ESAIM: COCV 29 (2023) 83
https://doi.org/10.1051/cocv/2023060

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

STACKELBERG METHOD TO STABILIZE GAME-BASED CONTROL SYSTEM

Yue Sun[1], Juanjuan Xu[2], Huanshui Zhang[3,*] and Renren Zhang[4]

Abstract. In this paper, we are concerned with the stabilization problem of the game-based control system. In particular, two players are involved in the system where one is to minimize the related cost function and the other is to stabilize the system. Different from the previous works, the new contribution is to derive the necessary and sufficient condition for the stabilization of the game-based control system by applying Stackelberg game method. The key technique is to explicitly solve the forward and backward difference equations (FBDEs) from the Stackelberg game and give the optimal feedback gain matrix of the leader by using the matrix maximum principle.

Mathematics Subject Classification. 49N10, 91A65.

Received July 28, 2022. Accepted August 13, 2023.

1. Introduction

In modern control theory, stabilizability is an essential and important concept, especially in system analysis and synthesis. During the past few years, there have been extensive studies. Brian D. O. Anderson, in 1971, has presented the stabilization of the time-invariant deterministic system in Section 3.2 of [1]. Under some additional assumptions, the stabilization of the closed-loop system was formed when the control law is resulted from optimizing an infinite-time performance index [1]. Subsequently, the stabilization problem of control systems has been extended to stochastic systems, and then extended to time-delay systems, successively; see [3, 8, 13, 18, 19, 23] and reference therein. In [23], the stabilization of discrete-time systems with delay and multiplicative noise was studied. It showed that the system is stabilizable in the mean-square sense if and only if an algebraic Riccati-ZXL equation has a particular solution.

The above mentioned results belong to classical control theory, that is, the controllers have the common goals of minimizing the performances and stabilizing the system. As the control theory is applied to more and more fields, the classical control theory will no longer meet the demand, where the controlled system has multiple objectives to realize, such as economic in social system [5], immune system in biological system [6], ecosystem [14], smart grid and intelligent transportation in engineering system and so on. For example, for the

Keywords and phrases: FBDEs, game-based control system, stabilization, Stackelberg game.

[1] School of Control Science and Engineering, Shandong University, Jinan, Shandong 250061, PR China.
[2] School of Control Science and Engineering, Shandong University, Jinan, Shandong 250061, PR China.
[3] College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao, Shandong 266590, PR China.
[4] Research Center for Mathematics and Interdisciplinary Sciences, Shandong University, Qingdao, Shandong 266237, PR China.

[*] Corresponding author: hszhang@sdu.edu.cn

virus and the immune system in [6], systemic therapy developed was often initially successful. However, with the execution of the treatment, cancer or virus almost invariably evolved resistance. There generates different objectives for the virus and the immune system, where the purpose of the immune system, which can be viewed as the macro-controller, is to keep the body stable, while the virus, which can be viewed as the the rational agent, has its own objective to pursue. Moreover, with the process of socio-economic modeling in recent years, it is often necessary to consider the strategic interactions between different decision makers and the government, especially in order to better achieve the objectives of economic planning and policies [9], where the government modeled as the macro-controller, stabilizes the market through macro-control, while the different decision makers modeled as the rational agents, want to maximize their own interests. The common feature of these systems mentioned above is that the controlled system contains multiple controllers, which can be divided into rational agent and macro-controller. The different status between the rational agent and macro-controller leads to the objectives to be pursued differently. In particular, the effective way to solve this kind of problem is the game method and such controlled systems can be modeled as the game-based systems. Most recently, a new control framework called game-based control system (GBCS) is introduced in [20], which is composed by a hierarchical decision-making structure, heterogeneous agents, different objectives for the plant and the control, and macro regulation.

Motivated by [20], we will consider the stabilization problem for game-based system with Stackelberg method, such as [6] and [9]. There has been extensive studies on the stabilization of Stackelberg game, such as [4, 15, 21, 24]. The infinite-horizon linear-quadratic Stackelberg games for discrete-time stochastic system with multiple decision makers was studied in [12]. Necessary conditions for the existence of the Stackelberg strategy set were derived in terms of the solvability of cross-coupled stochastic algebraic equations. The linear-quadratic Stackelberg differential games including time preference rates with an open-loop information structure was investigated in [10], and sufficient conditions to guarantee a predefined degree of stability were given based on the distribution of the eigenvalues in the complex plane. By using a memoryless state feedback representations, [11] developed an incentive strategy for discrete-time LQ state feedback Stackelberg games. It designed an incentive policy for the leader, while the followers rational reaction guaranteed system stability. [17] proposed an adaptive dynamic programming algorithm by solving the coupled partial differential equations and the Stackelberg feedback equilibrium solution was obtained to ensure the stability of the system according to the Lyapunov function. However, the necessary and sufficient condition for the stabilization problem of Stackelberg game is still a challenge.

In this paper, we consider the Stackelberg method to stabilize the game-based system. In particular, the rational controller modeled as the follower is to minimize its own cost function, while the macro-controller modeled as the leader is to stabilize the system, which is different from the classical control theory of Stackelberg game. In order to address the problem mentioned above, we consider the finite-horizon open-loop Stackelberg strategy firstly. In the optimization of the follower, based on the maximum principle, the non-homogeneous relationship between the costate and state is constructed and there derives a backward equation. By solving the FBDEs, a homogeneous relationship between the state and non-homogeneous terms in costate equation is derived. In the optimization of the leader, by using the matrix maximum principle, the optimization of the leader is converted into finding the optimal gain matrix, which minimizes the cost function and is subject to the state equation. It is noted that the analytical solution is the feedback form of the state and the calculation of which is more concise than the augmented state obtained in [16]. Finally, we derive the necessary and sufficient condition for the leader to stabilize the game-based system in stabilization.

The rest of the paper is organized as follows. Section 2 presents some preliminaries of the stabilization problem for the GBCS. The finite-time horizon results are given in Section 3. The stabilization of the GBCS is settled in Section 4. Numerical examples are given in Section 5. Conclusions are provided in Section 6.

Notation: R^n denotes the $n-$dimensional Euclidean space and A^T denotes the transpose of matrix A. A symmetric matrix $M > 0$ (reps. ≥ 0) means that it is positive definite (reps. positive semi-definite). If $f = f(x, y)$, f_x represents the partial derivative of $f(x, y)$ with respect to x.

2. Problem formulation

Considering the following discrete-time GBCS

$$x_{k+1} = Ax_k + B_1 u_k^1 + B_2 u_k^2, \tag{2.1}$$

where $x_k \in R^n$ is the state, $u_k^1 \in R^{m_1}$ and $u_k^2 \in R^{m_2}$ are the control inputs of the follower and the leader, respectively. A and B_i, $i = 1, 2$ are constant matrices of appropriate dimensions. The initial state is $x_0 \in R^n$, and the cost function to be minimized by u_k^1 is

$$J_1 = \sum_{k=0}^{\infty} (x_k^T Q_1 x_k + u_k^{1T} R_{11} u_k^1 + u_k^{2T} R_{12} u_k^2), \tag{2.2}$$

where Q_1, R_{12} are positive semi-definite matrices and R_{11} is positive definite matrix of compatible dimensions.

Referring from [7], by assuming the homogeneous linear relation between the three adjoint states, three coupling and asymmetric Riccati equations were proposed to design optimal strategies, where the analytical Stackelberg equilibrium had been expressed in feedback form. In this paper, we consider the admissible control set $\mathcal{U}_2$ for the leader in the feedback form to stabilize the system (2.1):

$$\mathcal{U}_2 = \{u_k^2 = K^2 x_k : K^2 \in R^{m_2 \times n}\}, \tag{2.3}$$

and the admissible control set of the follower is assumed to be $\mathcal{U}_1 = \{u_k^1 : x_0 \times [0, \infty) \longrightarrow R^{m_1}\}$.

Now, we are in the position to give the main problem.

Problem 2.1. Under the optimal controller $u_k^1 \in \mathcal{U}_1$ which minimizes the cost function J_1, find the necessary and sufficient conditions for $u_k^2 = K^2 x_k \in \mathcal{U}_2$ to stabilize the GBCS (2.1).

Remark 2.2. It's clear that Problem 2.1 is different from the classical control of Stackelberg game. The difficulty of Problem 2.1 is to obtain the necessary and sufficient condition for the stabilization of the game-based system, which is compared with the previous work [12], where only the solvable condition of the Stackelberg strategy in the infinite horizon was given.

In order to fundamentally solve the Problem 2.1 mentioned above, we convert the problem of the stabilization of the GBCS into an optimization problem based on the Stackelberg game, and denote the cost function of the leader as

$$J_2 = \sum_{k=0}^{\infty} (x_k^T Q_2 x_k + u_k^{1T} R_{21} u_k^1 + u_k^{2T} R_{22} u_k^2), \tag{2.4}$$

where Q_2 is positive semi-definite matrices, such that there exist some matrices C_2 satisfying $Q_2 = C_2^T C_2$ and R_{21}, R_{22} are positive definite matrices of compatible dimensions.

3. Finite-horizon results

It should be pointed out that the controller to minimize (2.4) will be the best stabilization controller for (2.1). Now we will find the controller u_k^1 and u_k^2 to minimize (2.1) and (2.4), respectively. To this end, we firstly consider the following problem in finite-time horizon.

