From Chaos to Order: Symmetry and Conservation Laws in Game Dynamics

Sai Ganesh Nagarajan, David Balduzzi
and Georgios Piliouras

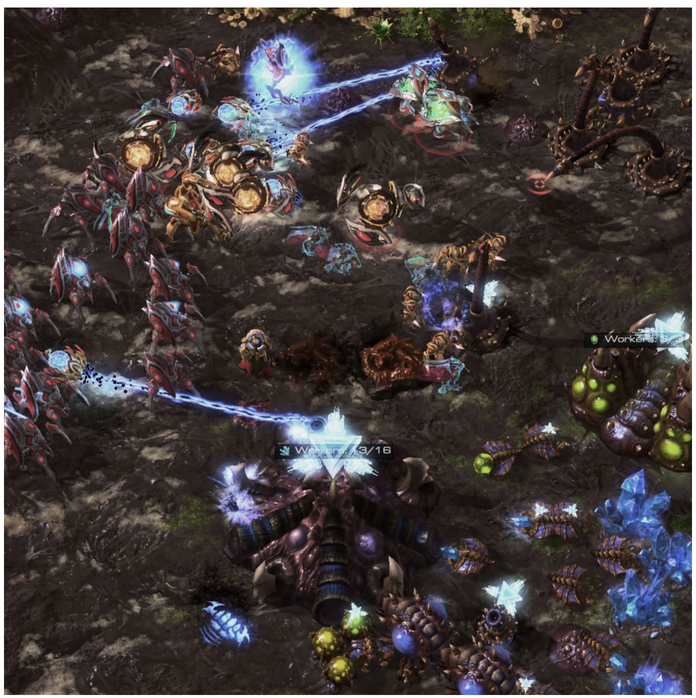
ICML 2020

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Google DeepMind

Introduction

Games are central in machine learning (ML) training [Goodfellow et al.,2014,Silver et al.,2017, Vinyals et al., 2019] in GANs, Starcraft, Alpha GO, Chess etc.

