Lower Bounds on the Best-Case Complexity of Solving a Class of Non-cooperative Games

Ehsan Nekouei * Tansu Alpcan * Girish N. Nair † Robin J. Evans *

* Dept. Electrical and Electronic Engineering, University of Melbourne, VIC 3010, Australia (e-mail: ehsan.nekouei, tansu.alpcan,girish.ronibje@unimelb.edu.au).

Abstract: This paper studies the complexity of solving the class $G$ of all $N$-player non-cooperative games with continuous action spaces that admit at least one Nash equilibrium (NE). We consider a distributed Nash seeking setting where agents communicate with a set of system nodes (SNs), over noisy communication channels, to obtain the required information for updating their actions. The complexity of solving games in the class $G$ is defined as the minimum number of iterations required to find a NE of any game in $G$ with $\epsilon$ accuracy. Using information-theoretic inequalities, we derive a lower bound on the complexity of solving the game class $G$ that depends on the Kolmogorov $2\epsilon$-capacity of the constraint set and the total capacity of the communication channels. We also derive a lower bound on the complexity of solving games in $G$, which depends on the volume and surface area of the constraint set.

1. INTRODUCTION

1.1 Motivation

Game theory offers a suite of analytical frameworks for investigating the interaction between rational decision-makers, hereafter called agents. In the past decade, game theory has found diverse applications across engineering disciplines ranging from power control in wireless networks to modeling the behavior of travelers in a transport system. The Nash equilibrium (NE) is the fundamental solution concept for non-cooperative games, in which a number of agents compete to maximize conflicting utility functions that are influenced by the action of others. At a NE, no agent benefits from a unilateral deviation from its NE strategy.

Finding a NE of a non-cooperative game is a fundamental research problem that lies at the heart of game theory literature. For non-cooperative games with continuous action spaces, various Nash seeking algorithms have been proposed in the literature, e.g., see Li and Basar (1987). In this paper, we investigate the intrinsic difficulty of finding a NE in such games. Using the notion of complexity from the convex optimization literature, and information-theoretic inequalities, we derive lower bounds on the minimum number of iterations required to find a NE within a desired accuracy, for any $N$-player, non-cooperative game in a given class.

1.2 Related Work

The work by Nemirovski and Yudin (1983) pioneers the investigation of complexity in convex optimization problems. In their model, an optimization algorithm sequentially queries the objective function of an optimization problem, and the oracle responds to the algorithm according to the queries and the objective function. They derived bounds on the minimum number of queries required to find the minimum/maximum of any function in a given function class. In Agarwal et al. (2009), lower bounds were derived on the complexity of convex optimization problems, under a stochastic first order oracle, for the class of Lipschitz functions. These results were extended to different function classes in Agarwal et al. (2012).

The paper Raginsky and Rakhlin (2011) considered an oracle-based model in which the algorithm only observes noisy versions of oracle’s response to its queries. Lower bounds were established on the complexity of convex optimization problems under first order as well as gradient-only oracles. In Jamieson et al. (2012), lower bounds were obtained on the complexity of convex optimization problems under stochastic zero-order stochastic oracle. The paper Duchi et al. (2012) studied the complexity of convex optimization problems under a zero-order stochastic oracle in which the optimization algorithm submits two queries at each iteration and the oracle responds to both queries. These results were extended to the case in which the algorithm makes queries about multiple points at each iteration in Duchi et al. (2015). In Singer and Vondrak (2015), the complexity of convex optimization problems was studied under an erroneous oracle model wherein the oracle’s responses to queries are subject to absolute/relative errors.

1.3 Contributions

The current paper studies the complexity of solving the game class $G$, which consists of all $N$-player non-cooperative games with a continuous action space such that all the games in $G$ admit at least one Nash equilibrium (NE). In our model, each agent communicates with a subset of system nodes (SNs) to obtain the required information for updating its action. The communication between agents and SNs is subject to communication noise, i.e., SNs will receive noisy versions of agents’ actions, and agents will receive noisy information from SNs.

We derive lower bounds on minimum number of iterations required to find a NE of any game in $G$. Our results indicate that the complexity of solving non-cooperative games to an $\epsilon$ accuracy depends on the Kolmogorov $2\epsilon$-capacity of the constraint set and the total capacity of communication channels.

* This work was supported by the Australian Research Council’s Discovery Projects funding scheme (DP140100819).
from SNs to agents which convey utility-related information to agents. We also derive a lower bound on the complexity of solving the game class \( \mathcal{G} \) in terms of the volume and surface area of the constraint set.

This paper is organized as follows. Section 2 discusses our modeling assumptions and problem formulation. Section 3 discusses our main results along with their interpretations, and Section 4 concludes the paper. All the proofs are relegated to appendices to aid the flow of the paper.