Problem 3.1. Find the unique open-loop Stackelberg strategy (u_k^1, u_k^2), subject to (2.1), and such that

$$J_N^1(u_k^{1*}(u_k^2), u_k^2) \le J_N^1(u_k^1(u_k^2), u_k^2),$$
$$J_N^2(u_k^{1*}(u_k^{2*}), u_k^{2*}) \le J_N^2(u_k^{1*}(u_k^2), u_k^2),$$

where

$$J_N^1 = \sum_{k=0}^{N}(x_k^T Q_1 x_k + u_k^{1T} R_{11} u_k^1 + u_k^{2T} R_{12} u_k^2) + x_{N+1}^T H_1 x_{N+1}, \tag{3.1}$$

$$J_N^2 = \sum_{k=0}^{N}(x_k^T Q_2 x_k + u_k^{1T} R_{21} u_k^1 + u_k^{2T} R_{22} u_k^2) + x_{N+1}^T H_2 x_{N+1}, \tag{3.2}$$

with $H_1 \geq 0$ and $H_2 \geq 0$ of compatible dimensions.

To solve the Problem 3.1, we introduce the following Riccati equations

$$P_k^1 = Q_1 + A^T P_{k+1}^1 A - A^T P_{k+1}^1 B_1 (\Gamma_{k+1}^1)^{-1} B_1^T P_{k+1}^1 A, \tag{3.3}$$

$$P_k^2 = Q_2 + A^T Y_{k+1}^T R_{21} Y_{k+1} A + A^T M_{k+1}^{1T} (\Upsilon_{k+1}^{-1})^T P_{k+1}^2 \Upsilon_{k+1}^{-1} M_{k+1}^1 A - M_{k+1}^{2T} (\Gamma_{k+1}^2)^{-1} M_{k+1}^2, \tag{3.4}$$

with terminal values $P_{N+1}^1 = H_1$ and $P_{N+1}^2 = H_2$, where

$$\begin{aligned}
\Gamma_{k+1}^1 &= R_{11} + B_1^T P_{k+1}^1 B_1,\\
\Gamma_{k+1}^2 &= R_{22} + B_2^T Y_{k+1}^T R_{21} Y_{k+1} B_2 + B_2^T M_{k+1}^{1T} (\Upsilon_{k+1}^{-1})^T P_{k+1}^2 \Upsilon_{k+1}^{-1} M_{k+1}^1 B_2,\\
M_{k+1}^1 &= I - B_1 (\Gamma_{k+1}^1)^{-1} B_1^T P_{k+1}^1,\\
M_{k+1}^2 &= B_2^T Y_{k+1}^T R_{21} Y_{k+1} A + B_2^T M_{k+1}^{1T} (\Upsilon_{k+1}^{-1})^T P_{k+1}^2 \Upsilon_{k+1}^{-1} M_{k+1}^1 A,\\
S_{k+1} &= (\Gamma_{k+1}^1)^{-1} B_1^T P_{k+1}^1,\\
\Upsilon_{k+1} &= I + B_1 (\Gamma_{k+1}^1)^{-1} B_1^T T_{k+1}, && (3.5)\\
Y_{k+1} &= S_{k+1} + (\Gamma_{k+1}^1)^{-1} B_1^T T_{k+1} \Upsilon_{k+1}^{-1} M_{k+1}^1,\\
K_k^1 &= -Y_{k+1}(A + B_2 K_k^2), && (3.6)\\
K_k^2 &= -(\Gamma_{k+1}^2)^{-1} M_{k+1}^2, && (3.7)\\
T_k &= A^T M_{k+1}^{1T} T_{k+1} \Upsilon_{k+1}^{-1} M_{k+1}^1 (A + B_2 K_k^2) + A^T M_{k+1}^{1T} P_{k+1}^1 B_2 K_k^2, && (3.8)
\end{aligned}$$

with the terminal value $T_{N+1} = 0$.

Remark 3.2. The invertibility of Γ_{k+1}^1 and Γ_{k+1}^2 are guaranteed by $R_{11} > 0$ and $R_{22} > 0$. In (3.5), it is assumed that Υ_{k+1} is invertible. Otherwise, the recursion stops.

We now give the main results in Theorem 3.3 to find the open-loop Stackelberg strategy stated in Problem 3.1.

Theorem 3.3. *If the Riccati equations (3.3)–(3.4) admit solutions such that $\Gamma_{k+1}^1 > 0$, $\Gamma_{k+1}^2 > 0$, and under the condition Υ_{k+1} is invertible, then Problem 3.1 has a unique open-loop Stackelberg strategy. And the optimal Stackelberg strategy for the follower and the leader are*

$$u_k^1 = K_k^1 x_k, \tag{3.9}$$

$$u_k^2 = K_k^2 x_k. \tag{3.10}$$

The optimal cost functions of the follower and the leader are given as

$$J_N^{1*} = x_0^T [P_0^1 + T_0^T + T_0 + \Xi] x_0, \tag{3.11}$$

$$J_N^{2*} = x_0^T P_0^2 x_0, \tag{3.12}$$

where

$$\begin{aligned}\Xi = \sum_{k=0}^{N}[&\Phi_{k+1,1}^T T_{k+1}^T M_{k+1}^1 B_2 K_k^2 \Phi_{k,1} - \Phi_{k+1,1}^T T_{k+1}^T B_1 (\Gamma_{k+1}^1)^{-1} B_1^T T_{k+1} \Phi_{k+1,1} \\ &+\Phi_{k,1}^T K_k^{2T}(R_{12} + B_2^T P_{k+1}^1 M_{k+1}^1 B_2) K_k^2 \Phi_{k,1} + \Phi_{k,1}^T K_k^{2T} B_2^T M_{k+1}^{1T} T_{k+1} \Phi_{k+1,1}],\end{aligned}$$

with

$$\begin{aligned}\Phi_{k+1,1} &= A_{k+1} A_k \dots A_1, \\ A_{k+1} &= \Upsilon_{k+1}^{-1} M_{k+1}^1 (A + B_2 K_k^2).\end{aligned}$$

Proof. The proof will be divided into three steps. The first step is to optimize the follower by the maximum principle. The second step is to optimize the leader by the matrix maximum principle. Finally, we will calculate the optimal cost functions of the two players.

3.1. The optimization of the follower

The optimization of the follower is considered firstly. By using the maximum principle, the non-homogeneous relationship between the state and the costate is formulated in this subsection.

Following from [22], and applying Pontryagin's maximum principle to the system (2.1) with cost function (3.1), the following costate equations is derived:

$$0 = R_{11} u_k^1 + B_1^T \lambda_k, \tag{3.13}$$

$$\lambda_{k-1} = A^T \lambda_k + Q_1 x_k, \tag{3.14}$$

with the terminal value $\lambda_N = H_1 x_{N+1}$.

The existence of u_k^2 in (2.1) leads to the relationship between λ_{k-1} and x_k are no longer homogeneous. Thus, we assume that

$$\lambda_k = P_{k+1}^1 x_{k+1} + \zeta_k, \tag{3.15}$$

with terminal value $\zeta_N = 0$.

Adding (3.15) into the equilibrium equation (3.13), then it can be rewritten as

$$\begin{aligned}0 &= R_{11} u_k^1 + B_1^T (P_{k+1}^1 x_{k+1} + \zeta_k) \\ &= (R_{11} + B_1^T P_{k+1}^1 B_1) u_k^1 + B_1^T P_{k+1}^1 A x_k + B_1^T P_{k+1}^1 B_2 u_k^2 + B_1^T \zeta_k,\end{aligned}$$

since $\Gamma_{k+1}^1 > 0$, thus, the control of the follower can be calculated as

$$u_k^1 = -(\Gamma_{k+1}^1)^{-1} (B_1^T P_{k+1}^1 A x_k + B_1^T P_{k+1}^1 B_2 u_k^2 + B_1^T \zeta_k). \tag{3.16}$$

Next, we shall prove that (3.15) is established for any $0 \le k \le N$ by the method of inductive hypothesis, where P_k^1 satisfies (3.3) and ζ_{k-1} satisfies the following equation,

$$\zeta_{k-1} = A^T M_{k+1}^{1T} \zeta_k + A^T M_{k+1}^{1T} P_{k+1}^1 B_2 u_k^2. \tag{3.17}$$

According to $\lambda_N = P_{N+1}^1 x_{N+1}$ and $\zeta_N = 0$, we can derive that (3.15) is established for $k = N$. Given any $s \ge 0$, assume that (3.15) is established for any $k \ge s$, where P_{k+1}^1 and ζ_k satisfy (3.3) and (3.17), respectively, we shall show that (3.15) also holds for $k = s - 1$.

By using (3.15) and (3.14), then it derives

$$\begin{aligned}\lambda_{s-1} &= A^T(P^1_{s+1}x_{s+1}+\zeta_s)+Q_1x_s\\ &= A^TP^1_{s+1}Ax_s + A^TP^1_{s+1}B_1u^1_s + A^TP^1_{s+1}B_2u^2_s + A^T\zeta_s + Q_1x_s\\ &= (Q_1 + A^TP^1_{s+1}A - A^TP^1_{s+1}B_1(\Gamma^1_{s+1})^{-1}B_1^TP^1_{s+1}A)x_s + A^TM^{1T}_{s+1}\zeta_s + A^TM^{1T}_{s+1}P^1_{s+1}B_2u^2_s\\ &= P^1_sx_s + \zeta_{s-1},\end{aligned}$$

which indicates that (3.15) holds for $k=s-1$ with P^1_s and ζ_{s-1} satisfying (3.3) and (3.17), respectively.