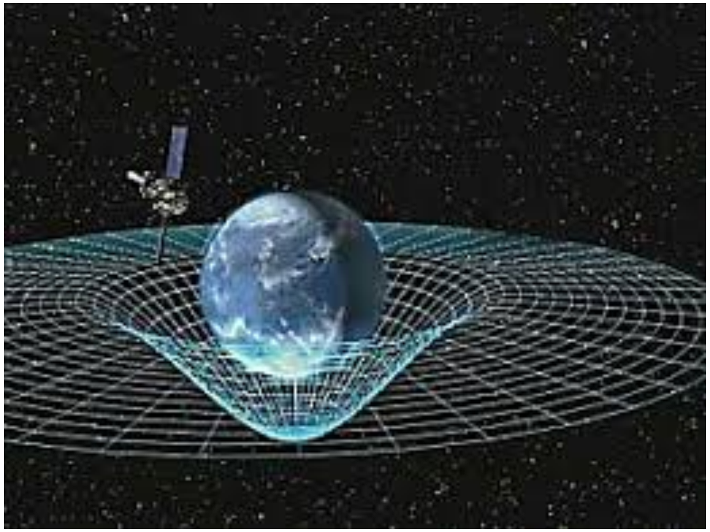


Non-convergence in Games

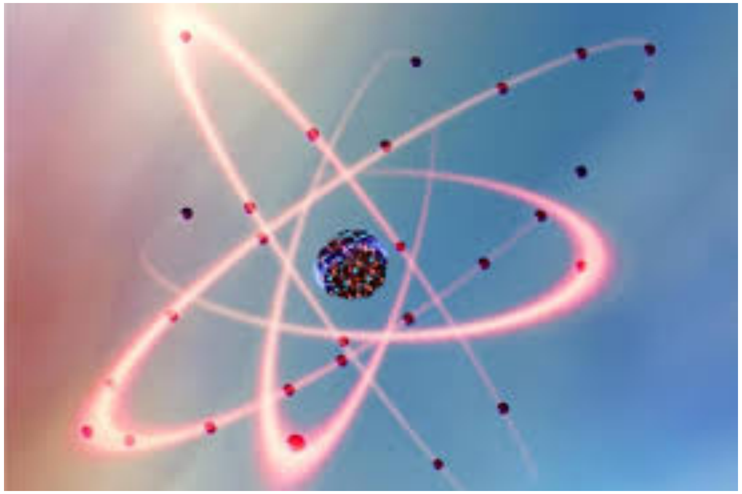
- General assumption in ML is that competition between learning algorithms forces the algorithm to improve.
- Games can be unpredictable [Galla et al.,2013, Piliouras et al.,2014] or formally chaotic [Palaiopanos et al.,2017, Sato et al., 2002].
- Recently, [Balduzzi et al.,2020] showed that continuous time Gradient ascent in smooth games can diverge. A similar divergence result for discrete time multiplicative weight updates in zero-sum games was shown by [Bailey et al.,2018].
- In continuous action/state multi-agent RL, policy gradient was shown to have no guarantees of local convergence in simple games [Mazumdar et al., 2020].

How can we control learning dynamics in large scale multi-agent systems?

A Physics Approach



Gravity

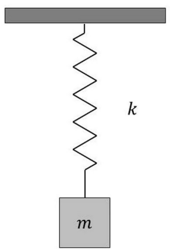


Quantum Mechanics

Complex systems in two vastly different scales!

But governed by fundamental laws of conservation and symmetry.

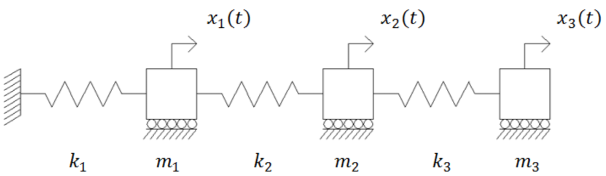
An Example with Springs



A simple 1 mass and 1 spring system.

- **Method 1**: Use **Newton's Laws**, using force diagrams to derive equations of motion.
- **Method 2**: Write down the **Hamiltonian or Lagrangian** and apply the **least action principle** to obtain the equations of motion.

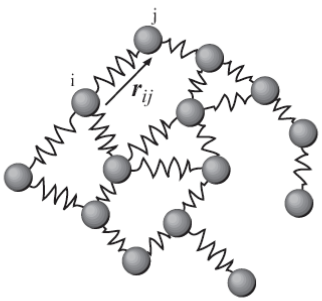
How does it Scale?



Can still salvage using Newton's laws and force diagrams.

A spring mass systems with 3 springs and 3 masses

Writing force diagrams becomes an uphill task and does not scale well!
Method 2 works better!



A network of spring mass systems

A Physicist's Checklist

- Make an **appropriate coordinate transform** dictated by the geometry of the problem.
- Identify the **conservation laws** in the system.
- Exploit the **symmetries** in the transformed system to obtain the required equations.

Main Question

Can we identify a class of games and learning dynamics that have conservation laws?

Network Game with Charges (NGC)

Column Player

Row player

	R	P	S
R	0,0	-1,1	1,-1
P	1,-1	0,0	-1,1
S	-1,1	1,-1	0,0

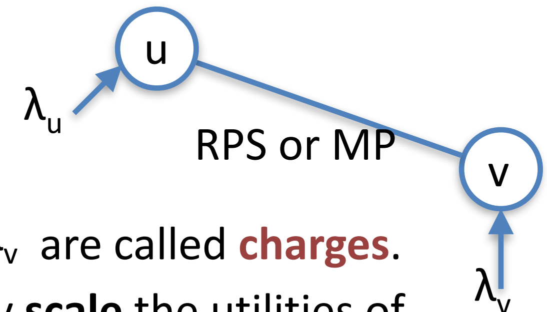
Rock-Paper-Scissors (RPS)

Column Player

Row player

	H	T
H	1,-1	-1,1
T	-1,1	1,-1

Matching Pennies (MP)



λ_u, λ_v are called **charges**. They **scale** the utilities of the respective players.

