Network Learning from Best-Response Dynamics in LQ Games

Yijun Chen[1], Kemi Ding[2], and Guodong Shi[1]

Abstract— **In this paper, we focus on network structure inference problem for linear-quadratic (LQ) games from best-response dynamics. An adversary is considered to have no knowledge of the game network structure but have the ability to observe all players' best-response actions and manipulate some players' actions. This work presents a comprehensive framework for network learning from best-response dynamics in LQ games. First of all, we establish theoretic results that characterize network structure identifiability and provide numerical examples to demonstrate the usefulness of our theoretic results. Next, in the face of the inherent stability and sparsity constraints for the game network structure, we propose an information-theoretic stable and sparse system identification algorithm for learning the network structure. Finally, the effectiveness of the proposed learning algorithm is tested in both synthesis networks and real networks. The connection between network structure inference problem and classical system identification theory is covered by our work, which advances the literature.**

I. Introduction

Game theory has been an essential tool in describing how self-interested individuals, referred to as players, develop rationality and competition [1]. In a game with multiple players, each player chooses her action to maximize her payoffs determined by her payoff function. The payoff function of each player relies on not only her own action, but also other players' actions, resulting in underlying competition. Many games are played over networks, in the sense that players are interconnected through a network [2]. In many aspects of our life, network games are extensively employed to model strategic interactions among interconnected players, such as online E-commerce in social networks [3], route planning in transportation networks [4], and games arising in wireless communication networks [5]. Significant progress has been made in computing equilibrium solution concepts such as Nash equilibrium (NE) and correlated equilibrium, and identifying the most influential players in network games [6]–[8]. All of these works are in the field of algorithmic game theory and are concerned with analyzing the properties of network games on known graphs.

Another vibrant stream of game literature focuses on network structure inference where players' behavior actions are collected to reconstruct the network structure. Network structure inference problem is also related to graphical learning problems in the fields of machine learning and signal processing. Many approaches have been proposed, differing in the applied models such as probabilistic graphical models [9], [10], physically-motivated models [11] and signal processing models [12], [13]. As for learning network structure in the game settings, there also has been a considerable amount of work. In the continuous-action setting, different methods were proposed according to different action and payoff function settings (e.g., non-parametric payoffs in a recent work [14]). As for the discrete-action setting, the work of [15] proposed a maximum-likelihood approach to learning linear influence games with binary actions. In [16], the authors proposed a l_1-regularized logistic regression model to learn linear influence given observations over the pure-strategy NEs. The work of [17] studied a learning problem of LQ games on networks, over which a number of independent games were played with all NE actions observed for learning the graph.

Problem of Interest In this paper, we consider a LQ game over a network where each player's payoffs is a LQ function of her own action and other players' actions. Each player participating in this game is assumed to play her best responses in a sequential decision process. Each player will decide her next action as the decision that maximizes her payoff given other players' current actions [28]. We assume that there is an adversary who does not know the game network structure but is capable of observing all players' best-response actions and manipulating some players' actions. The aim of the adversary is to infer the game network structure from its observations and manipulations.

The reasons that the problem is worth studying are as follows. First, this problem will be naturally a dynamic extension of network structure inference from static NEs in LQ games in [17]. Second, when network games are considered to study the behaviors of players, NE is not the only behavior model to explain players' strategic decision-making. Instead, best response is also a possible model to describe the players' strategic behavior in a sequential decision-making manner. In practice, players may not able to take the NE action at the beginning of networks games because players cannot get or anticipate complete information about other players' payoff functions [18].

There are two practical issues that need to be addressed. It is expected that players' best-response actions should converge to a NE, which means the network structure should be stable [19]; the empirical analysis has shown that real-world social networks are typically sparse [20]. As a result, network structure inference from best-response dynamics becomes a stable and sparse system identification problem. System

[1]Yijun Chen and Guodong Shi are with the Australian Centre for Field Robotics, the University of Sydney, Sydney 2004, Australia (email: yijun.chen@sydney.edu.au; guodong.shi@sydney.edu.au).

[2]Kemi Ding is with the Department of Mechanical and Energy Engineering, Southern University of Science and Technology, Shenzhen 518055, China (email: dingkm@sustech.edu.cn).

Manuscript 62 submitted to 2023 American Control Conference (ACC).
Received September 13, 2022.

identification of stable systems is a well-studied problem in the literature. In [21], a frequency domain solution to linear system identification of a stable system was provided in presence of undermodeling. In [22], sharp finite-time error analysis was derived based on ordinary least-squares, which was suitable for both stable and unstable systems. In [23], a method for projecting an arbitrary square matrix to the non-convex set of asymptotically stable matrices was proposed, whose projection was optimal in an information-theoretic sense. Rich results of sparse system identification also have been well established, most of which are based on general penalties [24]. A specific line of penalty type is the method of l_1-regularized penalization [25], [26].

Contributions In this paper, we have made the following contributions:

(i) We manage to establish theoretic results characterizing network structure identifiability from the view of the adversary: network structure identifiability is shown to be equivalent to certain controllability conditions;

(ii) Taking into account the intrinsic stability and sparsity properties of the game network structure, we propose a stable and sparse system identification framework for learning the network structure. In the algorithm, we first solve an l_1-regularized least square problem arising from players' best-response action data, and then project the solution to the set of stable matrices by applying the method in [23];

(iii) We conduct extensive experiments on synthetic and real-world networks to validate the effectiveness of the proposed learning algorithm.

Our results provide a complete answer to network structure inference for LQ games from best-response dynamics. By introducing stable and sparse system identification into network structure inference for LQ games, we cover the connection between network structure inference for LQ games and classical system identification theory, advancing the literature.

Organization The remainder of the paper is organized as follows. In Section II, we introduce LQ games, best-response dynamics, and the adversary model, followed by a problem description. In Section III, we investigate the identifiability problem under which conditions the network structure can be uniquely identified by the adversary from the perspective of control theory. In Section IV, we propose an information-theoretic algorithm to learn the network structure, assuming that the network structure is stable and sparse. In Section V, experiments for the proposed algorithm are conducted with a comprehensive discussion about the performance of effectiveness. This paper ends with concluding remarks in Section VI.

