This article has been accepted for publication in IEEE Transactions on Automatic Control. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAC.2025.3610251

Online Parameter Identification of Cost Functions in Generalized Nash Games

Jianguo Chen, Jinlong Lei, *Member, IEEE*, Yiguang Hong, *Fellow, IEEE*, and Hongsheng Qi, *Member, IEEE*

***Abstract*—This work studies the online parameter identification of cost functions in a generalized Nash game, where each player's cost function is influenced by an observable signal and some unknown parameters. A learner can sequentially observe the equilibria reached by the game system as the observable signal changes, and its goal is to identify the unknown parameters in the cost functions. We recast this problem as an online optimization and introduce a novel online parameter identification algorithm. To be specific, we construct a regularized loss function that balances between conservativeness and correctiveness, where the conservativeness term ensures that the renewed estimates do not deviate significantly from the current estimates, while the correctiveness term is captured by the Karush-Kuhn-Tucker conditions of the generalized Nash equilibrium. For which, the balance is adjusted by a learning rate parameter. We then prove that when the players' cost functions are linear with respect to the unknown parameters and the learning rate satisfies $\mu_k \propto 1/\sqrt{k}$, along with other assumptions, the regret bound of the proposed algorithm is $O(\sqrt{K})$ (k and K represent the current round of online algorithm and the total rounds of games played, respectively). Finally, we conduct numerical simulations on the parameter identification of a Nash-Cournot problem to demonstrate that the performance of our proposed online algorithm is comparable to that of the offline setting.**

***Index Terms*—Parameter identification, generalized Nash games, online learning, inverse game, regret bound.**

I. Introduction

A generalized Nash game is an n-person noncooperative game in which strategy sets are coupled across players (see e.g., [1], [2]). Its solution concept, the generalized Nash equilibrium (GNE), has found broad applications in power grids [3], natural gas markets [4], and autonomous driving [5], [6]. To address the computational challenges of GNE, various efficient algorithms have been developed, such as methods based on local Lagrangian functions and exact penalty terms [7] and fast solvers like *ALGAMES* for autonomous driving [8].

The computation of the GNE heavily relies on the access to the cost functions of all players [9] or distributed coordination between players [7]. In some particular settings, the learner can only observe the equilibrium behaviors of players in a given game while remaining unaware of the specific cost functions underlying the game [10]. For instance, in a competitive market consisting of multiple companies, the learner might be able to observe market pricing and product volumes (i.e., observable signals), but lack precise information about the production costs of companies (i.e., the cost functions). Despite the parametric uncertainty in cost functions, the learner might be able to identify these parameters through the observed equilibria. Since knowing the cost functions of the game is important for predicting players' future behaviors, parameter identification in the context of generalized Nash game has significant importance in various fields such as autonomous driving [6] and transportation problem [11].

Real-life scenarios are often faced with challenges where direct access to all equilibria is impossible. Instead, the learner may constantly be observing new equilibrium as observed signals of the environment evolve [12]. For instance, in a competitive market, the game is perpetually in progress, with companies continuously reaching new equilibria as the external market conditions change. In such dynamic situations, identifying parameters of the cost functions based on the streaming data becomes necessary. Thus, this article focuses on the topic regarding the online parameter identification of generalized Nash games.

A. Literature Review

The field of systems and control has witnessed extensive research on the parameter identification problem for both linear [13], [14] and nonlinear systems [15], [16]. However, there remains a dearth of investigation regarding parameter identification in game systems, with a particular emphasis on online identification. Some literatures have referred to the parameter identification in non-cooperative games as inverse games. To address the challenges associated with parameter identification in differential games, researchers have employed the Pontryagin's maximum principle derived from optimal control [17], [18]. Moreover, an inverse optimization method has been developed to estimate parameters within the cost functions of traffic flow games, which are modeled as generalized Nash games [11]. When it comes to matrix games, the inverse problem primarily involves estimating the cost matrix [10], [19]. Furthermore, inverse reinforcement learning has been utilized to explore the intricacies of inverse Markov games [20]. However, it is worth noting that most of the aforementioned works on inverse games assume the availability of simultaneously observed data, resulting in the proposal of offline algorithms.

The online method in [21] is specifically designed for discrete-time noncooperative dynamic games, which cannot be directly applied to generalized Nash games. LUCIDGames, proposed by [6], is based on the unscented Kalman filter and is only used for unconstrained parameter settings, in which the unknown parameters are treated as Gaussian variables. Nevertheless, to the best of our knowledge, the online parameter identification of cost functions in generalized Nash games remains an open problem.

B. Contributions

This work focuses on the online parameter identification of generalized Nash games. The main contributions are summarized as follows:

- We model the problem as an online convex optimization problem, and propose an online algorithm to learn parameters of cost functions in generalized Nash games. Specially, we delicately design a regularized loss function to balance between

The paper is sponsored by the National Key Research and Development Program of China under No. 2022YFA1004701, the National Natural Science Foundation of China under Nos. 62373283, 62173250, 62088101, and the Fundamental Research Funds for the Central Universities, and partially by Shanghai Municipal Science and Technology Major Project No. 2021SHZDZX0100.

Jianguo Chen and Hongsheng Qi are with the State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China; also with the University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: chenjianguo@amss.ac.cn, qihongsh@amss.ac.cn).

Jinlong Lei and Yiguang Hong are with the Department of Control Science and Engineering, Tongji University, Shanghai, 201804, China; also with State Key Laboratory of Autonomous Intelligent Unmanned Systems, and Frontiers Science Center for Intelligent Autonomous Systems, Ministry of Education, and the Shanghai Institute of Intelligent Science and Technology, Tongji University, Shanghai, 200092, China (e-mail: leijinlong@tongji.edu.cn, yghong@tongji.edu.cn).

conservativeness and correctiveness, where conservatism refers to the use of existing information, and correctness refers to the use of newly acquired information.
- We prove that when the players' cost functions are linear in the unknown parameters, the learning rate of the algorithm $\mu_k \propto 1/\sqrt{k}$ and other assumptions are satisfied, the regret bound of the online parameter identification algorithm is $O(\sqrt{K})$.

C. Paper Organization and Notations

Paper Organization: Section II formulates the online parameter identification of generalized Nash games. In Section III, we propose an online algorithm to identify parameters of cost functions in generalized Nash games, while the regret bound is established in Section IV with some proofs given in Appendices. Section V showcases numerical simulations on a natural gas market, and some concluding remarks are provided in Section VI.

Notations: $\mathcal{R}^n$ represents the n-dimensional vector space; $\mathcal{R}^{n\times m}$ represents the space of $n\times m$-dimensional matrices; $\|\cdot\|_2$ denotes the $\mathcal{L}_2$-norm; bold letters denote vectors or matrices; $\mathrm{col}(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_N)$ represents the column vector $(\boldsymbol{x}_1^T,\cdots,\boldsymbol{x}_N^T)^T$; vector $\boldsymbol{v}>0$ indicates that every element is positive, and vector $\boldsymbol{v}\geq 0$ indicates that every element is non-negative; $\boldsymbol{a}\perp\boldsymbol{b}$ signifies that the product of the corresponding elements of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ is equal to 0; $\langle\cdot,\cdot\rangle$ denotes the inner product between vectors; and $R_{\geq 0}$ denotes the set of non-negative real numbers.

II. Problem Statement

In this section, we present the problem statement concerning the online parameter identification of games.

A. Generalized Nash games

The generalized Nash game is composed of N players, denoted as $\mathcal{N}\triangleq\{1,\cdots,N\}$. Each player $v\in\mathcal{N}$ possesses control over its strategy $\boldsymbol{x}_v\in\boldsymbol{X}_v\subseteq\mathcal{R}^{n_v}$, where $\boldsymbol{X}_v$ represents the feasible set of decision variables for player v. For every player $v\in\mathcal{N}$, considering the parameter $\boldsymbol{\theta}_v\in\boldsymbol{\Theta}_v\subseteq\mathcal{R}^{n_v'}$, the observable signal $\boldsymbol{u}_v\in\boldsymbol{U}_v\subseteq\mathcal{R}^{n_v''}$ and the decision variables of other players $\boldsymbol{x}_{-v}\in\boldsymbol{X}_{-v}\triangleq\prod_{s\neq v}\boldsymbol{X}_s$, it selects a strategy $\boldsymbol{x}_v$ to minimize the optimization problem

$$\begin{aligned}&\min_{\boldsymbol{x}_v} f_v(\boldsymbol{x}_v;\boldsymbol{x}_{-v},\boldsymbol{u}_v,\boldsymbol{\theta}_v)\\ &s.t.\ \boldsymbol{x}_v\in\boldsymbol{X}_v(\boldsymbol{x}_{-v},\boldsymbol{u}_v),\end{aligned}\tag{1}$$

where $f_v:\boldsymbol{X}_v\times\boldsymbol{X}_{-v}\times\boldsymbol{U}_v\times\boldsymbol{\Theta}_v\mapsto\mathcal{R}$, with $n\triangleq\sum_{v=1}^N n_v$, represents the cost function of player v. The feasible set $\boldsymbol{X}_v(\boldsymbol{x}_{-v},\boldsymbol{u}_v)$ depends on the decision variables of other players and the signal. Here, we impose an assumption, which was widely used in literatures [22]–[24].

Assumption 1. *For every player $v\in\mathcal{N}$, $\boldsymbol{u}_v\in\boldsymbol{U}_v$, $\boldsymbol{\theta}_v\in\boldsymbol{\Theta}_v$ and $\boldsymbol{x}_{-v}\in\boldsymbol{X}_{-v}$, the cost function $f_v(\cdot;\boldsymbol{x}_{-v},\boldsymbol{u}_v,\boldsymbol{\theta}_v)$ is convex and continuously differentiable in $\boldsymbol{X}_v(\boldsymbol{x}_{-v},\boldsymbol{u}_v)$. Moreover, the sets $X_v(\boldsymbol{x}_{-v},\boldsymbol{u}_v)$ and $\boldsymbol{\Theta}_v$ are closed and convex.*

For simplicity, let $\boldsymbol{x}=\mathrm{col}(\boldsymbol{x}_1,\cdots,\boldsymbol{x}_N)$, $\boldsymbol{u}=\mathrm{col}(\boldsymbol{u}_1,\cdots,\boldsymbol{u}_N)\in\boldsymbol{U}$ and $\boldsymbol{\theta}=\mathrm{col}(\boldsymbol{\theta}_1,\cdots,\boldsymbol{\theta}_N)\in\boldsymbol{\Theta}\triangleq\prod_{v=1}^N\boldsymbol{\Theta}_v$. Here, $\boldsymbol{\Theta}$ is a closed and convex set from Assumption 1. The generalized Nash equilibrium (GNE) defined below is a crucial solution concept for the problem (1).

Definition 1. *Given $\boldsymbol{\theta}\in\boldsymbol{\Theta}$ and $\boldsymbol{u}\in\boldsymbol{U}$, let $\boldsymbol{x}^*=\mathrm{col}(\boldsymbol{x}_1^*,\cdots,\boldsymbol{x}_N^*)$ be a GNE. Then for every player $v\in\mathcal{N}$ and any $\boldsymbol{x}_v\in X_v(\boldsymbol{x}_{-v}^*,\boldsymbol{u}_v)$, the following holds.*

$$f_v(\boldsymbol{x}_v^*;\boldsymbol{x}_{-v}^*,\boldsymbol{u}_v,\boldsymbol{\theta}_v)\leq f_v(\boldsymbol{x}_v;\boldsymbol{x}_{-v}^*,\boldsymbol{u}_v,\boldsymbol{\theta}_v).$$

GNE refers to a scenario in the generalized Nash game where each player's current strategy minimizes its cost function, while the strategies of other players remain unchanged. According to [9, Theorem 4.1], the existence of a GNE can be guaranteed under Assumption 1.

One extensively studied coupling constraint set in the problem (1) is the jointly convex generalized Nash game (c.f. [9]).