2. SYSTEM MODEL

2.1 Basic Game-theoretic Definitions

Consider a non-cooperative game with \( N \) agents indexed over \( \mathcal{N} = \{1, \ldots, N\} \). Let \( x^i \) with \( i \in \mathcal{N} \), and \( x = [x^1, \ldots, x^N]^{\top} \) denote the action of the \( i \)th agent and the collection of all agents’ actions, respectively. The utility function of the \( i \)th agent is denoted by \( u_i(x^i, x^{-i}) \) where \( x^{-i} \) is the vector of other agents’ actions that affect the \( i \)th agent’s utility. The utility function of the \( i \)th agent quantifies the desirability of any point in the action space for the \( i \)th agent. The actions of agents are limited by \( l \) convex constraints denoted by \( C_i(x) \leq 0 \) where \( C_i(.) = [C_i^1(.) \ldots C_i^l(.)]^{\top} \) is a mapping from \( \mathbb{R}^N \) to \( \mathbb{R}^l \). The set of constraint functions is indexed over \( \mathcal{L} = \{1, \ldots, l\} \). Let \( \mathcal{S} \) denote the action space of agents, i.e., \( \mathcal{S} = \{x \in \mathbb{R}^N : s.t. C_i(x) \leq 0\} \). We assume that \( \mathcal{S} \) is a compact and convex subset of \( \mathbb{R}^N \).

In non-cooperative games, each agent is interested in maximizing its own utility function, irrespective of other agents. Since the maximizers of utility functions of agents do not necessarily coincide with each other, a trade-off is required. In this paper, the Nash equilibrium (NE) is considered as the solution concept of the non-cooperative game among agents. Let \( x_{NE} \in \mathcal{S} \) be the NE of the game among agents. Then, at the NE, no agent has incentive to unilaterally deviate its action from its NE strategy, i.e.,

\[
x_{NE} = \arg \max_{x^i \in \mathcal{S}} u_i(x^i, x_{NE}^{-i}), \forall i \in \mathcal{N},
\]

where \( x_{NE}^{-i} \) is the collection of NE strategies of agents which are coupled with the \( i \)th agent through constraints, and \( \mathcal{S}_{(x_{NE}^{-i})} \) is the set of possible actions of the \( i \)th agent given \( x_{NE}^{-i} \). The vector of all utility functions is denoted by \( U(X) = [u_1(x), \ldots, u_N(x)]^{\top} \). Let \( \mathcal{F} \) denote the class of functions from \( \mathbb{R}^N \) to \( \mathbb{R}^N \) such that any \( N \)-player non-cooperative game with the constraint set \( \mathcal{S} \) and utility function vector in \( \mathcal{F} \) admits at least one NE. By the class of non-cooperative games \( \mathcal{G} = \langle \mathcal{N}, \mathcal{S}, \mathcal{F} \rangle \), we mean the set of all games with \( N \) agents, the action space \( \mathcal{S} \), and the utility function vector in \( \mathcal{F} \), i.e., \( U(.) \in \mathcal{F} \).

2.2 Communication Model

In this paper, we consider a distributed Nash seeking set-up wherein, at each time-step, agents communicate with a set of system nodes (SNs) to obtain the required utility/constraint related information for updating their actions. Each SN computes utility-related information for a set of agents. It also evaluates a subset of constraint functions based on the received actions of agents. The number of SNs, denoted by \( K \), is less than or equal to the number of constraints plus the number of agents. Let \( \mathcal{N}_j \) denote the set of agents which rely on the \( j \)th SN for utility-related information, and let \( \mathcal{L}_j \) represent the set of constraints which are evaluated by the \( j \)th SN. The SNs are designed such that we have \( \mathcal{N} = \bigcup_j \mathcal{N}_j \) and \( \mathcal{L} = \bigcup_j \mathcal{L}_j \). An agent transmits its action to the \( j \)th SN if its action appears in the objective function of an agent in \( \mathcal{N}_j \), or if its action affects a constraint function in \( \mathcal{L}_j \). The \( j \)th SN transmits utility-related information to the agents in \( \mathcal{N}_j \). It also transmits the value of each constraint function in \( \mathcal{L}_j \) to any agent whose action affects that constraint. The communication between agents and SNs is performed over noisy communication channels, i.e., agents receive noisy information from SNs, and SNs receive noisy versions of agents’ actions.

The communication topology between agents and SNs follows from a bipartite digraph in which the agents and the SNs form two disjoint sets of vertices. There exists a directed edge, in the communication graph, from the \( i \)th agent to the \( j \)th SN if the \( j \)th SN requires the \( i \)th agent’s action for its computations. Similarly, there exists a directed edge from \( j \)th SN to the \( i \)th agent if the \( j \)th SN provides utility-related information to the \( i \)th agent or evaluates a constraint which involves the \( i \)th agent’s action. Fig. 1 shows a pictorial representation of the communication topology with \( K \) SNs and \( N \) agents. Agents communicate with SNs using frequency division multiplexing (FDM) or time division multiplexing (TDM) schemes, i.e., each agent broadcasts its action to its neighboring SNs in the communication graph using a dedicated time or frequency band. Similarly, SNs communicate with agents via FDM or TDM communication schemes.

2.3 Nash Seeking Iterative Algorithms

Let \( X_{i,k-1} \) denote the history of the \( i \)th agent’s actions from time 1 to \( k-1 \). We use \( Y_{i,k-1} \) and \( C_{i,k-1} \) to denote all the information received by the \( i \)th agent from SNs up to time \( k-1 \), where \( Y_{i,k-1} \) and \( C_{i,k-1} \) contain information regarding the \( i \)th agent’s utility function and the constraint functions, respectively. The \( i \)th agent at time \( k \) updates its action according to the update rule \( x_{i,k} = A_i^k \left( X_{i,k-1}, Y_{i,k-1}, C_{i,k-1}\right) \). Next, it transmits \( x_{i,k} \) to the its neighboring SNs in the communication graph.