Following from (3.16), then the state (2.1) can be rewritten as

$$x_{k+1} = M^1_{k+1}Ax_k + M^1_{k+1}B_2u^2_k - B_1(\Gamma^1_{k+1})^{-1}B_1^T\zeta_k. \tag{3.18}$$

3.2. The optimization of the leader

In this subsection, the optimization of the leader is considered, that is, minimizing (3.2), subject to (3.17)–(3.18), where u^1_k is given in (3.16). The feedback form of the leader, *i.e.*, $u^2_k = K^2_kx_k$, is used in the following optimization due to the equivalence between the open-loop and the feedback solutions for the leader, which is explained in Remark 3.4. We will iteratively solve the FBDEs (3.17)–(3.18) firstly. And then, by using the matrix maximum principle, the optimal feedback gain matrix of the leader is yielded.

Adding $u^2_k = K^2_kx_k$ into (3.17) and (3.18), respectively, then there derives the FBDEs

$$x_{k+1} = M^1_{k+1}(A+B_2K^2_k)x_k - B_1(\Gamma^1_{k+1})^{-1}B_1^T\zeta_k, \tag{3.19}$$

$$\zeta_{k-1} = A^TM^{1T}_{k+1}\zeta_k + A^TM^{1T}_{k+1}P^1_{k+1}B_2K^2_kx_k, \tag{3.20}$$

with the initial value x_0 and the terminal value $\zeta_N = 0$.

Then, by using inductive hypothesis, we will proof the FBDEs satisfy the homogeneous relationship

$$\zeta_{k-1} = T_kx_k, \tag{3.21}$$

and in this way, we have

$$x_{k+1} = \Upsilon^{-1}_{k+1}M^1_{k+1}(A+B_2K^2_k)x_k. \tag{3.22}$$

For $k=N$, it yields

$$\begin{aligned}x_{N+1} &= M^1_{N+1}(A+B_2K^2_N)x_N,\\ \zeta_{N-1} &= A^TM^{1T}_{N+1}P^1_{N+1}B_2K^2_Nx_N = T_Nx_N,\end{aligned}$$

which satisfy (3.21) and (3.22) for $k=N$.

We take any $n\ge 0$, and assume that ζ_{k-1} and x_{k+1} are as (3.21) and (3.22) for all $k\ge n+1$. We show that these conditions will also holds for $k=n$.

For $k=n$, it follows that

$$\begin{aligned}x_{n+1} &= M^1_{n+1}(A+B_2K^2_n)x_n - B_1(\Gamma^1_{n+1})^{-1}B_1^T\zeta_n\\ &= M^1_{n+1}(A+B_2K^2_n)x_n - B_1(\Gamma^1_{n+1})^{-1}B_1^TT_{n+1}x_{n+1},\end{aligned}$$

i.e., $\Upsilon_{n+1}x_{n+1} = M^1_{n+1}(A+B_2K^2_n)x_n$. Since Υ_{n+1} is invertible, then the state x_{n+1} is such that

$$x_{n+1} = \Upsilon^{-1}_{n+1}M^1_{n+1}(A+B_2K^2_n)x_n,$$

which is exactly (3.22) for $k = n$. And the costate equation ζ_{n-1} satisfies

$$\begin{aligned}\zeta_{n-1} &= A^T M_{n+1}^{1T}\zeta_n + A^T M_{n+1}^{1T} P_{n+1}^1 B_2 K_n^2 x_n \\ &= [A^T M_{n+1}^{1T} T_{n+1}\Upsilon_{n+1}^{-1} M_{n+1}^1 (A + B_2 K_n^2) + A^T M_{n+1}^{1T} P_{n+1}^1 B_2 K_n^2] x_n = T_n x_n,\end{aligned}$$

which is exactly (3.21) for $k = n$.

To this end, the control input of the follower can be written as

$$\begin{aligned}u_k^1 &= -(\Gamma_{k+1}^1)^{-1}(B_1^T P_{k+1}^1 A x_k + B_1^T P_{k+1}^1 B_2 u_k^2 + B_1^T \zeta_k) \\ &= -S_{k+1}(A + B_2 K_k^2)x_k - (\Gamma_{k+1}^1)^{-1} B_1^T T_{k+1}\Upsilon_{k+1}^{-1} M_{k+1}^1 (A + B_2 K_k^2) x_k \\ &= -Y_{k+1}(A + B_2 K_k^2)x_k. \end{aligned} \tag{3.23}$$

Now, we are in the position to calculate the optimal feedback gain matrix of the leader.

Adding (3.23) and $u_k^2 = K_k^2 x_k$ into (3.2), we have

$$J_N^2 = \sum_{k=0}^{N}[x_k^T Q_2 x_k + x_k^T (A + B_2 K_k^2)^T Y_{k+1}^T R_{21} Y_{k+1}(A + B_2 K_k^2)x_k + x_k^T K_k^{2T} R_{22} K_k^2 x_k] + x_{N+1}^T H_2 x_{N+1}. \tag{3.24}$$

Denote the matrix as $X_{k+1} = x_{k+1}x_{k+1}^T$. Then, combining it with (3.22), it derives that

$$X_{k+1} = \Upsilon_{k+1}^{-1} M_{k+1}^1 (A + B_2 K_k^2) X_k (A + B_2 K_k^2)^T M_{k+1}^{1T} (\Upsilon_{k+1}^{-1})^T. \tag{3.25}$$

Hence, the optimization of the leader can be converted into finding the optimal feedback gain matrix K_k^2, minimizes the cost function (3.24) and satisfies (3.25).

Denote the Hamiltonian function of the leader as

$$\begin{aligned}H_k^2 = Tr\Big(&[Q_2 + (A + B_2 K_k^2)^T Y_{k+1}^T R_{21} Y_{k+1}(A + B_2 K_k^2) + K_k^{2T} R_{22} K_k^2] X_k \\ &+ \Upsilon_{k+1}^{-1} M_{k+1}^1 (A + B_2 K_k^2) X_k (A + B_2 K_k^2)^T M_{k+1}^{1T} (\Upsilon_{k+1}^{-1})^T P_{k+1}^{2T}\Big),\end{aligned} \tag{3.26}$$

applying the matrix maximum principle, it yields

$$\begin{aligned}0 &= \frac{\partial H_k^2}{\partial K_k^2} \\ &= R_{22} K_k^2 X_k + R_{22} K_k^2 X_k^T + B_2^T Y_{k+1}^T R_{21} Y_{k+1} A X_k^T + B_2^T Y_{k+1}^T R_{21} Y_{k+1} A X_k + B_2^T Y_{k+1}^T R_{21} Y_{k+1} \\ &\quad \times B_2 K_k^2 X_k + B_2^T Y_{k+1}^T R_{21} Y_{k+1} B_2 K_k^2 X_k^T + B_2^T M_{k+1}^{1T} (\Upsilon_{k+1}^{-1})^T P_{k+1}^{2T} \Upsilon_{k+1}^{-1} M_{k+1}^1 A X_k \\ &\quad + B_2^T M_{k+1}^{1T} (\Upsilon_{k+1}^{-1})^T P_{k+1}^2 \Upsilon_{k+1}^{-1} M_{k+1}^1 A X_k^T + B_2^T M_{k+1}^{1T} (\Upsilon_{k+1}^{-1})^T P_{k+1}^{2T} \Upsilon_{k+1}^{-1} M_{k+1}^1 B_2 K_k^2 X_k \\ &\quad + B_2^T M_{k+1}^{1T} (\Upsilon_{k+1}^{-1})^T P_{k+1}^2 \Upsilon_{k+1}^{-1} M_{k+1}^1 B_2 K_k^2 X_k^T, \end{aligned} \tag{3.27}$$

$$\begin{aligned}P_k^2 &= \frac{\partial H_k^2}{\partial X_k} \\ &= Q_2 + (A + B_2 K_k^2)^T Y_{k+1}^T R_{21} Y_{k+1} (A + B_2 K_k^2) + K_k^{2T} R_{22} K_k^2 \\ &\quad + (A + B_2 K_k^2)^T M_{k+1}^{1T} (\Upsilon_{k+1}^{-1})^T P_{k+1}^2 \Upsilon_{k+1}^{-1} M_{k+1}^1 (A + B_2 K_k^2),\end{aligned} \tag{3.28}$$

with terminal value $P_{N+1}^2 = H_2$.

According to the symmetric of the matrices X_k and P_{k+1}^2, we have,

$$\begin{aligned}0 &= [R_{22} + B_2^T Y_{k+1}^T R_{21} Y_{k+1} B_2 + B_2^T M_{k+1}^{1T} (\Upsilon_{k+1}^{-1})^T P_{k+1}^2 \Upsilon_{k+1}^{-1} M_{k+1}^1 B_2] K_k^2 \\ &\quad + B_2^T Y_{k+1}^T R_{21} Y_{k+1} A + B_2^T M_{k+1}^{1T} (\Upsilon_{k+1}^{-1})^T P_{k+1}^2 \Upsilon_{k+1}^{-1} M_{k+1}^1 A \\ &= \Gamma_{k+1}^2 K_k^2 + M_{k+1}^2.\end{aligned}$$

Due to $\Gamma_{k+1}^2 > 0$, then the optimal gain matrix of the leader can be uniquely obtained, which satisfies (3.7). Next, we will prove the feedback gain matrix K_k^2 calculated in (3.27) by matrix maximum principle is optimal.