II. Problem Formulation

A. Linear Quadratic Game

A network game is associated with a graph. The nodes of the graph represent the players of the network game, and the links define the interdependency among the players in payoffs. In an n-player network game with LQ payoffs, each player $i \in \mathrm{V} := \{1, 2, \ldots, n\}$ chooses her action $\mathbf{x}_i \in \mathbb{R}$ that can maximize her payoffs J_i described by

$$J_i = \alpha_i \mathbf{x}_i - \frac{1}{2}\mathbf{x}_i^2 + \sum_{j=1}^{n} g_{ij}\mathbf{x}_i\mathbf{x}_j. \tag{1}$$

In (1), the first two terms characterize the benefit of player i by taking her own action $\mathbf{x}_i$, where the parameter $\alpha_i > 0$ is called the marginal benefit, capturing the level of selfishness of player i. The last term of this payoff function reflects the peer effect suffered by player i from the actions of other players: if $g_{ij} > 0$, players i and j are strategic complements (friends); if $g_{ij} < 0$, players i and j are strategic substitutes (opponents); if $g_{ij} = 0$, there is no influence on player i from player j.

Let $\mathbf{G} \in \mathbb{R}^{n\times n}$ be the matrix formed by the g_{ij}, i.e., the ij-entry of $\mathbf{G}$ is g_{ij}. The interaction graph $\mathrm{G} = (\mathrm{V}, \mathrm{E})$ underlying the game (1) is then defined as the induced graph of $\mathbf{G}$, where a directed link $(j, i) \in \mathrm{E}$ if and only if $g_{ij} \neq 0$. Note that $g_{ii} = 0$ as we suppose there exists no self-loop in the graph G.

B. Nash Equilibrium and Best-response Dynamics

Nash equilibrium A common solution concept in game theory is called NE, and it is a profile of actions under which no one will benefit by changing her action when others keep theirs unchanged. We denote the vector $\mathbf{x}^\star := [\mathbf{x}_1^\star, \ldots, \mathbf{x}_n^\star]^\top$ as the NE of the LQ game introduced previously. We adopt a critical assumption in the following.

Assumption 1: The spectral radius of the matrix $\mathbf{G}$, denoted by $\rho(\mathbf{G})$ is less than 1.

It guarantees the existence and uniqueness of pure-strategy NE $\mathbf{x}^\star$ in LQ games. Moreover, as derived in [8], the form of NE is $\mathbf{x}^\star = (\mathbf{I} - \mathbf{G})^{-1}\boldsymbol{\alpha}$, where $\boldsymbol{\alpha} := [\alpha_1, \ldots, \alpha_n]^\top$.

Best response For players participating in a game, generally, NE can not be played immediately when the game starts since the payoff functions are held in private. Rather, in practice, players' behaviors are better explained in a sequential decision process [27]. Let time be indexed in $\mathcal{T} := \{0, 1, 2, \ldots, T\}$, and let $\mathbf{x}_i(k)$ be player i's decision at any time $k \in \mathcal{T}$. Then, it is reasonable to assume that any rational player at any given time k will decide her action at the next time $k + 1$ as the decision that maximizes her payoff given other players' current decisions at time k. This is known as *best responses* [28]. As a result, in repeated plays for LQ games, the dynamics of $\mathbf{x}_i(k)$ obey

$$\begin{aligned}\mathbf{x}_i(k+1) &= \arg\max_{\mathbf{x}_i} J_i(\mathbf{x}_i, \mathbf{x}_{-i}(k)) \\ &= \alpha_i + \sum_{j=1}^{n} g_{ij}\mathbf{x}_j(k), \quad i \in \mathrm{V},\ t \in \mathcal{T}. \end{aligned} \tag{BR}$$

In practice, it is expected that the best-response dynamics (BR) should converge to a NE [19], which is guaranteed by Assumption 1 according to the work [8].

C. The Adversary Model

In this paper, we investigate the possibility of inferring the sensitive network structure from the perspective of an adversary. Specifically, an *adversary* refers to an individual who is intended to infer the relationship through injecting adversarial perturbations and monitoring the players' actions.

We assume that the adversary is able to observe all players' best-response actions and manipulate some players' actions. We term such players as *action-compromised* players, indexed in the set $\mathrm{M} \subseteq \mathrm{V}$. Then $\mathrm{V} \setminus \mathrm{M}$ contains *benign* players who just follows (BR). For $i \in \mathrm{M}$, we model the manipulation of the adversary as a perturbation $\mathbf{u}_i(k)$ over the best response actions of player i. The best response of the players in the presence of the action-compromised players becomes

$$\begin{aligned}\mathbf{x}_i(k+1) &= \arg\max_{\mathbf{x}_i} J_i(\mathbf{x}_i, \mathbf{x}_{-i}(k)) + \mathbf{I}_{\mathrm{M}}(i)\mathbf{u}_i(k) \\ &= \alpha_i + \sum_{j=1}^{n} g_{ij}\mathbf{x}_j(k) + \mathbf{I}_{\mathrm{M}}(i)\mathbf{u}_i(k), \quad i = 1, \ldots, n.\end{aligned} \tag{p-BR}$$

where $\mathbf{I}_{\mathrm{M}}(i)$ is the indicator function: $\mathbf{I}_{\mathrm{M}}(i) = 1$ if $i \in \mathrm{M}$; $\mathbf{I}_{\mathrm{M}}(i) = 0$ otherwise.

Real-world Interpretations Fake reviews and misinformation has been shown to be ubiquitous in the study of social bots in online social networks. Many works have characterized the role and the ability of social bots in the manipulation of social media [29], [30]. Therefore, the proposed adversary model captures this critical characteristic of real-world online social networks.