Let \( X_{1,k-1} \) and \( X_{i,k} \) denote all the information received by the \( j \)th SN at time \( k \) and up to time \( k \), respectively. At time \( k \), the \( j \)th SN computes \( y_{k} = C_j^k \left( X_{i,k} \right) \) for \( i \in \mathcal{N}_j \). Where \( X_{i,k} \) denotes those received actions that affect the \( j \)th agent’s utility function. Then, it transmits \( y_{k} \) to the \( i \)th agent for \( i \in \mathcal{N}_j \). For example, \( y_{k} \) can be equal to
\[ T_{\epsilon,\delta}(G, \mathcal{O}) = \inf \left\{ T \in \mathbb{N} : \exists A \quad \text{s.t.} \quad \sup_{U(\cdot) \in F^1} \inf \Pr \left( \left\| x_{\text{NE},U(\cdot)} - A_{T+1} \left( X_{1:T}, \hat{Y}_{1:T}, \hat{C}_{1:T} \right) \right\| \geq \epsilon \right) \leq \delta \right\} \] (1)

\[ y_k^i = \frac{\partial^2}{\partial x^2} u_i \left( x^i, x^{-i} \right) \bigg|_{x^i_k}. \] The \( j \)th SN also computes

\[ C_p \left( [\hat{X}_k^i]_p \right) \text{ for } p \in L_j \] where \([\hat{X}_k^i]_p\) is the collection of received actions at time \( k \) which affect \( C_p(\cdot) \). Finally, the \( j \)th SN transmits \( y_k^i \) to the \( j \)th agent. It also broadcasts \( C_p \left( [\hat{X}_k^i]_p \right) \) to the agents which their actions affect \( C_p(\cdot) \) for all \( p \in L_j \). The received utility-related information by the \( j \)th agent at time \( k \) is denoted by \( y_k^i \). Also, the collection of received constraint-related information by the \( j \)th agent at time \( k \) is denoted by \( \hat{C}_k^i \).

The \( k \)th step of the algorithm \( A \) is denoted by

\[ A_k \left( X_{1:k-1}, \hat{Y}_{1:k-1}, \hat{C}_{1:k-1} \right) = \left\{ A_k \left( X_{i:k-1}, \hat{Y}_{i:k-1}, \hat{C}_{i:k-1} \right) \right\} _i \]

where \( X_{1:k-1} = \left\{ X_{i:k-1} \right\} _i \), \( \hat{Y}_{1:k-1} = \left\{ \hat{Y}_{i:k-1} \right\} _i \), and \( \hat{C}_{1:k-1} = \left\{ \hat{C}_{i:k-1} \right\} _i \). We refer to

\[ A = \left\{ A_k \left( X_{1:k-1}, \hat{Y}_{1:k-1}, \hat{C}_{1:k-1} \right) \right\} _k \]

as the Nash seeking algorithm \( A \). We also refer to \( \mathcal{O} = \left\{ \mathcal{O}_k \left( \left[ X_{1:k-1}^i \right]_i, U(\cdot) \right) \right\} _{k,i} \) as the computation model of SNs.

### 2.4 The Complexity Criterion

Consider the class of games \( G \) and the computation model \( \mathcal{O} \). Then, the \((\epsilon, \delta)\)-complexity of solving the class of games \( G \) with the computation model \( \mathcal{O} \), denoted by \( T_{\epsilon,\delta}^*(G, \mathcal{O}) \), is defined in (1) where \( x_{\text{NE},U(\cdot)} \) is a NE of the non-cooperative game with the utility functions given by \( U(\cdot) \in F \). According to (1), the \((\epsilon, \delta)\)-complexity of solving the class of games \( G \) with the computation model \( \mathcal{O} \) is defined as the smallest positive integer \( T \) for which there exists an algorithm \( A \) such that, for any game in \( G \), the probability of \( \epsilon \) deviation of the algorithm’s output at time \( T + 1 \) from a NE of the game is less than \( \delta \).

Definition 1. A subset of \( S \) is 2\( \epsilon \)-distinguishable if the distance between any two of its distinct points is more than \( 2\epsilon \).

Next theorem establishes a lower bound on \( T_{\epsilon,\delta}^*(G, \mathcal{O}) \).

**Theorem 1.** The complexity of the class of \( N \)-player non-cooperative games \( G \) with the continuous action space \( S \) can be bounded as

\[ T_{\epsilon,\delta}^*(G, \mathcal{O}) \geq \frac{(1 - \delta) \log M_{2\epsilon}(S) - 1}{C_{\text{down}}} \] (2)

where \( C_{\text{down}} \) is the total capacity of channels transmitting \( y_k^i \) to agents, and \( \log M_{2\epsilon}(S) \) is the Kolmogorov 2\( \epsilon \)-capacity of the action space \( S \). Kolmogorov and Tikhomirov (1959) and \( M_{2\epsilon}(S) \) is the cardinality of maximal, 2\( \epsilon \)-distinguishable subsets of \( S \).