Following from (3.26), we derive

$$\frac{\partial^2 H_k^2}{\partial (K_k^2)^2} = 2X_k \otimes \Gamma_{k+1}^2. \tag{3.29}$$

Since $X_k \geq 0$, then $\frac{\partial^2 H_k^2}{\partial (K_k^2)^2} > 0$ if and only if $\Gamma_{k+1}^2 > 0$, *i.e.*, K_k^2 calculated in (3.27) by matrix maximum principle is optimal.

Based on the discussion above, we can derive that the optimal controllers of the follower and the leader are exactly (3.9) and (3.10).

By using (3.7), the Riccati equation of the leader can be written as

$$\begin{aligned}P_k^2 &= Q_2 + K_k^{2T} \Gamma_{k+1}^2 K_k^2 + K_k^{2T} M_{k+1}^2 + M_{k+1}^{2T} K_k^2 + A^T Y_{k+1}^T R_{21} Y_{k+1} A \\ &\quad + A^T M_{k+1}^{1T} (\Upsilon_{k+1}^{-1})^T P_{k+1}^2 \Upsilon_{k+1}^{-1} M_{k+1}^1 A \\ &= Q_2 + A^T Y_{k+1}^T R_{21} Y_{k+1} A + A^T M_{k+1}^{1T} (\Upsilon_{k+1}^{-1})^T P_{k+1}^2 \Upsilon_{k+1}^{-1} M_{k+1}^1 A - M_{k+1}^{2T} (\Gamma_{k+1}^2)^{-1} M_{k+1}^2,\end{aligned}$$

which is exactly (3.4), with terminal value $P_{N+1}^2 = H_2$.

Lastly, we will calculate the optimal cost functions of the two players. We shall calculate the optimal cost function of the follower firstly.

By applying (2.1) and (3.14), we get

$$\begin{aligned}x_k^T \lambda_{k-1} - x_{k+1}^T \lambda_k &= x_k^T (A^T \lambda_k + Q_1 x_k) - (A x_k + B_1 u_k^1 + B_2 u_k^2)^T \lambda_k \\ &= x_k^T Q_1 x_k - (B_1 u_k^1 + B_2 u_k^2)^T \lambda_k,\end{aligned}$$

Adding the above equation from $k = 0$ to $k = N$, we have

$$x_0^T \lambda_{-1} - x_{N+1}^T \lambda_N = \sum_{k=0}^{N} [x_k^T Q_1 x_k - (B_1 u_k^1 + B_2 u_k^2)^T \lambda_k].$$

Compared the above equation with the cost function (2.2) and based on the equilibrium condition (3.13), there derives

$$\begin{aligned}J_N^1 &= x_0^T \lambda_{-1} + \sum_{k=0}^{N} [u_k^{1T} (R_{11} u_k^1 + B_1^T \lambda_k) + u_k^{2T} (R_{12} u_k^2 + B_2^T \lambda_k)] \\ &= x_0^T \lambda_{-1} + \sum_{k=0}^{N} [u_k^{2T} (R_{12} u_k^2 + B_2^T \lambda_k)].\end{aligned} \tag{3.30}$$

And by using (3.15), (3.17) and (3.18), it yields

$$\begin{aligned} u_k^{2T}B_2^T\lambda_k &= u_k^{2T}B_2^T(P_{k+1}^1x_{k+1}+\zeta_k)\\ &= u_k^{2T}B_2^TP_{k+1}^1(M_{k+1}^1Ax_k+M_{k+1}^1B_2u_k^2-B_1(\Gamma_{k+1}^1)^{-1}B_1^T\zeta_k)+u_k^{2T}B_2^T\zeta_k\\ &= (\zeta_{k-1}-A^TM_{k+1}^{1T}\zeta_k)^Tx_k+u_k^{2T}B_2^TP_{k+1}^1M_{k+1}^1B_2u_k^2+u_k^{2T}B_2^TM_{k+1}^{1T}\zeta_k\\ &= \zeta_{k-1}^Tx_k-\zeta_k^T(x_{k+1}-M_{k+1}^1B_2u_k^2+B_1(\Gamma_{k+1}^1)^{-1}B_1^T\zeta_k)\\ &\quad +u_k^{2T}B_2^TP_{k+1}^1M_{k+1}^1B_2u_k^2+u_k^{2T}B_2^TM_{k+1}^{1T}\zeta_k\\ &= \zeta_{k-1}^Tx_k-\zeta_k^Tx_{k+1}+\zeta_k^TM_{k+1}^1B_2u_k^2-\zeta_k^TB_1(\Gamma_{k+1}^1)^{-1}B_1^T\zeta_k\\ &\quad +u_k^{2T}B_2^TP_{k+1}^1M_{k+1}^1B_2u_k^2+u_k^{2T}B_2^TM_{k+1}^{1T}\zeta_k. \end{aligned} \tag{3.31}$$

Combining (3.30) with (3.31), thus, we have

$$\begin{aligned} J_N^1 &= x_0^T\lambda_{-1}+\zeta_{-1}^Tx_0+\sum_{k=0}^N[\zeta_k^TM_{k+1}^1B_2u_k^2-\zeta_k^TB_1(\Gamma_{k+1}^1)^{-1}B_1^T\zeta_k\\ &\quad +u_k^{2T}(R_{22}+B_2^TP_{k+1}^1M_{k+1}^1B_2)u_k^2+u_k^{2T}B_2^TM_{k+1}^{1T}\zeta_k]. \end{aligned} \tag{3.32}$$

According to (3.22), we can yield that

$$\begin{aligned} x_{k+1} &= \Upsilon_{k+1}^{-1}M_{k+1}^1(A+B_2K_k^2)x_k = A_{k+1}x_k\\ &= A_{k+1}A_kx_{k-1} = \ldots = A_{k+1}A_k\ldots A_1x_0 = \Phi_{k+1,1}x_0. \end{aligned}$$

Adding it into (3.21) and $u_k^2=K_k^2x_k$, there follows $\zeta_k=T_{k+1}\Phi_{k+1,1}x_0$ and $u_k^2=K_k^2\Phi_{k,1}x_0$.
Then the optimal cost function of the follower is

$$\begin{aligned} J_N^{1*} &= x_0^TP_0^1x_0+x_0^T\zeta_{-1}+\zeta_{-1}^Tx_0+\sum_{k=0}^N\Big(x_0^T[\Phi_{k+1,1}^TT_{k+1}^TM_{k+1}^1B_2K_k^2\Phi_{k,1}-\Phi_{k+1,1}^TT_{k+1}^TB_1(\Gamma_{k+1}^1)^{-1}B_1^T\\ &\quad \times T_{k+1}\Phi_{k+1,1}+\Phi_{k,1}^TK_k^{2T}(R_{22}+B_2^TP_{k+1}^1M_{k+1}^1B_2)K_k^2\Phi_{k,1}+\Phi_{k,1}^TK_k^{2T}B_2^TM_{k+1}^{1T}T_{k+1}\Phi_{k+1,1}]x_0\Big)\\ &= x_0^T[P_0^1+T_0^T+T_0+\Xi]x_0. \end{aligned} \tag{3.33}$$

Next, we will consider the optimal cost function of the leader.
Adding

$$u_k^1 = -Y_{k+1}(A+B_2K_k^2)x_k = -Y_{k+1}Ax_k-Y_{k+1}B_2u_k^2$$

into the state (2.1), we have

$$x_{k+1}=(A-B_1Y_{k+1}A)x_k+(B_2-B_1Y_{k+1}B_2)u_k^2. \tag{3.34}$$

Denote the value function of the leader as

$$V_k^2 = x_{k+1}^TP_{k+1}^2x_{k+1}+x_k^TQ_2x_k+u_k^{1T}R_{21}u_k^1+u_k^{2T}R_{22}u_k^2. \tag{3.35}$$

Combining (3.34) with (3.35), the value function can be written as

$$\begin{aligned} V_k^2 &= u_k^{2T}[R_{22} + B_2^T Y_{k+1}^T R_{21} Y_{k+1} B_2 + B_2^T P_{k+1}^2 B_2 - B_2^T P_{k+1}^2 B_1 Y_{k+1} B_2 - B_2^T Y_{k+1}^T B_1^T P_{k+1}^2 B_2 \\ &\quad + B_2^T Y_{k+1}^T B_1^T P_{k+1}^2 B_1 Y_{k+1} B_2] u_k^2 + 2u_k^{2T}[B_2^T P_{k+1}^2 A - B_2^T P_{k+1}^2 B_1 Y_{k+1} A \\ &\quad - B_2^T Y_{k+1}^T B_1^T P_{k+1}^2 A + B_2^T Y_{k+1}^T B_1^T P_{k+1}^2 B_1 Y_{k+1} A + B_2^T Y_{k+1}^T R_{21} Y_{k+1} A] x_k \\ &\quad + x_k^T [Q_2 + A^T P_{k+1}^2 A - A^T P_{k+1}^2 B_1 Y_{k+1} A - A^T Y_{k+1}^T B_1^T P_{k+1}^2 A + A^T Y_{k+1}^T B_1^T P_{k+1}^2 B_1 Y_{k+1} A \\ &\quad + A^T Y_{k+1}^T R_{21} Y_{k+1} A] x_k \\ &= [u_k^2 + (\Gamma_{k+1}^2)^{-1} M_{k+1}^2 x_k]^T \Gamma_{k+1}^2 [u_k^2 + (\Gamma_{k+1}^2)^{-1} M_{k+1}^2 x_k] + x_k^T P_k^2 x_k, \end{aligned} \tag{3.36}$$

where the last equality is established because of

$$I - B_1 Y_{k+1} = M_{k+1}^1 - (I + B_1 (\Gamma_{k+1}^1)^{-1} B_1^T T_{k+1}) \Upsilon_{k+1}^{-1} M_{k+1}^1 + \Upsilon_{k+1}^{-1} M_{k+1}^1 = \Upsilon_{k+1}^{-1} M_{k+1}^1.$$

Substituting (3.36) from $k = 0$ to $k = N$ on both sides of the equation, it yields that

$$\begin{aligned} &x_{N+1}^T P_{N+1}^2 x_{N+1} - x_0^T P_0^2 x_0 + \sum_{k=0}^{N} (x_k^T Q_2 x_k + u_k^{1T} R_{21} u_k^1 + u_k^{2T} R_{22} u_k^2) \\ &= \sum_{k=0}^{N} [u_k^2 + (\Gamma_{k+1}^2)^{-1} M_{k+1}^2 x_k]^T \Gamma_{k+1}^2 [u_k^2 + (\Gamma_{k+1}^2)^{-1} M_{k+1}^2 x_k], \end{aligned}$$

according to the optimal control of the leader is $u_k^2 = -(\Gamma_{k+1}^2)^{-1} M_{k+1}^2 x_k$, and compared with (2.4), the optimal cost function of the leader is exactly (3.12).

