SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

December 2026, Vol. 69, Iss. 12, 222202:1–222202:18
https://doi.org/10.1007/s11432-025-5094-3

Optimal intrinsic formation using exogenous systems

Yueyue XU[1,2,3,4], Panpan ZHOU[2], Lin WANG[1,3,4*] & Xiaoming HU[2]

[1]*State Key Laboratory of Submarine Geoscience, Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, China*
[2]*Department of Mathematics, KTH Royal Institute of Technology, Stockholm 10044, Sweden*
[3]*Key Laboratory of System Control and Information Processing, Ministry of Education, Shanghai 200240, China*
[4]*Shanghai Key Laboratory for Perception and Control in Industrial Network Systems, Shanghai 200240, China*

Received 13 March 2025/Revised 21 August 2025/Accepted 15 October 2025/Published online 9 November 2026

Abstract This paper investigates the intrinsic formation problem of a multi-agent system using an exogenous system. The problem is formulated as an intrinsic infinite time-horizon linear quadratic optimal control problem, where no formation error information is incorporated in the performance index. Convergence to the formation is achieved by utilizing an exogenous system, which expands the steady-state formation space. We study both the forward and inverse problems: the former derives the formation from given system matrices, while the latter constructs the system matrices from a desired formation. For the forward problem, we provide the existence condition for a nonzero steady state and characterize the steady-state space. For the inverse problem, we design both the input matrix and the exogenous system so that the desired formation can be achieved. Finally, numerical simulations are provided to illustrate the effectiveness of the proposed results.

Keywords multi-agent system, intrinsic formation control, exogenous system, quadratic optimal control, linear system

Citation Xu Y Y, Zhou P P, Wang L, et al. Optimal intrinsic formation using exogenous systems. Sci China Inf Sci, 2026, 69(12): 222202, https://doi.org/10.1007/s11432-025-5094-3

1 Introduction

The formation control problem has been an interesting and attractive topic in the field of multi-agent coordination control for its great applications in, for instance, surveillance or exploration activity, transportation, and space missions [1, 2]. The formation control problem aims to design cooperative control strategies to steer a group of agents to form a desired geometric pattern.

In most of the existing literature on formation control, multi-agent systems range from single or double integrators [3, 4], to more general linear systems [5], and even to nonlinear systems such as unicycles [6] and Euler-Lagrange systems [7]. Among the formation control approaches, displacement-based method is one of the fundamental methods [8]. As far as we know, in the majority of research results that utilize the displacement-based method, the desired formation configuration is included in the designed control strategy, regardless of whether they formulate the problem as an optimization problem [9–11] or not [5,12]. The explicit desired formation can be easily stolen, which poses security risks for formation control. In recent years, several researchers have begun studying the formation control problem using an intrinsic approach [2,13–16], which does not require including any information about the desired formation in the controller or the performance index. Instead, the resulting formation pattern is determined by the specified inter-agent connectivity topology. To the best of our knowledge, the intrinsic approach was first used in [13] to study the formation control, where agents converge to a great circle. Later, Zhang et al. [14] used gain functions to achieve intrinsic formation control for reduced attitude without any prior information of the desired formation, but it was limited to achieving a tetrahedron formation. After that, Zhang et al. [15] extended this work to achieve all regular polyhedra formations. In [2], a differential game approach is used to achieve intrinsic control. But due to the symmetry requirements of the objective function for the game, the formations were restricted to Platonic solids.

Motivated by the observation that the existing literature on intrinsic formation has limitations in terms of formation diversity, this work formulates the intrinsic formation control problem as an infinite time-horizon linear quadratic optimal control problem, enabling a richer set of formation configurations. It is widely understood that, if the performance index does not incorporate the difference between the agent's current position and the target

* Corresponding author (email: wanglin@sjtu.edu.cn)

© Science China Press 2026

Table 1 Notations.

Notation	Definition
$\mathrm{Im}(A)$	Image space of matrix A
$\mathrm{Ker}(A)$	Kernel space of matrix A
$\mathrm{rank}(A)$	The rank of matrix A
I_n	$n \times n$ identity matrix
$\mathrm{diag}(X_1, ..., X_n)$	Block diagonal matrix with blocks $X_1, ..., X_n$
$\otimes$	Kronecker product

position, the optimal trajectory derived from a standard linear quadratic optimal control problem can at most converge to the null space of the system matrix. However, in many cases, the desired formation does not lie within that space. Thus, we incorporate an exogenous system in the quadratic optimal control and achieve essentially arbitrary formations in any spatial dimension, excluding only certain degenerate cases (e.g., those associated with eigenvectors of matrix A having eigenvalues with positive real parts or zero eigenvalues with nontrivial Jordan blocks). It is also worth noting that the existing work used static feedback control, while we study the dynamic feedback control in this paper by adding an exogenous system. Even though full state information is required, more formation configurations can be achieved. Li et al. [16] studied formation using an intrinsic optimal control approach for the first time by considering a two-dimensional case. However, in their design of exogenous systems only a single integrator was considered. By introducing general exogenous systems, our method can achieve (almost) arbitrary formations in spaces of any dimension. Furthermore, we can control only a subset of the players, relying on their inherent interactions to guide the entire system to the desired formation.

In this paper, we first address the forward problem, which includes determining the existence condition for a nonzero steady state and characterizing the steady-state space. We then investigate the inverse problem which aims to design the system to ensure convergence to the desired formation. To achieve this, we design the input matrix so that the desired formation lies within the maximum steady-state space while minimizing the number of control. Then, we design the exogenous system and performance index to ensure convergence to the specific desired formation. Finally, we specify how to set the initial state of the exogenous system so that any desired formation scaling can be achieved.

The main contributions of this article are summarized as follows.

• We formulate the intrinsic formation control problem as an infinite-horizon linear quadratic problem augmented with an exogenous system. Compared with prior intrinsic-formation studies [2, 13–16], this work enlarges the achievable steady-state formation space and enables (almost) arbitrary formations in any spatial dimension, except for standard degenerate cases.

• To achieve the desired formation, we address both the forward and inverse problems: the former characterizes the steady-state space, and the latter designs the system. We design the input matrix to use as few control inputs as possible and construct an exogenous system of low dimension, thereby preserving performance while reducing cost and computational complexity.

• The exogenous system enhances formation confidentiality through three mechanisms: (1) the desired pattern does not explicitly appear in either the controller or the performance index, thereby mitigating susceptibility to certain attacks [17]; (2) multiple exogenous systems can reproduce the same formation, improving stealth; and (3) by minimizing exogenous system input rather than the multi-agent input, agents may follow non-shortest, deception-oriented trajectories that hinder adversarial inference of the desired formation.

This paper is organized as follows. In Section 2, we first formulate the formation problem under the framework of intrinsic quadratic optimal control. Then we study the existence condition of a nonzero steady state and characterize the steady-state space in Section 3. Section 4 presents the design of the input matrix, exogenous system, performance index and initial state required to achieve the desired formation. In Section 5, simulations are given to illustrate the theoretical results. A summary of notations is provided in Table 1. Finally, Section 6 concludes the paper and discusses future research directions.

2 Problem formulation

Consider a multi-agent system with N agents, $x_i(t) \in \mathbb{R}^d$ is the state of the agent i, where d is the dimension of space. We can arrange the states of N agents into a column vector $x(t) \in \mathbb{R}^n$ where $n = N \times d$. Then we write the dynamics of multi-agent system in the compact form

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

where $u(t) \in \mathbb{R}^m$ is the control for the multi-agent system and the form of matrix A depends on the interaction patterns among the agents. For example, in a Laplacian networked system, the dynamics of the agent i is described by $\dot{x}_i(t) = \sum_{j\in\mathcal{N}_i}(x_j(t) - x_i(t)) + B_i u(t)$, where $\mathcal{N}_i$ is the set of agents that interact with the agent i, $B_i \in \mathbb{R}^{d\times m}$ is the input matrix. Then we have $A = -L \otimes I_d,\ B = \left[B_1^{\mathrm{T}},\ B_2^{\mathrm{T}},\ \cdots,\ B_N^{\mathrm{T}}\right]^{\mathrm{T}}$, where L is the Laplacian matrix.

An intrinsic formation refers to a formation pattern that emerges solely from the inter-agent connectivity topology and the geometric properties of the network, without embedding any desired formation information explicitly in the controller or performance index. Then a standard way to formulate the intrinsic formation problem under the framework of quadratic optimal control would be as follows:

$$\begin{aligned} \min_{u} \quad & J = \frac{1}{2}\int_0^\infty \left(x^{\mathrm{T}}Qx + \|u\|^2\right) dt, \\ \text{s.t.} \quad & \dot{x} = Ax + Bu, \end{aligned} \tag{2}$$

where the positive semi-definite matrix Q is to be designed.

However, directly minimizing the index J will drive the state at most towards the null space of the matrix A, which may not include the formations of interest. In [16], an integrator is added to the system to modify the control signal, and by this way some stripe patterns are achieved. However, the patterns achieved there are limited to two-color patterns on a two-dimensional plane. In this paper, we enable more general formations in spaces of any dimension by introducing a general exogenous system:

$$\begin{aligned} \dot{w} =&\ F_1 x + F_2 w + G v, \\ u =&\ H w + K v, \end{aligned} \tag{3}$$

where $w \in \mathbb{R}^k$ is the state of the exogenous system, $F_1 \in \mathbb{R}^{k\times n}, F_2 \in \mathbb{R}^{k\times k}, G \in \mathbb{R}^{k\times p}, H \in \mathbb{R}^{m\times k}, K \in \mathbb{R}^{m\times p}$ are matrices of the exogenous system to be designed, $v \in \mathbb{R}^p$ is the new control that substitutes u in the performance index of (2) and the performance index becomes

$$\bar{J} = \frac{1}{2}\int_0^\infty \left(x^{\mathrm{T}}Qx + \|v\|^2\right) dt. \tag{4}$$

The above optimization problems (1), (3) and (4) can be restated as the following linear-quadratic optimal control problem:

$$\min_{v}\ \bar{J} = \frac{1}{2}\int_0^\infty \left(\bar{x}^{\mathrm{T}}\bar{Q}\bar{x} + \|v\|^2\right) dt, \tag{5a}$$

$$\text{s.t. } \dot{\bar{x}} = \bar{A}\bar{x} + \bar{B}v, \tag{5b}$$

where

$$\begin{aligned} &\bar{x} = \begin{bmatrix} x \\ w \end{bmatrix}, \bar{A} = \begin{bmatrix} A & BH \\ F_1 & F_2 \end{bmatrix}, \\ &\bar{B} = \begin{bmatrix} BK \\ G \end{bmatrix}, \bar{Q} = \begin{bmatrix} Q & 0_{n\times k} \\ 0_{k\times n} & 0_{k\times k} \end{bmatrix}. \end{aligned} \tag{6}$$