Assumption 2. *There exists a closed convex set $\boldsymbol{X}(\boldsymbol{u})\subseteq\mathcal{R}^n$ associated with the signal $\boldsymbol{u}$ such that for each player v,*

$$X_v(\boldsymbol{x}_{-v},\boldsymbol{u}_v)=\{\boldsymbol{x}_v|(\boldsymbol{x}_v,\boldsymbol{x}_{-v})\in\boldsymbol{X}(\boldsymbol{u})\}.$$

Additionally, $X(\boldsymbol{u})$ is given by $\{\boldsymbol{x}\mid h_q(\boldsymbol{x},\boldsymbol{u})\leq 0, 1\leq q\leq m; g_j(\boldsymbol{x},\boldsymbol{u})=0, 1\leq j\leq p\}$, where m and p are the number of inequality and equality constraints, respectively. Moreover, $h_q(\cdot,\boldsymbol{u})$ and $g_j(\cdot,\boldsymbol{u})$ are continuously differentiable functions.

B. Online Parameter Identification of the Generalized Nash Game based on Noisy Equilibrium Observation

The computation of equilibrium in games typically necessitates the knowledge of all players' cost functions [9]. However, in real-world scenarios, the learner may not have direct access to the players' cost functions. Instead, the learner might observe the equilibrium solutions of these problems (see e.g., [6]). Let $\boldsymbol{y}$ represent the observed equilibrium of the game problem (1). Our objective is to estimate the unknown parameter $\boldsymbol{\theta}$ of the cost functions based on the observed equilibrium $\boldsymbol{y}$. This process is known as parameter identification in the game. In particular, we consider a scenario where the observed equilibrium is disturbed by noise, i.e., $\boldsymbol{y}=\boldsymbol{x}^*+\boldsymbol{\epsilon}$, where $\boldsymbol{x}^*$ denotes the true equilibrium, while $\boldsymbol{\epsilon}$ represents a random variable following a specific distribution. We focus on the online scenario where data are sequentially observed, namely, in the k-th round, the learner has access to a signal $\boldsymbol{u}^k$ and a noise-corrupted equilibrium $\boldsymbol{y}^k$. Let $\boldsymbol{\theta}^1$ denote an initial estimate of the unknown parameters, and $\boldsymbol{\theta}^{k+1}\in\Theta$ represent the estimate in the k-th round. Upon obtaining new observed data $(\boldsymbol{y}^k,\boldsymbol{u}^k)$ in the k-th round, the learner can update the estimate $\boldsymbol{\theta}^{k+1}\in\Theta$ using an online learning algorithm guided by a well-designed loss function $l(\boldsymbol{\theta};\boldsymbol{y}^k,\boldsymbol{u}^k)$.

Remark 1. *Our framework is applicable to various real-world scenarios. For example, Stackelberg games with multi-followers can be used to model the electric vehicles charging scheduling problem [25], where the learner (leader) can infer followers' cost functions using historical follower equilibria and signals (leader actions). In autonomous driving, vehicles can learn others' cost functions using historical equilibria (trajectories data) and signals (initial states and lane constraints of vehicles at different times) to improve interaction efficiency [6].*

The performance evaluation of such an online parameter identification algorithm of the generalized Nash game is assessed by the regret defined as follows.

Definition 2. *The regret of online parameter identification is*

$$R_K\triangleq\sum_{k=1}^K l(\boldsymbol{\theta}^k;\boldsymbol{y}^k,\boldsymbol{u}^k)-\sum_{k=1}^K l(\boldsymbol{\theta}_*^K;\boldsymbol{y}^k,\boldsymbol{u}^k),\tag{2}$$

where $\boldsymbol{\theta}_*^K\triangleq\arg\min_{\theta\in\boldsymbol{\Theta}}\sum_{k=1}^K l(\boldsymbol{\theta};\boldsymbol{y}^k,\boldsymbol{u}^k)$ represents the optimal inference in the whole batch setting. In particular, an online learning algorithm is said to have the no-regret property if $R_K=o(K)$.

The regret can reflect the performance of an online parameter identification algorithm by comparing the cumulative loss function with that in the whole batch learning.

III. Online Parameter Identification Algorithm

In this section, we first define the loss function based on Karush-Kuhn-Tucker (KKT) conditions of the GNE. Subsequently, we design an online parameter identification.

A. Loss Function Based on KKT Conditions

From [9], the generalized Nash game (1) satisfying Assumption 1 may admit multiple generalized Nash equilibria. Here we consider a special setting where the observed equilibria belong to the class of variational equilibrium (VE).

Remark 2. *The VE is characterized by equal optimal dual variables for common constraints across all players. This ensures identical shadow prices for all players, with significant economic implications [26]. Owing to this fairness property, VE is widely applied in real-world problems. For instance, in autonomous driving systems, VE in a trajectory game enables fair interactions among connected autonomous vehicles by ensuring equitable payoff reduction for collision avoidance [5].*

Definition 3. *Under Assumptions 1-2, $\boldsymbol{x}^*$ is a VE if, given $\boldsymbol{u} \in \boldsymbol{U}$ and $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, it satisfies the variational inequality $VI(\boldsymbol{X}, \boldsymbol{F})$, i.e.,*

$$\boldsymbol{F}(\boldsymbol{x}^*, \boldsymbol{u}, \boldsymbol{\theta})^T(\boldsymbol{y} - \boldsymbol{x}^*) \geq 0,\ \forall \boldsymbol{y} \in X(\boldsymbol{u}), \tag{3}$$

where $\boldsymbol{F}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\theta}) \triangleq (\nabla_{\boldsymbol{x}_v} f_v(\boldsymbol{x}_v; \boldsymbol{x}_{-v}, \boldsymbol{u}_v, \boldsymbol{\theta}_v))_{v=1}^N$.

With $X(\boldsymbol{u})$ defined in Assumptions 2, the KKT conditions corresponding to variational inequality (3) are as follows.

$$\begin{cases} \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{u}, \boldsymbol{\theta}) + \sum_{q=1}^m \lambda_q \nabla_{\boldsymbol{x}} h_q(\boldsymbol{x}, \boldsymbol{u}) + \sum_{j=1}^p \nu_j \nabla_{\boldsymbol{x}} g_j(\boldsymbol{x}, \boldsymbol{u}) = \boldsymbol{0} \\ \lambda_q \geq 0, h_q(\boldsymbol{x}, \boldsymbol{u}) \leq 0 \text{ and } \lambda_q h_q(\boldsymbol{x}, \boldsymbol{u}) = 0,\ \forall q, 1 \leq q \leq m \\ g_j(\boldsymbol{x}, \boldsymbol{u}) = 0,\ \forall j, 1 \leq j \leq p, \end{cases} \tag{4}$$

where $\lambda_q, q = 1, \cdots, m$ and $\nu_j, j = 1, \cdots, p$ are the dual variables corresponding to the inequality and equality constraints, respectively. Denote by $\boldsymbol{\lambda} = \mathrm{col}(\lambda_1, \cdots, \lambda_m)$ and $\boldsymbol{\nu} = \mathrm{col}(\nu_1, \cdots, \nu_p)$. Subsequently, we define the loss function $l(\boldsymbol{\theta}; \boldsymbol{y}, \boldsymbol{u})$ based on the KKT conditions (4).

Definition 4. *For a given signal $\boldsymbol{u} \in \boldsymbol{U}$ and the corresponding observed equilibrium $\boldsymbol{y}$, set $l(\boldsymbol{\theta}; \boldsymbol{y}, \boldsymbol{u}) \triangleq \min_{\boldsymbol{\lambda}, \boldsymbol{\nu}} \{L(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{\nu}; \boldsymbol{y}, \boldsymbol{u}) \mid \boldsymbol{\lambda} \geq 0\}$, where*

$$\begin{aligned} L(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{\nu}; \boldsymbol{y}, \boldsymbol{u}) &= \sum_{q=1}^m \chi(\lambda_q h_q(\boldsymbol{y}, \boldsymbol{u})) + \sum_{j=1}^p \chi(g_j(\boldsymbol{y}, \boldsymbol{u})) \\ &+ \chi\Bigg(\boldsymbol{F}(\boldsymbol{y}, \boldsymbol{u}, \boldsymbol{\theta}) + \sum_{q=1}^m \lambda_q \nabla_{\boldsymbol{y}} h_q(\boldsymbol{y}, \boldsymbol{u}) + \sum_{j=1}^p \nu_j \nabla_{\boldsymbol{y}} g_j(\boldsymbol{y}, \boldsymbol{u})\Bigg), \end{aligned}$$

and $\chi(\cdot) : \mathcal{R}^n \mapsto R_{\geq 0}$ is a penalty function with $\chi(\boldsymbol{0}) = 0$.

Remark 3. *When the penalty function is convex, L is also convex with respect to the dual variables $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$, as $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$ appear affinely in $\chi(\cdot)$. Therefore, the loss function l, defined via the minimization operation with respect to $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$, is both well-defined and unique.*

B. Online Parameter Identification Algorithm

In the k-th round, the learner obtains a new observation $(\boldsymbol{y}^k, \boldsymbol{u}^k)$ and possesses the estimate $\boldsymbol{\theta}^k$ from the previous $(k-1)$-round. The learner updates the estimate using the learning method proposed in [27] to strike a balance between conservativeness and correctiveness. This implies that the new estimate $\boldsymbol{\theta}^{k+1}$ should align with the new observation while preserving certain consistency with the previous estimate $\boldsymbol{\theta}^k$. In particular, we define the regularized loss function $G_k(\boldsymbol{\theta})$ as

$$G_k(\boldsymbol{\theta}) = \underbrace{D\left(\boldsymbol{\theta}, \boldsymbol{\theta}^k\right)}_{\text{conservativeness}} + \mu_k \underbrace{l\left(\boldsymbol{\theta}; \boldsymbol{y}^k, \boldsymbol{u}^k\right)}_{\text{correctiveness}}, \tag{5}$$

where $D(\cdot, \cdot) : \boldsymbol{\Theta} \times \boldsymbol{\Theta} \mapsto R_{\geq 0}$ denotes a distance function, μ_k is the learning rate, and $l(\boldsymbol{\theta}; \boldsymbol{y}^k, \boldsymbol{u}^k)$ is defined in Definition 4. The first term of (5) captures the 'conservativeness' by evaluating the distance between $\boldsymbol{\theta}$ and $\boldsymbol{\theta}^k$, while the second term measures the 'correctiveness' by assessing the concordance between $\boldsymbol{\theta}$ and the new observation $(\boldsymbol{y}^k, \boldsymbol{u}^k)$. The learning rate μ_k effectively balances these two aspects.

The parameter update involves two steps. Firstly, we get the optimal inference $\tilde{\boldsymbol{\theta}}^{k+1}$ by minimizing the regularized loss function. Secondly, we find a point $\boldsymbol{\theta}^{k+1}$ within $\boldsymbol{\Theta}$ that has the minimum distance from $\tilde{\boldsymbol{\theta}}^{k+1}$. The two steps are as follows:

$$\tilde{\boldsymbol{\theta}}^{k+1} = \arg\min_{\boldsymbol{\theta}} G_k(\boldsymbol{\theta}), \tag{6a}$$

$$\boldsymbol{\theta}^{k+1} = \arg\min_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} D\left(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}^{k+1}\right). \tag{6b}$$

The key computational cost of this update process lies in solving the optimization problem (6a). As a matter of fact, with the definition of $l(\boldsymbol{\theta}; \boldsymbol{y}, \boldsymbol{u})$ in Definition 4, we need to solve the following problem.

$$\begin{aligned} &\min_{\boldsymbol{\theta}}\ D\left(\boldsymbol{\theta}, \boldsymbol{\theta}^k\right) + \mu_k l\left(\boldsymbol{\theta}; \boldsymbol{y}^k, \boldsymbol{u}^k\right) \\ =& \min_{\boldsymbol{\theta}, \boldsymbol{\lambda} \geq 0, \boldsymbol{\nu}}\ D\left(\boldsymbol{\theta}, \boldsymbol{\theta}^k\right) + \mu_k L\left(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{\nu}; \boldsymbol{y}^k, \boldsymbol{u}^k\right). \end{aligned} \tag{7}$$

As such, the parameter update step (6a) is converted to

$$\begin{aligned} &\left(\tilde{\boldsymbol{\theta}}^{k+1}, \boldsymbol{\lambda}^k, \boldsymbol{\nu}^k\right) \\ &= \arg\min_{\boldsymbol{\theta}, \boldsymbol{\lambda} \geq 0, \boldsymbol{\nu}} \left(D\left(\boldsymbol{\theta}, \boldsymbol{\theta}^k\right) + \mu_k L\left(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{\nu}; \boldsymbol{y}^k, \boldsymbol{u}^k\right)\right). \end{aligned} \tag{8}$$

Therefore, we summarize the procedure of the online parameter identification of generalized Nash games in Algorithm 1.

