L$ and an auxiliary state $\breve{x}(t) = x(t) - \sum_{\ell=1}^L x^\ell(t)$. A similar decomposition is obtained for the control inputs. These $L+1$ components are decoupled both in dynamics and cost. Therefore, the optimal control input for each component can be obtained by solving decoupled Riccati equations.

The proposed approach requires solving $L+1$ Riccati equations, each of dimension $d_x \times d_x$. In contrast, a centralized solution requires solving a $nd_x \times nd_x$-dimensional Riccati equation. Thus, even when $L = n$, the proposed approach leads to considerable computational savings. These savings improve significantly when $L \ll n$, as is the case for many real-world networks.

References

[1] Y. Ho and S. Mitter, *Directions in Decentralized Control, Many-Person Optimization, and Large Scale Systems*. Plenum, New York, 1976.

[2] N. Sandell, P. Varaiya, M. Athans, and M. Safonov, "Survey of decentralized control methods for large scale systems," *IEEE Trans. Autom. Control*, vol. 23, no. 2, pp. 108–128, 1978.

[3] Y.-Y. Liu, J.-J. Slotine, and A.-L. Barabási, "Controllability of complex networks," *Nature*, vol. 473, no. 7346, pp. 167–173, 2011.

[4] G.-H. Yang and S.-Y. Zhang, "Structural properties of large-scale systems possessing similar structures," *Automatica*, vol. 31, no. 7, pp. 1011–1017, 1995.

[5] G. Yan, G. Tsekenis, B. Barzel, J.-J. Slotine, Y.-Y. Liu, and A.-L. Barabási, "Spectrum of controlling and observing complex networks," *Nature Physics*, vol. 11, no. 9, pp. 779–786, 2015.

[6] F. Pasqualetti, S. Zampieri, and F. Bullo, "Controllability metrics, limitations and algorithms for complex networks," *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 1, pp. 40–52, 2014.

[7] M. K. Sundareshan and R. M. Elbanna, "Qualitative analysis and decentralized controller synthesis for a class of large-scale systems with symmetrically interconnected subsystems," *Automatica*, vol. 27, no. 2, pp. 383–388, 1991.

[8] G.-H. Yang and S.-Y. Zhang, "Decentralized control of a class of large-scale systems with symmetrically interconnected subsystems," *IEEE Trans. Autom. Control*, vol. 41, no. 5, pp. 710–713, 1996.

[9] D. D. Šiljak and A. Zečević, "Control of large-scale systems: Beyond decentralized feedback," *Annual Reviews in Control*, vol. 29, no. 2, pp. 169–179, 2005.

[10] T. Bian, Y. Jiang, and Z.-P. Jiang, "Decentralized adaptive optimal control of large-scale systems with application to power systems," *IEEE Trans. Ind. Electron.*, vol. 62, no. 4, pp. 2439–2447, 2015.

[11] R. Olfati-Saber and R. M. Murray, "Consensus protocols for networks of dynamic agents," in *Proc. American Control Conf.*, vol. 2. IEEE, 2003, pp. 951–956.

[12] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proc. IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[13] K. H. Movric and F. L. Lewis, "Cooperative optimal control for multi-agent systems on directed graph topologies," *IEEE Trans. Autom. Control*, vol. 59, no. 3, pp. 769–774, 2013.

[14] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, "Synchronization in complex networks," *Physics Reports*, vol. 469, no. 3, pp. 93–153, 2008.

[15] H. Gang and Q. Zhilin, "Controlling spatiotemporal chaos in coupled map lattice systems," *Physical Review Letters*, vol. 72, no. 1, p. 68, 1994.

[16] R. Grigoriev, M. Cross, and H. Schuster, "Pinning control of spatiotemporal chaos," *Physical Review Letters*, vol. 79, no. 15, p. 2795, 1997.

[17] X. F. Wang and G. Chen, "Pinning control of scale-free dynamical networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 310, no. 3-4, pp. 521–531, 2002.