Finally, we will show the uniqueness solvable of the open-loop Stackelberg strategy.

Combining (2.1) with (3.3), we have

$$x_k^T P_k^1 x_k - x_{k+1}^T P_{k+1}^1 x_{k+1} = x_k^T Q_1 x_k + u_k^{1T} R_{11} u_k^1 - u_k^{1T} \Gamma_{k+1}^1 u_k^1 - u_k^{2T} B_2^T P_{k+1}^1 B_2 u_k^2 - \Psi_{k+1}, \tag{3.37}$$

where

$$\Psi_{k+1} = x_k^T A^T P_{k+1}^1 B_1 (\Gamma_{k+1}^1)^{-1} B_1^T P_{k+1}^1 A x_k + 2u_k^{1T} B_1^T P_{k+1}^1 A x_k + 2u_k^{1T} B_1^T P_{k+1}^1 B_2 u_k^2 + 2x_k^T A^T P_{k+1}^1 B_2 u_k^2.$$

Substituting (3.37) from $k = 0$ to $k = N$ on both sides, it follows

$$J_N^1 = x_0^T P_0^1 x_0 + \sum_{k=0}^{N} [u_k^{1T} \Gamma_{k+1}^1 u_k^1 + u_k^{2T} (R_{12} + B_2^T P_{k+1}^1 B_2) u_k^2 + \Psi_{k+1}],$$

subsequently, we have $\frac{\partial^2 J_N^1}{\partial (u_k^1)^2} = \Gamma_{k+1}^1$. As a result, $\Gamma_{k+1}^1 > 0$ guarantees the unique solution to the follower.

Following from (3.29), $\Gamma_{k+1}^2 > 0$ guarantees the unique solution to the leader. This completes the proof of Theorem 3.3. □

Remark 3.4. Noting that the feedback form of the leader, *i.e.*, $u_k^2 = K_k^2 x_k$, is equivalent to the open-loop solution in the optimization of the leader. Adding (3.16) into (3.2), J_N^2 is recalculated as

$$\begin{aligned} J_N^2 = \sum_{k=0}^{N} [&x_k^T (Q_2 + A^T P_{k+1}^1 B_1 S_{k+1}^{21} B_1^T P_{k+1}^1 A) x_k + u_k^{2T} (R_{22} + B_2^T P_{k+1}^1 B_1 S_{k+1}^{21} B_1^T P_{k+1}^1 B_2) u_k^2 \\ &+ \zeta_k^T B_1 S_{k+1}^{21} B_1^T \zeta_k + 2x_k^T A^T P_{k+1}^1 B_1 S_{k+1}^{21} B_1^T P_{k+1}^1 B_2 u_k^2 + 2x_k^T A^T P_{k+1}^1 B_1 S_{k+1}^{21} B_1^T \zeta_k \\ &+ 2u_k^{2T} B_2^T P_{k+1}^1 B_1 S_{k+1}^{21} B_1^T \zeta_k] + x_{N+1}^T H_2 x_{N+1} \doteq \sum_{k=0}^{N} f(k, x_k, u_k^2, \zeta_k) + x_{N+1}^T H_2 x_{N+1}, \end{aligned}$$

where $S_{k+1}^{21} = (\Gamma_{k+1}^1)^{-1} R_{21} (\Gamma_{k+1}^1)^{-1}$. Denoting the Hamiltonian function

$$\begin{aligned} H(k, x_k, u_k^2, \zeta_k) = f(k, x_k, u_k^2, \zeta_k) &+ \alpha_k^T (M_{k+1}^1 A x_k + M_{k+1}^1 B_2 u_k^2 - B_1 (\Gamma_{k+1}^1)^{-1} B_1^T \zeta_k) \\ &+ \beta_k^T (A^T M_{k+1}^{1T} \zeta_k + A^T M_{k+1}^{1T} P_{k+1}^1 B_2 u_k^2), \end{aligned}$$

we have the following results by using the maximum principle:

$$0 = \frac{\partial H(k, x_k, u_k^2, \zeta_k)}{\partial u_k^2} = H_{u^2}, \tag{3.38}$$

$$\alpha_{k-1} = \frac{\partial H(k, x_k, u_k^2, \zeta_k)}{\partial x_k} = H_x + H_{u^2} \frac{\partial u_k^2}{\partial x_k}, \tag{3.39}$$

$$\beta_{k+1} = \frac{\partial H(k, x_k, u_k^2, \zeta_k)}{\partial \zeta_k} = H_\zeta, \tag{3.40}$$

with $\alpha_{N+1} = H_2 x_{N+1}$ and $\beta_0 = 0$.

Due to $H_{u^2} = 0$, (3.38)–(3.40) are reduced to the following form:

$$0 = H_{u^2}, \tag{3.41}$$

$$\alpha_{k-1} = H_x, \tag{3.42}$$

$$\beta_{k+1} = H_\zeta, \tag{3.43}$$

where (3.41)–(3.43) are exactly open-loop maximum principle for the leader, which means the equivalence between the open-loop and the feedback solutions for the leader.

Remark 3.5. The advantage of the result in Theorem 3.3 is that the optimal open-loop Stackelberg strategy is a direct feedback form of x_k. Compared to [16], where the optimal open-loop Stackelberg strategy is designed by augmented the state, *i.e.*, the optimal controllers are the feedback form of ξ_k and $\xi_k = \begin{bmatrix} \beta_k \\ x_k \end{bmatrix}$, the calculation is more concise than that in [16].

4. Stabilization results

In this section, the stabilization of Stackelberg GBCS will be investigated.

Before giving the main results, we introduce the following AREs:

$$P^1 = Q_1 + A^T P^1 A - A^T P^1 B_1 (\Gamma^1)^{-1} B_1^T P^1 A, \tag{4.1}$$

$$P^2 = Q_2 + A^T Y^T R_{21} Y A + A^T M^{1T} (\Upsilon^{-1})^T P^2 \Upsilon^{-1} M^1 A - M^{2T} (\Gamma^2)^{-1} M^2. \tag{4.2}$$

where

$$
\begin{aligned}
\Gamma^1 &= R_{11} + B_1^T P^1 B_1, && (4.3)\\
\Gamma^2 &= R_{22} + B_2^T Y^T R_{21} Y B_2 + B_2^T M^{1T} (\Upsilon^{-1})^T P^2 \Upsilon^{-1} M^1 B_2,\\
M^1 &= I - B_1 (\Gamma^1)^{-1} B_1^T P^1,\\
M^2 &= B_2^T Y^T R_{21} Y A + B_2^T M^{1T} (\Upsilon^{-1})^T P^2 \Upsilon^{-1} M^1 A,\\
S &= (\Gamma^1)^{-1} B_1^T P^1,\\
\Upsilon &= I + B_1 (\Gamma^1)^{-1} B_1^T T, && (4.4)\\
Y &= S + (\Gamma^1)^{-1} B_1^T T \Upsilon^{-1} M^1,\\
K^1 &= -Y(A + B_2 K^2), && (4.5)\\
K^2 &= -(\Gamma^2)^{-1} M^2, && (4.6)\\
T &= A^T M^{1T} T \Upsilon^{-1} M^1 (A + B_2 K^2) + A^T M^{1T} P^1 B_2 K^2. && (4.7)
\end{aligned}
$$

Throughout the rest of this paper, the following standard assumption is made, readers may refer to [8].

Assumption 4.1. $(A, Q_2^{\frac{1}{2}})$ is observable.

Now, we will give the main results.

Theorem 4.2. *Under Assumption 4.1, $u_k^2 \in \mathcal{U}_2$ stabilizes the GBCS (2.1) in stabilization if and only if there exists a unique positive definite solution P^2 to the ARE (4.2). In this case, the stabilizing controller is given as*

$$u_k^2 = K^2 x_k, \tag{4.8}$$

while the optimal controller of minimizing J_1 is given as following, i.e.,

$$u_k^1 = K^1 x_k, \tag{4.9}$$

And the optimal cost functions satisfy

$$
\begin{aligned}
J_1^* &= x_0^T (P^1 + T + T^T + \Xi) x_0, && (4.10)\\
J_2^* &= x_0^T P^2 x_0, && (4.11)
\end{aligned}
$$

where

$$
\begin{aligned}
\Xi = \sum_{k=0}^{\infty} [&(\bar{A}^k)^T K^{2T} (R_{12} + B_2^T P^1 M^1 B_2) K^2 \bar{A}^k + (\bar{A}^k)^T K^{2T} B_2^T M^{1T} T \bar{A}^{k+1}\\
&+ (\bar{A}^{k+1})^T T^T M^1 B_2 K^2 \bar{A}^k - (\bar{A}^{k+1})^T T^T B_1 (\Gamma^1)^{-1} B_1^T T \bar{A}^{k+1}],
\end{aligned}
$$

with $\bar{A} = \Upsilon^{-1} M^1 (A + B_2 K^2)$.