Some instantiations of **charges**:

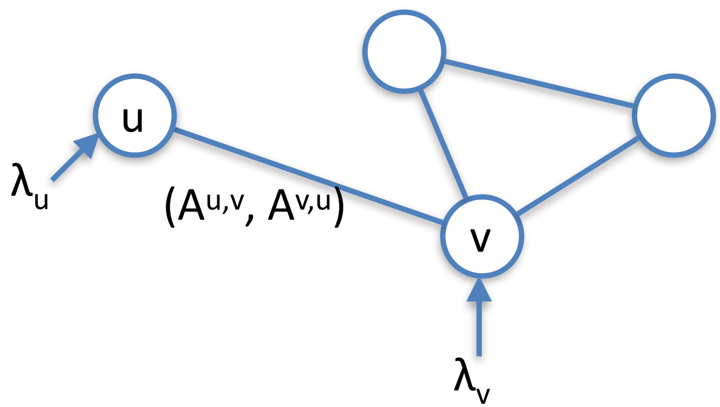
$\lambda_u = 1$ and $\lambda_v = 1$ - RPS or MP

$\lambda_u = -1$ and $\lambda_v = 1$- Coordination versions of RPS/MP

$\lambda_u = -1$ and $\lambda_v = -1$ - Switches row and column player.

NGC Formalism

Consider a graphical polymatrix game [Kearns et al., 2001]: Nodes are players, playing a game with each neighbor.



Bimatrix game between u and v is $(A^{u,v}, A^{v,u})$.
Charges : λ_u ,λ_v are elements in $R\setminus\{0\}$.
Utility of player u:

$$\lambda_u \left(\sum_{j \in \mathcal{N}(u)} A^{u,j}(s_u, s_{-u}) \right)$$

Special cases include :

Network Zero-sum, coordination games, hybrid varieties and other **large scale games** [Nagarajan et al.,2018, Szabo et al.,2007, Wang et al.,2015].

Learning Dynamics

Players use classic **no-regret algorithms** to update their mixed strategies via Follow the Regularized Leader (FTRL) [Hazan et al.,2016, Mertikopoulos et al., 2018] with possibly **different regularizers**.

$$y_i(t) = y_i(0) + \int_0^t v_i(x(s))ds$$

→ Accumulated payoff from 0 until time t.

$$x_i(t) = Q_i(y_i(t))$$

$$Q_i(y_i) = \arg\max_{x_i \in \mathcal{X}_i} \{< y_i, x_i > -h_i(x_i)\}$$

Choose the next mixed strategy by optimizing Q with regularizer h.

Applying h=negative entropy , leads to Replicator dynamics and h = squared euclidean norm, gives gradient descent.

Main Theorems (Informal)

- **Conservation Law**: When the agents play via **any FTRL** dynamics in a NGC, a notion of a "**linear combination of distances**" in the **payoff space** is **invariant** with respect to time.
- **Dimensionality Reduction**: The dynamics in some families of NGC, allows for many invariant functions and this in-turn guarantees that the trajectories lie in a **lower dimensional space**.
- **Periodicity**: The dynamics of a **bipartite NGC** with a base constant sum game that is 2-by-2 is **periodic** when the **charges** are of the **same sign**.

Conservation Laws in NGC

When the players play in a NGC, we show that the following quantity is **invariant** with respect to time :

$$H(y) = \sum_{i \in V} \lambda_i \left(h_i^*(y_i) - < y_i, x_i^* > \right)$$

Where, y(t) describes the evolution w.r.t to the **payoffs**.

$$h_i^*(y_i) = \max_{x_i \in \mathcal{X}_i} \left\{ < y_i, x_i > - h_i(x_i) \right\}.$$

$h_i^*(.)$ is the **convex conjugate** of $h_i(.)$ and x_i^* is the **fully mixed Nash equilibrium**.

Key idea is to take the time derivative of H(y) and show that it is 0.