[18] J.-S. Li, "Ensemble control of finite-dimensional time-varying linear systems," *IEEE Trans. Autom. Control*, vol. 56, no. 2, pp. 345–357, 2011.

[19] J. Lunze, "Dynamics of strongly coupled symmetric composite systems," *International Journal of Control*, vol. 44, no. 6, pp. 1617–1640, 1986.

[20] J. Grizzle and S. Marcus, "The structure of nonlinear control systems possessing symmetries," *IEEE Trans. Autom. Control*, vol. 30, no. 3, pp. 248–258, 1985.

[21] D. Madjidian and L. Mirkin, "Distributed control with low-rank coordination," *IEEE Trans. Control Netw. Syst.*, vol. 1, no. 1, pp. 53–63, 2014.

[22] J. Arabneydi and A. Mahajan, "Linear quadratic mean field teams: Optimal and approximately optimal decentralized solutions," *arXiv preprint arXiv:1609.00056v2*, 2017.

[23] S. C. Hamilton and M. E. Broucke, "Patterned linear systems," *Automatica*, vol. 48, no. 2, pp. 263–272, 2012.

[24] M. Huang, P. E. Caines, and R. P. Malhamé, "Individual and mass behaviour in large population stochastic wireless power control problems: centralized and nash equilibrium solutions," in *Proc. Conf. Decision and Control*, 2003, pp. 98–103.

[25] M. Huang, R. P. Malhamé, and P. E. Caines, "Large population stochastic dynamic games: closed-loop McKean-Vlasov systems and the nash certainty equivalence principle," *Communications in Information & Systems*, vol. 6, no. 3, pp. 221–252, 2006.

[26] J.-M. Lasry and P.-L. Lions, "Jeux à champ moyen. i–le cas stationnaire," *Comptes Rendus Mathématique*, vol. 343, no. 9, pp. 619–625, 2006.

[27] T. Li and J.-F. Zhang, "Asymptotically optimal decentralized control for large population stochastic multiagent systems," *IEEE Trans. Autom. Control*, vol. 53, no. 7, pp. 1643–1660, 2008.

[28] S. Gao and P. E. Caines, "Graphon control of large-scale networks of linear systems," *arXiv preprint arXiv:1807.03412*, 2018.

[29] J. Arabneydi and A. Mahajan, "Team-optimal solution of finite number of mean-field coupled LQG subsystems," in *Proc. Conf. Decision and Control*, Dec 2015, pp. 5308–5313.

[30] M. Udell and A. Townsend, "Why are big data matrices approximately low rank?" *SIAM Journal on Mathematics of Data Science*, vol. 1, no. 1, pp. 144–160, 2019.

[31] K. Zhou, H. Zha, and L. Song, "Learning social infectivity in sparse low-rank networks using multi-dimensional hawkes processes," in *Artificial Intelligence and Statistics*, 2013, pp. 641–649.

[32] S. Gao and P. E. Caines, "Spectral representations of graphons in very large network systems control," Accepted by *the 58th IEEE Conference on Decision and Control (CDC)*, December 2019.

[33] B. D. MacArthur, R. J. Sánchez-García, and J. W. Anderson, "Symmetry in complex networks," *Discrete Applied Mathematics*, vol. 156, no. 18, pp. 3525–3531, 2008.

[34] S. Gao, "Graphon control theory for linear systems on complex networks and related topics," Ph.D. dissertation, McGill University, 2019.

APPENDIX

A. Preliminary properties of the state decomposition

Lemma 1 *Let k be a positive integer k and $\ell, \ell' \in \{1, \dots, L\}$. Then, we have the following:*

(P1) $x^\ell(t)M = \lambda^\ell x^\ell(t)$ *and* $u^\ell(t)M = \lambda^\ell u^\ell(t)$.