**Proof.** The complete proof of this theorem is relegated to Appendix A to aid the flow of the paper. Here, we state the main steps which are used to establish the lower bound in Theorem 1.

1. Firstly, a lower bound on \( T_{\epsilon,\delta}^*(G, \mathcal{O}) \) is obtained by restricting the utility functions of agents to a finite subset of \( F \), called \( F' \), such that the NE corresponding to two distinct elements of \( F' \) are at least \( 2\epsilon \) apart (see subsection A.1 in Appendix A for more details).
2. Secondly, for a given algorithm \( A \), we construct a hypothesis test and we show that the deviation probability of algorithm \( A \) from a NE at time \( T + 1 \) can be lower bounded by the error probability of the constructed hypothesis test when the utility functions are selected from \( F' \) (see subsection A.2 in Appendix A for more details).
3. Thirdly, the generalized Fano inequality is used to obtain a lower bound on the error probability of the constructed hypothesis test (see subsection A.2 in Appendix A for more details).
4. Finally, information theoretic inequalities are used to obtain an upper bound on the mutual information term which appears in the generalized Fano inequality.

Theorem 1 establishes an algorithm-independent lower bound on the complexity of solving non-cooperative games. According to this theorem, \( T_{\epsilon,\delta}^*(G, \mathcal{O}) \) is lower bounded by the ratio of Kolmogorov 2\( \epsilon \)-capacity of the action space \( S \) to the total Shannon capacity of downlink channels. Note that Kolmogorov 2\( \epsilon \)-capacity can be interpreted as a measure of agents’ ambiguity about their NE strategies. Thus, as \( \log M_{2\epsilon}(S) \) becomes large, \( T_{\epsilon,\delta}^*(G, \mathcal{O}) \) is expected to increase since agents have to search in a bigger space to find their NE strategies. Based on Theorem 1, \( C_{\text{down}} \) has a reverse impact on \( T_{\epsilon,\delta}^*(G, \mathcal{O}) \). Note that \( C_{\text{down}} \) is an indication of the information transmission quality from SNs to agents. That is, as \( C_{\text{down}} \) decreases, agents will receive noisier information regarding their utility functions compared with a large value of \( C_{\text{down}} \).

Theorem 1 depends on the 2\( \epsilon \)-capacity of the constraint set \( \mathcal{S} \) which is usually hard to compute unless the action space of agents is restricted to special geometries. Next corollary

---

In this section, we present two lower bounds on \( T_{\epsilon,\delta}^*(G, \mathcal{O}) \). To this end, we first define the notion of 2\( \epsilon \)-distinguishable sets Kolmogorov and Tikhomirov (1959) which is used in the statement of Theorem 1.
establishes a lower bound on \( T_{c, \delta}^* (G, O) \) by lower bounding \( M_{2c} (S) \) using a result from lattice theory.

**Corollary 1.** The complexity of solving the class of \( N \)-player non-cooperative games \( G \) with continuous action space \( S \) can be lower bounded as

\[
T_{c, \delta}^* (G, O) \geq \frac{(1 - \delta)(N \log \frac{1}{2\epsilon} + \log (V(S) - cP(S))) - 1}{C_{\text{down}}}
\]

where \( V(S) \) and \( P(S) \) are the volume and the surface area of action space of agents, respectively.

**Proof.** Please see Appendix B.

According to Corollary 1, the lower bound on the complexity of solving non-cooperative games increases as the volume of the action space of agents becomes large. Also, for a given surface area of action space of agents, \( i.e., P(S) \), the volume of action space of agents can be upper bounded using the isoperimetric inequality for convex bodies (Gardner 2002) as follows:

\[
V(S) \leq \frac{V(B)}{(P(B))^\frac{N}{N-1}} P(S)^\frac{N}{N-1} \tag{3}
\]

where \( B \) is the closed unit ball in \( N \)-dimensional Euclidean space \( \mathbb{R}^N \). Note that the equality in (3) is achieved if and only if \( S \) is a ball in \( \mathbb{R}^N \) (Gardner 2002). Thus, for a given surface area of action space of agents \( P(S) \), the lower bound on the complexity of solving games in the class \( G \) increases as the action space of agents becomes closer to a ball in \( \mathbb{R}^N \) with the volume \( \frac{V(B)}{(P(B))^\frac{N}{N-1}} P(S)^\frac{N}{N-1} \).

Based on Corollary 1, the lower bound on \( T_{c, \delta}^* (G, O) \) increases at least linearly with the number of agents. This is due to the fact that the amount of uncertainty about the NE increases as the number of agents becomes large. Recall that \( \log M_{2c} (S) \) is a quantitative indicator of agents’ ambiguity regarding the NE, and it is a non-decreasing function of the number of agents. Furthermore, \( \epsilon \) has a logarithmic effect on \( T_{c, \delta}^* (G, O) \), i.e., the complexity of solving the class of games \( G \) increases according to \( \Omega (\log \frac{1}{\epsilon}) \) as \( \epsilon \) becomes small.