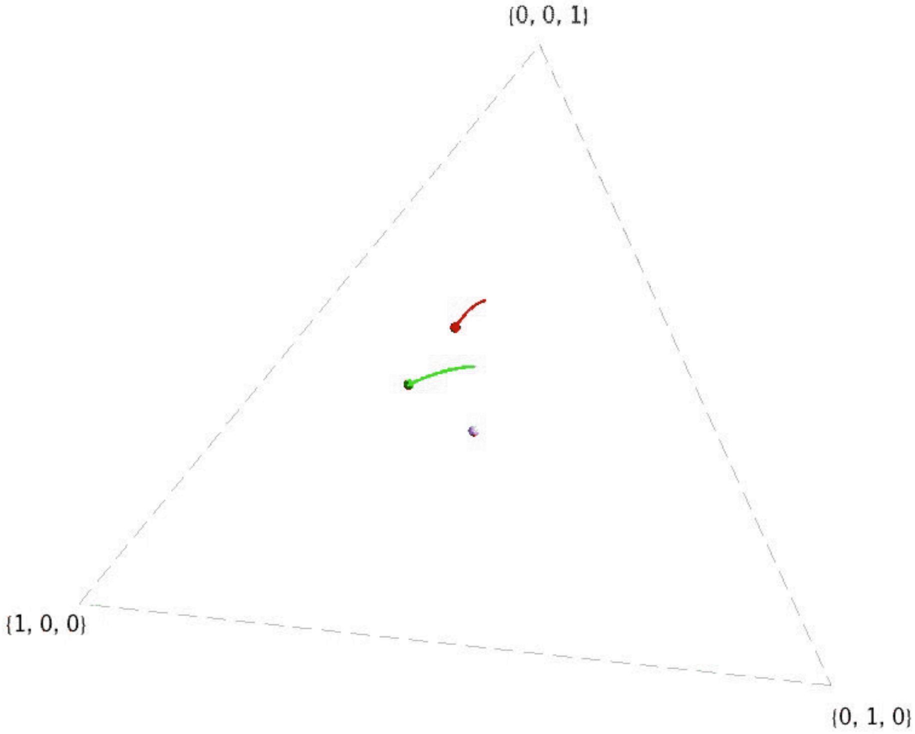
Facts about Conservation Laws

- **Interpretation**: A linear combination of “distances” (closely connected to the notion of **Bregman divergences**) from the Nash equilibrium strategy of each player, **scaled** by their respective **charge** is **invariant** in the **space of payoffs**.

- Some games can have **multiple conservation laws**.

- Conservation laws generally **constrain the dynamics** leading to **simple**, **non-chaotic** behavior.

A Visualisation of Conservation

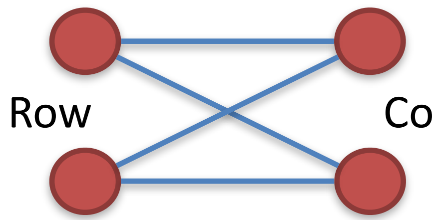


Replicator Dynamics on Rock-Paper-Scissors.

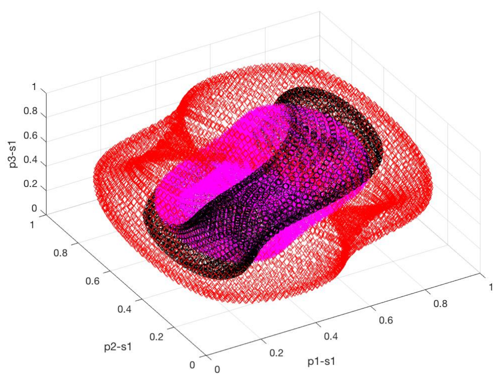
Complex Behavior in NGC

Consider the MP game being played along the edges of this bipartite network.

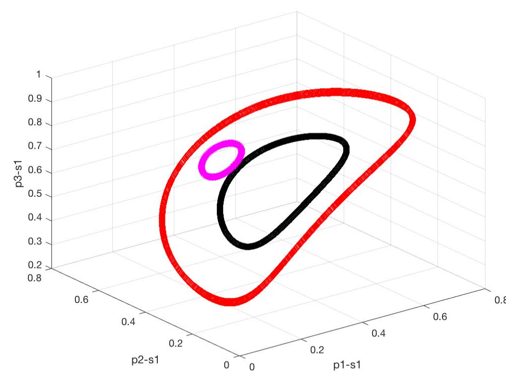
Bipartite network game in the absence of symmetries.