(P2) $x^\ell(t)M^k = (\lambda^\ell)^k x^\ell(t)$ *and* $u^\ell(t)M^k = (\lambda^\ell)^k u^\ell(t)$.

(P3) $x^\ell(t)G = q^\ell x^\ell(t)$ *and* $u^\ell(t)H = r^\ell u^\ell(t)$.

(P4) $\breve{x}(t)M = 0$ *and* $\breve{u}(t)M = 0$.

(P5) $\breve{x}(t)M^k = 0$ *and* $\breve{u}(t)M^k = 0$.

(P6) $\breve{x}(t)G = q_0\breve{x}(t)$ *and* $\breve{u}(t)H = r_0\breve{u}(t)$.

(P7) $x(t)G = q_0\breve{x}(t) + \sum_{\ell=1}^L q^\ell x^\ell(t)$ *and* $u(t)G = r_0\breve{u}(t) + \sum_{\ell=1}^L r^\ell u^\ell(t)$.

(P8) $\sum_{i\in\mathcal{N}} x_i^\ell(t)^\intercal Q x_i^{\ell'}(t) = \delta_{\ell\ell'} \sum_{i\in\mathcal{N}} x_i^\ell(t)^\intercal Q x_i^{\ell'}(t)$, *where $\delta_{\ell\ell'}$ is the Kronecker delta function.*

(P9) $\sum_{i\in\mathcal{N}} x_i(t)^\intercal Q x_i^\ell(t) = \sum_{i\in\mathcal{N}} x_i^\ell(t)^\intercal Q x_i^\ell(t)$ *and* $\sum_{i\in\mathcal{N}} u_i(t)^\intercal R u_i^\ell(t) = \sum_{i\in\mathcal{N}} u_i^\ell(t)^\intercal R u_i^\ell(t)$ □

PROOF We show the result for $\breve{x}(t)$. The result for $\breve{u}(t)$ follows from a similar argument.

Since $w^1, \dots, w^L$ are orthonormal, from (7) we have $w^\ell {w^\ell}^\intercal M = \lambda^\ell w^\ell {w^\ell}^\intercal$, which implies (P1). (P2) follows immediately from (P1) and (P3) follows from (P2).

(P4) follows immediately from the definition of $\breve{x}(t)$, (12) and (P1). (P5) follows immediately from (P4) and (P6) follows from (P5).

(P7) follows from (16), (P3) and (P6). To prove (P8), we observe that (8) implies that

$$\begin{aligned}\sum_{i\in\mathcal{N}} x_i^\ell(t)^\intercal Q x_i^{\ell'}(t) &= \sum_{i\in\mathcal{N}} w_i^\ell {w^\ell}^\intercal x(t)^\intercal Q x(t) w^{\ell'} {w_i^{\ell'}}^\intercal \\ &= \Big(\sum_{i\in\mathcal{N}} w_i^\ell w_i^{\ell'}\Big) {w^\ell}^\intercal x(t)^\intercal Q x(t) w^{\ell'}. \quad (25)\end{aligned}$$

Since $w^1, \dots, w^L$ is orthonormal, we get $\sum_{i\in\mathcal{N}} w_i^\ell w_i^{\ell'} = {w^\ell}^\intercal w^{\ell'} = \delta_{\ell\ell'}$. Substituting this in (25) completes the proof of (P8). To prove (P9) observe that

$$\begin{aligned}\sum_{i\in\mathcal{N}} x_i(t)^\intercal Q x_i^\ell(t) &= \sum_{i\in\mathcal{N}} x_i(t)^\intercal Q x(t) w^\ell w_i^\ell \\ &= \sum_{i\in\mathcal{N}} w_i^\ell x_i(t)^\intercal Q x(t) w^\ell \\ &= {w^\ell}^\intercal x(t)^\intercal Q x(t) w^\ell. \quad (26)\end{aligned}$$