D. Problem Definition

In this paper, we are interested in whether and how the underlying graph G (or equivalently, its adjacency matrix $\mathbf{G}$) is exposed to risks of being uniquely identified by the adversary. To be precise, the adversary holds information

$$\mathcal{I} : \left\{\mathbf{u}_i(k), i \in \mathrm{M}, k \in \mathcal{T}\right\} \bigcup \tag{2a}$$
$$\left\{\mathbf{x}_j(k), j \in \mathrm{V}, k \in \mathcal{T}\right\} \tag{2b}$$

with the $\mathbf{u}_i(k)$ and $\mathbf{x}_i(k)$ being produced by the perturbed best response (p-BR). The problems of interest from the perspective of the adversary lie in

- *Identifiability*: Whether it is possible to uniquely identify $\mathbf{G}$ from $\mathcal{I}$, perhaps with the help of strategically designed $\mathbf{u}_i(k), i \in \mathrm{M}$, $k \in \mathcal{T}$.
- *Learning*: How to build effective learning frameworks for estimating $\mathbf{G}$ from $\mathcal{I}$, perhaps with prior structural information such as stability and sparsity.

For the above-defined problems, there are three critical points worth mentioning. First, best-response dynamics is a good start to investigate the game network structure inference problem given observations from the players' best-response dynamics. Even though the form of best response dynamics is simple, the learning problem is still challenging. Second, the "identifiability" problem characterizes the fundamental limitation of the adversary's learning capability on a theoretic level; on the other hand, the "learning" problem focuses on designing computational algorithms that would instruct the adversary to infer the network structure on a practical level. These two parts are complementary and consist a thorough study of learning the network structure from best-response dynamics. Third, how inherent properties of stability and sparsity for the network structure can be integrated into this learning problem has been largely open under the dynamic setting of games.

III. Network Structure Identifiability

In this section, we look into identifiability conditions under which the network structure can be uniquely identified by the adversary. Different identifiability conditions are provided under two scenarios: the adversary injects no perturbation on players' actions, or the adversary injects perturbations on the actions of a subset of players.

For notational simplicity, we denote the cardinalities of the set M as m. The index of action-compromised players is sorted in ascending order with $\mathrm{M} := \{p_1, \ldots, p_m\}$. For example, if the network game has three action-compromised players (i.e., player 1, 3, 6), then $m = 3$ and $\mathrm{M} = \{p_1 = 1, p_2 = 3, p_3 = 6\}$.

We also denote the aggregated action profile of all players and the overall injected perturbations at time k by $\mathbf{x}(k) := [\mathbf{x}_1(k), \ldots, \mathbf{x}_n(k)]^\top \in \mathbb{R}^n$ and $\mathbf{u}(k) := [\mathbf{u}_i(k), i \in \mathrm{M}]^\top \in \mathbb{R}^m$, respectively.

We introduce $\mathbf{B} \in \mathbb{R}^{n \times m}$ to represent the set of M. Specifically, the (i, j)-th entry of matrix $\mathbf{B}$ satisfies that

$$\mathbf{B}_{i,j} = \begin{cases} 1, & i = p_j; \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

From classical control theory [31], for two matrices $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times r}$, the matrix pair (A, B) is said to be *controllable* if the $n \times nr$ controllability matrix (where the subscript n denotes the number of block columns)

$$\mathcal{C}_n(A, B) := \begin{bmatrix} B & AB & \ldots & A^{n-1}B \end{bmatrix}$$

has rank n.

Controllability is regarded as an critical property of a control system. Basically, in a control system, controllability is the ability of the system to shift from the initial state to a specified state by applying input in a finite amount of time. In the sequel, we establish a control system based on the best-response dynamics, and link the determination of identifiability for the adversary with the property of control system (i.e., controllability).

A. Main Results

We now present the following result that if the adversary has access to all players' actions but has no ability to add action perturbations (i.e., $\mathrm{M} = \emptyset$), network structure identifiability can be precisely characterized by a special form of controllability.

Theorem 1: Assume $\mathrm{M} = \emptyset$. Let the adversary has access to $\mathcal{I}$ from (p-BR). Then $\mathbf{G}$ can be uniquely identified by the adversary for sufficiently large T if and only if $(\mathbf{G}, \mathbf{x}(1) - \mathbf{x}(0))$ is controllable.

Proof. Regarding the action dynamics described in (p-BR), we define the action difference and the perturbation difference as $\mathbf{y}_i(k) := \mathbf{x}_i(k+1) - \mathbf{x}_i(k)$ and $\mathbf{v}_i(k) := \mathbf{u}_i(k+1) - \mathbf{u}_i(k)$ respectively, in order to eliminate the effect of marginal benefit α_i.

Adding the observation ports to the action dynamics (p-BR), we arrive at the compact form of an input-output linear invariant system:

$$\mathbf{y}(k+1) = \mathbf{G}\mathbf{y}(k) + \mathbf{B}\mathbf{v}(k), \tag{4a}$$

$$\mathbf{z}(k) = \mathbf{y}(k), \tag{4b}$$

in which the state $\mathbf{y}(k) := [\mathbf{y}_1(k), \dots, \mathbf{y}_n(k)]^\top$, the input signal $\mathbf{v}(k) : [\mathbf{u}_i(k+1) - \mathbf{u}_i(k), i \in \mathrm{M}] \in \mathbb{R}^m$, and $\mathbf{B} \in \mathbb{R}^{n\times m}$. The nonzero entries of matrices $\mathbf{B}$ correspond to the set of *action-compromised* players. Hence, the adversary has a trajectory of the input/output signal of this system (4a): $\mathcal{I} := \{\mathbf{v}(k), \mathbf{z}(k), 0 \le k \le T, k \in \mathbb{N}\}$.

When $\mathrm{M} = \emptyset$ and $\mathcal{I}$ is accessible, (4a) is degenerated as

$$\mathbf{z}(k+1) = \mathbf{y}(k+1) = \mathbf{G}\mathbf{y}(k).$$

Then, the single trajectory satisfies

$$Z = \mathbf{G}Y,$$

where $Z = [\mathbf{y}(1), \mathbf{y}(2), \dots, \mathbf{y}(T)] \in \mathbb{R}^{n\times T}$ and $Y = [\mathbf{y}(0), \mathbf{y}(1), \dots, \mathbf{y}(T-1)] \in \mathbb{R}^{n\times T}$.