Bipartite network game with symmetries.



Both systems are driven by **same conservation laws**!!

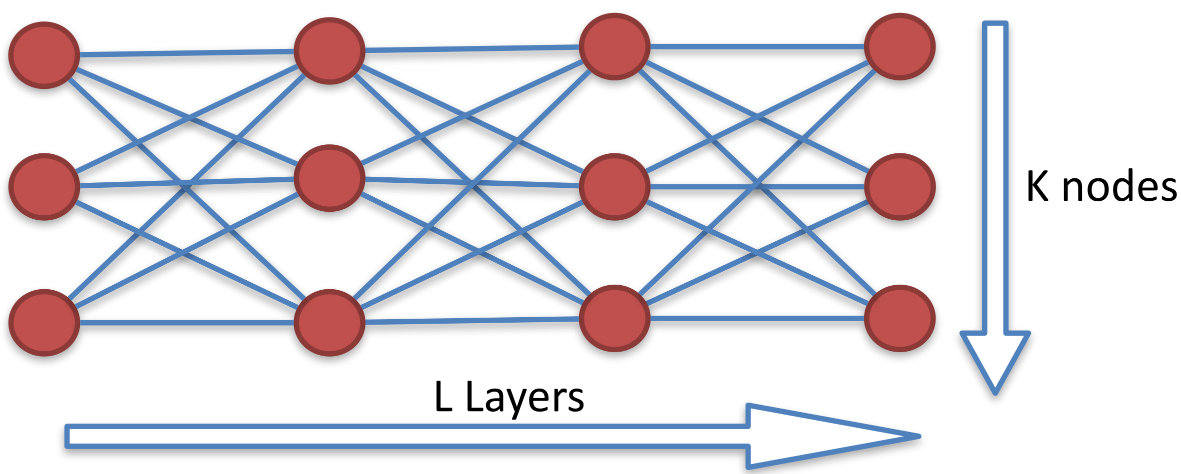


The trajectories of the mixed strategies over time is **chaotic**!

The trajectories of the mixed strategies over time is **periodic**!

Dimensionality Reduction

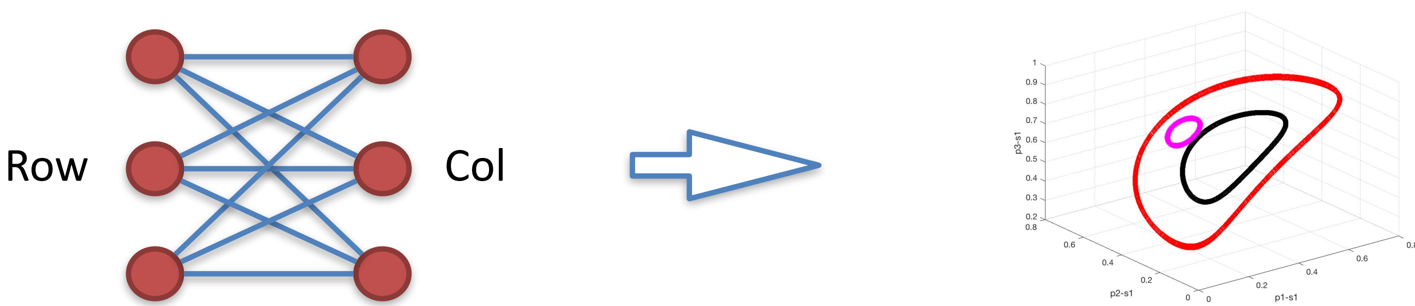
- Consider a network constant sum game with charges described by a base game A (n-by-n). Such games appear in [Szabo et al.,2007, Wang et al.,2015].
- Consider a graph with L layers, K nodes per layer. Each node is a player with charge λ_i.