Remark 4.3. The main contribution in Theorem 4.2 is that the stabilizable condition for the game-based system has been proposed, while only the solvable condition of the Stackelberg strategy in the infinite horizon was given in [12].

Proof. We give the necessity proof at first.

Necessity: Under Assumption 4.1, suppose there exists controllers u_k^1 and u_k^2 with constants K^1 and K^2 such that u_k^1 minimizes the cost function J_1 and u_k^2 stabilizes the system (2.1) in stabilization, respectively. We will show there exists a unique solution $P^2 > 0$ to the ARE (4.2).

To make the time horizon N explicit in the finite horizon case, we rewrite P_k^1, P_k^2, Γ_{k+1}^1, Γ_{k+1}^2, M_{k+1}^1, M_{k+1}^2, Y_{k+1}, S_{k+1}, Υ_{k+1}, K_k^1, K_k^2 and T_k in (3.3)–(3.8) as $P_k^1(N)$, $P_k^2(N)$, $\Gamma_{k+1}^1(N)$, $\Gamma_{k+1}^2(N)$, $M_{k+1}^1(N)$, $M_{k+1}^2(N)$, $Y_{k+1}(N)$, $S_{k+1}(N)$, $\Upsilon_{k+1}(N)$, $K_k^1(N)$, $K_k^2(N)$ and $T_k(N)$. To consider the infinite horizon case, the terminal weighting matrices in (3.1) and (3.2) are set to be zero, *i.e.*, $H_1 = H_2 = 0$.

The proof will be divided into three parts. The first two steps will show $P_k^1(N)$ and $P_k^2(N)$ are convergent. Then we will prove $\lim_{N\to\infty} P_k^2(N) = P^2 > 0$.

Noted from Theorem 3.3, we can conclude that Problem 3.1 admits a unique solution.

Since $P_k^1(N)$ satisfies the standard Riccati equation

$$P_k^1(N) = Q_1 + A^T P_{k+1}^1(N)A - A^T P_{k+1}^1(N)B_1(\Gamma_{k+1}^1(N))^{-1}B_1^T P_{k+1}^1(N)A,$$

and based on the monotone bounded theorem, we can yield that $\lim_{N\to\infty} P_k^1(N) = P^1$.

Next, we will show $\lim_{N\to\infty} P_k^2(N) = P^2 > 0$.

Rewritten the Riccati equation (3.4) as

$$\begin{aligned}
P_k^2(N) = {} & Q_2 + K_k^{2T}(N)R_{22}K_k^2(N) + [Y_{k+1}(N)A - Y_{k+1}(N)B_2K_k^2(N)]^T R_{21}[Y_{k+1}(N)A - Y_{k+1}(N)B_2K_k^2(N)] \\
& + [\Upsilon_{k+1}^{-1}(N)M_{k+1}^1(N)A - \Upsilon_{k+1}^{-1}(N)M_{k+1}^1(N)B_2K_k^2(N)]^T P_{k+1}^2(N)[\Upsilon_{k+1}^{-1}(N)M_{k+1}^1(N)A - \Upsilon_{k+1}^{-1}(N) \\
& \times M_{k+1}^1(N)B_2K_k^2(N)].
\end{aligned}$$

According to the terminal condition $H_2 = 0$, then by using induction method, we have $P_k^2(N) \geq 0$.

Following from (3.12), it derives that,

$$x_0^T P_0^2(N)x_0 = J_2^*(N) \leq J_2^*(N+1) = x_0^T P_0^2(N+1)x_0, \tag{4.12}$$

the arbitrary of x_0 implies that $P_0^2(N) \leq P_0^2(N+1)$, *i.e.*, $P_0^2(N)$ increases with respect to N. Next the boundedness of $P_0^2(N)$ is to be proved as below.

Since the controllers $u_k^2 = K^2x_k$ stabilizes system (2.1), while $u_k^1 = K^1x_k$ minimizes the cost function (2.2), thus we have

$$\lim_{k\to\infty} x_k = \lim_{k\to\infty}(A + B_1K^1 + B_2K^2)x_{k-1} = 0, \tag{4.13}$$

and then we can yield that $\lim_{k\to\infty} x_k^T x_k = 0$.

Following from [2], it follows that there exist constant c_1 satisfying: $\sum\limits_{k=0}^{\infty} x_k^T x_k \leq c_1 x_0^T x_0$. Therefore, noting from (2.4) that Q_2, $K^{1T}R_{21}K^1$ and $K^{2T}R_{22}K^2$ are both bounded, there exists constant c_2 such that

$$\begin{aligned}
J_2 &= \sum_{k=0}^{\infty}(x_k^T Q_2 x_k + u_k^{1T}R_{21}u_k^1 + u_k^{2T}R_{22}u_k^2) \\
&= \sum_{k=0}^{\infty}[x_k^T(Q_2 + K^{1T}R_{21}K^1 + K^{2T}R_{22}K^2)x_k] \leq c_2 x_0^T x_0.
\end{aligned}$$

Thus for any $N > 0$, from Theorem 3.3 we know that

$$x_0^T P_0^2(N)x_0 = J_2^*(N) \leq J_2 \leq c_2 x_0^T x_0,$$

which indicates that $P_0^2(N)$ is bounded.

Also we know that $P_0^2(N)$ is monotonically increasing, hence $P_0^2(N)$ is convergent, that is $\lim_{N\to\infty} P_0^2(N) = P^2$. According to $P_k^2(N) = P_0^2(N-k)$, we have

$$\lim_{N\to\infty} P_k^2(N) = \lim_{N\to\infty} P_0^2(N-k) = P^2.$$

What should be noted is that we do not use the convergence of the $T_k(N)$ in the proof of the convergence of $P_k^2(N)$, however, according to the Riccati equation (3.4), the existence of the convergence for $P_k^2(N)$ means that $T_k(N)$ is also convergence, thus, we assume $\lim_{N\to\infty} T_k(N) = T$.

Taking limitation of N on both sides of (3.3)–(3.8), therefore, (4.3)–(4.7) can be, respectively, obtained. And P^1, P^2 satisfy the Riccati equations (4.1) and (4.2).

Next, we will show that there exist a positive integer N_0 such that $P_0^2(N_0)$ is positive definite. Suppose this is not the case. Since $P_0^2(N) \geq 0$, then we can get an non-empty set

$$X_N = \{x \in R^n : x \neq 0, x^T P_0^2(N_0)x = 0\}. \tag{4.14}$$

By using (4.12), which means that

$$x_0^T P_0^2(N+1)x_0 = 0, \Rightarrow x_0^T P_0^2(N)x_0 = 0,$$

i.e., $X_{N+1} \subseteq X_N$. Each X_N is a non-empty finite-dimensional set, so

$$1 \leq \cdots \leq dim(X_1) \leq dim(X_0) \leq n,$$

where dim represents the dimension of the set. Thus, there must exist N_1, such that for any $N \geq N_1$,

$$dim(X_N) = dim(X_{N_1})$$

which yields that $X_N = X_{N_1}$, and thus

$$\bigcap_{N\geq 0} X_N = X_{N_1} \neq 0.$$

So there exists a nonzero vector $x \in X_{N_1}$, such that

$$x^T P_0^2(N)x = 0, \forall N \geq 0.$$

Let x_0 be equal to x. Then the optimal value of (2.4) is as

$$J_2^*(N) = \sum_{k=0}^{N} [x_k^{*T} Q_2 x_k^* + (u_k^{1*})^T R_{21} u_k^{1*} + (u_k^{2*})^T R_{22} u_k^{2*}] = x^T P_0^2(N)x = 0,$$

where x_k^*, u_k^{1*} and u_k^{2*} represent the optimal state and the optimal controllers, respectively. Note that $R_{21} > 0$, $R_{22} > 0$ and $Q_2 = C_2^T C_2 \geq 0$. It follows that:

$$C_2 x_k^* = 0, u_k^{1*} = u_k^{2*} = 0, N \geq 0.$$

The observability of $(A, Q_2^{\frac{1}{2}})$ given in Assumption 4.1 indicates that $x_0 = x = 0$, which is a contradiction with $x \neq 0$, *i.e.*, there exists N_0 such that $P_0^2(N) > 0$ for $N > N_0$.

The uniqueness is shown in the following paragraph.