Algorithm 1 Online parameter identification algorithm

Require: $\theta^1 \in \Theta$
1: Let $k \leftarrow 1$
2: **While** $k \leq K$ **do**
3: Observe $(\boldsymbol{y}^k, \boldsymbol{u}^k)$
4: Solve $\min_{\boldsymbol{\theta}, \boldsymbol{\lambda} \geq 0, \boldsymbol{\nu}}\ D\left(\boldsymbol{\theta}, \boldsymbol{\theta}^k\right) + \mu_k L\left(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{\nu}; \boldsymbol{y}^k, \boldsymbol{u}^k\right)$ to get $\tilde{\boldsymbol{\theta}}^{k+1}$
5: Solve $\min_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} D\left(\boldsymbol{\theta}, \tilde{\boldsymbol{\theta}}^{k+1}\right)$ to get $\boldsymbol{\theta}^{k+1}$
6: $k \leftarrow k + 1$
7: **end**

The problem (8) is convex under certain conditions, for example, the penalty and distance functions are the square of Euclidean distance and $\boldsymbol{\theta}$ is linearly structured in the cost functions. This can be proved by the convexity of $l\left(\cdot; \boldsymbol{y}^k, \boldsymbol{u}^k\right)$, which is validated in the next section. Therefore, step 4 can be solved by some well-known optimization algorithms, like Trust-Region Constrained Algorithm and Sequential Quadratic Programming [28].

IV. Theoretical Analysis

In this section, we show that under certain conditions the regret bound (2) produced by Algorithm 1 is $O(\sqrt{K})$.

A. Assumptions and Preliminary Results

For simplicity, we define

$$\begin{aligned}
&\boldsymbol{F_\theta} \triangleq \boldsymbol{F}(\boldsymbol{y},\boldsymbol{u},\boldsymbol{\theta}) \in \mathcal{R}^n,\\
&\nabla \boldsymbol{h} \triangleq [\nabla_{\boldsymbol{y}} h_1(\boldsymbol{y},\boldsymbol{u}),\cdots,\nabla_{\boldsymbol{y}} h_m(\boldsymbol{y},\boldsymbol{u})] \in \mathcal{R}^{n\times m},\\
&\nabla \boldsymbol{g} \triangleq [\nabla_{\boldsymbol{y}} g_1(\boldsymbol{y},\boldsymbol{u}),\cdots,\nabla_{\boldsymbol{y}} g_p(\boldsymbol{y},\boldsymbol{u})] \in \mathcal{R}^{n\times p},\\
&\boldsymbol{H} \triangleq \mathrm{diag}(h_1(\boldsymbol{y},\boldsymbol{u}),\cdots,h_m(\boldsymbol{y},\boldsymbol{u})) \in \mathcal{R}^{m\times m},\\
&\boldsymbol{g} \triangleq [g_1(\boldsymbol{y},\boldsymbol{u}),\cdots,g_p(\boldsymbol{y},\boldsymbol{u})]^T \in \mathcal{R}^p,\\
&\boldsymbol{\lambda} \triangleq (\lambda_1,\cdots,\lambda_m)^T \in \mathcal{R}^m, \boldsymbol{\nu} \triangleq (\nu_1,\cdots,\nu_p)^T \in \mathcal{R}^p.
\end{aligned} \tag{9}$$

Here, the players' cost functions in the game are linear with respect to the unknown parameters $\boldsymbol{\theta}$, which is widely adopted in many applications, such as autonomous driving [5] and smart grid cost allocation [29]. It also aligns with prior works on inverse game theory [19], [21].

Assumption 3. *For every $v \in \mathcal{N}$, $f_v(\boldsymbol{x}_v;\boldsymbol{x}_{-v},\boldsymbol{u}_v,\boldsymbol{\theta}_v)$ is linear with respect to $\boldsymbol{\theta}_v$, i.e., there exists auxiliary functions $\tilde{\boldsymbol{f}}_v(\boldsymbol{x}_v;\boldsymbol{x}_{-v},\boldsymbol{u}_v) : \boldsymbol{X}_v \times \boldsymbol{X}_{-v} \times \boldsymbol{U}_v \mapsto \mathcal{R}^{n'_v}$ and $\breve{f}_v(\boldsymbol{x}_v;\boldsymbol{x}_{-v},\boldsymbol{u}_v) : \boldsymbol{X}_v \times \boldsymbol{X}_{-v} \times \boldsymbol{U}_v \mapsto \mathcal{R}$, which do not depend on the parameter $\boldsymbol{\theta}$ (Abbreviated as f_v, $\tilde{\boldsymbol{f}}_v$, $\breve{f}_v$ in the following), such that $f_v = \boldsymbol{\theta}_v^T \tilde{\boldsymbol{f}}_v + \breve{f}_v$.*

Then we introduce an assumption regarding the boundedness of certain variables.

Assumption 4. *(1) $\boldsymbol{y}$, $\boldsymbol{u}$ and $\boldsymbol{\theta}$ are bounded, i.e., there exists a real number $B_1 > 0$, such that for every $\boldsymbol{u} \in \boldsymbol{U}, \boldsymbol{\theta} \in \boldsymbol{\Theta}$ and $\boldsymbol{y}$, we have $\max\{\|\boldsymbol{u}\|_2, \|\boldsymbol{y}\|_2, \|\boldsymbol{\theta}\|_2\} < B_1$.*
(2) Furthermore, there exists $B_2 > 0$ such that for every $v \in \mathcal{N}$, $\boldsymbol{u} \in \boldsymbol{U}, \boldsymbol{y}$ and $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, we have

$$\max\left\{\left\|\nabla_{\boldsymbol{x}_v}\tilde{\boldsymbol{f}}_v\right\|_2, \left\|\nabla_{\boldsymbol{x}_v}\breve{f}_v\right\|, \|\boldsymbol{H}\|_2, \|\nabla\boldsymbol{h}\|_2, \|\nabla\boldsymbol{g}\|_2\right\} < B_2.$$

Under Assumptions 3 and 4, we can establish the boundedness and Lipschitz continuity of $\boldsymbol{F_\theta}$ defined by (9).

Proposition 1. *Let Assumptions 3 and 4 hold. Then there exist constants $M_1, M_2 > 0$ such that for every $\boldsymbol{u} \in \boldsymbol{U}$ and $\boldsymbol{y}$,*

$$\|\boldsymbol{F_\theta}\|_2 \le M_1, \forall \boldsymbol{\theta} \in \boldsymbol{\Theta}, \tag{10}$$
$$\|\boldsymbol{F}_{\boldsymbol{\theta}_1} - \boldsymbol{F}_{\boldsymbol{\theta}_2}\|_2 \le M_2 \|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|_2, \forall \boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \boldsymbol{\Theta}. \tag{11}$$

Proof. From the linearity of f_v with respect to $\boldsymbol{\theta}_v$, we can obtain the structure of $\boldsymbol{F_\theta}$ as follows.

$$\boldsymbol{F}_\theta = \begin{bmatrix} \nabla_{\boldsymbol{x}_1}\tilde{\boldsymbol{f}}_1 & & \\ & \ddots & \\ & & \nabla_{\boldsymbol{x}_N}\tilde{\boldsymbol{f}}_N \end{bmatrix}^T \boldsymbol{\theta} + \begin{bmatrix} \nabla_{\boldsymbol{x}_1}\breve{f}_1 \\ \vdots \\ \nabla_{\boldsymbol{x}_N}\breve{f}_N \end{bmatrix}, \tag{12}$$

where $\boldsymbol{\theta} = \mathrm{col}(\boldsymbol{\theta_1},\cdots,\boldsymbol{\theta_N})$. By the boundedness of $\|\boldsymbol{\theta}\|_2$, $\left\|\nabla_{\boldsymbol{x}_v}\tilde{\boldsymbol{f}}_v\right\|_2$ and $\left\|\nabla_{\boldsymbol{x}_v}\breve{f}_v\right\|_2, v \in \mathcal{N}$ from Assumption 4, according to the structure of $\boldsymbol{F_\theta}$ in (12), there exists a positive constant M_1 such that (10) holds.

From (12), we have that for every $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \boldsymbol{\Theta}$,

$$\boldsymbol{F}_{\boldsymbol{\theta}_1} - \boldsymbol{F}_{\boldsymbol{\theta}_2} = \begin{bmatrix} \nabla_{\boldsymbol{x}_1}\tilde{\boldsymbol{f}}_1 & & \\ & \ddots & \\ & & \nabla_{\boldsymbol{x}_N}\tilde{\boldsymbol{f}}_N \end{bmatrix}^T (\boldsymbol{\theta_1} - \boldsymbol{\theta_2}).$$

Since $\left\|\nabla_{\boldsymbol{x}_v}\tilde{\boldsymbol{f}}_v\right\|_2 < B_2, v \in \mathcal{N}$ from Assumption 4, (11) holds for a positive constant M_2. □

In the following, we set the distance function as $D(\boldsymbol{x},\boldsymbol{y}) = \frac{1}{2}\|\boldsymbol{x}-\boldsymbol{y}\|_2^2$ and the penalty function as the square of the $\mathcal{L}_2$-norm. Thus, the loss function defined in Definition 4 is given by

$$l(\boldsymbol{\theta};\boldsymbol{y},\boldsymbol{u}) = \min_{\boldsymbol{\lambda}\ge 0,\boldsymbol{\nu}} \{\|\boldsymbol{F_\theta} + \nabla\boldsymbol{h}\boldsymbol{\lambda} + \nabla\boldsymbol{g}\boldsymbol{\nu}\|_2^2 + \|\boldsymbol{H}\boldsymbol{\lambda}\|_2^2 + \|\boldsymbol{g}\|_2^2\}, \tag{13}$$

and the regularized loss function (5) can be rewritten as

$$G_k(\boldsymbol{\theta}) = \frac{1}{2}\left\|\boldsymbol{\theta} - \boldsymbol{\theta}^k\right\|_2^2 + \mu_k l(\boldsymbol{\theta};\boldsymbol{y}^k,\boldsymbol{u}^k), \tag{14}$$

where $l(\boldsymbol{\theta};\boldsymbol{y}^k,\boldsymbol{u}^k)$ is defined by (13).

An assumption concerning the linear independence between gradients of constraint functions is presented as follows.

Assumption 5. *$[\nabla\boldsymbol{h}, \nabla\boldsymbol{g}] \in \mathcal{R}^{n\times(m+p)}$ is a column full rank matrix, where $\nabla\boldsymbol{h}$ and $\nabla\boldsymbol{g}$ are defined in (9).*

Assumption 5 is valid in certain practical settings. For instance, consider the resource constraint in a market, which can be expressed as $\boldsymbol{W} \triangleq \{\boldsymbol{y'} \mid \boldsymbol{A'}\boldsymbol{y'} \le \boldsymbol{b'}\}$, where $\boldsymbol{A'} \in \mathcal{R}^{m\times n}$ is a matrix with full row rank and $\boldsymbol{b'} \in \mathcal{R}^m$. Note that Assumptions 3-5 are useful for ensuring that the loss function defined by (13) is both Lipschitz continuous and convex with respect to the unknown parameter θ. These properties are particularly useful for establishing the regret bound of Algorithm 1. Then we provide two propositions.

Proposition 2. *Under Assumptions 1-5, the loss function (13) and is uniformly C-Lipschitz continuous in Θ for every $\boldsymbol{u} \in \boldsymbol{U}$ and $\boldsymbol{y}$, namely,*

$$|l(\boldsymbol{\theta}_1;\boldsymbol{y},\boldsymbol{u}) - l(\boldsymbol{\theta}_2;\boldsymbol{y},\boldsymbol{u})| \le C\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|_2, \forall \boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \boldsymbol{\Theta}.$$

Proof. See Appendix A. □

Proposition 3. *Under Assumptions 1-5, the loss function* (13) *is convex in $\boldsymbol{\Theta}$ for every $\boldsymbol{u} \in \boldsymbol{U}$ and $\boldsymbol{y}$.*

Proof. See Appendix B. □

B. Regret Bound

The following theorem reveals that the regret bound of the online parameter identification algorithm is $O(\sqrt{K})$.