This equation can be decoupled as several linear equations: $Y^\top[\mathbf{G}^\top]_i = [Z^\top]_i, \forall i \in \mathrm{V}$ with $[\cdot]_i$ representing the i-th column of a matrix. Then, $\mathbf{G}$ is uniquely constructable from the trajectory $\mathcal{T}$ if and only if each of these linear equations has a unique solution, i.e., $\mathrm{rank}(Y) = \mathrm{rank}([Y^\top, [Z^\top]_i]) = n, \forall i \in \mathrm{V}$.

Note that $Y = [\mathbf{y}(0), \mathbf{y}(1), \dots, \mathbf{y}(T-1)] = [\mathbf{y}(0), \mathbf{G}\mathbf{y}(0), \dots, \mathbf{G}^{T-1}\mathbf{y}(0)]$, of which the vectors span a $\mathbf{G}$-cyclic subspace of $\mathbb{R}^n$, denoted by $H(\mathbf{y}(0); \mathbf{G})$. This subspace is an invariant subspace for $\mathbf{G}$, in the sense that $\mathbf{G}H(\mathbf{y}(0); \mathbf{G}) \subseteq H(\mathbf{y}(0); \mathbf{G})$, which validates $\mathrm{rank}(Y) = \mathrm{rank}([Y^\top, [Z^\top]_i])$. Moreover, the condition $\mathrm{rank}(Y) = n$ holds if and only if $(\mathbf{G}, \mathbf{y}(0))$ is controllable. $\square$

The next result characterizes the ability to identify $\mathbf{G}$ by linear feedback perturbations, namely $\mathbf{u}(k) = \mathbf{K}\mathbf{x}(k)$ for $\mathbf{K} \in \mathbb{R}^{m\times n}$.

For a matrix A, we denote the image of matrix A by $\mathrm{Im}(A)$. It is the span of the vectors of the matrix A or linear transformation A.

We now introduce a lemma coming from [31, Lemma 2.2].

Lemma 1: Let $\mathbf{0} \neq \mathbf{y}(0) \in Im(\mathbf{B})$. If $(\mathbf{G}, \mathbf{B})$ is controllable, there exists $\mathbf{K}$ such that $(\mathbf{G} + \mathbf{B}\mathbf{K}, \mathbf{y}(0))$ is controllable.

Theorem 2: Assume $\mathrm{M} \neq \emptyset$. Let the adversary has access to $\mathcal{I}$ from (p-BR). Suppose $\mathbf{u}(k)$ is generated by $\mathbf{u}(k) = \mathbf{K}\mathbf{x}(k)$. Then there exists $\mathbf{K}$ such that $\mathbf{G}$ can be uniquely identified if the following conditions hold:

(i) $\mathbf{0} \neq \mathbf{x}(1) - \mathbf{x}(0) \in \mathrm{Im}(\mathbf{B})$;

(ii) $(\mathbf{G}, \mathbf{B})$ is controllable.

Proof. When $\mathrm{M} \neq \emptyset$, (4a) is degenerated as

$$\mathbf{z}(k+1) = \mathbf{y}(k+1) = \mathbf{G}\mathbf{y}(k) + \mathbf{B}\mathbf{v}(k).$$

Setting $\mathbf{v}(k) = \mathbf{K}\mathbf{y}(k)$, there holds

$$\mathbf{y}(k+1) = (\mathbf{G} + \mathbf{B}\mathbf{K})\mathbf{y}(k).$$

According to Theorem 1, $\mathbf{G}$ can be uniquely identified by the adversary from the information $\mathcal{I}$ if and only if $(\mathbf{G} + \mathbf{B}\mathbf{K}, \mathbf{y}(0))$ is controllable.

Since conditions (i) and (ii) are satisfied, Lemma 1 holds that there exists $\mathbf{K}$ such that $(\mathbf{G}+\mathbf{B}\mathbf{K}, \mathbf{y}(0))$ is controllable. Therefore, there exists $\mathbf{K}$ such that $\mathbf{G}$ can be uniquely identified. $\square$

B. Numerical Examples

In what follows, we provide numerical examples to illustrate the usefulness of the conditions in Theorems 1 and 2.

Consider a 4-player LQ game. The LQ game is associated with an interaction graph G shown in Fig. 1. The corresponding adjacent matrix is

$$\mathbf{G} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ g_{21} & 0 & 0 & 0 \\ g_{31} & 0 & 0 & g_{34} \\ g_{41} & 0 & 0 & 0 \end{bmatrix}.$$

The same setting is used in the following three numerical examples.

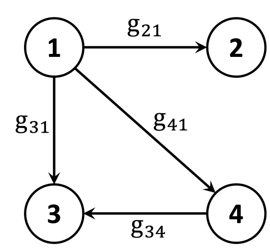


Fig. 1: An example of 4-player network structure.

Example 1: Suppose no perturbation is injected into the players' actions, i.e., $\mathrm{M} = \emptyset$. We can compute that $\mathrm{rank}(\mathcal{C}_n(\mathbf{G}, \mathbf{x}(1) - \mathbf{x}(0)))$ is always less or equal to 3 no matter what value $\mathbf{x}(1) - \mathbf{x}(0)$ takes. According to Theorem 1, the adversary cannot identify the network structure from its information $\mathcal{I}$. $\square$

In the following two examples, we consider two ways of adversary perturbations shown in Fig. 2

Example 2: Suppose the adversary is designed to influence the actions of players 1 and 2, i.e., $\mathrm{M} = \{1, 2\}$ (see blue perturbations in Fig. 2). We then obtain $\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}^\top$. It is obvious that the controllability condition in Theorem 2 holds since $\mathrm{rank}(\mathcal{C}_n(\mathbf{G}, \mathbf{B})) = 4$. Moreover, to guarantee the satisfactory of the first condition in Theorem 2, the initial state can be designed as $\mathbf{x}(0) =$

Manuscript 62 submitted to 2023 American Control Conference (ACC).
Received September 13, 2022.