Since P^1 satisfies the standard Riccati equation (4.1), then P^1 is uniquely solvable. Suppose Z^2 is another solution to (4.1), (4.2) and (4.3)–(4.7) satisfying $Z^2 > 0$, *i.e.*,

$$\begin{aligned}
P^1 &= Q_1 + A^T P^1 A - A^T P^1 B_1 (\Gamma^1)^{-1} B_1^T P^1 A,\\
Z^2 &= Q_2 + A^T Y^T R_{21} Y A - M^{2T} (\Gamma^2)^{-1} M^2 + A^T M^{1T} (\Upsilon^{-1})^T Z^2 \Upsilon^{-1} M^1 A.\\
\Gamma^1 &= R_{11} + B_1^T P^1 B_1,\\
\Gamma^2 &= R_{22} + B_2^T Y^T R_{21} Y B_2 + B_2^T M^{1T} (\Upsilon^{-1})^T Z^2 \Upsilon^{-1} M^1 B_2,\\
M^1 &= I - B_1 (\Gamma^1)^{-1} B_1^T P^1,\\
M^2 &= B_2^T Y^T R_{21} Y A + B_2^T M^{1T} (\Upsilon^{-1})^T Z^2 \Upsilon^{-1} M^1 A,\\
S &= (\Gamma^1)^{-1} B_1^T P^1,\\
\Upsilon &= I + B_1 (\Gamma^1)^{-1} B_1^T T,\\
Y &= S + (\Gamma^1)^{-1} B_1^T T \Upsilon^{-1} M^1,\\
K^1 &= -Y(A + B_2 K^2),\\
K^2 &= -(\Gamma^2)^{-1} M^2,\\
T &= A^T M^{1T} P^1 B_2 K^2 + A^T M^{1T} T \Upsilon^{-1} M^1 (A + B_2 K^2).
\end{aligned}$$

According to the uniqueness of the optimal cost function (4.11), then we can obtain

$$J_2^* = x_0^T P^2 x_0 = x_0^T Z^2 x_0. \tag{4.15}$$

Since x_0 is the arbitrary initial state, thus (4.15) implies that $P^2 = Z^2$, the uniqueness of the solution to (4.2) has been obtained.

Sufficiency: Under Assumption 4.1, suppose $P^2 > 0$ is the solution to (4.2), we shall show that $u_k^1 \in \mathcal{U}_1$ minimizes the cost function J_1 and $u_k^2 \in \mathcal{U}_2$ stabilizes the system (2.1) in stabilization and minimizes the cost function J_2.

Denote the Lyapunov function $\bar{H}_k^2$ as

$$\bar{H}_k^2 = x_k^T P^2 x_k. \tag{4.16}$$

By using (4.8) and (4.9), we have

$$u_k^1 = -YAx_k - YB_2 u_k^2, \tag{4.17}$$

adding (4.17) into (2.1), it derives

$$\begin{aligned}
x_{k+1} &= Ax_k - B_1 Y A x_k - B_1 Y B_2 u_k^2 + B_2 u_k^2\\
&= (I - B_1 Y) A x_k + (I - B_1 Y) B_2 u_k^2\\
&= \Upsilon^{-1} M^1 A x_k + \Upsilon^{-1} M^1 B_2 u_k^2.
\end{aligned} \tag{4.18}$$

Then, combining (4.16) and (4.18), there follows

$$\begin{aligned}
\bar{H}_k^2 - \bar{H}_{k+1}^2 &= -[u_k^2 + (\Gamma^2)^{-1} M^2 x_k]^T \Gamma^2 [u_k^2 + (\Gamma^2)^{-1} M^2 x_k] + x_k^T Q_2 x_k + u_k^{1T} R_{21} u_k^1 + u_k^{2T} R_{22} u_k^2\\
&= x_k^T Q_2 x_k + u_k^{1T} R_{21} u_k^1 + u_k^{2T} R_{22} u_k^2 \geq 0,
\end{aligned} \tag{4.19}$$

where $u_k^2 = K^2 x_k$ for $k \geq 0$ has been imposed in the last identity.

Thus, we can conclude that $\bar{H}_k^2$ is monotonically decreasing with respect to k. And Since $P^2 > 0$, then $\bar{H}_k^2 \geq 0$ is bounded below. According to monotone bounded theorem, $\bar{H}_k^2$ is convergent. We will show $\lim_{k\to\infty} x_k = 0$ in the following paragraph.

Actually, let m be any nonnegative integer, by adding from $k = m$ to $k = m + N$ on both sides of (4.19) and taking limitation of m, it holds that

$$0 = \lim_{m\to\infty} [\bar{H}_m^2 - \bar{H}_{m+N+1}^2] = \lim_{m\to\infty} \sum_{k=m}^{m+N} (x_k^T Q_2 x_k + u_k^{1T} R_{21} u_k^1 + u_k^{2T} R_{22} u_k^2) \geq 0. \tag{4.20}$$

Since the coefficient matrices in (2.1) and (2.4) are time invariant, then via a time-shift of length m, it yields that

$$\sum_{k=m}^{m+N} (x_k^T Q_2 x_k + u_k^{1T} R_{21} u_k^1 + u_k^{2T} R_{22} u_k^2) \geq x_m^T P_m^2 (m+N) x_m = x_m^T P_0^2(N) x_m \geq 0. \tag{4.21}$$

Taking limitation on both sides of (4.21) and using (4.20), we know that

$$\lim_{m\to\infty} x_m^T P_0^2(N) x_m = 0. \tag{4.22}$$

Moreover, there exists integer N_0 such that for any $N > N_0$, $P_0^2(N_0) > 0$, thus (4.22) implies that

$$\lim_{m\to\infty} x_m^T x_m = 0. \tag{4.23}$$

According to (4.13), we have

$$\lim_{m\to\infty} x_{m-1}^T (A + B_1 K^1 + B_2 K^2)^T (A + B_1 K^1 + B_2 K^2) x_{m-1} = 0, \tag{4.24}$$

which implies that $\lim_{m\to\infty} x_m = 0$ due to the boundedness of $A + B_1 K^1 + B_2 K^2$.

Therefore, the controllers (4.8) stabilizes system (2.1) in stabilization.

To complete the proof the Theorem 4.2, we shall show that the optimal controllers (4.9) in Theorem 4.2 minimizes the cost function (2.2).

Denote the Lyapunov function

$$\bar{H}_k^1 = x_k^T P^1 x_k + x_k^T \zeta_{k-1}. \tag{4.25}$$

Then combining (2.1) and (4.25), we have

$$\begin{aligned}
&\bar{H}_k^1 - \bar{H}_{k+1}^1 \\
&= x_k^T P^1 x_k + x_k^T \zeta_{k-1} - x_{k+1}^T P^1 x_{k+1} + x_{k+1}^T \zeta_k - \zeta_k^T x_{k+1} + \zeta_k^T x_{k+1} \\
&= x_k^T Q_1 x_k + u_k^{1T} R_{11} u_k^1 + u_k^{2T} R_{12} u_k^2 - [u_k^1 + (\Gamma^1)^{-1}(B_1^T P^1 A x_k + B_1^T P^1 B_2 u_k^2 + B_1^T \zeta_k)]\Gamma^1 \\
&\quad \times [u_k^1 + (\Gamma^1)^{-1}(B_1^T P^1 A x_k + B_1^T P^1 B_2 u_k^2 + B_1^T \zeta_k)] - 2u_k^{2T} B_2^T P^1 M^1 A x_k - x_k^T A^T M^{1T} \zeta_k - \zeta_k^T M^1 A x_k \\
&\quad - u_k^{2T} B_2^T M^{1T} \zeta_k - \zeta_k^T M^1 B_2 u_k^2 - u_k^{2T}(R_{12} + B_2^T P^1 M^1 B_2) u_k^2 + x_k^T \zeta_{k-1} + \zeta_k^T x_{k+1} + \zeta_k^T B_1 (\Gamma^1)^{-1} B_1^T \zeta_k,
\end{aligned} \tag{4.26}$$

where

$$\begin{aligned}
-2u_k^{2T}B_2^TP^1M^1Ax_k &= -u_k^{2T}B_2^TP^1M^1Ax_k - x_k^TA^TM^{1T}P^1B_2u_k^2 \\
&= -(\zeta_{k-1} - A^TM^{1T}\zeta_k)^Tx_k - x_k^T(\zeta_{k-1} - A^TM^{1T}\zeta_k) \\
&= -\zeta_{k-1}^Tx_k - x_k^T\zeta_{k-1} + x_k^TA^TM^{1T}\zeta_k + \zeta_k^TM^1Ax_k.
\end{aligned} \tag{4.27}$$

Adding (4.27) into (4.26), then it yields that

$$\begin{aligned}
&x_k^TQ_1x_k + u_k^{1T}R_{11}u_k^1 + u_k^{2T}R_{12}u_k^2 \\
= &H_k^1 - H_{k+1}^1 + [u_k^1 + (\Gamma^1)^{-1}(B_1^TP^1Ax_k + B_1^TP^1B_2u_k^2 + B_1^T\zeta_k)]\Gamma^1[u_k^1 + (\Gamma^1)^{-1}(B_1^TP^1Ax_k \\
&+B_1^TP^1B_2u_k^2 + B_1^T\zeta_k)] + u_k^{2T}(R_{12} + B_2^TP^1M^1B_2)u_k^2 + u_k^{2T}B_2^TM^{1T}\zeta_k + \zeta_k^TM^1B_2u_k^2 \\
&-\zeta_k^TB_1(\Gamma^1)^{-1}B_1^T\zeta_k + \zeta_{k-1}^Tx_k - \zeta_k^Tx_{k+1}.
\end{aligned} \tag{4.28}$$

Thus, it holds

$$\begin{aligned}
&\sum_{k=0}^{N}(x_k^TQ_1x_k + u_k^{1T}R_{11}u_k^1 + u_k^{2T}R_{12}u_k^2) \\
= &x_0^TP^1x_0 + x_0^T\zeta_{-1} - H_{N+1}^1 + \sum_{k=0}^{N}[u_k^1 + (\Gamma^1)^{-1}(B_1^TP^1Ax_k + B_1^TP^1B_2u_k^2 + B_1^T\zeta_k)]\Gamma^1 \\
&\times[u_k^1 + (\Gamma^1)^{-1}(B_1^TP^1Ax_k + B_1^TP^1B_2u_k^2 + B_1^T\zeta_k)] + \zeta_{-1}^Tx_0 + \sum_{k=0}^{N}[u_k^{2T}(R_{12} + B_2^TP^1M^1B_2)u_k^2 \\
&+u_k^{2T}B_2^TM^{1T}\zeta_k + \zeta_k^TM^1B_2u_k^2 - \zeta_k^TB_1(\Gamma^1)^{-1}B_1^T\zeta_k].
\end{aligned} \tag{4.29}$$

By using (4.23), it holds that

$$\lim_{N\to\infty}\bar{H}_{N+1}^1 = 0. \tag{4.30}$$

Taking limitation of $N \to \infty$ on both sides of (4.29) we know the optimal controller minimizes (2.2) is exactly (4.9).