Q is positive semi-definite and can be rewritten as $Q = C^{\mathrm{T}}C$, where $C \in \mathbb{R}^{q\times n}$. Define $\bar{C} = [C \quad 0_{q\times k}]$, then $\bar{Q} = \bar{C}^{\mathrm{T}}\bar{C}$. If the optimal control exists, then it can be expressed as

$$v^* = -\bar{B}^{\mathrm{T}}P\bar{x}, \tag{7}$$

where $P \in \mathbb{R}^{(n+k)\times(n+k)}$ is the smallest real symmetric positive semi-definite solution to the algebraic Riccati equation

$$\bar{A}^{\mathrm{T}}P + P\bar{A} - P\bar{B}\bar{B}^{\mathrm{T}}P + \bar{Q} = 0. \tag{8}$$

We say that P is the smallest if for any other real symmetric positive semi-definite solution P_0 of (8), the difference $P_0 - P$ is positive semi-definite. The closed-loop system under the control strategy (7) is as follows:

$$\dot{\bar{x}} = (\bar{A} - \bar{B}\bar{B}^{\mathrm{T}}P)\bar{x} \triangleq \bar{A}_{cl}\bar{x}. \tag{9}$$

For the multi-agent formation problem, we define the desired formation space as $\mathrm{Im}(X^{df})$, where $X^{df} \in \mathbb{R}^{n\times z}$. A desired specific formation is any vector $x_{df} \in \mathbb{R}^n$ in desired formation space $\mathrm{Im}(X^{df})$. Given the above information, we give the purpose of this work:

• Design the input matrix B with the minimal number of control inputs, such that the desired formation space $\mathrm{Im}(X^{df})$ is achievable;

• Design the exogenous system matrices (H, F_1, F_2, K, G) of minimal dimension, the exogenous system initial state $w(0)$, and the performance index matrix Q, such that the state x following the optimal control defined in (5) and (6) converges to a specific formation x_{df}.

To achieve this objective, we divide the study into a forward problem and an inverse problem. The former provides a steady-state description, which forms the basis for the inverse problem, while the latter focuses on system design.

3 Forward problem: steady-state description

In this section, we will address the forward problem, namely describing the steady-state formation as solution to the optimal control problem (5). Note that since the formations we are interested in are in general non-consensus formations, $(\bar{C}, \bar{A})$ is not detectable, which differs from the standard case.

3.1 Existence condition of a nonzero steady state

In order to give the existence condition of a nonzero steady state, we first cite the following lemma that provides the existence condition of an optimal control for (5).

Lemma 1 ([18,19]). Supposing $(\bar{A}, \bar{B})$ is stabilizable and $\bar{Q}$ is positive semi-definite, then there exists a smallest real symmetric positive semi-definite solution P for the algebraic Riccati equation (8). Furthermore,

$$v(t) = -\bar{B}^{\mathrm{T}} P \bar{x}(t)$$

is the optimal solution of the linear quadratic optimal control problem (5). The closed-loop matrix $\bar{A}_{cl}$ is stable if and only if $(\bar{C}, \bar{A})$ is detectable.

By Lemma 1, to ensure the existence of the optimal control, we assume the following.

Assumption 1. The pair $(\bar{A}, \bar{B})$ is stabilizable.

Besides, in order to assure that the optimal trajectory converges to a nonzero stationary steady state, we make the following assumption.

Assumption 2. The pair $(\bar{C}, \bar{A})$ is undetectable and the only undetectable eigenvalue of $\bar{A}$ for the pair $(\bar{C}, \bar{A})$ is 0. And its algebraic multiplicity and geometric multiplicity are equal.

This assumption is based on the fact that undetectable eigenvalues cannot be stabilized by optimal control, and it guarantees that under optimal control, the closed-loop system (9) is marginally stable. That is, all eigenvalues of $\bar{A}_{cl}$ are non-positive, and every Jordan block corresponding to the eigenvalue 0 is of size 1×1.

3.2 Steady-state space

Let x_{ss} and w_{ss} denote the steady state of the multi-agent system and the exogenous system, respectively. Setting the combined state $\bar{x}_{ss} = \begin{bmatrix} x_{ss} \\ w_{ss} \end{bmatrix}$, then the following conditions hold:

(I) $P\bar{x}_{ss} = 0$;
(II) $\bar{Q}\bar{x}_{ss} = 0$, i.e., $\bar{C}\bar{x}_{ss} = 0$;
(III) $\bar{A}\bar{x}_{ss} = 0$.

Condition (I) is implied by the existence of optimal control while condition (II) can be derived from the fact that $\mathrm{Ker}(P) \subseteq \mathrm{Ker}(\bar{C})$. Condition (III) ensures that the steady state is stationary. Although condition (I) is well known, it does not seem to be easy to locate a proof. For the sake of completeness, we provide a brief proof as follows.

Lemma 2. If the steady state $\bar{x}_{ss}$ exists for the closed-loop system (9), then there must be $P\bar{x}_{ss} = 0$, where P is the solution of the algebraic Riccati equation (8).

Proof. Since the steady state $\bar{x}_{ss}$ exists for the closed-loop system (9), the optimal solution of the linear quadratic optimal control problem (5) exists and the performance index in (5) is bounded for all feasible control. Considering the finite horizon index

$$\bar{J}_T = \frac{1}{2}\int_t^{t+T} \left(\bar{x}^{\mathrm{T}}\bar{Q}\bar{x} + \|v\|^2\right) dt, \tag{10}$$

then the optimal cost of (10) is $\bar{x}^{\mathrm{T}}(t)P(T)\bar{x}(t)$. Thus $\bar{J}_T \geqslant \bar{x}^{\mathrm{T}}(t)P(T)\bar{x}(t)$ for any feasible control. Since $\forall T > 0, \bar{J}_T \to 0$ when $t \to \infty$, otherwise the infinite integral will not converge, we have

$$\lim_{t\to\infty} \bar{x}^{\mathrm{T}}(t)P(T)\bar{x}(t) = 0.$$

Thus

$$\lim_{T\to\infty}\lim_{t\to\infty} \bar{x}^{\mathrm{T}}(t)P(T)\bar{x}(t) = \lim_{t\to\infty} \bar{x}^{\mathrm{T}}(t)P\bar{x}(t) = 0,$$

where P is the smallest solution of the algebraic Riccati equation (8). Thus

$$\lim_{t\to\infty} P\bar{x}(t) = P\bar{x}_{ss} = 0,$$

which completes the proof.

The following result gives the expression of the steady state $\bar{x}_{ss}$.

Lemma 3 (Steady state)**.** Under Assumptions 1 and 2, the state of the closed-loop system (9) will converge to a steady state $\bar{x}_{ss}$ which satisfies

$$\bar{x}_{ss} = \sum_{i=1}^{s} (\phi_i^{\mathrm{T}}\bar{x}(0))\phi_i,$$

where $\phi_i \in \mathbb{R}^{(n+k)\times 1}$ $(i = 1, 2, \cdots, s)$ are orthonormal eigenvectors of the zero eigenvalue of the closed-loop matrix $\bar{A}_{cl}$, and s is the algebraic multiplicity of the zero-eigenvalue.

Proof. Based on Assumptions 1 and 2, $\bar{A}_{cl}$ is marginally stable, meaning that each zero eigenvalue is associated with a first-order Jordan block, and all other eigenvalues have negative real parts. Let $\mathrm{diag}(J_{cl}, 0_{s\times s}) = V^{-1}\bar{A}_{cl}V$ denote the Jordan canonical form of $\bar{A}_{cl}$. Here, J_{cl} comprises all Jordan blocks corresponding to eigenvalues with negative real parts, while $0_{s\times s}$ is an $s\times s$ zero matrix representing the Jordan blocks associated with zero eigenvalues. For the closed-loop system (9), straight calculation gives

$$\begin{aligned}\bar{x}_{ss} &= \lim_{t\to\infty} e^{\bar{A}_{cl}t}\bar{x}(0) = \lim_{t\to\infty} V e^{\mathrm{diag}(J_{cl},0)t}V^{-1}\bar{x}(0)\\ &= \lim_{t\to\infty} V\,\mathrm{diag}(e^{J_{cl}t}, I_s)V^{-1}\bar{x}(0) = V\,\mathrm{diag}(0, I_s)V^{-1}\bar{x}(0).\end{aligned}$$

By standard construction of the V matrix, we then have $\bar{x}_{ss} = \sum_{i=1}^{s}\phi_i\phi_i^{\mathrm{T}}\bar{x}(0) = \sum_{i=1}^{s}(\phi_i^{\mathrm{T}}\bar{x}(0))\phi_i$, and the proof is complete.

Here $\phi_i^{\mathrm{T}}\bar{x}(0) \in \mathbb{R}$ represents the weight of ϕ_i in $\bar{x}_{ss}$ and the steady state $\bar{x}_{ss}$ is a linear combination of the right eigenvectors of the closed-loop matrix $\bar{A}_{cl}$ corresponding to the zero eigenvalue, i.e., $\bar{x}_{ss} \in \mathrm{Ker}(\bar{A}_{cl})$.

Specifically, $\mathrm{Ker}(\bar{A}_{cl})$ is the steady-state space. For (9), if $\bar{x}(0) \in \mathrm{Ker}(\bar{A}_{cl})$, there is $\bar{x}(t) = \bar{x}(0), \forall t \geqslant 0$. Therefore, $\forall\phi \in \mathrm{Ker}(\bar{A}_{cl})$, ϕ must be a steady state and satisfy conditions (I)–(III), i.e.,

$$\mathrm{Ker}(\bar{A}_{cl}) \subseteq \mathrm{Ker}\left(\begin{bmatrix} P \\ \bar{C} \\ \bar{A}\end{bmatrix}\right) = \mathrm{Ker}\left(\begin{bmatrix} P \\ \bar{A}\end{bmatrix}\right),$$

where $\mathrm{Ker}(P) \subseteq \mathrm{Ker}(\bar{C})$ is used.

Because $\mathrm{Ker}(\bar{A}_{cl})$ is related to the solution of the Riccati equation, it is challenging to analyze. Then we define another space $\mathrm{Ker}([\ \bar{A}^{\mathrm{T}}\ \bar{C}^{\mathrm{T}}\ ]^{\mathrm{T}})$. We will prove the equivalence of the steady-state space $\mathrm{Ker}(\bar{A}_{cl})$ and $\mathrm{Ker}([\ \bar{A}^{\mathrm{T}}\ \bar{C}^{\mathrm{T}}\ ]^{\mathrm{T}})$ as follows.