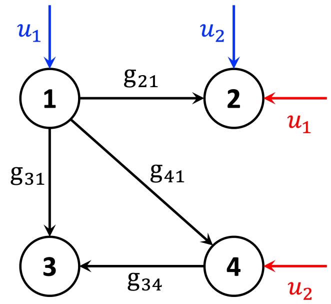


Fig. 2: Two ways of adversary perturbations highlighted in blue and red.

$(\mathbf{I}-\mathbf{G})^{-1}(\boldsymbol{\alpha}-\mathbf{e}_1)$, where $\mathbf{e}_1$ is a unit vector only with the first entry to be one, and the inverse exists according to Assumption 1. According Theorem 2, $\mathbf{G}$ is learnable. □

Example 3: Suppose the adversary is designed to influence the actions of players 2 and 4, i.e., $\mathrm{M} = \{2,4\}$ (see red perturbations in Fig. 2). We then obtain $\mathbf{B} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^\top$. The controllability condition in Theorem 2 does not hold because $\operatorname{rank}(\mathcal{C}_n(\mathbf{G},\mathbf{B})) = 3$. The pair $(\mathbf{G},\mathbf{B})$ is uncontrollable. Clearly, the unidentifiability of matrix $\mathbf{G}$ is a direct consequence caused by which subset of players is chosen by the adversary to perturb. □

From Examples 1, 2 and 3, it reveals that although the adversary cannot identify the network structure in Fig 1 without perturbations, strategically injecting perturbations on a subset of players can benefit the adversary in identifying the network structure. Moreover, the subset of players it chooses could affect the ability of network structure identification.

IV. Network Structure Learning

In this section, we develop a framework for identifying the network structure for LQ games $\mathbf{G}$ from players' best response action data.

A. The Framework: Stable and Sparse System Identification

Let $\mathcal{S}$ be the set of stable matrices in $\mathbb{R}^{n\times n}$, i.e., $\mathcal{S} := \{A \in \mathbb{R}^{n\times n} : \rho(A) < 1\}$. Assumption 1 requires that $\mathbf{G} \in \mathcal{S}$. Moreover, in practice, $\mathbf{G}$ is typically a sparse matrix. Sparsity is a fundamental characteristic of numerous biological, social, and technological networks references [20]. In summary, the adversary attempts to learn a stable and sparse graph matrix $\mathbf{G}$ from its information $\mathcal{I}$ raising from (p-BR).

Upon the action dynamics (p-BR), we have already arrived at the action difference dynamics (4a) in the proof of Theorem 1:

$$\mathbf{y}(k+1) = \mathbf{G}\mathbf{y}(k) + \mathbf{B}\mathbf{v}(k).$$

Note that the practical observations of the $\mathbf{x}(k)$ (and thus, $\mathbf{y}(k)$) captured by the adversary are subject to the influence of noises. That is $\mathbf{y}^{\mathrm{m}}(k) = \mathbf{y}(k) + \mathbf{e}_k$. It will be the actually observed actions, where $\mathbf{e}_k \in \mathbb{R}^n, k \in \mathcal{T}$ are stationary random noises with zero mean and co-variance $\mathbf{S}_n$. So the adversary has access to

$$Z = \left[\mathbf{y}^{\mathrm{m}}(1) \quad \mathbf{y}^{\mathrm{m}}(2) \quad \cdots \quad \mathbf{y}^{\mathrm{m}}(T)\right]; \tag{5}$$
$$Y = \left[\mathbf{y}^{\mathrm{m}}(0) \quad \mathbf{y}^{\mathrm{m}}(1) \quad \cdots \quad \mathbf{y}^{\mathrm{m}}(T-1)\right]; \tag{6}$$
$$V = \left[\mathbf{v}(0) \quad \mathbf{v}(1) \quad \cdots \quad \mathbf{v}(T-1)\right]. \tag{7}$$

We propose the following Stable Sparse System Identification (SSSI) for learning the network structure from full player action observations:

$$\mathbf{G}_{\mathrm{SSSI}} = \arg \inf_{\mathbf{G}\in\mathcal{S}} \frac{1}{2}||Z - \mathbf{G}Y - \mathbf{B}V||_F^2 + \theta||\mathbf{G}||_1. \tag{SSSI}$$

B. An Information-Theoretic Algorithm

As $\mathcal{S}$ is an open and nonconvex set, solving (SSSI) is numerically challenging. We propose to adopt the recently developed algorithm for stable system identification in [23], where an information-theoretic projection onto $\mathcal{S}$ was used to soften the computational complexity by solving Riccati equations. This leads to Algorithm 1.

Algorithm 1 Information-theoretic Projection SSSI Algorithm

Input: $Z, Y, V, \mathbf{B}$

Output: Network structure $\mathbf{G}$

1: Solve the regularized least square solution

$$\hat{\mathbf{G}} = \arg\min_{\mathbf{G}} \frac{1}{2}||Z - \mathbf{G}Y - \mathbf{B}V||_F^2 + \theta||\mathbf{G}||_1; \tag{8}$$

2: Random generate $\delta \geq 0$;

3: Solve the algebraic Riccati equation (DARE) with the unique solution $\mathbf{P}_\delta$:

$$\mathbf{P} = \mathbf{I} + \hat{\mathbf{G}}^\top\mathbf{P}\hat{\mathbf{G}} - \hat{\mathbf{G}}^\top\mathbf{P}(\mathbf{P} + (2\delta\mathbf{S}_n)^{-1})^{-1}\mathbf{P}\hat{\mathbf{G}}^\top; \tag{9}$$

4: Compute $\mathbf{L}_\delta = -(\mathbf{P}_\delta + (2\delta\mathbf{S}_n)^{-1})^{-1}\mathbf{P}_\delta\hat{\mathbf{G}}$;

5: Return $\mathbf{G}_{\mathrm{SSSI}} = \hat{\mathbf{G}} + \mathbf{L}_\delta$.

V. Experiments

In this section, experiments are designed to evaluate the effectiveness of Algorithm 1 in both synthesis networks and real networks.

A. Synthesis Networks

In this subsection, we examine the performance of Algorithm 1 on three types of synthetic networks generated using the Erdős–Rényi (ER), the Watts–Strogatz (WS), and the Barabási–Albert (BA) models. The networks under evaluation have $n = 100$ nodes.