From (25), we get that the expression in (26) is equal to $\sum_{i\in\mathcal{N}} x_i^\ell(t)^\intercal Q x_i^\ell(t)$. ■

Lemma 2 *Let P, x, and y be defined in* (18). *Let P_i denote the i-th column of P. Then, we can write*

$$\langle x, y\rangle_P = \sum_{i\in\mathcal{N}} x_i^\intercal y P_i \quad \textit{or} \quad \langle x, y\rangle_P = \sum_{j\in\mathcal{N}} P_j^\intercal x^\intercal y_j.$$ □

PROOF The result follows immediately from the definition of $\langle x, y\rangle_P$. ■

B. Proof for Proposition 2

We consider the terms depending on $x(t)$. The term depending on $u(t)$ may be simplified in a similar manner.

From (16) and linearity of $\langle\cdot,\cdot\rangle_G$ in both arguments, we get

$$\begin{aligned}\langle x(t), Qx(t)\rangle_G &= \Big\langle \breve{x}(t) + \sum_{\ell=1}^L x^\ell(t), Q\Big(\breve{x}(t) + \sum_{\ell=1}^L x^\ell(t)\Big)\Big\rangle_G \\ &= \langle \breve{x}(t), Q\breve{x}(t)\rangle_G + 2\Big\langle \sum_{\ell=1}^L x^\ell(t), Q\breve{x}(t)\Big\rangle_G \\ &\quad + \Big\langle \sum_{\ell=1}^L x^\ell(t), Q\Big(\sum_{\ell=1}^L x^\ell(t)\Big)\Big\rangle_G. \quad (27)\end{aligned}$$

From Lemma 2 and (P6), the first term of (27) simplifies to

$$\langle \breve{x}(t), Q\breve{x}(t)\rangle_G = q_0 \sum_{i\in\mathcal{N}} \breve{x}_i(t)^\intercal Q \breve{x}_i(t), \quad (28)$$

and the second term simplifies to

$$\begin{aligned}\Big\langle \sum_{\ell=1}^L x^\ell(t), Q\breve{x}(t)\Big\rangle_G &= q_0 \sum_{i\in\mathcal{N}} \sum_{\ell=1}^L x_i^\ell(t)^\intercal Q \breve{x}_i(t) \\ &= q_0 \sum_{\ell=1}^L \sum_{i\in\mathcal{N}} x_i^\ell(t)^\intercal Q \Big(x_i(t) - \sum_{\ell'=1}^L x_i^{\ell'}(t)\Big) \\ &\overset{(a)}{=} q_0 \sum_{\ell=1}^n \sum_{i\in\mathcal{N}} \Big(x_i^\ell(t)^\intercal Q x_i^\ell(t) - x_i^\ell(t)^\intercal Q x_i^\ell(t)\Big) \\ &= 0, \quad (29)\end{aligned}$$

where (a) follows from (P8) and (P9). From Lemma 2 and (P3), the third term of (27) simplifies to

$$\begin{aligned}&\Big\langle \sum_{\ell=1}^L x^\ell(t), Q\Big(\sum_{\ell=1}^L x^\ell(t)\Big)\Big\rangle_G \\ &= \sum_{i\in\mathcal{N}} \sum_{\ell=1}^L x_i^\ell(t)^\intercal Q \Big(\sum_{\ell'=1}^L q^{\ell'} x_i^{\ell'}(t)\Big) \\ &= \sum_{\ell=1}^L \sum_{i\in\mathcal{N}} x_i^\ell(t)^\intercal Q \Big(\sum_{\ell'=1}^L q^{\ell'} x_i^{\ell'}(t)\Big) \\ &\overset{(b)}{=} \sum_{\ell=1}^L \sum_{i\in\mathcal{N}} q^\ell x_i^\ell(t)^\intercal Q x_i^\ell(t), \quad (30)\end{aligned}$$

where (b) follows from (P8). We get the result by substituting (28)–(30) in (27).