Towards Characterizing the First-order Query Complexity of Learning (Approximate) Nash Equilibria in Zero-sum Matrix Games

Hédi Hadiji[(2),(3)], Sarah Sachs[(2)], Tim van Erven[(2)] and Wouter M. Koolen[(1),(4)]

(1): Centrum Wiskunde & Informatica (2): University of Amsterdam, Korteweg-de Vries Institute for Mathematics (3): Université Paris-Saclay (4): University of Twente

Summary

Nash equilibria have been central to game theory since Von Neumann, but how fast can we compute them?

- The best upper bound is achieved by letting a minimizing and a maximizing online learning algorithm play against each other. But is there a better way?
- Our contributions:
  1. Fascinating insights into why **all existing techniques to prove lower bounds must fail!**
  2. The **first non-trivial lower bound.**

Setup and Definitions

We study two-player zero-sum simultaneous-play matrix games.

Minimax problem: For a $K \times K$ pay-off matrix M with entries in $[-1, +1]$ and randomized strategies $p, q \in \Delta_K$:

$$\min_p \max_q p^\top M q = \bar{p}^\top M \bar{q} = \max_q \min_p p^\top M q.$$

Nash-Equilibrium: $(\bar{p}, \bar{q})$ is a Nash-equilibrium if neither player can gain by changing their play:

$$p^\top M \bar{q} \geqslant \bar{p}^\top M \bar{q} \geqslant \bar{p}^\top M q \qquad \forall p, q.$$

ε-Nash-Equilibrium: Players cannot gain more than ε:

$$p^\top M \bar{q} + \varepsilon \geqslant \bar{p}^\top M \bar{q} \geqslant \bar{p}^\top M q - \varepsilon \qquad \forall p, q.$$

Query Model: Players learn expected pay-off for all their actions under randomized play of opponent:

$$(p_t, q_t) \mapsto (M^\top p_t, M q_t).$$

Question: What is the smallest number of queries $T(\varepsilon)$ we need to find an ε-Nash-equilibrium?

Why Do All Existing Lower Bound Techniques Fail?

Discrete entries are too easy!

Theorem 1. *Suppose we know in advance that all entries in M come from a known countable alphabet $\mathcal{A}$ (with at least two elements). E.g. all entries in M are either -1 or $+1$. Then it is possible to fully identify M with 1 query, and hence $T(\varepsilon) \leqslant 1$.*

This **rules out all lower bound approaches** that have been successful for other query models:

- Reasoning about specific classes of binary payoff matrices [FGGS15]
- Reducing from submodular optimization over the hypercube by encoding it as a binary matrix [HK16]
- Randomly generating a matrix with binary entries

Computing Exact Nash Equilibria

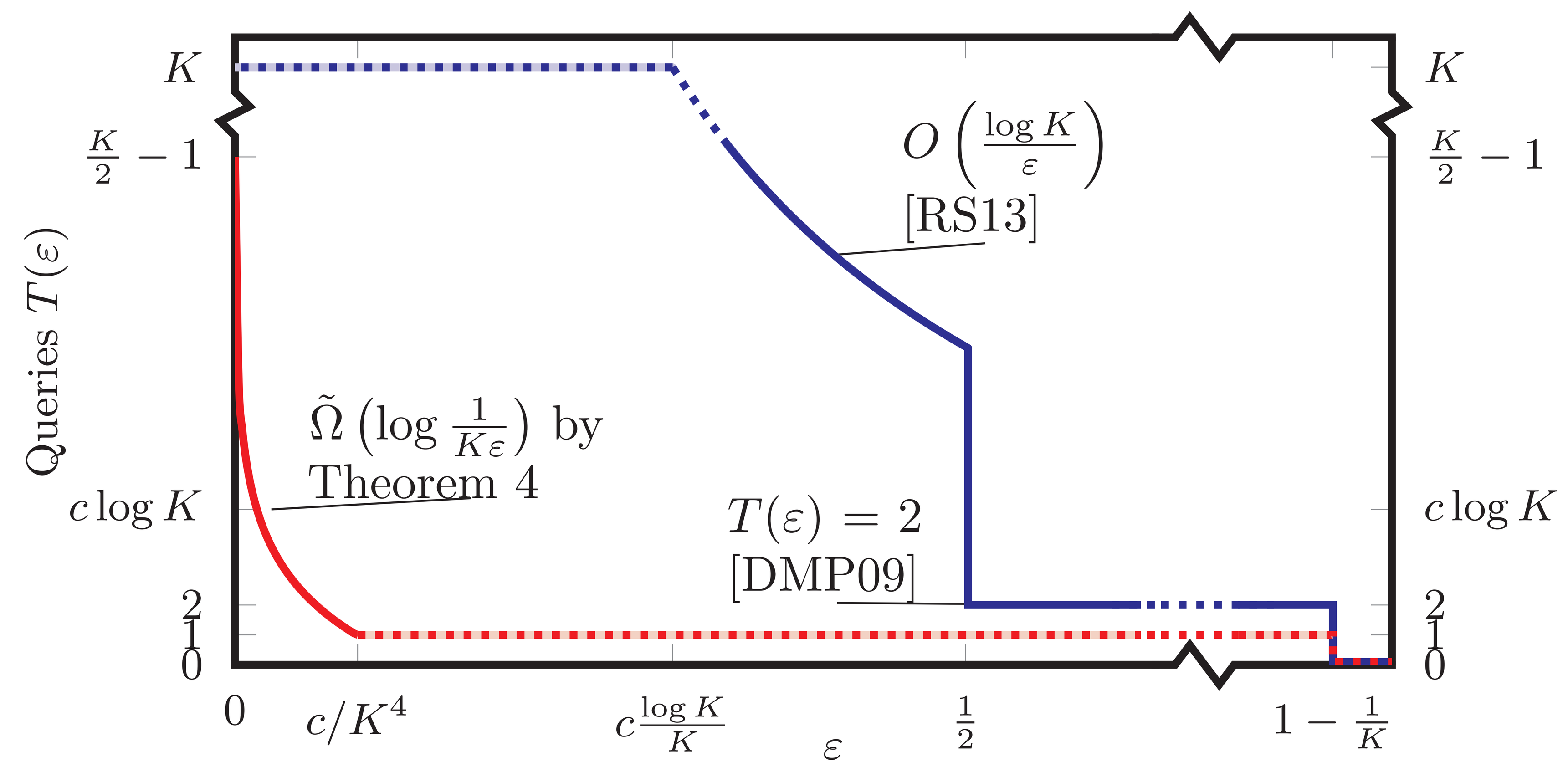
But continuous-valued entries are potentially much harder!