Network Setup In the ER graph, each link takes place with probability $p_{er} = 0.1$ independently with all the other links; in the WS graph, each node's average degree and the random rewiring process is set to be is $k_{ws} = 5$ and $p_{ws} = 0.2$,

Manuscript 62 submitted to 2023 American Control Conference (ACC). Received September 13, 2022.

respectively; in the BA graph, a new node at each time step is created to connect to $m_{ba} = 2$ existing nodes via preferential attachment. After the realization of each graph, a nonzero random number is selected between -5 and 5 as the weight for each link. Each entry of the adjacency matrix $\mathbf{G}$ is then divided by the largest absolute value of its eigenvalues to ensure the stability $\mathbf{G}$.

Data Generation A set M representing action-compromised players is selected and a matrix $\mathbf{B}$ is created according to (3). We then generate $\mathbf{x}(0)$, $\mathbf{u}(k), k \in \mathcal{T}$, and $\boldsymbol{\alpha}$ by considering $\mathbf{x}(0), \mathbf{u}(k), \boldsymbol{\alpha} \sim \mathcal{N}(0, \mathbf{I})$, and simulate the dynamics (p-BR) to obtain $\mathbf{x}(k), k \in \mathcal{T}$. The observation noises follow a normal distribution $\mathbf{e}_k \sim \xi * \mathcal{N}(0, \mathbf{I})$ where ξ represents the noise intensity level. Upon Eq. (5)-(7), we finally obtain three matrices Z, Y, V.

Baselines We consider two baseline approaches: the stable least square ((SLS)) and stable L_2-regularized least square ((SL2LS)). The method of the stable least square is defined by the optimization problem

$$\inf_{\mathbf{G} \in \mathcal{S}} \frac{1}{2} ||Z - \mathbf{G}Y - \mathbf{B}V||_F^2. \qquad \text{(SLS)}$$

The method of the stable L_2-regularized least square solves

$$\inf_{\mathbf{G} \in \mathcal{S}} \frac{1}{2} ||Z - \mathbf{G}Y - \mathbf{B}V||_F^2 + \beta ||\mathbf{G}||_F^2. \qquad \text{(SL2LS)}$$

Solving (SLS) and (SL2LS) is numerically challenging. As in Algorithm 1, we propose to solve $\hat{\mathbf{G}}_{\mathsf{SLS}} = \arg\min_{\mathbf{G}} \frac{1}{2} ||Z - \mathbf{G}Y - \mathbf{B}V||_F^2$ and $\hat{\mathbf{G}}_{\mathsf{SL2LS}} = \arg\min_{\mathbf{G}} \frac{1}{2} ||Z - \mathbf{G}Y - \mathbf{B}V||_F^2 + \beta ||\mathbf{G}||_F^2$ and project them to their nearest stable ones by applying Eq. (9).

Results By applying Algorithm 1 and baselines to the respective network settings, we obtain the network structures $\mathbf{G}_{\mathrm{learn}}$. We then compare the outcomes against the ground truth by the relative error

$$\mathsf{Err} := \frac{||\mathbf{G}_{\mathrm{learn}} - \mathbf{G}_{\mathrm{Truth}}||_F}{||\mathbf{G}_{\mathrm{Truth}}||_F}. \qquad (10)$$

For (SSSI) and (SL2LS) baselines, we provide the results using the parameters θ and β that yield the greatest average performance across 50 randomly generated graph instances.

First and foremost, we are to discover how the trajectory length affects learning performance. We fix noise level to be $\xi = 0.1$, use 6 different values of $T = \{15, 20, 25, 30, 35, 40\}$ and follow the aforementioned data generation process. The learning performance of the three methods versus trajectory length on the ER, WS, and BA networks is illustrated in Fig. 3.

We next examine the robustness of the three methods in the face of various levels of noise during the observations. Let the trajectory length be $T = 40$ and the noise intensity level take values in $\xi = \{0.05, 0.1, 0.15, 0.2, 0.25, 0.3\}$. The learning performance of the three methods versus noise intensities on the ER, WS, and BA networks is shown in Fig. 4. Clearly, in terms of reconstructing the weights for the ground truth links, the (SSSI) outperforms the other two baseline methods significantly. Moreover, the BA networks appear to be more difficult to learn under (SSSI) in all cases, compared to the ER graphs and WS graphs.

B. Real Networks

In this section, we investigate the performance of Algorithm 1 for learning unweighted real networks where $g_{ij}, \forall i, j \in \mathrm{V}$, is either 0 or 1. In other words, we try to solve the classification task of learning the real networks. Since the real networks used are sparse, we use the area under the curve (AUC) for the evaluation.

Indian Villages Dataset We consider learning social networks in 75 villages in southern rural India [32]. In each village network, nodes represent households and links represent the connection between two households. The average number of nodes and links per village network is 198.72 and 1782.99, respectively. In order to guarantee the stability of $\mathbf{G}$, each element of each village network's binary adjacency matrix is then normalized by the largest absolute value of its eigenvalues.

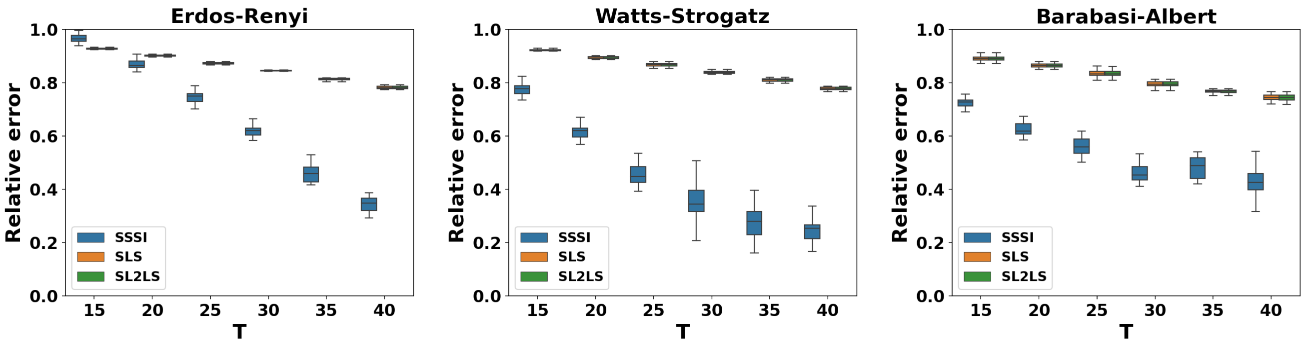
Results We follow the same data generation process in Section V-A. We apply Algorithm 1 to the respective settings to learn village networks and compare the results against the ground truth in a binary classification scenario.