### 4. CONCLUSION

In this paper, we have studied the complexity of solving the game class \( G \) which is comprised of all \( N \)-player non-cooperative games with a continuous action space such that any game in \( G \) admits at least a Nash equilibrium. In our set-up, agents obtain the required information for updating their actions by communicating with a set of SNs over noisy communication channels. We derived a lower bound on the complexity of solving the game class \( G \) to an \( \epsilon \) accuracy which depends on the Kolmogorov 2\( \epsilon \)-capacity of the constraint set and the total capacity of the communication channels which convey utility-related information to agents. We also derived a lower bound on the complexity of solving the game class \( G \) which depends on the volume and surface area of the constraint set.

### Appendix A. PROOF OF THEOREM 1

#### A.1 Restricting the Class of Utility Functions

The first step in deriving the lower bound on \( T_{c, \delta}^* (G, O) \) is to restrict our analysis to an appropriately chosen, finite subset of \( F \). Let \( F' \) be a finite subset of \( F \) such that any \( N \)-player non-cooperative game with the action space \( S \) and the utility functions in \( F' \) admits a unique Nash equilibrium, and let \( |F'| \) denote the size of \( F' \). The set \( F' \) can be represented as an indexed family of utility functions in \( F \), i.e., \( F' = \{ U_m (\cdot) \in F, m = 1, \ldots, |F'| \} \). Let \( x_{NE,m} \) and \( x_{NE,m} \) denote the NE strategy of the \( i \)th agent and the vector of NE strategies of all agents, respectively, when the utility functions of agents are specified by \( U_m (\cdot) \in F' \). The set of Nash equilibria due to different utility functions in \( F' \) is denoted by \( \{ x_{NE,m} \} \) and \( T_{c, \delta}^* (G, O) \) is lower bounded by (A.1).

The number of elements in the set \( F' \) plays an important role in our final results, \( e.g., \) see Theorem 1. Here, we impose a structure on the set \( F' \) by assuming that the Euclidean distance between any two Nash equilibria due to distinct elements of \( F' \) is more than \( 2\epsilon \). That is, we have \( |x_{NE,m} - x_{NE,n}| > 2\epsilon \) for \( U_m (\cdot), U_n (\cdot) \in F' \) and \( m \neq n \). This property allows us to lower bound the \( \epsilon \) deviation probability of any algorithm from the NE by the error probability of a hypothesis test which we construct in the next subsection (see Lemma 1 in this appendix and its proof for more details).

Now, we construct the set \( F' \) as follows. Let \( S^* \) be a 2\( \epsilon \)-distinguishable subset of \( S \) with \( |S^*| = M_{2c} (S) \) (see Definition 1 for more information about 2\( \epsilon \)-distinguishable subsets of \( S \)). The set \( F' \) is constructed such that the set of Nash equilibria of elements of \( F' \) coincides with the set \( S^* \), \( i.e., S^* = \{ x_{NE,m}, m = 1, \ldots, |F'| \} \). To this end, let \( \lambda_i = [a_{ij}]_{i,j} \) be a symmetric, negative definite \( N \)-by-\( N \) matrix. Also, let \( b_m = A x_{NE,m} \) where \( x_{NE,m} \) is a member of \( S^* \). Then, \( U_m (x) \in F' \) is defined as \( U_m (x) = (u^n_m (x)) \), wherein \( u^n_m (x) = 2 \frac{\epsilon}{\sqrt{2}} (x')^2 + \frac{\epsilon}{2} \sum_{j \neq i} a_{ij} x_j - b_m \), is the utility function of the \( i \)th agent, and \( b_m \) is the \( i \)th element of \( b_m \). It is straightforward to verify that \( x_{NE,m} \) is a NE of the game among agents when the utility functions of agents are given by \( U_m (x) \).

Here, we make the technical assumption that, for each \( U_m (x) \), the game among agents admits a unique NE. This condition can be satisfied by imposing more structure on the constraint set \( S \), \( e.g., \) see Rosen (1965). The set \( F' \) will be the collection of the constructed utility functions \( U_m (x) \)s. Clearly, we have \( |F'| = M_{2c} (S) \).

#### A.2 A Genie-aided Hypothesis Test

Given any algorithm \( A \), in this subsection, we construct a genie-aided hypothesis test which operates based on \( A \). Then, we show that if the utility functions of agents are selected from \( F' \), then the probability of \( \epsilon \) deviation from the Nash equilibrium for each algorithm is lower bounded by the error probability of its corresponding hypothesis test. Under the proposed test, first, a genie selects a utility function from the set \( F' \) uniformly at random. We use \( U_M (\cdot) \in F' \) to denote the selected utility function where \( M \) is a random variable uniformly distributed over the set \( \{ 1, \ldots, |F'| \} \). Then, SNs are notified about the \( U_M (\cdot) \). At time \( k \in \{ 1, \ldots, T \} \), the \( i \)th agent updates its action using the algorithm \( A \) according to \( x^t_i = A^k_{i'} \left( X^t_{i'1:k-1}, Y^t_{i'1:k-1}, \hat{C}^t_{i'1:k-1} \right) \). At time \( T + 1 \), the genie will estimate the NE among agents, using \( \left( X^t_i \right)_{1:T}, Y^t_i \right)_{1:T}, \hat{C}^t_i \right)_{1:T} \), according to the following decision rule:
\[ T_{\epsilon, \delta}(G, O) \geq \inf \{ T \in \mathbb{N} : \exists A \text{ s.t.} \sup_{m \in \{1, \ldots, |F'|\}} \Pr \left( \|x_{NE,m} - A_{T+1} (X_{T+1}, \hat{Y}_{T+1}, \hat{C}_{1:T})\| \geq \epsilon \right) \leq \delta \} \quad (A.1) \]