Theorem 1. *Let Assumptions 1-5 hold. Consider Algorithm 1, where $\mu_k = \mu_1/\sqrt{k}$ with $\mu_1 > 0$. Then the regret bound defined in* (2) *is $O(\sqrt{K})$.*

Proof. Due to (6a) and the convexity of the loss function (13) from Proposition 3, we can conclude that $\boldsymbol{0} \in \partial G_k(\tilde{\boldsymbol{\theta}}^{k+1})$, where $\partial G_k(\tilde{\boldsymbol{\theta}}^{k+1})$ is the subgradient set of $G_k(\cdot)$ at $\tilde{\boldsymbol{\theta}}^{k+1}$. Therefore, by recalling $D(\boldsymbol{x},\boldsymbol{y}) = \frac{1}{2}\|\boldsymbol{x}-\boldsymbol{y}\|_2^2$, there exists a subgradient $\boldsymbol{s} \in \partial l(\tilde{\boldsymbol{\theta}}^{k+1};\boldsymbol{y}^k,\boldsymbol{u}^k)$ such that

$$\tilde{\boldsymbol{\theta}}^{k+1} - \boldsymbol{\theta}^k + \mu_k\boldsymbol{s} = \boldsymbol{0}. \tag{15}$$

From the convexity of the loss function,

$$l(\boldsymbol{\theta}_*^K;\boldsymbol{y}^k,\boldsymbol{u}^k) \ge l(\tilde{\boldsymbol{\theta}}^{k+1};\boldsymbol{y}^k,\boldsymbol{u}^k) + \boldsymbol{s}^T(\boldsymbol{\theta}_*^K - \tilde{\boldsymbol{\theta}}^{k+1}). \tag{16}$$

Then by using (15),

$$\begin{aligned}
&-\mu_k\boldsymbol{s}^T\left(\boldsymbol{\theta}_*^K - \tilde{\boldsymbol{\theta}}^{k+1}\right) = \left(\tilde{\boldsymbol{\theta}}^{k+1} - \boldsymbol{\theta}^k\right)^T\left(\boldsymbol{\theta}_*^K - \tilde{\boldsymbol{\theta}}^{k+1}\right)\\
&=\frac{1}{2}\left\|\boldsymbol{\theta}^k - \boldsymbol{\theta}_*^K\right\|_2^2 - \frac{1}{2}\left\|\tilde{\boldsymbol{\theta}}^{k+1} - \boldsymbol{\theta}_*^K\right\|_2^2 - \frac{1}{2}\left\|\tilde{\boldsymbol{\theta}}^{k+1} - \boldsymbol{\theta}^k\right\|_2^2.
\end{aligned} \tag{17}$$

According to the Lipschitz continuity of the loss function from Proposition 2,

$$\begin{aligned}
&\mu_k\left(l\left(\boldsymbol{\theta}^k;\boldsymbol{y}^k,\boldsymbol{u}^k\right)-l\left(\tilde{\boldsymbol{\theta}}^{k+1};\boldsymbol{y}^k,\boldsymbol{u}^k\right)\right)-\tfrac{1}{2}\left\|\tilde{\boldsymbol{\theta}}^{k+1}-\boldsymbol{\theta}^k\right\|_2^2\\
\le& C\mu_k\left\|\boldsymbol{\theta}^k-\tilde{\boldsymbol{\theta}}^{k+1}\right\|_2-\tfrac{1}{2}\left\|\tilde{\boldsymbol{\theta}}^{k+1}-\boldsymbol{\theta}^k\right\|_2^2\\
=&2\left[\tfrac{1}{2}\left\|\boldsymbol{\theta}^k-\tilde{\boldsymbol{\theta}}^{k+1}\right\|_2\left(C\mu_k-\tfrac{1}{2}\left\|\boldsymbol{\theta}^k-\tilde{\boldsymbol{\theta}}^{k+1}\right\|_2\right)\right]\\
\le&2\left[\frac{\tfrac{1}{2}\left\|\boldsymbol{\theta}^k-\tilde{\boldsymbol{\theta}}^{k+1}\right\|_2+C\mu_k-\tfrac{1}{2}\left\|\boldsymbol{\theta}^k-\tilde{\boldsymbol{\theta}}^{k+1}\right\|_2}{2}\right]^2=\frac{C^2\mu_k^2}{2}.
\end{aligned}\tag{18}$$

Because $\boldsymbol{\theta}^{k+1}\in\boldsymbol{\Theta},\boldsymbol{\theta}_*^K\in\boldsymbol{\Theta}$, according to the generalized Pythagorous inequality (see [30, Theorem 2.5.1]), we have $\frac{1}{2}\left\|\tilde{\boldsymbol{\theta}}^{k+1}-\boldsymbol{\theta}^{k+1}\right\|_2^2\le\frac{1}{2}\left\|\tilde{\boldsymbol{\theta}}^{k+1}-\boldsymbol{\theta}_*^K\right\|_2^2-\frac{1}{2}\left\|\boldsymbol{\theta}^{k+1}-\boldsymbol{\theta}_*^K\right\|_2^2$. Thus,

$$\begin{aligned}
&\frac{1}{2}\left\|\boldsymbol{\theta}^k-\boldsymbol{\theta}_*^K\right\|_2^2-\frac{1}{2}\left\|\tilde{\boldsymbol{\theta}}^{k+1}-\boldsymbol{\theta}_*^K\right\|_2^2\\
\le&\frac{1}{2}\left\|\boldsymbol{\theta}^k-\boldsymbol{\theta}_*^K\right\|_2^2-\frac{1}{2}\left\|\tilde{\boldsymbol{\theta}}^{k+1}-\boldsymbol{\theta}_*^K\right\|_2^2+\frac{1}{2}\left\|\tilde{\boldsymbol{\theta}}^{k+1}-\boldsymbol{\theta}^{k+1}\right\|_2^2\\
\le&\frac{1}{2}\left\|\boldsymbol{\theta}^k-\boldsymbol{\theta}_*^K\right\|_2^2-\frac{1}{2}\left\|\boldsymbol{\theta}^{k+1}-\boldsymbol{\theta}_*^K\right\|_2^2.
\end{aligned}\tag{19}$$

Therefore,

$$\begin{aligned}
&\mu_k l(\boldsymbol{\theta}^k;\boldsymbol{y}^k,\boldsymbol{u}^k)-\mu_k l(\boldsymbol{\theta}_*^K;\boldsymbol{y}^k,\boldsymbol{u}^k)\\
\overset{(16)}{\le}&\mu_k l(\boldsymbol{\theta}^k;\boldsymbol{y}^k,\boldsymbol{u}^k)-\mu_k l(\tilde{\boldsymbol{\theta}}^{k+1};\boldsymbol{y}^k,\boldsymbol{u}^k)-\mu_k\boldsymbol{s}^T(\boldsymbol{\theta}_*^K-\tilde{\boldsymbol{\theta}}^{k+1})\\
\overset{(17)}{=}&\mu_k l(\boldsymbol{\theta}^k;\boldsymbol{y}^k,\boldsymbol{u}^k)-\mu_k l(\tilde{\boldsymbol{\theta}}^{k+1};\boldsymbol{y}^k,\boldsymbol{u}^k)+\frac{1}{2}\left\|\boldsymbol{\theta}^k-\boldsymbol{\theta}_*^K\right\|_2^2\\
&-\frac{1}{2}\left\|\tilde{\boldsymbol{\theta}}^{k+1}-\boldsymbol{\theta}_*^K\right\|_2^2-\frac{1}{2}\left\|\tilde{\boldsymbol{\theta}}^{k+1}-\boldsymbol{\theta}^k\right\|_2^2\\
\overset{(18)}{\le}&\frac{C^2\mu_k^2}{2}+\frac{1}{2}\left\|\boldsymbol{\theta}^k-\boldsymbol{\theta}_*^K\right\|_2^2-\frac{1}{2}\left\|\tilde{\boldsymbol{\theta}}^{k+1}-\boldsymbol{\theta}_*^K\right\|_2^2\\
\overset{(19)}{\le}&\frac{C^2\mu_k^2}{2}+\frac{1}{2}\left\|\boldsymbol{\theta}^k-\boldsymbol{\theta}_*^K\right\|_2^2-\frac{1}{2}\left\|\boldsymbol{\theta}^{k+1}-\boldsymbol{\theta}_*^K\right\|_2^2.
\end{aligned}\tag{20}$$

After dividing both sides of inequality (19) by μ_k and summing over k, we derive

$$\begin{aligned}
R_K=&\sum_{k=1}^K l(\boldsymbol{\theta}^k;\boldsymbol{y}^k,\boldsymbol{u}^k)-\sum_{k=1}^K l(\boldsymbol{\theta}_*^K;\boldsymbol{y}^k,\boldsymbol{u}^k)\\
=&\underbrace{\sum_{k=1}^K\frac{C^2}{2}\mu_k}_{\mathrm{I}}+\underbrace{\sum_{k=1}^K\frac{1}{2\mu_k}\left(\left\|\boldsymbol{\theta}^k-\boldsymbol{\theta}_*^K\right\|_2^2-\left\|\boldsymbol{\theta}^{k+1}-\boldsymbol{\theta}_*^K\right\|_2^2\right)}_{\mathrm{II}}.
\end{aligned}\tag{21}$$

Recall $\mu_k=\mu_1/\sqrt{k}$. For part I, using $\sum_{k=1}^K\frac{1}{\sqrt{k}}\le 1+\int_1^K\frac{1}{\sqrt{k}}dt\le 2\sqrt{K}$, we have $\sum_{k=1}^K\frac{C^2}{2}\mu_k\le\mu_1C^2\sqrt{K}$. For part II, since $\|\theta\|_2<B_1$ from Assumption 4,

$$\begin{aligned}
&\sum_{k=1}^K\frac{1}{2\mu_k}\left(\left\|\boldsymbol{\theta}^k-\boldsymbol{\theta}_*^K\right\|_2^2-\left\|\boldsymbol{\theta}^{k+1}-\boldsymbol{\theta}_*^K\right\|_2^2\right)\\
=&\frac{1}{2\mu_1}\sum_{k=1}^K\sqrt{k}\left(\left\|\boldsymbol{\theta}^k-\boldsymbol{\theta}_*^K\right\|_2^2-\left\|\boldsymbol{\theta}^{k+1}-\boldsymbol{\theta}_*^K\right\|_2^2\right)\\
=&\frac{1}{2\mu_1}\sum_{k=2}^K\left\|\boldsymbol{\theta}^k-\boldsymbol{\theta}_*^K\right\|_2^2(\sqrt{k}-\sqrt{k-1})\\
&+\frac{1}{2\mu_1}\left[\left\|\boldsymbol{\theta}^1-\boldsymbol{\theta}_*^K\right\|_2^2-\sqrt{K}\left\|\boldsymbol{\theta}^{K+1}-\boldsymbol{\theta}_*^K\right\|_2^2\right]\\
\le&\frac{1}{\mu_1}\left[2B_1^2(\sqrt{K}-1)+2B_1^2\right]=\frac{1}{\mu_1}2B_1^2\sqrt{K}.
\end{aligned}$$

Therefore, $R_K\le\left(\frac{2B_1^2}{\mu_1}+\mu_1C^2\right)\sqrt{K}$ from (21). Thus, the conclusion follows. □

V. Numerical Simulations

This section applies Algorithm 1 to identify the parameters of cost functions in a Nash-Cournot model, which is widely adopted for modeling market competition (see e.g., [3], [31]) and can be seen as a generalized Nash game.

A. Simulation Model

Consider an energy market with N companies engaged in selling natural gas. Each company v competes in the market by determining its output x_v. When $x_v\ge 0$, company v sells x_v units of natural gas to the market; when $x_v<0$, it buys $|x_v|$ units of natural gas from the market. Additionally, a minimum gas demand of $q_k>0$ must be met, i.e., $\sum_{v=1}^N x_v\ge q_k$. Suppose that the market price of natural gas is influenced by the total amount of natural gas $\sum_{v=1}^N x_v$ in the market, and is set as $p_k\triangleq a_k+b_k\sum_{v=1}^N x_v$, where $a_k>0,b_k<0$ are observable variables affected by the market. Take $a_k+b_kq_k>0$ to make the price meaningful. The profit of company v is given by $f_v\left(x_v,\mathbf{x}_{-v},(a_k,b_k),\theta_v\right)\triangleq p_kx_v-\theta_vx_v$, where $\theta_v\ge 0$ denotes company v's production cost per unit of natural gas. The goal of every company v is to maximize its profit, namely, it selects a strategy x_v to minimize the following optimization problem:

$$\min_{x_v}-f_v\left(x_v,\mathbf{x}_{-v},(a_k,b_k),\theta_v\right),\quad s.t.\ \sum_{v=1}^N x_v\ge q_k.$$

This game is a jointly convex generalized Nash game. Let (a_k,b_k,q_k) be an observable signal that adjusts with the changes of the market, and $\theta_v\in\{\theta_v:\theta_v\ge 0\},v\in\mathcal{N}$ be unknown parameters that we aim to estimate online.