Proposition 1 (Steady-state space)**.** Consider the optimal control problem (5) under Assumptions 1 and 2. Then the steady-state space $\mathrm{Ker}(\bar{A}_{cl})$ satisfies

$$\mathrm{Ker}(\bar{A}_{cl}) = \mathrm{Ker}\left([\ \bar{A}^{\mathrm{T}}\ \bar{C}^{\mathrm{T}}\ ]^{\mathrm{T}}\right). \tag{11}$$

Proof. From conditions (II) and (III) for a steady state, we immediately see that $\mathrm{Ker}(\bar{A}_{cl}) \subseteq \mathrm{Ker}([\ \bar{A}^{\mathrm{T}}\ \bar{C}^{\mathrm{T}}\ ]^{\mathrm{T}})$. Next, we prove the reverse inclusion. $\forall\bar{x} \in \mathrm{Ker}([\ \bar{A}^{\mathrm{T}}\ \bar{C}^{\mathrm{T}}\ ]^{\mathrm{T}})$, we have $\bar{Q}\bar{x} = 0$. By left-multiplying the algebraic Riccati equation (8) by $\bar{x}^{\mathrm{T}}$ and right-multiplying by $\bar{x}$, we can obtain $\bar{B}^{\mathrm{T}}P\bar{x} = 0$. Thus $\bar{A}_{cl}\bar{x} = (\bar{A} - \bar{B}\bar{B}^{\mathrm{T}}P)\bar{x} = 0$, which implies $\mathrm{Ker}([\ \bar{A}^{\mathrm{T}}\ \bar{C}^{\mathrm{T}}\ ]^{\mathrm{T}}) \subseteq \mathrm{Ker}(\bar{A}_{cl})$. Combining both inclusions, we conclude that $\mathrm{Ker}(\bar{A}_{cl}) = \mathrm{Ker}([\ \bar{A}^{\mathrm{T}}\ \bar{C}^{\mathrm{T}}\ ]^{\mathrm{T}})$, which completes the proof.

Remark 1. Specifically, any nonzero vector in $\mathrm{Ker}([\,\bar{A}^{\mathrm{T}}\ \bar{C}^{\mathrm{T}}]^{\mathrm{T}})$ is an eigenvector of $\bar{A}$ corresponding to the undetectable eigenvalue 0 for the pair $(\bar{C},\bar{A})$. Thus, the steady state of the linear-quadratic optimal control problem (5) is an eigenvector of $\bar{A}$ corresponding to the undetectable eigenvalue 0.

3.3 Maximal steady-state space

Consider the steady-state space defined in (11),

$$\mathrm{Ker}\left(\begin{bmatrix}\bar{A}\\ \bar{C}\end{bmatrix}\right)=\mathrm{Ker}\left(\begin{bmatrix}A & BH\\ F_1 & F_2\\ C & 0\end{bmatrix}\right).$$

In the multi-agent system described by (1), the dynamics (including matrices A and B) is given. In contrast, the control strategy specified in (3) and (7) can be designed, involving the exogenous system (with matrices H, F_1, and F_2) and the performance index (matrix C). Regardless of the control design, the steady state must satisfy

$$Ax_{ss}=-BH\,w_{ss}. \tag{12}$$

This condition confines the steady state to a particular subspace, which we define as the *maximal steady-state space*. Discussing this space is crucial because it characterizes the theoretical limits of the formation patterns that can be achieved through subsequent control design. By understanding the maximal steady-state space, we can determine the full range of feasible formations and thereby guide the design process to realize as many formations as possible.

Without loss of generality, we assume that the system matrices in (1) have already been transformed into the Jordan normal form via a linear transformation, i.e.,

$$A=\begin{bmatrix}J(\lambda_1) & & & \\ & J(\lambda_2) & & \\ & & \ddots & \\ & & & J(\lambda_l)\end{bmatrix},B=\begin{bmatrix}B_{\lambda_1}\\ B_{\lambda_2}\\ \vdots\\ B_{\lambda_l}\end{bmatrix}, \tag{13}$$

where $\lambda_i\neq\lambda_j,\forall i\neq j\in\mathcal{I},\ \mathcal{I}=\{1,2,\cdots,l\};\ J(\lambda_i)\in\mathbb{R}^{\sigma_i\times\sigma_i},\ B_{\lambda_i}\in\mathbb{R}^{\sigma_i\times m}$, σ_i is the algebraic multiplicity of λ_i. We group the Jordan blocks corresponding to the nonzero eigenvalues into A_r and explicitly write out each Jordan block corresponding to the eigenvalue 0. Then Eq. (13) can be rewritten as follows:

$$A=\begin{bmatrix}A_r & & & & \\ & J_1(0) & & & \\ & & J_2(0) & & \\ & & & \ddots & \\ & & & & J_{\alpha_0}(0)\end{bmatrix},\quad B=\begin{bmatrix}B_r\\ B_0^1\\ B_0^2\\ \vdots\\ B_0^{\alpha_0}\end{bmatrix}, \tag{14}$$

where $A_r\in\mathbb{R}^{r_0\times r_0}$ is the block diagonal matrix composed of all Jordan blocks corresponding to the nonzero eigenvalues, $B_r\in\mathbb{R}^{r_0\times m}$, α_0 is the geometric multiplicity of the eigenvalue 0,

$$J_j(0)=\begin{bmatrix}0 & 1 & 0 & \cdots & 0\\ 0 & 0 & 1 & \cdots & 0\\ \vdots & & \ddots & \ddots & \vdots\\ 0 & 0 & & 0 & 1\\ 0 & 0 & & & 0\end{bmatrix}\in\mathbb{R}^{r_j\times r_j},\ B_0^j\triangleq\begin{bmatrix}B_0^j(1:r_j-1)\\ B_0^j(r_j)\end{bmatrix}\in\mathbb{R}^{r_j\times m},\ \text{for } j=1,2,\ldots,\alpha_0, \tag{15}$$

$\sum_{j=1}^{\alpha_0}r_j$ is the algebraic multiplicity of the eigenvalue 0, $B_0^j(1:r_j-1)\in\mathbb{R}^{(r_j-1)\times m}$ is the submatrix of B_0^j formed by the first (r_j-1) rows of B_0^j, and $B_0^j(r_j)\in\mathbb{R}^{1\times m}$ is the last row of B_0^j.

With matrices A and B given as (14) and (15), we can derive the maximal steady-state space in the following theorem.

Theorem 1 (Maximal steady-state space)**.** Consider the optimal control problem (5), where $A \in \mathbb{R}^{n\times n},\ B \in \mathbb{R}^{n\times m}$ are given as (14) and (15). Under Assumptions 1 and 2, for any linear exogenous system (namely for any matrices F_1, F_2, G, H, K), the steady state of the multi-agent system x_{ss} must belong to

$$\mathrm{Im}(\widetilde{X}) = \mathrm{Im}\left(\left[\begin{array}{c:c} \begin{bmatrix} A_r^{-1}B_rU_{B_{\mathrm{last}}} \\ \widetilde{X}_1 \\ \widetilde{X}_2 \\ \vdots \\ \widetilde{X}_{\alpha_0} \end{bmatrix} & U_A \end{array}\right]\right), \tag{16}$$

where

$$\widetilde{X}_i = \begin{bmatrix} 0 \\ B_0^i\,(1:r_i-1)\,U_{B_{\mathrm{last}}} \end{bmatrix},\ \text{for } i=1,2,\cdots,\alpha_0,\ B_{\mathrm{last}} = \left[B_0^{1\mathrm{T}}(r_1)\ B_0^{2\mathrm{T}}(r_2)\ \cdots\ B_0^{\alpha_0\mathrm{T}}(r_{\alpha_0})\right]^{\mathrm{T}}, \tag{17}$$

$U_{B_{\mathrm{last}}}$ and U_A are the matrices whose columns form a standard orthonormal basis for $\mathrm{Ker}(B_{\mathrm{last}})$ and $\mathrm{Ker}(A)$, respectively, and $\widetilde{X}$ is a matrix whose columns are linearly independent.

Proof. For any exogenous system, the steady state must satisfy $Ax_{ss} = -BHw_{ss}$ by (12). Since H and w_{ss} belong to the exogenous system, they can be arbitrarily chosen as long as they are finite. Defining $\hat{w} = -Hw_{ss} \in \mathbb{R}^{m\times 1}$, then $\hat{w}$ can also be arbitrarily selected and the condition is equal to

$$Ax_{ss} = B\hat{w}. \tag{18}$$

Define $x_{ss} = \left[x_r^{\mathrm{T}}\ x_1^{\mathrm{T}}\ x_2^{\mathrm{T}}\ \cdots\ x_{\alpha_0}^{\mathrm{T}}\right]^{\mathrm{T}}$, where $x_r \in \mathbb{R}^{r_0},\ x_i \in \mathbb{R}^{r_i},\ \text{for } i=1,2,\ldots,\alpha_0$. Then Eq. (18) is equal to

$$A_rx_r = B_r\hat{w},\ J_i(0)x_i = B_0^i\hat{w},\ \text{for } i=1,2,\cdots,\alpha_0.$$

Thus

$$x_r = A_r^{-1}B_r\hat{w},\ x_i = \begin{bmatrix} * \\ B_0^i(1:r_i-1)\hat{w} \end{bmatrix}, \text{for } i=1,2,\cdots,\alpha_0, \tag{19}$$

where $* \in \mathbb{R}$ is arbitrary, and $\hat{w}$ satisfies $B_0^i(r_i)\hat{w} = 0,\ \text{for } i=1,2,\cdots,\alpha_0$.

Defining $B_{\mathrm{last}} = \left[B_0^{1\mathrm{T}}(r_1)\ B_0^{2\mathrm{T}}(r_2)\ \cdots\ B_0^{\alpha_0\mathrm{T}}(r_{\alpha_0})\right]^{\mathrm{T}}$, then we have $B_{\mathrm{last}}\hat{w} = 0$. Thus

$$\hat{w} \in \mathrm{Im}(U_{B_{\mathrm{last}}}), \tag{20}$$

where $U_{B_{\mathrm{last}}}$ is the matrix whose columns form a standard orthonormal basis for $\mathrm{Ker}(B_{\mathrm{last}})$. In (19), because $* \in \mathbb{R}$ is arbitrary, we have

$$x_i \in \mathrm{Im}\left(\begin{bmatrix} 0 & 1 \\ B_0^i(1:r_i-1)U_{B_{\mathrm{last}}} & 0_{(r_i-1)\times 1} \end{bmatrix}\right),\ \text{for } i=1,2,\cdots,\alpha_0. \tag{21}$$

Combining (19)–(21), we have

$$\begin{bmatrix} x_r \\ x_1 \\ x_2 \\ \vdots \\ x_{\alpha_0} \end{bmatrix} \in \mathrm{Im}\left(\left[\begin{array}{c:c:c:c} A_r^{-1}B_rU_{B_{\mathrm{last}}} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ B_0^1\,(1:r_1-1)\,U_{B_{\mathrm{last}}} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ B_0^2\,(1:r_2-1)\,U_{B_{\mathrm{last}}} & 0 & 0 & \cdots\ 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 \\ B_0^{\alpha_0}\,(1:r_{\alpha_0}-1)\,U_{B_{\mathrm{last}}} & 0 & 0 & 0 \end{array}\right]\right), \tag{22}$$

where the columns on the right-hand side (each containing a single "1" and zeros elsewhere) form exactly the Kernel space of A. By collecting these vectors into a matrix U_A, Eq. (22) becomes equivalent to (16), which completes the proof.