Thus, the optimal cost function of (2.2) is

$$J_1^* = x_0^T(P^1 + T + T^T + \Xi)x_0^T, \tag{4.31}$$

which is exactly (4.10).

And applying similar procedure to the Lyapunov function of $\bar{H}_k^2$ in (4.16), the controller (4.8) minimizes the cost function (2.4) can be immediately obtained by (4.19), and the optimal cost function is exactly (4.11). This completes the proof. □

Remark 4.4. In Theorem 4.2, a necessary and sufficient stabilization condition has been obtained for the game-based system. The proposed method can be extended to solve the game-based systems with multiplicative noises by applying optimization. We also note that the direct application of the method to the system with input delay in the leader's control, *i.e.*, $x_{k+1} = Ax_k + B_1u_k^1 + B_2u_{k-d}^2$, may lose efficiency since $u_{k-d}^2 = K^2x_k$ makes no sense due to the causality. In this case, a predictor-like controller $u_{k-d}^2 = K^2\hat{x}_{k|k-d}$ as in [23], where $\hat{x}_{k|k-d}$ is the predictor of x_k based on the information from 0 to $k-d$, can be used to solve the problem combining with the proposed Stackelberg method in this paper. This topic deserves further in-depth research.

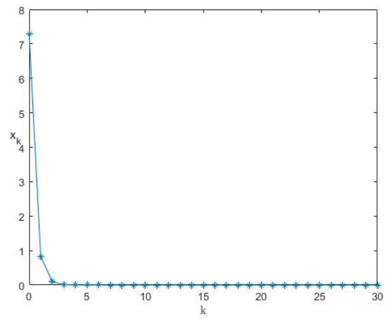

FIGURE 1. Trajectory of x_k with controllers $u_k^1 = -0.4268x_k$ and $u_k^2 = -0.0330x_k$.

5. Numerical examples

Consider the system (2.1) and the cost functions (2.2) and (2.4) with $A = 1$, $B_1 = 2$, $B_2 = 1$, $Q_1 = 1$, $Q_2 = 1$, $R_{11} = 0.6$, $R_{12} = 1$, $R_{21} = 1$, $R_{22} = 6.2$, $x_0 = 7.3$. By solving (4.1)–(4.2) and (4.3)–(4.7), we have $\Upsilon = 0.9965 > 0$, $P^1 = 1.1325 > 0$, $P^2 = 1.2044 > 0$, $K^1 = -0.4268$ and $K^2 = -0.0330$. According to Theorem 4.2, there exist $u_k^2 = -0.0330x_k$ stabilizing the system (2.1) in stabilization and $u_k^1 = -0.4268x_k$ minimizing the cost function J_1. As shown in Figure 1, the regulated state is stable.

6. Conclusion

In this paper, we have investigated the stabilization problem of the game-based system by Stackelberg method, where the follower is designed to minimise its own cost function and the leader is designed to stabilize the system. The analytical solution of the optimal open-loop Stackelberg strategy in finite horizon is derived. The contribution is that we present the necessary and sufficient condition for the stabilization of the game-based system. The result is of significance to the control problems of the virus and the immune system in [6], socio-economic modeling in [9] and so on.

Acknowledgements. This work was supported by the National Natural Science Foundation of China under Grants 61821004, 62250056, and the Natural Science Foundation of Shandong Province (ZR2021ZD14, ZR2021JQ24), and Science and Technology Project of Qingdao West Coast New Area (2019-32, 2020-20, 2020-1-4), High-level Talent Team Project of Qingdao West Coast New Area (RCTD-JC-2019-05), Key Research and Development Program of Shandong Province (2020CXGC01208).

References

[1] D. Anderson and B. Moore, Linear Optimal Control. Prentice-Hall, Englewood Cliffs, NJ (1971).

[2] A. Bouhtouri, D. Hinrichsen and A. Pritchard, H_∞ type control for discrete-time stochastic systems. *Int. J. Robust Nonlinear Control* **9** (1999) 923–948.

[3] F. Chen, Stability and stabilization for discrete-time Markovian jump quadratic systems with incomplete knowledge of transition probabilities, in *Proceedings of the 39th Chinese Control Conference* (2020) 825–830.

[4] B. Chen and A. Zadrozny, An anticipative feedback solution for the infinite-horizon, linear-quadratic, dynamic, Stackelberg game. *J. Econ. Dyn. Control* **26** (2002) 1397–1416.

[5] E. Dockner, S. Jorgensen, N. Long and G. Sorger, Differential Games in Economics and Management Science. Cambridge University Press, Cambridge, UK (2000).

[6] R. Eftimie, L. Bramson and D. Earn, Interactions between the immune system and cancer: a brief review of non-spatial mathematical models. *Bull. Math. Biol.* **73** (2011) 2–32.

[7] G. Freiling, G. Jank and R. Lee, Existence and uniqueness of open-loop Stackelberg equilibria in linear-quadratic differential games. *J. Optim. Theory Appl.* **110** (2001) 515–544.

[8] Y. Huang, W. Zhang and H. Zhang, Infinite horizon linear quadratic optimal control for discrete-time stochastic systems. *Asian J. Control* **10** (2008) 608–615.

[9] M. Johnson, T. Hiramatsu, N. Fitz-Coy and W. Dixon, Asymptotic Stackelberg optimal control design for an uncertain Euler Lagrange system. IEEE Conf. Decis. Control, Atlanta, GA, USA (2020) 6686–6891.

[10] M. Jungers, On linear-quadratic Stackelberg games with time preference rates. *IEEE Trans. Automatic Control* **53** (2008) 621–625.
[11] M. Li, J. Cruz and A. Simaan, An Approach to discrete-time incentive feedback Stackelberg games. *IEEE Trans. Syst. Cybernet.* **32** (2002) 472–481.
[12] H. Mukaidania and H. Xu, Infinite horizon linear-quadratic Stackelberg games for discrete-time stochastic systems. *Automatica* **76** (2017) 301–308.
[13] M. Rami and X. Zhou, Linear matrix inequalities, Riccati equations, and indefinite stochastic linear quadratic controls. *IEEE Trans. Automatic Control* **45** (2001) 1131–1143.
[14] K. Trejo, J. Clempner and A. Poznyak, Adapting strategies to dynamic environments in controllable Stackelberg security games, in *Proceedings of IEEE 55th Conference on Decision and Control* (2016) 5484–5489.
[15] F. Wei, W. Wang and J. Wang, Dynamics and stability analysis of a Stackelberg mixed duopoly game with price competition in insurance market. *Discrete Dyn. Nat. Soc.* **2021** (2021) 1–18.
[16] J. Xu, H. Zhang and T. Chai, Necessary and sufficient condition for two-player Stackelberg strategy. *IEEE Trans. Automatic Control* **60** (2015) 1356–1361.
[17] K. Xu and X. Zhao, Adaptive dynamic programming for a class of two-player Stackelberg differential games. Int. Conf. Syst. Sci. Eng., Kagawa, Japan (2020) 1–6.
[18] W. Zhang and B. Chen, On stabilizability and exact observability of stochastic systems with their applications. *Automatica* **24** (2004) 87–94.
[19] H. Zhang and J. Xu, Control for Itô stochastic systems with input delay. *IEEE Trans. Automatic Control* **62** (2017) 350–365.
[20] R. Zhang and L. Guo, Controllability of stochastic game-based control systems. *SIAM J. Control Optim.* **57** (2019) 3799–3826.
[21] H. Zhang, L. Huang and H. Xie, Infinite horizon stochastic H_2/H_∞ control for discrete-time systems with state and disturbance dependent noise. *Automatica* **44** (2008) 2306–2316.
[22] H. Zhang, H. Wang and L. Li, Adapted and casual maximum principle and analytical solution to optimal control for stochastic multiplicativenoise systems with multiple input-delays, in *Proceedings of the 51st IEEE Conference Decision Control* (2012) 2122–C2127.
[23] H. Zhang, L. Li, J. Xu and M. Fu, Linear quadratic regulation and stabilization of discrete-time systems with delay and multiplicative noise. *IEEE Trans. Automatic Control* **60** (2015) 2599–2613.
[24] L. Zhang, C. Yang and H. Wang, The influence of entanglement on complex dynamics of a quantum Stackelberg duopoly with heterogeneous expectations. *Quant. Inform. Process.* **21** (2022) 1–14.
[25] M. Zhu and S. Martinez, Stackelberg-game analysis of correlated attacks in cyber-physical systems. Am. Control Conf., San Francisco, CA, USA (2011) 4063–4068.