where \( \hat{x}_{NE} \) is the estimated NE strategies by genie. Let \( x_{NE,M} \) be the vector of NE strategies of agents when their utility functions are given by \( U_M (\cdot) \). The error event \( E_A \) is defined as \( E_A = \{ x_{NE,M} \neq \hat{x}_{NE} \} \). Finally, the genie declares an error if the event \( E_A \) occurs.

**Lemma 1.** Let \( \Pr (E_A) \) denote the error probability under the proposed genie-aided hypothesis test. Then,
\[
\Pr (E_A) \leq \sup_{m \in \{1, \ldots, |F'|\}} \Pr \left( \|x_{NE,m} - A_{T+1} (X_{T+1}, \hat{Y}_{T+1}, \hat{C}_{1:T})\| \geq \epsilon \right) \]
where \( x_{NE,m} \) is the NE of the game among agents with the utility function vector \( U_m (\cdot) \).

**Proof.** We show that the error event \( E_A \) implies
\[
\|x_{NE,M} - A_{T+1} (X_{T+1}, Y_{T+1}, \hat{C}_{1:T})\| > \epsilon
\]
by contraposition. That is, we show if the following inequality holds
\[
\|x_{NE,m} - A_{T+1} (X_{T+1}, Y_{T+1}, \hat{C}_{1:T})\| \leq \epsilon, \quad (A.3)
\]
then, we have \( \hat{x}_{NE} = x_{NE,M} \). Assume that the inequality (A.3) holds. For \( x_{NE,m} \neq x_{NE,M} \), we have
\[
2\epsilon < \|x_{NE,m} - x_{NE,M}\| < \|x_{NE,m} - A_{T+1} (X_{T+1}, Y_{T+1}, \hat{C}_{1:T})\| + \|x_{NE,M} - A_{T+1} (X_{T+1}, Y_{T+1}, \hat{C}_{1:T})\| + \epsilon
\]
where \((a)\) follows from the construction of the set \( F' \). Thus, \( x_{NE,m} \) cannot be the solution of optimization problem (A.2). Therefore, we have
\[
\Pr (E_A) \leq \Pr \left( \|x_{NE,m} - A_{T+1} (X_{T+1}, Y_{T+1}, \hat{C}_{1:T})\| \geq \epsilon \right) = \frac{1}{|F'|} \sum_{m} \Pr \left( \|x_{NE,m} - A_{T+1} (X_{T+1}, Y_{T+1}, \hat{C}_{1:T})\| \geq \epsilon \right) \leq \sup_{m \in \{1, \ldots, |F'|\}} \Pr \left( \|x_{NE,m} - A_{T+1} (X_{T+1}, Y_{T+1}, \hat{C}_{1:T})\| \geq \epsilon \right)
\]
which completes the proof.

Now, we apply the generalized Fano inequality to obtain a lower bound on \( \Pr (E_A) \).

**Lemma 2.** Let the random variable \( M \in \{1, \ldots, |F'|\} \) encode the choice of utility functions from the set \( F' \). Also, let the random variable \( M \in \{1, \ldots, |F'|\} \) encode the estimate of NE by the genie. Then, \( \Pr (E_A) \geq 1 - \frac{1 + |M| \delta}{\log |F'|} \).

**A.3 Applying information theoretic inequalities**

First note that \( M, (X_{T+1}, \hat{Y}_{T+1}, \hat{C}_{1:T}) \) and \( M \) form a Markov chain as follows: \( M \rightarrow (X_{T+1}, \hat{Y}_{T+1}, \hat{C}_{1:T}) \rightarrow \hat{M} \). Therefore, we have \( I \left[ M; \hat{M} \right] \leq I \left[ M; X_{T+1}, \hat{Y}_{T+1}, \hat{C}_{1:T} \right] \). Using the chain rule for mutual information, \( I \left[ M; X_{T+1}, \hat{Y}_{T+1}, \hat{C}_{1:T} \right] \) can be expanded as
\[
I \left[ M; X_{T+1}, \hat{Y}_{T+1}, \hat{C}_{1:T} \right] = \sum_{i=1}^{T} I \left[ M; X_i, \hat{Y}_i, \hat{C}_i \mid X_{i-1}, \hat{Y}_{i-1}, \hat{C}_{1:i-1} \right]
\]
where \( X_i = [x_i^1, \ldots, x_i^N]^T, x_i^j \) is the action of the \( j \)-th agent at time \( i \), \( \hat{Y}_i = [\hat{y}_i^1, \ldots, \hat{y}_i^N]^T, \hat{y}_i^j \) is the received utility-related information by the \( j \)-th agent at time \( i \), \( \hat{C}_i = [\hat{C}_i, \ldots, \hat{C}_i^N]^T \) and \( \hat{C}_i^j \) is the collection of constraint-related information received by the \( j \)-th agent at time \( i \).