First, we examine AUC versus trajectory length on 75 village networks. We fix noise level to be $\xi = 0.3$, set 4 different values of $T = \{10, 30, 50, 70\}$ and obtain the results in Fig. 5a. Clearly, AUC increases as trajectory length increases.

We then discover how noise level affects the robustness of Algorithm 1. We let trajectory length fixed to be $T = 120$ and noise level take values in $\xi = \{0.15, 0.2, 0.25, 0.3\}$ and present the results in Fig. 5b. It is straightforward to see that AUC decreases as the noise level increases.

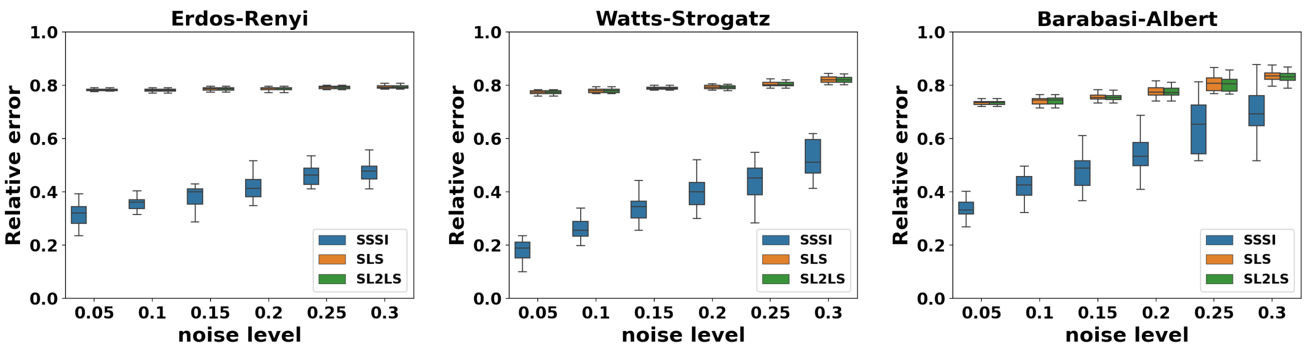
VI. Conclusions

In this work, we considered network structure inference problem for LQ games from best-response dynamics. A complete framework for network learning for LQ games from best-response dynamics is presented. First, we provided theoretic results that characterized network structure identifiability and demonstrated its usefulness with numerical examples. Then, given the intrinsic stability and sparsity constraint of network structure, we developed a stable and sparse system identification algorithm for learning the network structure. Lastly, the effectiveness of the proposed learning algorithm was evaluated in both synthesis networks and real networks. This work opened the door for understanding network structure inference problem for network games based on best-response dynamics from the perspective of classical system identification theory. We bridged the connection between these two fields by introducing stable and sparse system identification into network structure inference for LQ games. Future work may include extensions of this line of research to partial action observations of the adversary and general payoff functions of players.



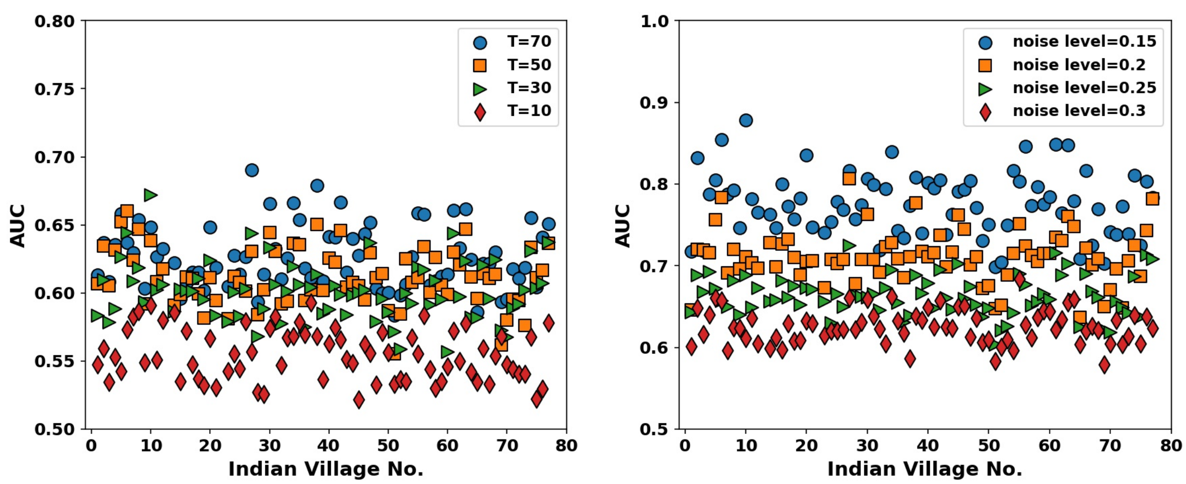
(a) Erdos Renyi (b) Watts Strogatz (c) Barabasi Albert

Fig. 3: The relative errors from (SSSI), (SLS), (SL2LS) versus trajectory length over the ER, WS and BA networks.



(a) Erdos Renyi (b) Watts Strogatz (c) Barabasi Albert

Fig. 4: The relative errors from (SSSI), (SLS), (SL2LS) versus noise intensity level over the ER, WS and BA networks.



(a) AUC versus trajectory length (b) AUC versus noise level

Fig. 5: Learning performance (AUC) of Algorithm 1 on 75 Indian village network datasets.

References

[1] J. Von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton University Press, 2007.