Theorem 2. *If the entries in M can take any values in $[-1, +1]$, then the number of queries required to fully identify M is exactly K.*

And exactly computing a Nash Equilibrium is essentially as hard as identifying the full matrix!

Theorem 3. *The number of queries required to compute an exact Nash Equilibrium is at least $T(0) \geqslant \frac{K}{2} - 1$.*

Approximate Nash Equilibria



Upper (blue) and lower (red) bounds on the query complexity of computing an ε-equilibrium for $K \times K$ matrix games.

Theorem 4 (Lower Bound). *For any $\varepsilon \leqslant 1/(e\, 2^{11} K^4)$, the number of queries required to compute an ε-Nash Equilibrium is at least*

$$T(\varepsilon) \geqslant \left(\frac{-\log(2^{11}K^4\varepsilon)}{\log(2^{11/2}K^{5/2}) + \log(-\log(2^{11}K^4\varepsilon))} - 1\right) \wedge \left(K - 3\right) = \tilde{\Omega}\left(\log \frac{1}{K\varepsilon}\right).$$

Proof Sketch for Exact Nash

Idea: construct adversary **answering** queries by the learner so as to **delay** revealing the equilibrium for as long as possible.

- We restrict ourselves to a ball $\mathcal{B}_{\|\cdot\|_{1,\infty}}\left(\frac{I_K}{2}, \frac{1}{16K^2}\right)$ of matrices around the (scaled) identity.
  - For any such matrix, the unique exact Nash equilibrium is a pair of **fully mixed** strategies.
  - Querying a fully mixed equilibrium strategy yields feedback proportional to $\mathbf{1}$ ⇐ "equaliser".
- Based on the feedback to the queries so far, a subset of consistent matrices remains: every round adds $2K$ equality constraints.
- We show that there is a common Nash equilibrium for the consistent subset **only if** the vector $\mathbf{1}$ is in the span of the feedback.
- Our adversary keeps $\mathbf{1}$ out of the span of the feedback for $\frac{K}{2} - 1$ rounds. ⇐ "dimension-as-a-resource"
- Technique: maintain consistent matrix M_t. Upon query (p_{t+1}, q_{t+1}), rank-one update $M_{t+1} = M_t + \bar{p}_{t+1} u_t^\top$, with $\bar{p}_{t+1}$ the rejection of p_{t+1} from the span of $p_1, \ldots, p_t$, and u_t orthogonal to $\mathbf{1}$ and
  - the span of $q_1, \ldots, q_t$ (consistency),
  - the past feedback $M_t^\top p_t$ (proof artifact),
  - the forced feedback part $M_t^\top p_{t+1}$.

References

[DMP09] Constantinos Daskalakis, Aranyak Mehta, and Christos Papadimitriou. A note on approximate Nash equilibria. *Theoretical Computer Science*, 410(17):1581–1588, 2009.

[FGGS15] John Fearnley, Martin Gairing, Paul W Goldberg, and Rahul Savani. Learning equilibria of games via payoff queries. *Journal of Machine Learning Research*, 16:1305–1344, 2015.

[HK16] Elad Hazan and Tomer Koren. The computational power of optimization in online learning. In *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing*, STOC '16, 2016.

[RS13] Sasha Rakhlin and Karthik Sridharan. Optimization, learning, and games with predictable sequences. *Advances in Neural Information Processing Systems*, 26, 2013.