We can expand \( I \left[ M; X_i, \hat{Y}_i, \hat{C}_i \mid X_{i-1}, \hat{Y}_{i-1}, \hat{C}_{1:i-1} \right] \) as
\[
I \left[ M; X_i, \hat{Y}_i, \hat{C}_i \mid X_{i-1}, \hat{Y}_{i-1}, \hat{C}_{1:i-1} \right] = I \left[ M; X_i \mid X_{i-1}, \hat{Y}_{i-1}, \hat{C}_{1:i-1} \right] + I \left[ M; \hat{Y}_i \mid X_{i}, \hat{Y}_{i-1}, \hat{C}_{1:i} \right]
\]
where \((a)\) follows from the facts that
\[
I \left[ M; X_i \mid X_{i-1}, \hat{Y}_{i-1}, \hat{C}_{1:i-1} \right] = 0
\]
and
\[
I \left[ M; \hat{C}_i \mid X_{i-1}, \hat{Y}_{i-1}, \hat{C}_{1:i-1} \right] = 0
\]
since \( M \rightarrow (X_{i-1}, \hat{Y}_{i-1}, \hat{C}_{1:i-1}) \rightarrow X_i \) and \( (M, X_{i-1}, \hat{Y}_{i-1}, \hat{C}_{1:i-1}) \rightarrow X_i \rightarrow \hat{C}_i \) as \( \hat{C}_i \) only depends on \( X_i \) and the communication noise at time \( i \). Thus, we have
\[
I \left[ M; X_{i+1}, \hat{Y}_{i+1}, \hat{C}_{i+1} \right] = \sum_{i=1}^{T} I \left[ M; \hat{Y}_i \mid X_{i+1}, \hat{Y}_{i-1}, \hat{C}_{1:i} \right]
\]
Now, \( I \left[ M; \hat{Y}_i \mid X_{i+1}, \hat{Y}_{i-1}, \hat{C}_{1:i} \right] \) can be upper bounded as
\[
I \left[ M; \hat{Y}_i \mid X_{i+1}, \hat{Y}_{i-1}, \hat{C}_{1:i} \right] \leq I \left[ M, Y_i; \hat{Y}_i \mid X_{i+1}, \hat{Y}_{i-1}, \hat{C}_{1:i} \right]
\]
\[
= I \left[ Y_i; \hat{Y}_i \mid X_{i+1}, \hat{Y}_{i-1}, \hat{C}_{1:i} \right] + I \left[ M; \hat{Y}_i \mid X_{i+1}, \hat{Y}_{i-1}, Y_i, \hat{C}_{1:i} \right]
\]
where \( Y_i = [y_i^1, \ldots, y_i^N]^T \). Now, \( I \left[ M; \hat{Y}_i \mid X_{i+1}, \hat{Y}_{i-1}, Y_i, \hat{C}_{1:i} \right] \) can be written as...
since we have \( \{ M, X_{1,i}, Y_{1,i-1}, C_{1,i} \} \rightarrow Y_i \rightarrow \hat{Y}_i. \) Now, we expand I \( \{ Y_i; \hat{Y}_i \} \rightarrow X_{1,i}, X_{1,i-1}, C_{1,i} \) as follows

\[
I \left[ Y_i; \hat{Y}_i \right] = h \left[ Y_i \, | \, X_{1,i}, X_{1,i-1}, \hat{C}_{1,i} \right] - h \left[ \hat{Y}_i \, | \, X_{1,i}, X_{1,i-1}, C_{1,i} \right]
\]

where (a) follows from the Markov chain \( X_{1,i}, \hat{Y}_{1,i-1}, \hat{C}_{1,i} \rightarrow Y_i \rightarrow \hat{Y}_i \) and (b) follows from the fact that conditioning reduces entropy. Combining (A.4)-(A.5), we have

\[
I \left[ M; X_{1:T}, \hat{Y}_{1:T}, \hat{C}_{1:T} \right] \leq \sum_{i=1}^{T} I \left[ Y_i; \hat{Y}_i \right]
\]

(5.6)

where \( C_{\text{down}} = \max_{p_{Y|X}} \mathbb{E} \left[ \left\| Y_{i+1} - \hat{Y}_{i+1} \right\| \right] \leq \alpha \) is the total capacity of channels from SNs to agents which convey the utility-related information and \( \alpha \) is the total transmission power constraint on these channels.

Now, for a given \( \epsilon \) and \( \delta \), consider any algorithm \( A \) for which after \( T_{\epsilon, \delta} (O, A) \) time steps, we have

\[
\sup_{U \in \mathcal{F}_i} \inf \Pr \left( \left\| X_{T_{\epsilon, \delta}} - \hat{X}_{T_{\epsilon, \delta}} \right\| \geq \epsilon \right) \leq \delta.
\]

Combining Lemmas 1 and 2 and (A.6), we have

\[
T_{\epsilon, \delta} (O, A) \geq \left( 1 - \delta \right) \log M_{2\epsilon} (S) - 1 \left/ C_{\text{down}} \right.
\]

which completes the proof.