[2] M. O. Jackson and Y. Zenou, "Games on networks," in *Handbook of Game Theory with Economic Applications*. Elsevier, 2015, vol. 4, pp. 95–163.

[3] Y. Leng, Y. Chen, X. Dong, J. Wu, and G. Shi, "Privacy risks of social interaction structure: Network learning in quadratic games," *Available at SSRN 3875878*, 2021.

[4] L. Bianco, M. Caramia, S. Giordani, and V. Piccialli, "A game-theoretic approach for regulating hazmat transportation," *Transportation Science*, vol. 50, no. 2, pp. 424–438, 2016.

[5] D. Niyato and E. Hossain, "Competitive pricing for spectrum sharing

Manuscript 62 submitted to 2023 American Control Conference (ACC).
Received September 13, 2022.

in cognitive radio networks: Dynamic game, inefficiency of nash equilibrium, and collusion," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 1, pp. 192–202, 2008.

[6] D. Fudenberg and J. Tirole, *Game theory*. MIT Press, 1991.

[7] C. H. Papadimitriou and T. Roughgarden, "Computing correlated equilibria in multi-player games," *Journal of the ACM*, vol. 55, no. 3, pp. 1–29, 2008.

[8] C. Ballester, A. Calvó-Armengol, and Y. Zenou, "Who's who in networks. wanted: The key player," *Econometrica*, vol. 74, no. 5, pp. 1403–1417, 2006.

[9] J. Friedman, T. Hastie, and R. Tibshirani, "Sparse inverse covariance estimation with the graphical lasso," *Biostatistics*, vol. 9, no. 3, pp. 432–441, 2008.

[10] S. Wong and C. Butz, "Constructing the dependency structure of a multiagent probabilistic network," *IEEE Transactions on Knowledge and Data Engineering*, vol. 13, no. 3, pp. 395–415, 2001.

[11] M. Gomez-Rodriguez, D. Balduzzi, and B. Schölkopf, "Uncovering the temporal dynamics of diffusion networks," in *Proceedings of the 28th International Conference on Machine Learning*, 2011, pp. 561–568.

[12] X. Dong, D. Thanou, M. Rabbat, and P. Frossard, "Learning graphs from data: A signal representation perspective," *IEEE Signal Processing Magazine*, vol. 36, no. 3, pp. 44–63, 2019.

[13] K. Ouyang, Y. Liang, Y. Liu, Z. Tong, S. Ruan, D. Rosenblum, and Y. Zheng, "Fine-grained urban flow inference," *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 6, pp. 2755–2770, 2022.

[14] A. Barik and J. Honorio, "Provable sample complexity guarantees for learning of continuous-action graphical games with nonparametric utilities," *arXiv preprint:2004.01022*, 2020.

[15] J. Honorio and L. E. Ortiz, "Learning the structure and parameters of large-population graphical games from behavioral data." *Journal of Machine Learning Research*, vol. 16, no. 1, pp. 1157–1210, 2015.

[16] A. Ghoshal and J. Honorio, "From behavior to sparse graphical games: Efficient recovery of equilibria," in *Proceedings of the 54th Annual Allerton Conference on Communication, Control, and Computing*, 2016, pp. 1220–1227.

[17] Y. Leng, X. Dong, J. Wu, and A. Pentland, "Learning quadratic games on networks," in *Proceedings of the 37th International Conference on Machine Learning*, 2020, pp. 5820–5830.

[18] J. De Martí and Y. Zenou, "Network games with incomplete information," *Journal of Mathematical Economics*, vol. 61, pp. 221–240, 2015.

[19] M. O. Jackson, *Social and economic networks*. Princeton University Press, 2010.

[20] A.-L. Barabási, *Network science*. Cambridge University Press, 2016.

[21] J. K. Tugnait and C. Tontiruttananon, "Identification of linear systems via spectral analysis given time-domain data: Consistency, reduced-order approximation, and performance analysis," *IEEE Transactions on Automatic Control*, vol. 43, no. 10, pp. 1354–1373, 1998.

[22] T. Sarkar and A. Rakhlin, "Near optimal finite time identification of arbitrary linear dynamical systems," in *Proceedings of the 36th International Conference on Machine Learning*, 2019, pp. 5610–5618.

[23] W. Jongeneel, T. Sutter, and D. Kuhn, "Efficient learning of a linear dynamical system with stability guarantees," *arXiv preprint:2102.03664*, 2021.

[24] A. Chiuso and G. Pillonetto, "System identification: A machine learning perspective," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 2, pp. 281–304, 2019.

[25] Y. Chen, Y. Gu, and A. O. Hero, "Sparse lms for system identification," in *Proceedings of the 34th IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, pp. 3125–3128.

[26] R. Meng, R. C. de Lamare, and V. H. Nascimento, "Sparsity-aware affine projection adaptive algorithms for system identification," in *Proceedings of the 2nd Sensor Signal Processing for Defence Conference*, 2011, pp. 1–5.

[27] M. L. Littman, "Markov games as a framework for multi-agent reinforcement learning," in *Proceedings of the 11th International Conference on Machine Learning*, 1994, p. 157–163.

[28] D. Fudenberg, F. Drew, D. K. Levine, and D. K. Levine, *The theory of learning in games*. MIT Press, 1998.

[29] C. Cheng, Y. Luo, and C. Yu, "Dynamic mechanism of social bots interfering with public opinion in network," *Physica A: Statistical Mechanics and its Applications*, vol. 551, p. 124163, 2020.

[30] E. Ferrara, "Bots, elections, and social media: a brief overview," *Disinformation, Misinformation, and Fake News in Social Media*, pp. 95–114, 2020.

[31] W. M. Wonham, *Linear Multivariable Control: A Geometric Approach*. Springer, 1985.

[32] M. O. Jackson, T. Rodriguez-Barraquer, and X. Tan, "Social capital and social quilts: Network patterns of favor exchange," *American Economic Review*, vol. 102, no. 5, pp. 1857–97, 2012.

Manuscript 62 submitted to 2023 American Control Conference (ACC).
Received September 13, 2022.