B. Simulation Setting and Results

Consider a market with three companies, where the unknown parameter vector $\boldsymbol{\theta}=[\theta_1,\theta_2,\theta_3]^T$ is taken as $[10,7.5,6]^T$. In each round k, we independently sample $\boldsymbol{u}^k=(a_k,b_k,q_k)$ from uniform distributions with ranges $[15,15\times 120\%]$, $[1,1\times 120\%]$, and $[5,5\times 120\%]$, respectively. The observed equilibrium result $\boldsymbol{y}^k=\boldsymbol{x}^k+\boldsymbol{\epsilon}^k$ incorporates noise, where $\boldsymbol{x}^k$ represents the Nash-Cournot equilibrium under the signal $\boldsymbol{u}^k$, and $\boldsymbol{\epsilon}^k$ is a random vector drawn from a multivariate standard normal distribution. Subsequently, we set $K=100$ and run the online parameter estimation algorithm and the algorithm in the whole batch setting. The effectiveness of this online algorithm is evaluated using the following metrics.

1) Time(s): the time required for the algorithm per round.
2) R_k/k: the average regret of the algorithm.
3) $\|\boldsymbol{\theta}^k-\boldsymbol{\theta}_*^k\|_2$: the deviation between the estimates $\boldsymbol{\theta}^k$ and $\boldsymbol{\theta}_*^k$. Here, $\boldsymbol{\theta}^k$ denotes the estimated unknown parameter in the online parameter algorithm at the k-th round, while $\boldsymbol{\theta}_*^k$ represents the corresponding estimate by using data from the previous k rounds in the whole batch setting.

Set $\mu_k=\mu_1/\sqrt{k}$ with $\mu_1=0.1$ in Algorithm 1, and compare the results with those obtained under the whole batch setting, as illustrated in Figure 1. Notably, while the whole batch setting can run independently at each round, our online algorithm offers a significant advantage in execution time with online data, whereas the whole batch method's execution time grows almost exponentially with the round. Moreover, both the average regret and $\|\boldsymbol{\theta}^k-\boldsymbol{\theta}_*^k\|_2$ all exhibit the convergence towards 0, indicating that the performance of our proposed online algorithm (Algorithm 1) closely approximates that

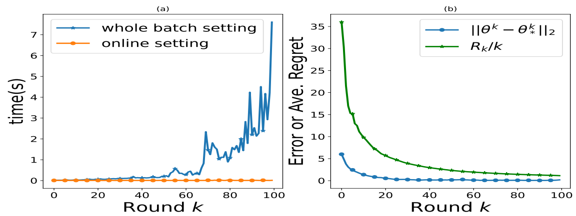

Fig. 1. Algorithm performance of the online parameter identification in a natural gas market. (a) The vertical coordinate is the time of algorithm execution. (b) The vertical coordinate indicates the deviation between $\boldsymbol{\theta}^k$ and $\boldsymbol{\theta}_*^k$ and the average regret for each round.

of the whole batch setting. These results highlight the effectiveness and efficiency of the online algorithm.

The learning rate μ_k plays an important role in Algorithm 1 by balancing the existing information and new observations. We set different learning rates $\mu_k = \mu_1/\sqrt{k}$ with $\mu_1 = 0.1, 0.3$ and 0.5 to evaluate the algorithm performance. Figure 2 (a) illustrates that a higher learning rate leads to a lower regret bound, probably because a larger learning rate indicates a greater emphasis on new observations. Besides, Figure 2 (b) demonstrates that a larger learning rate facilitates faster convergence of $\boldsymbol{\theta}^k$ to $\boldsymbol{\theta}_*^k$, albeit with greater fluctuations.

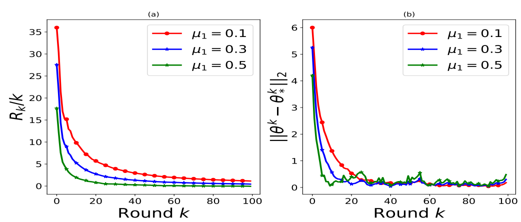

Fig. 2. The effect of learning rates on the average regret and $\|\boldsymbol{\theta}^k - \boldsymbol{\theta}_*^k\|_2$.

VI. Conclusion

In this work, an online algorithm was designed to identify the parameters in the cost functions of the generalized Nash game. Furthermore, it was proven that when the cost function is linear in the unknown parameters, and the learning rate of the algorithm $\mu_k \propto 1/\sqrt{k}$ along with other assumptions are satisfied, the regret bound is $O(\sqrt{K})$. Finally, numerical simulations of a Nash-Cournot problem were also implemented to demonstrate that the performance of the online algorithm is comparable to the whole batch setting after some rounds. For future direction, it is of interest to study the parameter identification of more complex and practical game systems, like dynamic game systems with nonlinear quadratic cost functions.

A. Proof of Proposition 2

To prove Proposition 2, we first give the following lemma.

Lemma 1. *For every $\boldsymbol{\theta}_1 \in \boldsymbol{\Theta}$ and $\boldsymbol{\theta}_2 \in \boldsymbol{\Theta}$, let $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ and $(\boldsymbol{\lambda}^\#, \boldsymbol{\nu}^\#)$ be the optimal solutions of the optimization problem*

$$\min_{\boldsymbol{\lambda}\geq \mathbf{0},\boldsymbol{\nu}} L\left(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{\nu}; \boldsymbol{y}, \boldsymbol{u}\right), \tag{22}$$

corresponding to $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, respectively. Under Assumptions 1 - 5, for every $\boldsymbol{u} \in \boldsymbol{U}$ and $\boldsymbol{y}$, there exist $a, d > 0$ such that $\left\|\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^\#\right\|_2 \leq a\left\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\right\|_2, \left\|\boldsymbol{\nu}^* - \boldsymbol{\nu}^\#\right\|_2 \leq a\left\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\right\|_2$ and $\max\left\{\left\|\boldsymbol{\lambda}^*\right\|_2, \left\|\boldsymbol{\lambda}^\#\right\|_2, \left\|\boldsymbol{\nu}^*\right\|_2, \left\|\boldsymbol{\nu}^\#\right\|_2\right\} \leq d$.

Proof. By recalling $L\left(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{\nu}; \boldsymbol{y}, \boldsymbol{u}\right)$ under (13) and the notation of (9), we have

$$\begin{aligned} L\left(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{\nu}; \boldsymbol{y}, \boldsymbol{u}\right) &= \|\boldsymbol{F}_{\boldsymbol{\theta}} + \nabla \boldsymbol{h}\boldsymbol{\lambda} + \nabla \boldsymbol{g}\boldsymbol{\nu}\|_2^2 + \|\boldsymbol{H}\boldsymbol{\lambda}\|_2^2 + \|\boldsymbol{g}\|_2^2 \\ &= \begin{bmatrix}\boldsymbol{\lambda}\\ \boldsymbol{\nu}\end{bmatrix}^T \begin{bmatrix}\nabla \boldsymbol{h}^T\nabla \boldsymbol{h} + \boldsymbol{H}^2 & \nabla \boldsymbol{h}^T \nabla \boldsymbol{g}\\ \nabla \boldsymbol{g}^T\nabla \boldsymbol{h} & \nabla \boldsymbol{g}^T\nabla \boldsymbol{g}\end{bmatrix} \begin{bmatrix}\boldsymbol{\lambda}\\ \boldsymbol{\nu}\end{bmatrix} \\ &\quad + 2\boldsymbol{F}_{\boldsymbol{\theta}}^T \begin{bmatrix}\nabla \boldsymbol{h} & \nabla \boldsymbol{g}\end{bmatrix} \begin{bmatrix}\boldsymbol{\lambda}\\ \boldsymbol{\nu}\end{bmatrix} + \boldsymbol{F}_{\boldsymbol{\theta}}^T\boldsymbol{F}_{\boldsymbol{\theta}} + \boldsymbol{g}^T\boldsymbol{g}, \end{aligned} \tag{23}$$

where $\boldsymbol{H}^2 = diag\left(\boldsymbol{h}_1^2(\boldsymbol{y}, \boldsymbol{u}), \cdots, \boldsymbol{h}_m^2(\boldsymbol{y}, \boldsymbol{u})\right)$. Therefore, the optimization problem (22) can be abbreviated as

$$\min_{\boldsymbol{z}} \boldsymbol{z}^T\boldsymbol{A}\boldsymbol{z} + 2\boldsymbol{F}_{\boldsymbol{\theta}}^T\boldsymbol{b}\boldsymbol{z} + \tilde{c}, \ s.t. \ \boldsymbol{B}\boldsymbol{z} \geq \mathbf{0}, \tag{24}$$

where $\boldsymbol{z} = \begin{bmatrix}\boldsymbol{\lambda}\\ \boldsymbol{\nu}\end{bmatrix}$, $\boldsymbol{A} = \begin{bmatrix}\nabla \boldsymbol{h}^T\nabla \boldsymbol{h} + \boldsymbol{H}^2 & \nabla \boldsymbol{h}^T\nabla \boldsymbol{g}\\ \nabla \boldsymbol{g}^T\nabla \boldsymbol{h} & \nabla \boldsymbol{g}^T\nabla \boldsymbol{g}\end{bmatrix}$, $\boldsymbol{b} = \begin{bmatrix}\nabla \boldsymbol{h} & \nabla \boldsymbol{g}\end{bmatrix}$, $\tilde{\boldsymbol{c}} = \boldsymbol{F}_{\boldsymbol{\theta}}^T\boldsymbol{F}_{\boldsymbol{\theta}} + \boldsymbol{g}^T\boldsymbol{g}$ and $\boldsymbol{B} = \begin{bmatrix}\boldsymbol{I}_m & \boldsymbol{O}_p\end{bmatrix}$.

We now show that matrix $\boldsymbol{A}$ is positive definite. Because $[\nabla \boldsymbol{h}, \nabla \boldsymbol{g}]$ is a column full rank matrix from Assumption 5, we have $\nabla \boldsymbol{h}\boldsymbol{w}_1 + \nabla \boldsymbol{g}\boldsymbol{w}_2 \neq 0$ for every non-zero vector $\boldsymbol{w} = \left[\boldsymbol{w}_1^T, \boldsymbol{w}_2^T\right]^T \in \mathcal{R}^{m+p}$. Thus, $\boldsymbol{w}^T\boldsymbol{A}\boldsymbol{w} = \|\nabla \boldsymbol{h}\boldsymbol{w}_1 + \nabla \boldsymbol{g}\boldsymbol{w}_2\|_2^2 + \|\boldsymbol{H}\boldsymbol{w}_1\|_2^2 > 0$ for every non-zero vector $\boldsymbol{w}$. Therefore, $\boldsymbol{A}$ is positive definite.