Define the space $\mathrm{Im}(\widetilde{X})$ as the maximal steady-state space, to which the steady state x_{ss} of the multi-agent system must belong. The design of the exogenous system cannot force the state to converge to a formation outside $\mathrm{Im}(\widetilde{X})$; it can only determine which specific formation within $\mathrm{Im}(\widetilde{X})$ the state converges to. When A has no zero eigenvalues, $\mathrm{Im}(\widetilde{X}) = \mathrm{Im}(A^{-1}B)$.

The following corollary considers the case in which all Jordan blocks associated with the zero eigenvalue of A are 1-dimensional. In this case, rewrite (14) as

$$A = \begin{bmatrix} A_r & 0 \\ 0 & 0_{(n-r_0)\times(n-r_0)} \end{bmatrix}, \quad B = \begin{bmatrix} B_r \\ B_0 \end{bmatrix}, \tag{23}$$

where $A_r \in \mathbb{R}^{r_0\times r_0}$ is the block diagonal matrix composed of all Jordan blocks corresponding to the nonzero eigenvalues, $B_r \in \mathbb{R}^{r_0\times m}$.

Corollary 1. Consider the optimal control problem (5), where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$ are given as (23) and all Jordan blocks associated with the zero eigenvalue of A are 1-dimensional. Under Assumptions 1 and 2, for any linear exogenous system (namely for any matrices F_1, F_2, G, H, K), the steady state of the multi-agent system x_{ss} must belong to

$$\mathrm{Im}(\widetilde{X}) = \mathrm{Im}\left(\left[\begin{array}{c:c} A_r^{-1}B_rU_{B_0} & 0 \\ \hdashline 0 & I_{n-r_0} \end{array}\right]\right),$$

where U_{B_0} is the matrix whose columns form a standard orthonormal basis for $\mathrm{Ker}(B_0)$, $\widetilde{X}$ is a matrix whose columns are linearly independent.

4 Inverse problem: system design

In Section 3, we established the existence condition for a nonzero steady state and discussed the steady-state space. However, in practice, our focus shifts to designing the system to ensure that the multi-agent system achieves the desired formation, thereby addressing the inverse problem. Designing the system requires three steps.

In step 1, for any desired formation, we construct a matrix B such that the corresponding maximal steady-state space encompasses the desired formation while ensuring the controllability of the pair (A, B). Most importantly, our construction guarantees that B has the minimal rank, which directly minimizes the number of required control inputs.

In step 2, we design the exogenous system—comprising matrices H, F_1, F_2, K, and G, along with the performance index matrix C—to ensure that the state converges to the desired formation under optimal control. Notably, our construction attains the minimal possible dimension for the exogenous system.

In step 3, we specify the initial state $w(0)$ of the exogenous system to achieve the desired formation scaling.

4.1 Step 1: design of the input matrix B

Theorem 1 characterizes the maximal steady-state space $\mathrm{Im}(\widetilde{X})$, which is fundamentally determined by the input matrix B. In this step, we design B to meet two essential conditions: (1) the maximal steady-state space $\mathrm{Im}(\widetilde{X})$ must encompass the desired formation, and (2) the pair (A, B) must be controllable. Although the construction of B is not unique, our approach deliberately constructs B with the minimal rank, thereby minimizing the number of required control inputs.

Define a collection of z desired formations as the space $\mathrm{Im}(X^{df})$, where $X^{df} = [x_{df}^1 \quad x_{df}^2 \quad \cdots \quad x_{df}^z] \in \mathbb{R}^{n\times z}$ is a matrix, and $x_{df}^i \in \mathbb{R}^n (i = 1, 2, \cdots, z)$ represents a desired formation vector. The reason for considering a collection of desired formations rather than a single desired formation is that we can design one B and then use different exogenous systems to achieve various formations within the collection, eliminating the need to redesign B each time. Because $\mathrm{Im}(\widetilde{X})$ must comprise $\mathrm{Im}(U_A)$ by (16), in order to make $\mathrm{Im}(X^{df}) \subseteq \mathrm{Im}(\widetilde{X})$, we can remove the components of $\mathrm{Im}(X^{df})$ that belong to $\mathrm{Im}(U_A)$ and define the residual desired formation space as

$$\mathrm{Im}(X^{rf}) = \mathrm{Im}(X^{df}) \cap (\mathrm{Im}(U_A))^{\perp}, \tag{24}$$

where $(\mathrm{Im}(U_A))^{\perp}$ represents the orthogonal complement of $\mathrm{Im}(U_A)$, $X^{rf} \in \mathbb{R}^{n\times z_0}$ and its column vectors are linearly independent. Define

$$X^{rf} = \left[X_r^{rf\mathrm{T}}\ X_1^{rf\mathrm{T}}\ X_2^{rf\mathrm{T}}\ \cdots\ X_{\alpha_0}^{rf\mathrm{T}}\right]^{\mathrm{T}}, \tag{25}$$

where $X_r^{rf} \in \mathbb{R}^{r_0\times z_0},\ X_i^{rf} \in \mathbb{R}^{r_i\times z_0}$.

In order to make (A, B) controllable, we give a lemma for the controllability of the Jordan normal form system first.

Lemma 4 ([20])**.** For the Jordan normal form system (13), the necessary and sufficient condition for the complete controllability is

$$\mathrm{rank}(B_{\lambda_i}) = \alpha_i, \forall i \in \mathcal{I}, \tag{26}$$

where $\mathcal{I} = \{1, 2, \cdots, l\}$, α_i is the geometric multiplicity of the eigenvalue λ_i of A.

Based on the above lemma, we propose a design for the matrix B that ensures $\mathrm{Im}(X^{df}) \subseteq \mathrm{Im}(\widetilde{X})$ and that the pair (A, B) is controllable. Although the design of B is not unique, the matrix B we designed achieves the minimal rank among all matrices satisfying these conditions.

Theorem 2 (Design B)**.** Consider the multi-agent system (1), where $A \in \mathbb{R}^{n\times n}$ is given as A in (14). For any desired formation space $\mathrm{Im}(X^{df})$, design the matrix B as $B = [B^{df}\ B^c]$,

$$B^{df} \triangleq \begin{bmatrix} Ar\, X_r^{rf} \\ B_0^1 \\ B_0^2 \\ \vdots \\ B_0^{\alpha_0} \end{bmatrix},\ B_0^j \triangleq \begin{bmatrix} X_j^{rf}\,(2:r_j) \\ 0 \end{bmatrix} \text{for } j = 1, 2, \cdots, \alpha_0,$$

where X_j^{rf} is defined in (25), $X_j^{rf}(2:r_j)$ is formed by the last $(r_j - 1)$ rows of X_j^{rf}. For each eigenvalue λ_i of A (with algebraic multiplicity σ_i), we define the submatrices $B_{\lambda_i}^{df}$ and $B_{\lambda_i}^c$ to have σ_i rows. These submatrices are obtained by partitioning the previously defined matrix B^{df} (and the corresponding matrix B^c) as follows:

$$\begin{bmatrix} B_{\lambda_1}^{df} & B_{\lambda_1}^c \\ B_{\lambda_2}^{df} & B_{\lambda_2}^c \\ \vdots & \vdots \\ B_{\lambda_l}^{df} & B_{\lambda_l}^c \end{bmatrix} = \left[B^{df}\ B^c\right]. \tag{27}$$

Then B^c is chosen such that

$$\mathrm{rank}([B_{\lambda_i}^{df}\ B_{\lambda_i}^c]) = \alpha_i, \text{for } i \in \mathcal{I},$$

where $\mathcal{I} = \{1, 2, \ldots, l\}$, and α_i is the geometric multiplicity of the eigenvalue λ_i of the matrix A. Plus, the number of columns in B^c is given by $\max_{i\in\mathcal{I}}\{\alpha_i - \mathrm{rank}(B_{\lambda_i}^{df})\}$.

With such matrix B, there is $\mathrm{Im}(X^{df}) \subseteq \mathrm{Im}(\widetilde{X})$, where $\widetilde{X}$ is determined by (16) and (A, B) is controllable. Moreover, the matrix B achieves the minimal rank among all matrices satisfying these conditions.

Proof. In (16), define matrix $\widetilde{X}_L$ such that $[\widetilde{X}_L\ \ U_A] = \widetilde{X}$. According to the definition of the residual desired formation space $\mathrm{Im}(X^{rf})$ in (24), in order to make $\mathrm{Im}(X^{df}) \subseteq \mathrm{Im}(\widetilde{X})$, we only need to make $\mathrm{Im}(X^{rf}) \subseteq \mathrm{Im}(\widetilde{X}_L)$. Thus, we set $\widetilde{X}_L = X^{rf}$. Then by (16) and (25), we have

$$\widetilde{X}_L = \begin{bmatrix} A_r^{-1}B_r U_{B_{\mathrm{last}}} \\ \widetilde{X}_1 \\ \widetilde{X}_2 \\ \vdots \\ \widetilde{X}_{\alpha_0} \end{bmatrix} = X^{rf} = \begin{bmatrix} X_r^{rf} \\ X_1^{rf} \\ X_2^{rf} \\ \vdots \\ X_{\alpha_0}^{rf} \end{bmatrix},$$

where

$$\widetilde{X}_j = \begin{bmatrix} 0 \\ B_0^j\,(1:r_j-1)\,U_{B_{\mathrm{last}}} \end{bmatrix},\ \text{for } j = 1, 2, \cdots, \alpha_0.$$

Without loss of generality, we can choose B which satisfies

$$B_{\text{last}} = 0,\ U_{B_{\text{last}}} = I, \tag{28}$$

and

$$B_r = A_r X_r^{rf}, \tag{29}$$

$$B_0^j\,(1 : r_j - 1) = X_j^{rf}\,(2 : r_j)\,. \tag{30}$$

From (28), we have $B_0^j\,(r_j) = 0$, for $j = 1, 2, \cdots, \alpha_0$. And combining (29) and (30), we have

$$B = \begin{bmatrix} A_r\, X_r^{rf} \\ B_0^1 \\ B_0^2 \\ \vdots \\ B_0^{\alpha_0} \end{bmatrix}, B_0^j = \begin{bmatrix} X_j^{rf}\,(2 : r_j) \\ 0 \end{bmatrix}, j = 1, 2, \cdots, \alpha_0. \tag{31}$$

Then $\text{Im}(X^{df}) \subseteq \text{Im}(\widetilde{X})$.

