Beating Price of Anarchy and Gradient Descent without Regret in Potential Games

Iosif Sakos

Singapore University of Technology and Design

Stefanos Leonardos

King’s College London

Stelios A. Stavroulakis

UC Irvine

William Overman

Stanford University

Ioannis Panageas

UC Irvine

Georgios Piliouras

Singapore University of Technology and Design

Quality of Learning Outcomes in Potential Games

- Multi-agent learning occurs on **highly non-convex** landscapes.
- Even if learning dynamics equilibrate, their fixed points may be of **poor performance**.

How do we **measure** the performance of learning dynamics?

Are some learning dynamics **better than** others?

Limitations of Static Performance Metrics

$$\mathrm{PoA} = \frac{\max_x \mathrm{SW}(x)}{\min_{x^* \in \mathrm{NE}} \mathrm{SW}(x^*)}$$

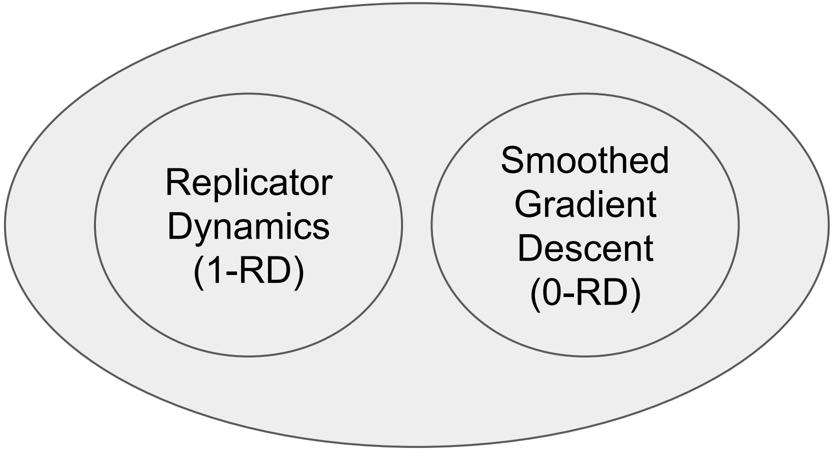
- The PoA **does not depend** on the learning process.
- The socially-worst equilibrium might be reachable by **only a few initializations**.
- The PoA can be **unbounded** even for simple classes of games.

Average Price of Anarchy (APoA)

$$\mathrm{APoA} = \frac{\max_x \mathrm{SW}(x)}{\mathbb{E}[\mathrm{SW}(x^*)]}$$

- The APoA **depends on the learning process**.
- The APoA **requires equilibration** for almost all initial conditions.

Convergence of QRD in Potential Games (Theorem 3.2)



Q-Replicator Dynamics (QRD)

- QRD equilibrate in almost **all potential games**
- QRD equilibrate for almost **all initial conditions**

Average Price of Anarchy Analysis for QRD

In symmetric 2x2 potential games, where the **payoff-dominant and risk dominant equilibria coincide**:

- **Gradient Descent outperforms Replicator Dynamics** in terms of APoA (Theorem 4.6)
- The **APoA of Gradient Descent is upper bounded by 2** (Theorem 4.8)

Thank you!

For further results and discussion visit our poster