For every $\boldsymbol{\theta}_1 \in \boldsymbol{\Theta}$ and $\boldsymbol{\theta}_2 \in \boldsymbol{\Theta}$, since $\boldsymbol{z}^* = \mathrm{col}(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ and $\boldsymbol{z}^\# = \mathrm{col}(\boldsymbol{\lambda}^\#, \boldsymbol{\nu}^\#)$ are the optimal solutions of the optimization problem (24) corresponding to $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, respectively. According to the KKT conditions,

$$\boldsymbol{A}\boldsymbol{z}^* + \boldsymbol{b}^T\boldsymbol{F}_{\boldsymbol{\theta}_1} - \frac{1}{2}\boldsymbol{B}^T\boldsymbol{\beta}_1 = 0 \tag{25a}$$

$$\mathbf{0} \leq \boldsymbol{B}\boldsymbol{z}^* \perp \boldsymbol{\beta}_1 \geq \mathbf{0}, \boldsymbol{\beta}_1 \in \mathcal{R}^{m+p} \tag{25b}$$

$$\boldsymbol{A}\boldsymbol{z}^\# + \boldsymbol{b}^T\boldsymbol{F}_{\boldsymbol{\theta}_2} - \frac{1}{2}\boldsymbol{B}^T\boldsymbol{\beta}_2 = 0 \tag{25c}$$

$$\mathbf{0} \leq \boldsymbol{B}\boldsymbol{z}^\# \perp \boldsymbol{\beta}_2 \geq \mathbf{0}, \boldsymbol{\beta}_2 \in \mathcal{R}^{m+p}. \tag{25d}$$

Because $\boldsymbol{A}$ is a positive definite matrix, its smallest eigenvalue denoted by $\lambda_{\min}(\boldsymbol{A})$ is positive. Then

$$\|\boldsymbol{z}^* - \boldsymbol{z}^\#\|_2^2 \leq \frac{1}{\lambda_{\min}(\boldsymbol{A})}\left\|\boldsymbol{z}^* - \boldsymbol{z}^\#\right\|_2 \|\boldsymbol{b}\|_2 \left\|\boldsymbol{F}_{\boldsymbol{\theta}_1} - \boldsymbol{F}_{\boldsymbol{\theta}_2}\right\|_2. \tag{26}$$

(A detailed proof of inequation (26) can be found in the supplementary material.) Furthermore, we obtain from Proposition 1 that

$$\begin{aligned} &\left\|\boldsymbol{z}^* - \boldsymbol{z}^\#\right\|_2 \leq \frac{1}{\lambda_{\min}(\boldsymbol{A})}\|\boldsymbol{b}\|_2\|\boldsymbol{F}_{\boldsymbol{\theta}_1} - \boldsymbol{F}_{\boldsymbol{\theta}_2}\|_2 \\ &\leq \frac{M_2}{\lambda_{\min}(\boldsymbol{A})}\|\boldsymbol{b}\|_2\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|_2 < \frac{2M_2B_2}{\lambda_{\min}(\boldsymbol{A})}\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|_2, \end{aligned} \tag{27}$$

where the last inequality holds since $\|\boldsymbol{b}\|_2 \leq \|\nabla \boldsymbol{h}\|_2 + \|\nabla \boldsymbol{g}\|_2 < 2B_2$ from Assumption 4(2). Thus, $\left\|\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^\#\right\|_2 \leq \left\|\boldsymbol{z}^* - \boldsymbol{z}^\#\right\|_2 \leq a\left\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\right\|_2$ and $\left\|\boldsymbol{\nu}^* - \boldsymbol{\nu}^\#\right\|_2 \leq \left\|\boldsymbol{z}^* - \boldsymbol{z}^\#\right\|_2 \leq a\left\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\right\|_2$ with $a \triangleq \frac{2M_2B_2}{\lambda_{\min}(\boldsymbol{A})}$.

From (25a) and (25b), we have

$$\boldsymbol{z}^{*T}\boldsymbol{A}\boldsymbol{z}^* = -\boldsymbol{z}^{*T}\boldsymbol{b}^T\boldsymbol{F}_{\boldsymbol{\theta}_1} + \frac{1}{2}\boldsymbol{z}^{*T}\boldsymbol{B}^T\boldsymbol{\beta}_1 = -\boldsymbol{z}^{*T}\boldsymbol{b}^T\boldsymbol{F}_{\boldsymbol{\theta}_1}.$$

Similar to the derivation of (26) and (27), we have $\|\boldsymbol{z}^*\|_2 \leq \frac{1}{\lambda_{\min}(\boldsymbol{A})}\left\|\boldsymbol{b}^T\boldsymbol{F}_{\boldsymbol{\theta}_1}\right\|_2 \leq \frac{2M_1B_2}{\lambda_{\min}(\boldsymbol{A})}$, where the last inequality is obtained by using $\|\boldsymbol{b}\|_2 < 2B_2$ and $\left\|\boldsymbol{F}_{\boldsymbol{\theta}}\right\| < M_1$ in Proposition 1. And because $\|\boldsymbol{\lambda}^*\|_2 \leq \|\boldsymbol{z}^*\|_2$ and $\|\boldsymbol{\nu}^*\|_2 \leq \|\boldsymbol{z}^*\|_2$, $\|\boldsymbol{\lambda}^*\|_2, \|\boldsymbol{\nu}^*\|_2 \leq d$, where $d = \frac{2M_1B_2}{\lambda_{\min}(\boldsymbol{A})}$. Then the boundedness of $\boldsymbol{\lambda}^\#$ and $\boldsymbol{\nu}^\#$ can be obtained in the same way. Thus, the conclusion follows. □

Proof of Proposition 2: For every $\boldsymbol{u} \in \boldsymbol{U}, \boldsymbol{y}$ and every $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \boldsymbol{\Theta}$, let $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$ and $\left(\boldsymbol{\lambda}^\#, \boldsymbol{\nu}^\#\right)$ be optimal solutions of the optimization problem (22) corresponding to $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$, respectively. Then $l(\boldsymbol{\theta}_1; \boldsymbol{y}, \boldsymbol{u}) = L(\boldsymbol{\theta}_1, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*; \boldsymbol{y}, \boldsymbol{u})$ and $l(\boldsymbol{\theta}_2; \boldsymbol{y}, \boldsymbol{u}) = L(\boldsymbol{\theta}_2, \boldsymbol{\lambda}^\#, \boldsymbol{\nu}^\#; \boldsymbol{y}, \boldsymbol{u})$, where $L(\cdot; \boldsymbol{y}, \boldsymbol{u})$ is given in (23).

Without loss of generality, suppose $l(\boldsymbol{\theta}_1;\boldsymbol{y},\boldsymbol{u}) \geq l(\boldsymbol{\theta}_2;\boldsymbol{y},\boldsymbol{u})$. Then

$$\begin{aligned}
&|l(\boldsymbol{\theta}_1;\boldsymbol{y},\boldsymbol{u}) - l(\boldsymbol{\theta}_2;\boldsymbol{y},\boldsymbol{u})| = l(\boldsymbol{\theta}_1;\boldsymbol{y},\boldsymbol{u}) - l(\boldsymbol{\theta}_2;\boldsymbol{y},\boldsymbol{u})\\
=&\|\boldsymbol{F}_{\boldsymbol{\theta}_1} + \nabla\boldsymbol{h}\boldsymbol{\lambda}^* + \nabla\boldsymbol{g}\boldsymbol{\nu}^*\|_2^2 - \|\boldsymbol{F}_{\boldsymbol{\theta}_2} + \nabla\boldsymbol{h}\boldsymbol{\lambda}^{\#} + \nabla\boldsymbol{g}\boldsymbol{\nu}^{\#}\|_2^2\\
&+\|\boldsymbol{H}\boldsymbol{\lambda}^*\|_2^2 - \|\boldsymbol{H}\boldsymbol{\lambda}^{\#}\|_2^2\\
=&\Big\langle \boldsymbol{F}_{\boldsymbol{\theta}_1} - \boldsymbol{F}_{\boldsymbol{\theta}_2} + \nabla\boldsymbol{h}(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{\#}) + \nabla\boldsymbol{g}(\boldsymbol{\nu}^* - \boldsymbol{\nu}^{\#}),\\
&\quad \boldsymbol{F}_{\boldsymbol{\theta}_1} + \nabla\boldsymbol{h}\boldsymbol{\lambda}^* + \nabla\boldsymbol{g}\boldsymbol{\nu}^* + \boldsymbol{F}_{\boldsymbol{\theta}_2} + \nabla\boldsymbol{h}\boldsymbol{\lambda}^{\#} + \nabla\boldsymbol{g}\boldsymbol{\nu}^{\#}\Big\rangle\\
&+\langle \boldsymbol{H}(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{\#}), \boldsymbol{H}(\boldsymbol{\lambda}^* + \boldsymbol{\lambda}^{\#})\rangle\\
\leq&\Big(\|\boldsymbol{F}_{\boldsymbol{\theta}_1} - \boldsymbol{F}_{\boldsymbol{\theta}_2}\|_2 + \|\nabla\boldsymbol{h}(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{\#})\|_2 + \|\nabla\boldsymbol{g}(\boldsymbol{\nu}^* - \boldsymbol{\nu}^{\#})\|_2\Big)\\
&\cdot\Big(\|\boldsymbol{F}_{\boldsymbol{\theta}_1} + \nabla\boldsymbol{h}\boldsymbol{\lambda}^* + \nabla\boldsymbol{g}\boldsymbol{\nu}^*\|_2 + \|\boldsymbol{F}_{\boldsymbol{\theta}_2} + \nabla\boldsymbol{h}\boldsymbol{\lambda}^{\#} + \nabla\boldsymbol{g}\boldsymbol{\nu}^{\#}\|_2\Big)\\
&+\|\boldsymbol{H}(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{\#})\|_2\|\boldsymbol{H}(\boldsymbol{\lambda}^* + \boldsymbol{\lambda}^{\#})\|_2,
\end{aligned} \tag{28}$$

where the last inequality holds by the Cauchy-Schwartz inequality. According to Assumption 4, Proposition 1, and Lemma 1, there exists a real number $E > 0$ such that

$$\max\Big\{\|\boldsymbol{H}(\boldsymbol{\lambda}^* + \boldsymbol{\lambda}^{\#})\|_2, \|\boldsymbol{F}_{\boldsymbol{\theta}_1} + \nabla\boldsymbol{h}\boldsymbol{\lambda}^* + \nabla\boldsymbol{g}\boldsymbol{\nu}^*\|_2, \|\boldsymbol{F}_{\boldsymbol{\theta}_2} + \nabla\boldsymbol{h}\boldsymbol{\lambda}^{\#} + \nabla\boldsymbol{g}\boldsymbol{\nu}^{\#}\|_2\Big\} \leq E.$$

This together with (28) implies

$$\begin{aligned}
&|l(\boldsymbol{\theta}_1;\boldsymbol{y},\boldsymbol{u}) - l(\boldsymbol{\theta}_2;\boldsymbol{y},\boldsymbol{u})| \leq E\|\boldsymbol{H}(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{\#})\|_2 +\\
&2E\big(\|\boldsymbol{F}_{\boldsymbol{\theta}_1} - \boldsymbol{F}_{\boldsymbol{\theta}_2}\|_2 + \|\nabla\boldsymbol{h}(\boldsymbol{\lambda}^* - \boldsymbol{\lambda}^{\#})\|_2 + \|\nabla\boldsymbol{g}(\boldsymbol{\nu}^* - \boldsymbol{\nu}^{\#})\|_2\big).
\end{aligned}$$

Then by recalling the Lipschitz continuity of $\boldsymbol{F}_{\boldsymbol{\theta}}$ in (11), using Lemma 1 and Assumption 4(2), we have

$$|l(\boldsymbol{\theta}_1;\boldsymbol{y},\boldsymbol{u}) - l(\boldsymbol{\theta}_2;\boldsymbol{y},\boldsymbol{u})| \leq C\|\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\|_2,$$

where $C = E(2M_2 + 5aB_2)$. □

B. Proof of Proposition 3

To prove Proposition 3, we give the following lemma.

Lemma 2. *Under Assumption 5, we have*

(1) $\nabla\boldsymbol{g}^T\nabla\boldsymbol{g}$ is invertible;

(2)

$$\boldsymbol{R} \triangleq \boldsymbol{I} - \nabla\boldsymbol{g}(\nabla\boldsymbol{g}^T\nabla\boldsymbol{g})^{-1}\nabla\boldsymbol{g}^T \tag{29}$$

is positive semidefinite and moreover, both $\nabla\boldsymbol{h}^T\boldsymbol{R}\nabla\boldsymbol{h}$ and $\nabla\boldsymbol{h}^T\boldsymbol{R}\nabla\boldsymbol{h} + \boldsymbol{H}^2$ are invertible; and

(3)

$$\boldsymbol{Q} \triangleq \boldsymbol{R} - \boldsymbol{R}\nabla\boldsymbol{h}(\nabla\boldsymbol{h}^T\boldsymbol{R}\nabla\boldsymbol{h} + \boldsymbol{H}^2)^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R} \tag{30}$$

is a positive semidefinite matrix.