Appendix B. PROOF OF COROLLARY 1

Let \( D \) be a diagonal matrix with diagonal entries equal to \( 2\epsilon \). Let \( L \) be the lattice generated by \( D \mathbb{Z} \), i.e., \( L = \{ Dz \mid z \in \mathbb{Z}^N \} \). Note that the elements of \( L \) are at least \( 2\epsilon \) apart. Now, let \( G (S, L) \) be the number of lattice points of \( L \) in \( S \), i.e., \( G (S, L) = | L \cap S | \). Clearly, \( M_{2\epsilon} (S) \) is lower bounded by \( G (S, L) \). Now, we use the following result from Schnell (1959) to obtain a lower bound on \( G (S, L) \) in terms of \( \epsilon \), volume and surface area of \( S \).

\[ G (S, L) \geq \frac{1}{\text{det} (L)} \left( V (S) - \frac{\lambda_N (L)}{2} \right) \] (B.1)

where \( V (S) \) is the volume of \( S \), \( \text{P}(S) \) is the surface area of \( S \) and \( \lambda_N (L) \) is the successive minima of \( L \) defined as the smallest \( \rho \) such that there exist \( N \) linearly independent elements of lattice, \( \{ l_1, \ldots, l_N \in L \mid \{ 0 \} \} \) such that \( \| l_i \| \leq \rho \) Fischlin and Seifert (1999).

For the lattice \( L = D \mathbb{Z}^N \), we have \( \text{det} (L) = (2\epsilon)^N \) and \( \lambda_N (L) = 2\epsilon \). Thus, \( M_{2\epsilon} (S) \) can be lower bounded as

\[
M_{2\epsilon} (S) \geq \left( \frac{1}{2\epsilon} \right)^N \left( V (S) - \epsilon P (S) \right)
\]

REFERENCES

Agarwal, A., Bartlett, P.L., Ravikumar, P., and Wainwright, M.J. (2012). Information-theoretic lower bounds on the oracle complexity of stochastic convex optimization. IEEE Transactions on Information Theory, 58(5), 3235–3249.

Agarwal, A., Wainwright, M.J., Bartlett, P.L., and Ravikumar, P.K. (2009). Information-theoretic lower bounds on the oracle complexity of convex optimization. In Advances in Neural Information Processing Systems 22, 1–9. Curran Associates, Inc.

Duchi, J.C., Jordan, M.I., Wainwright, M.J., and Wibisono, A. (2012). Finite sample convergence rates of zero-order stochastic optimization methods. In Advances in Neural Information Processing Systems 25, 1439–1447. Curran Associates, Inc.

Duchi, J.C., Jordan, M.I., Wainwright, M.J., and Wibisono, A. (2015). Optimal rates for zero-order convex optimization: The power of two function evaluations. IEEE Transactions on Information Theory, 61(5), 2788–2806.

Fischlin, R. and Seifert, J. (1999). Cryptography and Coding, chapter Tensor-based trapdoors for CVP and their application to public key cryptography, 244–257. Springer, Berlin Heidelberg.

Gardner, R.J. (2002). The brunn-minkowski inequality. Bulletin of the American Mathematical Society, 39(3), 355–405.

Jamieson, K.G., Nowak, R., and Recht, B. (2012). Query complexity of derivative-free optimization. In Advances in Neural Information Processing Systems 25, 2672–2680. Curran Associates, Inc.

Kolmogorov, A. and Tikhomirov, V. (1959). \( \epsilon \)-entropy and \( \epsilon \)-capacity of sets in function spaces. Uspekhi Matematicheskikh Nauk, 14(2), 3–86.

Li, S. and Basar, T. (1987). Distributed algorithms for the computation of noncooperative equilibria. Automatica, 23(4), 523–533.

Nemirovski, A., and Yudin, D. (1983). Problem complexity and method efficiency in optimization. Wiley.

Raginsky, M. and Rakhlin, A. (2011). Information-based complexity, feedback and dynamics in convex programming. IEEE Transactions on Information Theory, 57(10), 7036–7056.

Rosen, J.B. (1965). Existence and uniqueness of equilibrium points for concave n-person games. Econometrica, 33(3), 520–534.

Schnell, U. (1959). Lattice inequalities for convex bodies and arbitrary lattices. Monatshefte für Mathematik, 1993(116), 331–337.

Singer, Y. and Vondrak, J. (2015). Information-theoretic lower bounds for convex optimization with erroneous oracles. In Advances in Neural Information Processing Systems 28, 3204–3212. Curran Associates, Inc.
Author/s:
Nekouei, E; Alpcan, T; Nair, GN; Evans, RJ

Title:
Lower Bounds on the Best-Case Complexity of Solving a Class of Non-cooperative Games

Date:
2016

Citation:
Nekouei, E., Alpcan, T., Nair, G. N. & Evans, R. J. (2016). Lower Bounds on the Best-Case Complexity of Solving a Class of Non-cooperative Games. Proceedings of the 6th IFAC Workshop on Distributed Estimation and Control in Networked Systems, 49, (22), pp.244-249. Elsevier. https://doi.org/10.1016/j.ifacol.2016.10.404.

Persistent Link:
http://hdl.handle.net/11343/241547

File Description:
Accepted version