Then building on the condition that $\text{Im}(X^{df}) \subseteq \text{Im}(\widetilde{X})$, we further require that the pair (A, B) is controllable. Define B in (31) as B^{df}, which makes the desired formation $\text{Im}(X^{df})$ belong to $\text{Im}(\widetilde{X})$. According to Lemma 4, for (A, B^{df}) to be controllable, it is necessary that $\text{rank}(B_{\lambda_i}^{df}) = \alpha_i, \forall i \in \mathcal{I}$, where $B_{\lambda_i}^{df}$ is defined in (27). If $\exists\, i \in \mathcal{I}$ such that $\alpha_i > \text{rank}(B_{\lambda_i}^{df})$, then (A, B^{df}) is not controllable. Thus we need to increase the rank of B^{df} at least by $\max_{i\in\mathcal{I}} \alpha_i - \text{rank}(B_{\lambda_i}^{df})$. This means the minimal rank of B satisfying $\text{Im}(X^{df}) \subseteq \text{Im}(\widetilde{X})$ and (A, B) controllable is

$$\text{rank}(B^{df}) + \max_{i\in\mathcal{I}}\{\alpha_i - \text{rank}(B_{\lambda_i}^{df})\}. \tag{32}$$

Let B^c be a matrix whose number of columns is $\max_{i\in\mathcal{I}}\{\alpha_i - \text{rank}(B_{\lambda_i}^{df})\}$. Define $B = \begin{bmatrix} B^{df} B^c \end{bmatrix}$. We can always find B^c such that $\text{rank}([B_{\lambda_i}^{df}\ B_{\lambda_i}^c]) = \alpha_i, \forall i \in \mathcal{I}$. Hence (A, B) is controllable. Since the rank of B equals (32), B achieves the minimal rank among all matrices satisfying $\text{Im}(X^{df}) \subseteq \text{Im}(\widetilde{X})$ and that (A, B) is controllable, which completes the proof.

Then we consider several special cases for Theorem 2, where B^{df} in B has a simpler expression.

(1) When the desired formation space $\text{Im}(X^{df})$ belongs to $\text{Im}(U_A)$, there is $X^{rf} = 0$ and we have $B^{df} = 0$.

(2) When A has no zero eigenvalues, we have $B^{df} = AX^{df}$.

(3) When all Jordan blocks associated with the zero eigenvalue of A are 1-dimensional, we have $B^{df} = \begin{bmatrix} A_r X_r^{rf} \\ 0 \end{bmatrix}$.

4.2 Step 2: design the exogenous system and the performance index

For any desired formation space $\text{Im}(X^{df})$, an input matrix B can be designed with minimal rank such that $\text{Im}(X^{df}) \subseteq \text{Im}(\widetilde{X})$ and the pair (A, B) is controllable. Subsequently, to ensure that the system state converges to a specific formation $\text{Im}(x_{df})$ (with $x_{df} \in \mathbb{R}^n$) within the maximal steady-state space $\text{Im}(\widetilde{X})$, the exogenous system and the performance index must be appropriately designed.

It is important to note that the design of the exogenous system is not unique; the same formation may be achieved through different control strategies implemented by various exogenous systems, as illustrated in Case 3 of Section 5 (illustrative examples). This flexibility prevents unauthorized parties from deducing the desired formation based on historical control inputs, thereby enhancing system security. Here, we propose a design that prioritizes minimizing the dimension of the exogenous system.

Theorem 3 (Design exogenous system and performance index)**.** Consider the optimal control problem (5), where matrices $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$ are given as (14). Consider a desired formation vector $x_{df} \neq 0 \in \mathbb{R}^n$ such that $\text{Im}(x_{df}) \subseteq \text{Im}(\widetilde{X}) \setminus \{V_+(A), V_{mg(0)\geqslant 2}(A)\}$, where $\widetilde{X}$ is defined in (16), $V_+(A)$ and $V_{\text{mg}(0)\geqslant 2}(A)$ denote, respectively, the sets of eigenvectors of A corresponding to eigenvalues with positive real parts and those associated with the zero eigenvalue having Jordan blocks of size at least 2. Then the control strategy described by (3) and (7) can drive the state of the system (1) to $\text{Im}(x_{df})$ under the following exogenous system and the performance index matrix C:

$$H = U_{B_{last}} \in \mathbb{R}^{m\times k}, \tag{33a}$$

$$[F_1\ F_2] = [0_{k\times n}\ 0_{k\times k}], \tag{33b}$$

$$\left\{K \in \mathbb{R}^{m\times m}, G \in \mathbb{R}^{k\times m} \middle\| \left(\begin{bmatrix} A & BH \\ F_1 & F_2 \end{bmatrix}, \begin{bmatrix} BK \\ G \end{bmatrix}\right) \text{ is stabilizable}\right\}, \tag{33c}$$

$$C = [U_{x_{df}^{\mathrm{T}}}]^{\mathrm{T}} \in \mathbb{R}^{(n-1)\times n}, \tag{33d}$$

where B_{last} is defined in (17), and $U_{B_{\text{last}}}$ and $U_{x_{df}^{\mathrm{T}}}$ are matrices whose columns form a standard orthonormal basis for $\mathrm{Ker}(B_{\text{last}})$ and $\mathrm{Ker}(x_{df}^{\mathrm{T}})$.

Proof. The proof is presented in three parts.

Part 1: This step shows the design of H in (33a) ensures that any formation in $\mathrm{Im}(\widetilde{X})$ is possible to be achievable. The effect of H is reflected in $BHw_{ss} = -Ax_{ss}$. Since B is full column rank, we have

$$Hw_{ss} = -(B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}Ax_{ss}. \tag{34}$$

If H is chosen such that $\forall x_{ss} \in \mathrm{Im}(\widetilde{X})$, Eq. (34) is satisfied, then x_{ss} is possible to be achievable. Thus H should satisfy $\mathrm{Im}(-(B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}A\widetilde{X}) \subseteq \mathrm{Im}(H)$. When $\mathrm{Im}(H) = \mathrm{Im}((B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}A\widetilde{X})$, H achieves the minimal rank. Considering $AU_A = 0$, we have

$$A\widetilde{X} = \begin{bmatrix} A_r & & & & \\ & J_1(0) & & & \\ & & J_2(0) & & \\ & & & \ddots & \\ & & & & J_{\alpha_0}(0) \end{bmatrix} \begin{bmatrix} A_r^{-1}B_rU_{B_{last}} \\ \widetilde{X}_1 \\ \widetilde{X}_2 \\ \vdots \\ \widetilde{X}_{\alpha_0} \end{bmatrix}.$$

Compute

$$J_i(0)\widetilde{X}_i = J_i(0)\begin{bmatrix} 0 \\ B_0^i\,(1:r_i-1)\,U_{B_{\text{last}}} \end{bmatrix} = \begin{bmatrix} B_0^i\,(1:r_i-1)\,U_{B_{\text{last}}} \\ 0 \end{bmatrix} = B_0^iU_{B_{\text{last}}},$$

for $i = 1, 2, \ldots, \alpha_0$, where we use $B_0^i(r_i)U_{B_{last}} = 0$. Then we have $A\widetilde{X} = B\,U_{B_{last}}$. Finally, we have $\mathrm{Im}(H) = \mathrm{Im}((B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}A\widetilde{X}) = \mathrm{Im}\big((B^{\mathrm{T}}B)^{-1}B^{\mathrm{T}}B\,U_{B_{last}}\big) = \mathrm{Im}(U_{B_{last}})$. Since $U_{B_{last}}$ is the matrix whose columns form a standard orthonormal basis for $\mathrm{Ker}(B_{last})$, choosing $H = U_{B_{last}}$ minimizes the number of columns in H. As the dimension of the exogenous system is equal to the number of columns in H, this choice makes the exogenous system have the minimal dimension.

Part 2: In this step, we demonstrate that our design of F_1, F_2, and C in (33b) and (33d) renders $\begin{bmatrix} x_{df} \\ w_0 \end{bmatrix}$, the unique (up to a scalar multiple) nonzero vector in the steady-state space $\mathrm{Ker}(\bar{A}_{cl})$, where w_0 satisfies $Ax_{df} + BHw_0 = 0$. By Theorem 1, $\mathrm{Ker}(\bar{A}_{cl}) = \mathrm{Ker}([\,\bar{A}^{\mathrm{T}}\ \bar{C}^{\mathrm{T}}\,]^{\mathrm{T}})$. Therefore, it suffices to prove that $\mathrm{Ker}([\,\bar{A}^{\mathrm{T}}\ \bar{C}^{\mathrm{T}}\,]^{\mathrm{T}}) = \mathrm{Im}\left(\begin{bmatrix} x_{df} \\ w_0 \end{bmatrix}\right)$.

Define $\begin{bmatrix} x \\ w \end{bmatrix} \in \mathrm{Ker}([\,\bar{A}^{\mathrm{T}}\ \bar{C}^{\mathrm{T}}\,]^{\mathrm{T}})$, $x \in \mathbb{R}^n$, $w \in \mathbb{R}^k$. When $[F_1\ F_2] = [0\ 0]$, $C = [U_{x_{df}^{\mathrm{T}}}]^{\mathrm{T}}$, we have

$$Ax + BHw = 0, \tag{35a}$$

$$[U_{x_{df}^{\mathrm{T}}}]^{\mathrm{T}}x = 0. \tag{35b}$$

Because $U_{x_{df}^{\mathrm{T}}}$ is a matrix whose columns form a standard orthonormal basis for $\mathrm{Ker}(x_{df}^{\mathrm{T}})$, the unique (up to a scalar multiple) nonzero solution of (35b) is x_{df}. Thus if w_0 satisfies $Ax_{df} + BHw_0 = 0$, $\begin{bmatrix} x_{df} \\ w_0 \end{bmatrix}$ is one solution for (35). Another possible solution is when $x = 0, \exists w \neq 0$ which satisfies $BHw = 0$. But such w does not exist because B and H are full column rank matrices. Thus $\mathrm{Ker}([\,\bar{A}^{\mathrm{T}}\ \bar{C}^{\mathrm{T}}\,]^{\mathrm{T}}) = \mathrm{Im}\left(\begin{bmatrix} x_{df} \\ w_0 \end{bmatrix}\right)$.