Proof. (1) Because $\nabla\boldsymbol{g}$ is a column full rank matrix, by the Singular Value Decomposition (SVD), we have

$$\nabla\boldsymbol{g} = \boldsymbol{U}\begin{bmatrix}\boldsymbol{\Sigma}_p\\ \boldsymbol{0}\end{bmatrix}\boldsymbol{V}^T, \tag{31}$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are orthogonal matrices, and $\boldsymbol{\Sigma}_p$ is a p dimensional diagonal matrix with non-zero diagonal elements. Thus, $\nabla\boldsymbol{g}^T\nabla\boldsymbol{g} = \boldsymbol{V}\boldsymbol{\Sigma}_p^2\boldsymbol{V}^T$ is invertible and positive definite.

(2) Putting (31) into $\boldsymbol{R}$, we have

$$\boldsymbol{R} = \boldsymbol{U}\begin{bmatrix}\boldsymbol{0} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{I}_{n-p}\end{bmatrix}\boldsymbol{U}^T. \tag{32}$$

Thus, $\boldsymbol{R}$ is a positive semidefinite matrix. Note that $\boldsymbol{R}^2 = \boldsymbol{R}$.

We can show that $\boldsymbol{R}\nabla\boldsymbol{h}$ is a column full rank matrix by contradiction. (A detailed proof of this conclusion can be found in the supplementary material.) Similar to the method of proving that $\nabla\boldsymbol{g}^T\nabla\boldsymbol{g}$ is invertible and positive definite, $\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{R}\nabla\boldsymbol{h}$ is invertible and positive definite. Thus, $\nabla\boldsymbol{h}^T\boldsymbol{R}\nabla\boldsymbol{h}$ is invertible and positive definite because $\boldsymbol{R}^2 = \boldsymbol{R}$. Furthermore, $\nabla\boldsymbol{h}^T\boldsymbol{R}\nabla\boldsymbol{h} + \boldsymbol{H}^2$ is also invertible and positive definite.

(3) Denote $\boldsymbol{Q}_1 \triangleq \boldsymbol{R} - \boldsymbol{R}\nabla\boldsymbol{h}(\nabla\boldsymbol{h}^T\boldsymbol{R}\nabla\boldsymbol{h})^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}$. Because $\boldsymbol{R}^2 = \boldsymbol{R}$, we have

$$\begin{aligned}
\boldsymbol{Q}_1 &= \boldsymbol{R}^2 - \boldsymbol{R}^2\nabla\boldsymbol{h}(\nabla\boldsymbol{h}^T\boldsymbol{R}^2\nabla\boldsymbol{h})^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}^2\\
&= \boldsymbol{R}\Big\{\boldsymbol{I} - (\boldsymbol{R}\nabla\boldsymbol{h})\Big[(\boldsymbol{R}\nabla\boldsymbol{h})^T(\boldsymbol{R}\nabla\boldsymbol{h})\Big]^{-1}(\boldsymbol{R}\nabla\boldsymbol{h})^T\Big\}\boldsymbol{R}.
\end{aligned} \tag{33}$$

Since $\boldsymbol{R}\nabla\boldsymbol{h}$ is column full rank, similarly to the method of proving that $\boldsymbol{R}$ is positive semidefinite in (32), we can also show that $\boldsymbol{I} - (\boldsymbol{R}\nabla\boldsymbol{h})\left[(\boldsymbol{R}\nabla\boldsymbol{h})^T(\boldsymbol{R}\nabla\boldsymbol{h})\right]^{-1}(\boldsymbol{R}\nabla\boldsymbol{h})^T$ is positive semidefinite. Thus, $\boldsymbol{Q}_1$ is positive semidefinite.

We can prove that $\boldsymbol{Q}$ is positive semidefinite by showing $\boldsymbol{y}^T(\boldsymbol{Q} - \boldsymbol{Q}_1)\boldsymbol{y} \geq 0$ for every $\boldsymbol{y} \in \mathcal{R}^n$. (A detailed proof of this conclusion can be found in the supplementary material.) □

Proof of Proposition 3: Define $\boldsymbol{k} \triangleq \left(\nabla\boldsymbol{g}^T\nabla\boldsymbol{g}\right)^{-1}\nabla\boldsymbol{g}^T(\boldsymbol{F}_{\boldsymbol{\theta}} + \nabla\boldsymbol{h}\boldsymbol{\lambda})$, where $\nabla\boldsymbol{g}^T\nabla\boldsymbol{g}$ is invertible from Lemma 2(1). Then $L(\boldsymbol{\theta},\boldsymbol{\lambda},\boldsymbol{\nu};\boldsymbol{y},\boldsymbol{u})$ can be rewritten as

$$\begin{aligned}
&\|\boldsymbol{F}_{\boldsymbol{\theta}} + \nabla\boldsymbol{h}\boldsymbol{\lambda} + \nabla\boldsymbol{g}\boldsymbol{\nu}\|_2^2 + \|\boldsymbol{H}\boldsymbol{\lambda}\|_2^2 + \|\boldsymbol{g}\|_2^2\\
=&\boldsymbol{\nu}^T\nabla\boldsymbol{g}^T\nabla\boldsymbol{g}\boldsymbol{\nu} + 2(\boldsymbol{F}_{\boldsymbol{\theta}} + \nabla\boldsymbol{h}\boldsymbol{\lambda})^T\nabla\boldsymbol{g}\boldsymbol{\nu}\\
&+\boldsymbol{F}_{\boldsymbol{\theta}}^T\boldsymbol{F}_{\boldsymbol{\theta}} + 2\boldsymbol{\lambda}^T\nabla\boldsymbol{h}^T\boldsymbol{F}_{\boldsymbol{\theta}} + \boldsymbol{\lambda}^T\nabla\boldsymbol{h}^T\nabla\boldsymbol{h}\boldsymbol{\lambda} + \|\boldsymbol{H}\boldsymbol{\lambda}\|_2^2 + \|\boldsymbol{g}\|_2^2\\
=&(\boldsymbol{\nu} + \boldsymbol{k})^T\nabla\boldsymbol{g}^T\nabla\boldsymbol{g}(\boldsymbol{\nu} + \boldsymbol{k}) - \boldsymbol{k}^T\nabla\boldsymbol{g}^T\nabla\boldsymbol{g}\boldsymbol{k}\\
&+\boldsymbol{\lambda}^T(\boldsymbol{H}^2 + \nabla\boldsymbol{h}^T\nabla\boldsymbol{h})\boldsymbol{\lambda} + 2\boldsymbol{\lambda}^T\nabla\boldsymbol{h}^T\boldsymbol{F}_{\boldsymbol{\theta}} + \boldsymbol{F}_{\boldsymbol{\theta}}^T\boldsymbol{F}_{\boldsymbol{\theta}} + \boldsymbol{g}^T\boldsymbol{g}.
\end{aligned} \tag{34}$$

Note that

$$-\boldsymbol{k}^T\nabla\boldsymbol{g}^T\nabla\boldsymbol{g}\boldsymbol{k} \tag{35}$$
$$= -(\boldsymbol{F}_{\boldsymbol{\theta}} + \nabla\boldsymbol{h}\boldsymbol{\lambda})^T\nabla\boldsymbol{g}(\nabla\boldsymbol{g}^T\nabla\boldsymbol{g})^{-1}\nabla\boldsymbol{g}^T(\boldsymbol{F}_{\boldsymbol{\theta}} + \nabla\boldsymbol{h}\boldsymbol{\lambda})$$
$$\overset{(29)}{=} (\boldsymbol{F}_{\boldsymbol{\theta}} + \nabla\boldsymbol{h}\boldsymbol{\lambda})^T(\boldsymbol{R} - \boldsymbol{I})(\boldsymbol{F}_{\boldsymbol{\theta}} + \nabla\boldsymbol{h}\boldsymbol{\lambda}) \tag{36}$$
$$\begin{aligned}
&= \boldsymbol{F}_{\boldsymbol{\theta}}^T(\boldsymbol{R} - \boldsymbol{I})\boldsymbol{F}_{\boldsymbol{\theta}} + 2\boldsymbol{\lambda}^T\nabla\boldsymbol{h}^T(\boldsymbol{R} - \boldsymbol{I})\boldsymbol{F}_{\boldsymbol{\theta}}\\
&\quad+\boldsymbol{\lambda}^T\nabla\boldsymbol{h}^T(\boldsymbol{R} - \boldsymbol{I})\nabla\boldsymbol{h}\boldsymbol{\lambda}.
\end{aligned} \tag{37}$$

Define $\boldsymbol{P} \triangleq \nabla\boldsymbol{h}^T\boldsymbol{R}\nabla\boldsymbol{h} + \boldsymbol{H}^2$, which is invertible from Lemma 2(2). Then by substituting (35) into (34), we obtain

$$\begin{aligned}
&L(\boldsymbol{\theta},\boldsymbol{\lambda},\boldsymbol{\nu};\boldsymbol{y},\boldsymbol{u}) = (\boldsymbol{\nu} + \boldsymbol{k})^T\nabla\boldsymbol{g}^T\nabla\boldsymbol{g}(\boldsymbol{\nu} + \boldsymbol{k}) + \boldsymbol{g}^T\boldsymbol{g}\\
&\quad+\boldsymbol{\lambda}^T\boldsymbol{P}\boldsymbol{\lambda} + 2\boldsymbol{\lambda}^T\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}} + \boldsymbol{F}_{\boldsymbol{\theta}}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}}\\
&\overset{(30)}{=}\Big[\boldsymbol{\lambda} + \boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}}\Big]^T\boldsymbol{P}\Big[\boldsymbol{\lambda} + \boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}}\Big]\\
&\quad+(\boldsymbol{\nu} + \boldsymbol{k})^T\nabla\boldsymbol{g}^T\nabla\boldsymbol{g}(\boldsymbol{\nu} + \boldsymbol{k}) + \boldsymbol{F}_{\boldsymbol{\theta}}^T\boldsymbol{Q}\boldsymbol{F}_{\boldsymbol{\theta}} + \boldsymbol{g}^T\boldsymbol{g}.
\end{aligned} \tag{38}$$

Thus,

$$\begin{aligned}
&l(\boldsymbol{\theta};\boldsymbol{y},\boldsymbol{u}) = \min_{\boldsymbol{\lambda}\geq\boldsymbol{0},\boldsymbol{\nu}} L(\boldsymbol{\theta},\boldsymbol{\lambda},\boldsymbol{\nu};\boldsymbol{y},\boldsymbol{u})\\
&\overset{(38)}{=}\min_{\boldsymbol{\lambda}\geq\boldsymbol{0},\boldsymbol{\nu}}\Big[\boldsymbol{\lambda} + \boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}}\Big]^T\boldsymbol{P}\Big[\boldsymbol{\lambda} + \boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}}\Big]\\
&\quad+(\boldsymbol{\nu} + \boldsymbol{k})^T\nabla\boldsymbol{g}^T\nabla\boldsymbol{g}(\boldsymbol{\nu} + \boldsymbol{k}) + \boldsymbol{F}_{\boldsymbol{\theta}}^T\boldsymbol{Q}\boldsymbol{F}_{\boldsymbol{\theta}} + \boldsymbol{g}^T\boldsymbol{g}\\
&\overset{\boldsymbol{v}=-\boldsymbol{k}}{=}\min_{\boldsymbol{\lambda}\geq\boldsymbol{0}}\Big[\boldsymbol{\lambda} + \boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}}\Big]^T\boldsymbol{P}\Big[\boldsymbol{\lambda} + \boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}}\Big]\\
&\quad+\boldsymbol{F}_{\boldsymbol{\theta}}^T\boldsymbol{Q}\boldsymbol{F}_{\boldsymbol{\theta}} + \boldsymbol{g}^T\boldsymbol{g}.
\end{aligned}$$

For every $\theta_1, \theta_2 \in \Theta$ and every $\alpha, \beta \geq 0, \alpha+\beta = 1$, let $\boldsymbol{\lambda}^*_{\alpha\boldsymbol{\theta}_1+\beta\boldsymbol{\theta}_2}$, $\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_1}$ and $\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_2}$ be the optimal solutions of the optimization problem

(22) corresponding to $\alpha\boldsymbol{\theta}_1+\beta\boldsymbol{\theta}_2, \boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, respectively. Then

$$\begin{aligned}
&l(\alpha\boldsymbol{\theta}_1+\beta\boldsymbol{\theta}_2;\boldsymbol{y},\boldsymbol{u}) = \left[\boldsymbol{\lambda}^*_{\alpha\boldsymbol{\theta}_1+\beta\boldsymbol{\theta}_2}+\boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\alpha\boldsymbol{\theta}_1+\beta\boldsymbol{\theta}_2}\right]^T\\
&\cdot\boldsymbol{P}\left[\boldsymbol{\lambda}^*_{\alpha\boldsymbol{\theta}_1+\beta\boldsymbol{\theta}_2}+\boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\alpha\boldsymbol{\theta}_1+\beta\boldsymbol{\theta}_2}\right]\\
&+\boldsymbol{F}^T_{\alpha\boldsymbol{\theta}_1+\beta\boldsymbol{\theta}_2}\boldsymbol{Q}\boldsymbol{F}_{\alpha\boldsymbol{\theta}_1+\beta\boldsymbol{\theta}_2}+\boldsymbol{g}^T\boldsymbol{g}\\
\le&\left[\alpha\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_1}+\beta\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_2}+\boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\alpha\boldsymbol{\theta}_1+\beta\boldsymbol{\theta}_2}\right]^T\\
&\cdot\boldsymbol{P}\left[\alpha\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_1}+\beta\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_2}+\boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\alpha\boldsymbol{\theta}_1+\beta\boldsymbol{\theta}_2}\right]\\
&+\boldsymbol{F}^T_{\alpha\boldsymbol{\theta}_1+\beta\boldsymbol{\theta}_2}\boldsymbol{Q}\boldsymbol{F}_{\alpha\boldsymbol{\theta}_1+\beta\boldsymbol{\theta}_2}+\boldsymbol{g}^T\boldsymbol{g}\\
=&\left[\alpha(\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_1}+\boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}_1})+\beta(\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_2}+\boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}_2})\right]^T\\
&\cdot\boldsymbol{P}\left[\alpha(\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_1}+\boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}_1})+\beta(\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_2}+\boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}_2})\right]\\
&+(\alpha\boldsymbol{F}_{\boldsymbol{\theta}_1}+\beta\boldsymbol{F}_{\boldsymbol{\theta}_2})^T\boldsymbol{Q}(\alpha\boldsymbol{F}_{\boldsymbol{\theta}_1}+\beta\boldsymbol{F}_{\boldsymbol{\theta}_2})+\boldsymbol{g}^T\boldsymbol{g}.
\end{aligned}\tag{39}$$

where the first inequality holds because $\boldsymbol{\lambda}^*_{\alpha\boldsymbol{\theta}_1+\beta\boldsymbol{\theta}_2}$ is an optimal solution that minimizes the loss function and $\boldsymbol{\lambda}=\alpha\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_1}+\beta\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_2}$ does not make the function value smaller, and the last equality holds because $\boldsymbol{F}_{\alpha\boldsymbol{\theta}_1+\beta\boldsymbol{\theta}_2}=\alpha\boldsymbol{F}_{\boldsymbol{\theta}_1}+\beta\boldsymbol{F}_{\boldsymbol{\theta}_2}$ by (12). Therefore, for every $\boldsymbol{\theta}_1,\boldsymbol{\theta}_2\in\boldsymbol{\Theta}$ and every $\alpha,\beta\ge 0, \alpha+\beta=1$, we obtain

$$\begin{aligned}
&\alpha l(\boldsymbol{\theta}_1;\boldsymbol{y},\boldsymbol{u})+\beta l(\boldsymbol{\theta}_2;\boldsymbol{y},\boldsymbol{u})-l(\alpha\boldsymbol{\theta}_1+\beta\boldsymbol{\theta}_2;\boldsymbol{y},\boldsymbol{u})\\
\overset{(39)}{\ge}&\alpha(\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_1}+\boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}_1})^T\boldsymbol{P}(\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_1}+\boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}_1})\\
&+\beta(\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_2}+\boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}_2})^T\boldsymbol{P}(\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_2}+\boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}_2})\\
&-\left[\alpha(\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_1}+\boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}_1})+\beta(\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_2}+\boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}_2})\right]^T\\
&\cdot\boldsymbol{P}\left[\alpha(\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_1}+\boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}_1})+\beta(\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_2}+\boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}_2})\right]\\
&+\alpha\boldsymbol{F}^T_{\boldsymbol{\theta}_1}\boldsymbol{Q}\boldsymbol{F}_{\boldsymbol{\theta}_1}+\beta\boldsymbol{F}^T_{\boldsymbol{\theta}_2}\boldsymbol{Q}\boldsymbol{F}_{\boldsymbol{\theta}_2}\\
&-(\alpha\boldsymbol{F}_{\boldsymbol{\theta}_1}+\beta\boldsymbol{F}_{\boldsymbol{\theta}_2})^T\boldsymbol{Q}(\alpha\boldsymbol{F}_{\boldsymbol{\theta}_1}+\beta\boldsymbol{F}_{\boldsymbol{\theta}_2})\\
=&\alpha\beta(\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_1}+\boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}_1}-\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_2}-\boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}_2})^T\\
&\cdot\boldsymbol{P}(\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_1}+\boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}_1}-\boldsymbol{\lambda}^*_{\boldsymbol{\theta}_2}-\boldsymbol{P}^{-1}\nabla\boldsymbol{h}^T\boldsymbol{R}\boldsymbol{F}_{\boldsymbol{\theta}_2})\\
&+\alpha\beta(\boldsymbol{F}_{\boldsymbol{\theta}_1}-\boldsymbol{F}_{\boldsymbol{\theta}_2})^T\boldsymbol{Q}(\boldsymbol{F}_{\boldsymbol{\theta}_1}-\boldsymbol{F}_{\boldsymbol{\theta}_2})\ge 0
\end{aligned}$$

where the last inequality holds because $\boldsymbol{Q}$ and $\boldsymbol{P}$ are positive semidefinite matrices according to Lemma 2.

Therefore, the loss function is convex with respect to the parameter $\boldsymbol{\theta}$. □

References

[1] A. A. Kulkarni and U. V. Shanbhag, "Revisiting generalized Nash games and variational inequalities," *Journal of Optimization Theory and Applications*, vol. 154, pp. 175–186, 2012.

[2] P. T. Harker, "Generalized Nash games and quasi-variational inequalities," *European Journal of Operational Research*, vol. 54, no. 1, pp. 81–94, 1991.

[3] J. Contreras, M. Klusch, and J. Krawczyk, "Numerical solutions to Nash-Cournot equilibria in coupled constraint electricity markets," *IEEE Transactions on Power Systems*, vol. 19, no. 1, pp. 195–206, 2004.

[4] I. Abada, S. Gabriel, V. Briat, and O. Massol, "A generalized Nash–Cournot model for the northwestern European natural gas markets with a fuel substitution demand function: The GaMMES model," *Networks and Spatial Economics*, vol. 13, pp. 1–42, 2013.

[5] Z. Liu, J. Lei, P. Yi, and Y. Hong, "An interaction-fair semi-decentralized trajectory planner for connected and autonomous vehicles," *Autonomous Intelligent Systems*, vol. 5, no. 1, pp. 1–20, 2025.

[6] S. Le Cleac'h, M. Schwager, and Z. Manchester, "LUCIDGames: Online unscented inverse dynamic games for adaptive trajectory prediction and planning," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 5485–5492, 2021.

[7] S. Liang, S. Liu, Y. Hong, and J. Chen, "Distributed Nash equilibrium seeking with order-reduced dynamics based on consensus exact penalty," *Control Theory and Technology*, vol. 21, no. 3, pp. 363–373, 2023.

[8] S. L. Cleac'h, M. Schwager, and Z. Manchester, "Algames: A fast solver for constrained dynamic games," *arXiv preprint arXiv:1910.09713*, 2019.

[9] F. Facchinei and C. Kanzow, "Generalized Nash equilibrium problems," *Annals of Operations Research*, vol. 175, no. 1, pp. 177–211, 2010.

[10] J. Wu, W. Shen, F. Fang, and H. Xu, "Inverse game theory for Stackelberg games: the blessing of bounded rationality," *Advances in Neural Information Processing Systems*, vol. 35, pp. 32 186–32 198, 2022.

[11] S. Allen, S. A. Gabriel, and J. P. Dickerson, "Using inverse optimization to learn cost functions in generalized Nash games," *Computers & Operations Research*, vol. 142, p. 105721, 2022.

[12] Y. Gao, A. Peysakhovich, and C. Kroer, "Online market equilibrium with application to fair division," *Advances in Neural Information Processing Systems*, vol. 34, pp. 27 305–27 318, 2021.

[13] C. D. Sousa and R. Cortesao, "Physical feasibility of robot base inertial parameter identification: A linear matrix inequality approach," *The International Journal of Robotics Research*, vol. 33, no. 6, pp. 931–944, 2014.

[14] J. Na, X. Ren, and Y. Xia, "Adaptive parameter identification of linear siso systems with unknown time-delay," *Systems & Control Letters*, vol. 66, pp. 43–50, 2014.

[15] R. Ghanem and F. Romeo, "A wavelet-based approach for model and parameter identification of non-linear systems," *International Journal of Non-Linear Mechanics*, vol. 36, no. 5, pp. 835–859, 2001.

[16] A. W. Smyth, S. F. Masri, A. G. Chassiakos, and T. K. Caughey, "On-line parametric identification of mdof nonlinear hysteretic systems," *Journal of engineering mechanics*, vol. 125, no. 2, pp. 133–142, 1999.

[17] K. Cao and L. Xie, "Game-theoretic inverse reinforcement learning: A differential Pontryagin's maximum principle approach," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 11, pp. 9506–9513, 2023.

[18] T. L. Molloy, J. Inga, M. Flad, J. J. Ford, T. Perez, and S. Hohmann, "Inverse open-loop noncooperative differential games and inverse optimal control," *IEEE Transactions on Automatic Control*, vol. 65, no. 2, pp. 897–904, 2020.

[19] Y. Yu, J. Salfity, D. Fridovich-Keil, and U. Topcu, "Inverse matrix games with unique quantal response equilibrium," *IEEE Control Systems Letters*, vol. 7, pp. 643–648, 2023.

[20] X. Lin, P. A. Beling, and R. Cogill, "Multiagent inverse reinforcement learning for two-person zero-sum games," *IEEE Transactions on Games*, vol. 10, no. 1, pp. 56–68, 2018.

[21] Z. Zhang, Y. Li, and C. Yu, "Online inverse identification of noncooperative dynamic games," in *2021 IEEE International Conference on Unmanned Systems (ICUS)*, 2021, pp. 408–413.

[22] G. Debreu, "A social equilibrium existence theorem," *Proceedings of the national academy of sciences*, vol. 38, no. 10, pp. 886–893, 1952.

[23] Y. Zhu, W. Yu, W. Ren, G. Wen, and J. Gu, "Generalized Nash equilibrium seeking via continuous-time coordination dynamics over digraphs," *IEEE Transactions on Control of Network Systems*, vol. 8, no. 2, pp. 1023–1033, 2021.

[24] A. Fischer, M. Herrich, and K. Schönefeld, "Generalized Nash equilibrium problems-recent advances and challenges," *Pesquisa Operacional*, vol. 34, no. 3, pp. 521–558, 2014.

[25] Y. Dai, Y. Qi, L. Li, B. Wang, and H. Gao, "A dynamic pricing scheme for electric vehicle in photovoltaic charging station based on Stackelberg game considering user satisfaction," *Computers & Industrial Engineering*, vol. 154, p. 107117, 2021.

[26] A. A. Kulkarni and U. V. Shanbhag, "On the variational equilibrium as a refinement of the generalized Nash equilibrium," *Automatica*, vol. 48, no. 1, pp. 45–55, 2012.

[27] B. Kulis and P. L. Bartlett, "Implicit online learning," in *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, 2010, pp. 575–582.

[28] J. Nocedal and S. J. Wright, *Numerical optimization*. Springer, 1999.

[29] F. Lian, A. Chakrabortty, and A. Duel-Hallen, "Game-theoretic multi-agent control and network cost allocation under communication constraints," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 2, pp. 330–340, 2017.

[30] Y. Censor and S. A. Zenios, *Parallel optimization: Theory, algorithms, and applications*. Oxford University Press, 1997.

[31] S. W. Salant, "Imperfect competition in the international energy market: a computerized Nash-Cournot model," *Operations research*, vol. 30, no. 2, pp. 252–280, 1982.