Part 3: This step shows that Assumptions 1 and 2 are satisfied. Because K and G in (33c) ensure that the system is stabilizable, Assumption 1 is satisfied. Then we prove Assumption 2. If there exists an undetectable eigenvalue λ for $\bar{A}$, then $\exists\phi\neq 0$ which satisfies $\bar{A}\phi=\lambda\phi,\ \bar{C}\phi=0$. Because $\bar{A}=\begin{bmatrix}A & BH\\ 0 & 0\end{bmatrix}$, $\bar{C}=\begin{bmatrix}C & 0\end{bmatrix}$, define $\phi=\begin{bmatrix}\phi_1\\ \phi_2\end{bmatrix}$, and then we have

$$A\phi_1+BH\phi_2=\lambda\phi_1, \tag{36}$$

$$\lambda\phi_2=0, \tag{37}$$

$$C\phi_1=0. \tag{38}$$

Because $C=[U_{x_{df}^{\mathrm{T}}}]^{\mathrm{T}}$ in (38), we have

$$\phi_1=x_{df}\text{ or }0. \tag{39}$$

When $\lambda\neq 0$, because of (37), $\phi_2=0$. Because $\phi\neq 0$, ϕ_1 can only be x_{df} to satisfy (39). Then Eq. (36) becomes $Ax_{df}=\lambda x_{df}$. Because x_{df} cannot be chosen from $V_+(A)$, $\mathrm{Re}(\lambda)\leqslant 0$. Thus $\bar{A}$ has no undetectable eigenvalue with positive real part.

When $\lambda=0$, we have

$$A\phi_1+BH\phi_2=0, \tag{40}$$

$$C\phi_1=0.$$

Thus ϕ_1 also satisfies (39). If $\phi_1=0$, because BH is column full rank in (40), we have $\phi_2=0$, which contradicts $\phi\neq 0$. Thus we have $\phi_1=x_{df}$. Because $x_{df}\in\mathrm{Im}(\widetilde{X})$, there must exist ϕ_2 satisfying $Ax_{df}+BH\phi_2=0$. Define $\phi_2=w_0$, then $\bar{A}$ must have undetectable eigenvalue 0 with eigenvector $\begin{bmatrix}x_{df}\\ w_0\end{bmatrix}$. Then we need to prove the algebraic multiplicity and geometric multiplicity are equal for the undetectable eigenvalue 0. If the condition is not satisfied, $\exists\hat{\phi}\neq 0$ such that $\bar{A}\hat{\phi}=\phi$ and $\bar{C}\hat{\phi}=0$, where ϕ satisfies $\bar{A}\phi=0$ and $\bar{C}\phi=0$. Because $\mathrm{Ker}([\,\bar{A}^{\mathrm{T}}\ \bar{C}^{\mathrm{T}}\,]^{\mathrm{T}})=\mathrm{Im}\left(\begin{bmatrix}x_{df}\\ w_0\end{bmatrix}\right)$ as proved in **Part 2**, ϕ can only be $\begin{bmatrix}x_{df}\\ w_0\end{bmatrix}$, where w_0 satisfies

$$Ax_{df}+BHw_0=0. \tag{41}$$

Define $\hat{\phi}=\begin{bmatrix}\hat{\phi}_1\\ \hat{\phi}_2\end{bmatrix}$, and then we have

$$A\hat{\phi}_1+BH\hat{\phi}_2=x_{df}, \tag{42}$$

$$w_0=0, \tag{43}$$

$$C\hat{\phi}_1=0. \tag{44}$$

Because $C=[U_{x_{df}^{\mathrm{T}}}]^{\mathrm{T}}$ in (44), we have

$$\hat{\phi}_1=x_{df}\text{ or }0. \tag{45}$$

Because of (41) and (43), we have

$$Ax_{df}=0. \tag{46}$$

Then we take (45) into (42) and get

$$BH\hat{\phi}_2=x_{df}. \tag{47}$$

Because $H=U_{B_{last}}$, we can prove that $\mathrm{Im}(BH)\subseteq\mathrm{Im}(A)$. Then $\exists\,\varphi\in\mathbb{R}^n$ which satisfies $BH\hat{\phi}_2=A\varphi$. Considering (47), $A\varphi=x_{df}$. Because x_{df} is the eigenvector of 0 for A by (46), φ is a generalized-eigenvector of 0 for A. Because x_{df} cannot be chosen from $V_{\mathrm{mg}(0)\geqslant 2}(A)$, such φ does not exist. Thus, there does not exist $\hat{\phi}_2$, which indicates that $\bar{A}$ does not have any generalized-eigenvector of 0 corresponding to undetectable subspace. Therefore, Assumption 2 is satisfied.

Based on the proofs in **Parts 1–3**, we conclude that $\bar{A}_{cl}$ is marginally stable and the state of the closed-loop system (9) converges to $\mathrm{Im}\left(\begin{bmatrix}x_{df}\\ w_0\end{bmatrix}\right)$. Consequently, the state of system (1) will converge to $\mathrm{Im}(x_{df})$. This completes the proof.

Remark 2. In Theorem 3, we set $H = U_{B_{last}}$. It is possible when $\mathrm{Ker}(B_{last}) = \{0\}$. In this case, $\mathrm{Im}(\widetilde{X}) = \mathrm{Im}(U_A) = \mathrm{Ker}(A)$, which means the steady state is driven at most towards the null space of the matrix A. Actually, such formation can be achieved by the original optimal control problem (2), which does not involve an exogenous system. Actually, Theorem 3 is compatible with this case, because the exogenous system design in Theorem 3 is $H = F_1 = F_2 = G = 0 \in \mathbb{R}$, $\{K \in \mathbb{R}^{m\times m} \mid (A, BK) \text{ is stabilizable}\}$, and $C = [U_{x_{df}^{\mathrm{T}}}]^{\mathrm{T}}$. If (A, B) is controllable, we may choose $K = I_m$. With this design, the optimal control problem with the exogenous system (5) reduces to the original optimal control problem (2).

4.3 Step 3: design the initial state of the exogenous system

Based on Lemma 3, the steady state of the optimal control problem (5) is related to the initial state of the combined system, $\bar{x}(0)$. Typically, the initial state of the multi-agent system, $x(0)$, is fixed, whereas the initial state of the exogenous system, $w(0)$, can be any bounded value. Hence, by selecting an appropriate $w(0)$, one can achieve the desired formation scaling. The following theorem describes the method for determining $w(0)$.

Theorem 4 (Design the initial state of the exogenous system)**.** Suppose the exogenous system and the performance index are designed such that the state of system (1) will converge to $\mathrm{Im}(x_{df})$ under the control strategy described by (3) and (7), where $x_{df} \in \mathbb{R}^n$. Then the closed-loop matrix $\bar{A}_{\mathrm{cl}}$ has only one eigenvector $\phi \in \mathbb{R}^{(n+k)\times 1}$ of the zero eigenvalue, which satisfies $\bar{A}_{cl}\phi = 0$ and $\phi^{\mathrm{T}}\phi = 1$. Denote $\phi = \begin{bmatrix}\phi_x \\ \\ \phi_w\end{bmatrix}$, where $\phi_x \in \mathbb{R}^n$, $\phi_w \in \mathbb{R}^k$. Given the initial state of the multi-agent system $x(0)$ and a desired formation scaling $d \cdot x_{df}$ with $d \neq 0$, this scaling is achieved when the initial state of the exogenous system $w(0)$ satisfies

$$\phi_w^{\mathrm{T}} w(0) = d\frac{\|x_{df}\|}{\|\phi_x\|} - \phi_x^{\mathrm{T}} x(0), \tag{48}$$

where $\|x_{df}\|$ and $\|\phi_x\|$ represent the norms of the vectors x_{df} and ϕ_x, respectively.

Proof. Because the closed-loop matrix $\bar{A}_{\mathrm{cl}}$ has only one eigenvector ϕ of the zero eigenvalue, according to Lemma 3,

$$\bar{x}_{ss} = \left(\phi^{\mathrm{T}}\bar{x}(0)\right)\phi = \left(\phi_x^{\mathrm{T}} x(0) + \phi_w^{\mathrm{T}} w(0)\right)\phi. \tag{49}$$

Considering $\bar{x}_{ss} = \begin{bmatrix}x_{ss} \\ w_{ss}\end{bmatrix}$ and $\phi = \begin{bmatrix}\phi_x \\ \\ \phi_w\end{bmatrix}$, we have

$$x_{ss} = \left(\phi_x^{\mathrm{T}} x(0) + \phi_w^{\mathrm{T}} w(0)\right)\phi_x. \tag{50}$$

Because $w(0)$ satisfies (48), there is $x_{ss} = d\frac{\|x_{df}\|}{\|\phi_x\|}\cdot\phi_x = d\cdot x_{df}$. This equation holds because x_{df} and ϕ_x are linearly dependent. Therefore, the scaling is achieved and the proof is complete.

Algorithm 1 summarizes the design of system to achieve a desired formation $d \cdot x_{df}$, which includes three steps. In the first step, the algorithm checks whether an input matrix B is provided. If it is provided, the algorithm skips constructing B. If not, then B is constructed according to Theorem 2, which ensures $\mathrm{Im}(X^{df}) \subseteq \mathrm{Im}(\widetilde{X})$ and makes (A, B) controllable. In the second step, it constructs the exogenous system and performance index according to Theorem 3 if $\mathrm{Im}(x_{df}) \subseteq \mathrm{Im}(\widetilde{X})\backslash\{V_+(A), V_{mg(0)\geqslant 2}(A)\}$. Otherwise, it terminates and states "Infeasible", which means that the formation $d \cdot x_{df}$ cannot be achieved. In the third step, the algorithm checks if $\phi_w \neq 0$ or $d\frac{\|x_{df}\|}{\|\phi_x\|} = \phi_x^{\mathrm{T}} x(0)$. If true, it computes the initial state $w(0)$ according to Theorem 4. Otherwise, the algorithm terminates and states "Infeasible", which means that the formation $d \cdot x_{df}$ cannot be achieved.

5 Illustrative examples

In this section, we will illustrate the effectiveness of our main results. Consider the state evolution equation of four agents in three-dimensional space

$$\dot{\hat{x}} = \hat{A}\hat{x} + \hat{B}u = (-L \otimes I_3)\hat{x} + \hat{B}u, \tag{51}$$

Algorithm 1 System design to achieve a desired formation $d \cdot x_{df}$.

Input System matrix A, collection of desired formations $\mathrm{Im}(X^{df})$, desired formation with scaling $d \cdot x_{df}$, initial state $x(0)$ and input matrix B (may be omitted);
Step 1: Construct B if not provided;
$B \leftarrow$ getInputMatrixB();
if B is not provided **then**
  Compute $\mathrm{Im}(X^{rf}) = \mathrm{Im}(X^{df}) \cap (\mathrm{Im}(U_A))^{\perp}$;
  Compute B^{df} and B^{c} according to Theorem 2;
  Construct $B = \begin{bmatrix} B^{df} & B^{c} \end{bmatrix}$;
  return B;
else
  Skip constructing B;
end if
Step 2: Construct the exogenous system and the performance index;
Compute $\widetilde{X}$ according to (16);
if $\mathrm{Im}(x_{df}) \subseteq \mathrm{Im}(\widetilde{X}) \setminus \{V_{+}(A), V_{mg(0)\geqslant 2}(A)\}$ **then**
  Construct $H = U_{B_{last}}$;
  Construct matrix $[F_1\ F_2] = [0_{k\times n}\ 0_{k\times k}]$;
  Construct matrices K and G such that $(\bar{A}, \bar{B})$ is stabilizable;
  Construct matrix $C = [U_{x_{df}^{\mathrm{T}}}]^{\mathrm{T}}$;
  return H, F_1, F_2, K, G and C;
else
  return "Infeasible";
end if
Step 3: Construct the initial state $w(0)$;
Compute the closed-loop matrix $\bar{A}_{cl}$ and ϕ_x, ϕ_w;
if $\phi_w \neq 0$ or $d\frac{\|x_{df}\|}{\|\phi_x\|} = \phi_x^{\mathrm{T}} x(0)$; **then**
  $w(0)$ is selected satisfying: $\phi_w^{\mathrm{T}} w(0) = d\frac{\|x_{df}\|}{\|\phi_x\|} - \phi_x^{\mathrm{T}} x(0)$;
  return $w(0)$;
else
  return "Infeasible";
end if

where $L = \begin{bmatrix} 2 & -1 & 0 & -1 \\ -1 & 3 & -2 & 0 \\ 0 & -2 & 5 & -3 \\ -1 & 0 & -3 & 4 \end{bmatrix}$. We introduce a linear transformation to bring $\hat{A}$ into its Jordan normal form A, $A = \mathrm{diag}(0, -2.4746, -3.3691, -8.1563) \otimes I_3$, all of whose Jordan blocks associated with the zero eigenvalue are 1-dimensional. Under the Jordan normal form, we use theories presented in the aforementioned paper to compute required matrices and states. Then we apply an inverse linear transformation to obtain the corresponding matrices and states for system (51). It is worth noting that the exogenous system (including the state w and the matrices H, F_1, F_2, K, G) remains unchanged before and after the linear transformation. This can be interpreted as the exogenous system undergoing an identity transformation. For simplicity, we omit the results for the Jordan normal form and present only those for system (51).

Case 1 is an example where only part of the agents are controlled. Consider $\hat{B} = \begin{bmatrix} I_9 \\ 0_{3\times 9} \end{bmatrix}$, where only agents 1–3 are controlled directly. We compute the maximum steady-state space $\mathrm{Im}(\widetilde{X})$ according to (16) and select a desired formation within $\mathrm{Im}(\widetilde{X})$ as follows:

$$x_{df}^{1} = \begin{bmatrix} 1 & 1 & 0 & 10/3 & 10/3 & 0 & 7/3 & 7/3 & 0 & 2 & 2 & 0 \end{bmatrix}^{\mathrm{T}},$$

which corresponds to a line formation. The initial state of system (51) is selected arbitrarily as $\hat{x}(0) = [1.5\ 1\ 0.5\ 3.3\ 3\ 0.3\ 3\ 2\ 0.2\ 2\ 1.5\ 0]^{\mathrm{T}}$. Then we use Algorithm 1 to design the exogenous system, the performance index matrix C and the initial state of exogenous system, where $w(0) \approx \begin{bmatrix} 0.2684 & 0 \end{bmatrix}^{\mathrm{T}}$. Under these conditions, the results are shown in Figure 1. In Figure 1(a), the states of the agents stabilize quickly; after a short time we have $\hat{x} \approx \begin{bmatrix} 1 & 1 & 0 & 3.3 & 3.3 & 0 & 2.3 & 2.3 & 0 & 2 & 2 & 0 \end{bmatrix}^{\mathrm{T}} \approx x_{df}^{1}$. This is consistent with Figure 1(b), where the formation tracking error $\|\hat{x} - x_{df}^{1}\|_2$ decreases rapidly to nearly zero over time. Figure 1(c) shows the trajectories of agents, where the agents ultimately form a line formation.

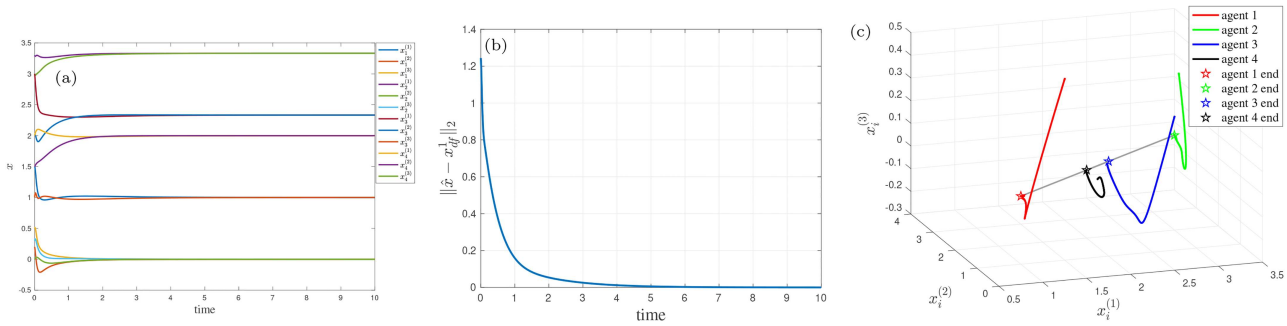

Figure 1 (Color online) Results for linear formation x_{df}^1. (a) Evolution of states; (b) formation tracking error $\|\hat{x} - x_{df}^1\|_2$; (c) trajectories of agents for linear formation x_{df}^1.

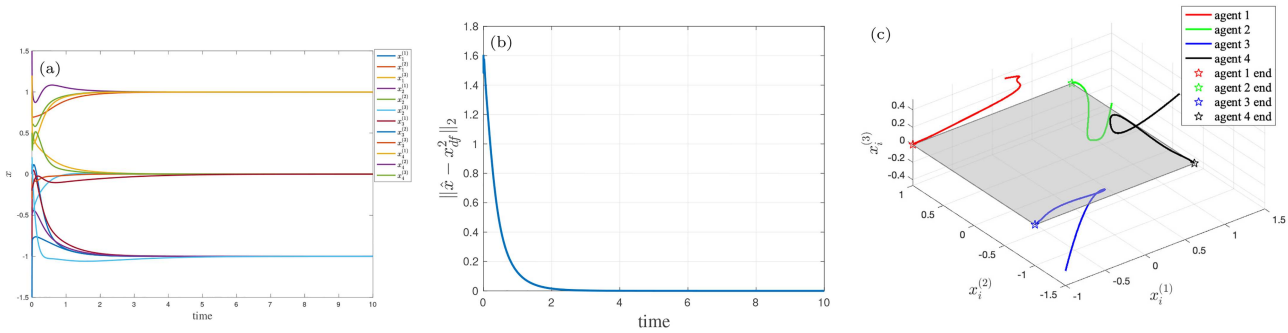

Figure 2 (Color online) Results for square formation x_{df}^2. (a) Evolution of states; (b) formation tracking error $\|\hat{x} - x_{df}^2\|_2$; (c) trajectories of agents for square formation x_{df}^2.

Case 2 is an example where the input matrix $\hat{B}$ is not provided. Consider a collection of two desired formations which includes the square formation x_{df}^2 and the tetrahedron formation x_{df}^3, where

$$x_{df}^2 = \begin{bmatrix} -1 & 1 & 0 & 1 & 1 & 0 & -1 & -1 & 0 & 1 & -1 & 0 \end{bmatrix}^\top,$$
$$x_{df}^3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 1.5 & 0 & -1 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}^\mathrm{T}.$$

We define $X^{df} = \left[x_{df}^2, x_{df}^3\right]$ and compute $\hat{B}$ according to Algorithm 1 which makes $\mathrm{Im}(X^{df})$ belong to the maximum steady-state space. Then we design different exogenous systems to achieve x_{df}^2 and x_{df}^3, respectively.

We first consider x_{df}^2 and choose the initial state arbitrarily as

$$\hat{x}(0) = \begin{bmatrix} 0 & 0.8 & 0.5 & 1.5 & 1 & -0.4 & -1 & -1.5 & -0.3 & 1.2 & -0.5 & 0.5 \end{bmatrix}^\mathrm{T}.$$

Then we use Algorithm 1 to design the exogenous system, determine the performance index matrix C and get the initial state of exogenous system $w(0) \approx \begin{bmatrix} 0.0345 & 0 \end{bmatrix}^\mathrm{T}$. Figure 2(a) shows that the states of the agents tend to stabilize; after a short time we have $\hat{x} \approx \begin{bmatrix} -1 & 1 & 0 & 1 & 1 & 0 & -1 & -1 & 0 & 1 & -1 & 0 \end{bmatrix}^\mathrm{T} = x_{df}^2$. This aligns with Figure 2(b), which shows the formation tracking error $\|\hat{x} - x_{df}^2\|_2$ quickly decaying to nearly zero. This is also consistent with the trajectories depicted in Figure 2(c), where the agents eventually form a square formation.

Similarly, we design another exogenous system to achieve the tetrahedron formation x_{df}^3. The initial state is set arbitrarily as $\hat{x}(0) = \begin{bmatrix} 0 & 2 & 0 & 1 & 0 & -1 & 0 & -2 & -1 & 0 & 0 & 0 \end{bmatrix}^\mathrm{T}$. Applying Algorithm 1 yields $w(0) \approx \begin{bmatrix} -2.3559 & 0 \end{bmatrix}^\mathrm{T}$. Figure 3(a) shows that after a short time we obtain $\hat{x} \approx \begin{bmatrix} 1 & 0 & 0 & 0 & 1.5 & 0 & -1 & 0 & 0 & 0 & 0 & 2 \end{bmatrix}^\mathrm{T} = x_{df}^3$. Consistent with Figure 3(b), the formation tracking error $\|\hat{x} - x_{df}^3\|_2$ drops swiftly toward zero. This is also corroborated by the trajectories in Figure 3(c), where the agents ultimately form a tetrahedron formation.

Moreover, in Figures 1(c), 2(c), and 3(c), the agents do not travel from their initial positions to the desired formation along the shortest path. Instead, they follow more circuitous routes, which helps to confuse potential

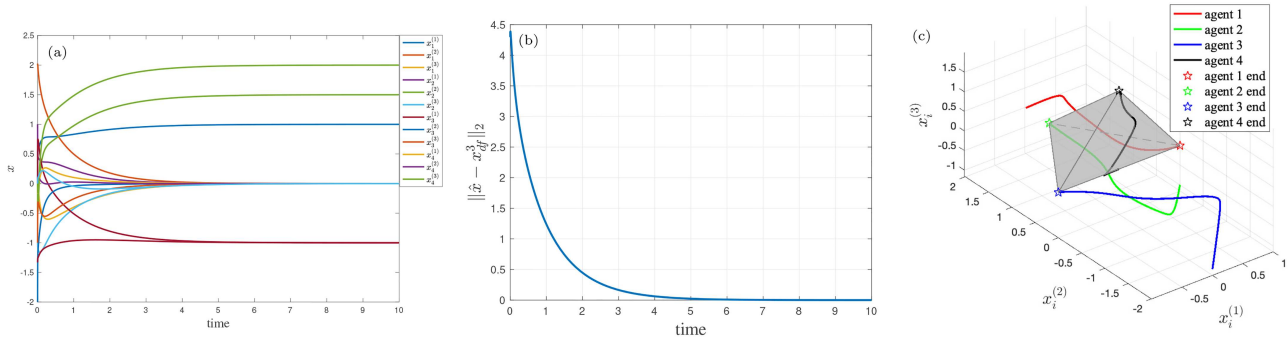

Figure 3 (Color online) Results for tetrahedron formation x_{df}^3. (a) Evolution of states; (b) formation tracking error $\|\hat{x}-x_{df}^3\|_2$; (c) trajectories of agents for tetrahedron formation x_{df}^3.

adversaries who might attempt to intercept or seize the formation. This arises because, in the optimal control problem (5), the objective is to minimize the input of the exogenous system rather than that of the multi-agent system.

In **Case 3**, we use different exogenous systems to achieve the same formation. In the tetrahedron formation example of Case 2, we obtained the exogenous system $\mathbf{S}_1$ by Algorithm 1 as follows:

$$H=\begin{bmatrix}1&0&0&0&0\\0&1&0&0&0\end{bmatrix}^{\mathrm{T}},\ [F_1\ F_2]=[0_{2\times 12}\ 0_{2\times 2}],$$

$$K=\begin{bmatrix}4&11&16&1&6\\1&4&2&8&12\\16&13&6&2&20\\4&10&4&2&9\\8&3&11&16&14\end{bmatrix},\ G=\begin{bmatrix}15&13&19&5&10\\9&2&3&16&16\end{bmatrix},$$

where the choices of K and G are not unique as long as they ensure that the system is stabilizable. We can achieve the same formation using a different exogenous system $\mathbf{S}_2$:

$$H=\begin{bmatrix}1.1&-0.8&0&0&0\\0.17&0.71&0&0&0\end{bmatrix}^{\mathrm{T}},\ [F_1\ \ F_2]=[0_{2\times 12}\ 0_{2\times 2}],$$

$$K=\begin{bmatrix}-18&-14&-10&-13&-6\\6&2&14&3&-7\\-3&18&18&-13&-2\\17&7&-5&-9&-11\\-3&-4&-18&-2&-19\end{bmatrix},G=\begin{bmatrix}6&12&-7&10&-16\\-14&-4&10&-13&-13\end{bmatrix}.$$

Exogenous systems $\mathbf{S}_1$ and $\mathbf{S}_2$ employ different control strategies for system (51) because the solutions of the Riccati equation P are different. To achieve the same formation scaling, the initial conditions for the two exogenous systems are distinct. Specifically, for $\mathbf{S}_1$, $w(0)\approx\begin{bmatrix}-2.3559&0\end{bmatrix}^{\mathrm{T}}$, while for $\mathbf{S}_2$, $w(0)\approx\begin{bmatrix}-0.6682&0\end{bmatrix}^{\mathrm{T}}$. Under the control generated by the exogenous system $\mathbf{S}_2$, we obtain $\hat{x}\approx\begin{bmatrix}1&0&0&0&1.5&0&-1&0&0&0&0&2\end{bmatrix}^{\mathrm{T}}=x_{df}^3$ only after a short time, which is the same as the tetrahedron formation in Case 2 with exogenous system $\mathbf{S}_1$. Figures 4(a) and (b) show the trajectories of the agents under exogenous system $\mathbf{S}_1$ and $\mathbf{S}_2$, respectively. These results demonstrate that by employing different control strategies, the agents follow distinct paths to reach the same formation.

6 Conclusion

In this paper, we study the intrinsic formation problem of multi-agent systems by introducing an exogenous system within an infinite time-horizon linear quadratic optimal control framework. By formulating the problem without

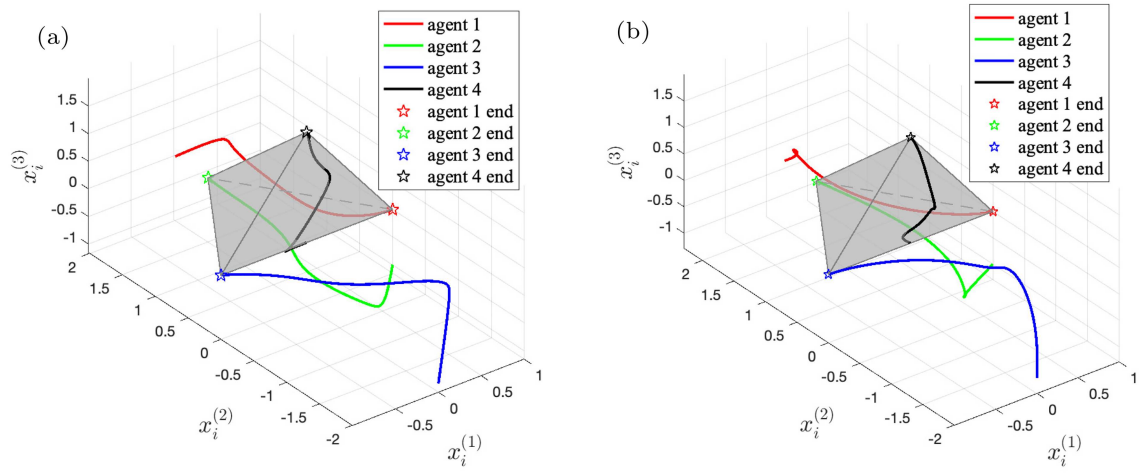

Figure 4 (Color online) Comparison of agents' trajectories under different exogenous systems for the formation x_{df}^3. (a) Under exogenous system $\mathbf{S}_1$; (b) under exogenous system $\mathbf{S}_2$.

explicitly incorporating the desired formation into the performance index or controller, we enhance the security of the target formation. For the forward problem, we identify the existence condition of a nonzero steady state and characterize the steady-state space. For the inverse problem, we design the input matrix B, exogenous system matrices and performance index to achieve the desired formation while minimizing the number of control inputs and the dimension of the exogenous system. We also design the initial state of the exogenous system to achieve the desired formation scaling. Numerical simulations validate the effectiveness of our methods. Our approach extends intrinsic control methods to the realization of almost arbitrary formations in any dimension—with the sole exception of eigenvectors of A corresponding to eigenvalues with positive real parts and those associated with the zero eigenvalue having Jordan blocks of size at least 2—thereby overcoming limitations of previous studies confined to limited patterns.

The proposed intrinsic-formation framework can be applied to multiple multi-agent missions. For UAV swarms, different exogenous systems can be used to generate different control while achieving the same target formation, thereby confusing adversaries attempting to steal the target formation. For deep-space satellite formations, intrinsic formation sustains the required geometries under long delays and weak observability, reducing reliance on ground-issued formation commands. In the future, we plan to extend our framework to more complex agent dynamics and incorporate uncertainties to enhance robustness, including exploration of nonlinear behaviors and higher-order systems. Besides, we can also adopt distributed control strategies which represent a promising approach to improve scalability and efficiency, especially in large-scale multi-agent networks. Finally, applying differential game methods could broaden the applicability of our approach, enabling the design of control laws in competitive environments.

Acknowledgements This work was supported in part by National Natural Science Foundation of China (Grant Nos. 62373245, 12426311), National Key Research and Development Program of China (Grant No. 2023YFB4706800), and Shanghai Oriental Talent Program (Grant No. BJJY2025063).

References

1 Zhang X, Yang Q, Xiao F, et al. Linear formation control of multi-agent systems. Automatica, 2025, 171: 111935
2 Li Y, Hu X. A differential game approach to intrinsic formation control. Automatica, 2022, 136: 110077
3 Zhao S, Li Z, Ding Z. Bearing-only formation tracking control of multiagent systems. IEEE Trans Automat Contr, 2019, 64: 4541–4554
4 Vu H M, Trinh M H, Van Tran Q, et al. Distance-based formation tracking of single- and double-integrator agents. IEEE Trans Automat Contr, 2023, 69: 1332–1339
5 Dong X, Hu G. Time-varying formation control for general linear multi-agent systems with switching directed topologies. Automatica, 2016, 73: 47–55
6 Roza A, Maggiore M, Scardovi L. A smooth distributed feedback for formation control of unicycles. IEEE Trans Automat Contr, 2019, 64: 4998–5011
7 Zhou P, Chen B M. Formation-containment control of Euler-Lagrange systems of leaders with bounded unknown inputs. IEEE Trans Cybern, 2020, 52: 6342–6353
8 Oh K K, Park M C, Ahn H S. A survey of multi-agent formation control. Automatica, 2015, 53: 424–440
9 Wang J N, Xin M. Integrated optimal formation control of multiple unmanned aerial vehicles. IEEE Trans Contr Syst Technol, 2012, 21: 1731–1744

10 Chang L, Gao S, Wang Z. Optimal control of pattern formations for an SIR reaction-diffusion epidemic model. J Theor Biol, 2022, 536: 111003
11 Liu M, Li Y, Zhu L, et al. Formation control for mixed-order UAVs-USVs-UUVs systems under cooperative and optimal control. JMSE, 2023, 11: 704
12 Coogan S, Arcak M. Scaling the size of a formation using relative position feedback. Automatica, 2012, 48: 2677–2685
13 Song W, Markdahl J, Zhang S, et al. Intrinsic reduced attitude formation with ring inter-agent graph. Automatica, 2017, 85: 193–201
14 Zhang S, Song W, He F, et al. Intrinsic tetrahedron formation of reduced attitude. Automatica, 2018, 87: 375–382
15 Zhang S, He F, Hong Y, et al. An intrinsic approach to formation control of regular polyhedra for reduced attitudes. Automatica, 2020, 111: 108619
16 Li T, Li Y, Liu Z, et al. Pattern formation using an intrinsic optimal control approach. 2025. ArXiv:2505.01302
17 Yang Y, Xiao Y, Li T. Attacks on formation control for multiagent systems. IEEE Trans Cybern, 2021, 52: 12805–12817
18 Kučera V. A review of the matrix Riccati equation. Kybernetika, 1973, 9: 42–61
19 Trentelman H L, Stoorvogel A A, Hautus M. Control Theory for Linear Systems. London: Springer, 2001
20 Chen C T. Linear System Theory and Design. New York: Holt, Rinehart and Winston, 1984