- Then, the dynamics of the mixed strategies effectively lies in the **lower-dimensional space** containing L*(n-1) variables.
- Note that in general, L is relatively small compared to K.

Special Cases-Periodic Orbits

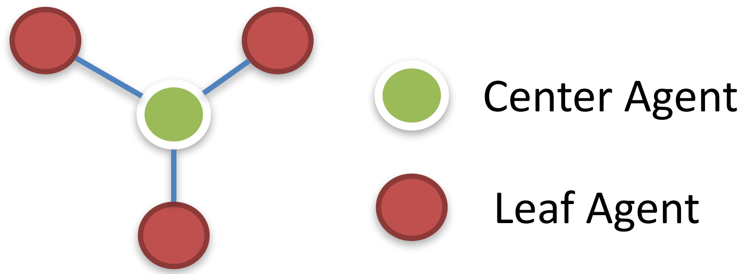
- Certain regularizers such as **negative entropy** (**replicator dynamics**), lead to a **closed form** solution for the reduction.
- **Periodic orbits** for **bipartite** network zero-sum game with **charges (same sign)** when the base game is 2-by-2.

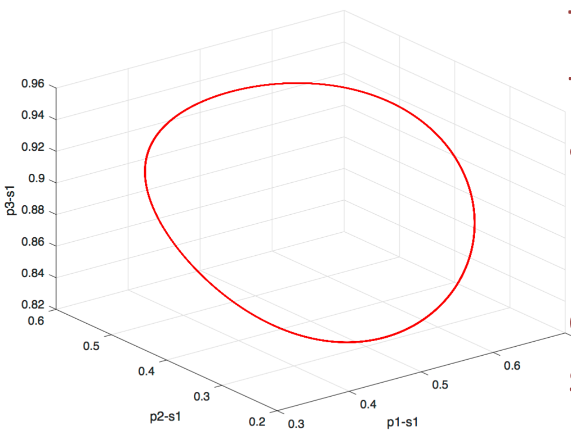


- **Proof Sketch**: Apply the result of **dimensionality reduction** to this case and then use **Poincaré-Bendixon theorem** and **Poincaré-Recurrence theorem** [Mertikopoulos et al., 2018].

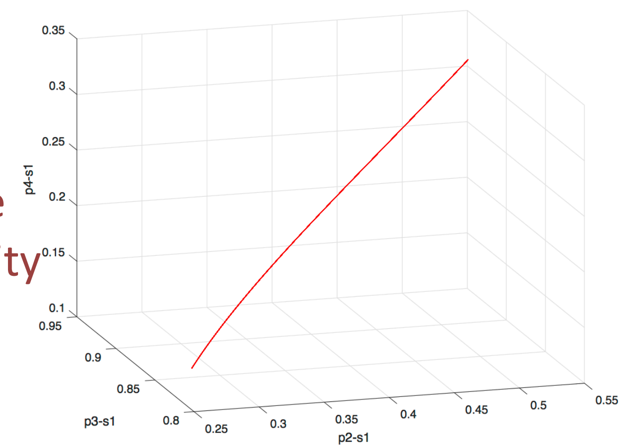
Cooperation from Competition

Star configuration.
MP game played
on the edges and all
charges are set to 1.





The trajectories involving the **Center** agent always results in a periodic orbit. But for the **Leaf** agents, the probability of playing a particular strategy **always moves concurrently**!



The trajectories of the mixed strategies involving the **Center** agent.

The trajectories of the mixed strategies of the **Leaf** agents.

Reverse Engineering the Game

- Given a **fully mixed Nash**, can we obtain a base **constant-sum game** matrix A, with **value** c, that implements the **conservation laws**?
- For the sake of illustration, consider two players, such that, $(x_0^*, x_1^*, x_2^*, y_0^*, y_1^*, y_2^*)$ is a fully mixed Nash equilibrium. We construct a sparse constant sum game that has the given Nash equilibrium profile.

$$\begin{pmatrix} a(c) & \frac{c-x_1^*}{x_0^*} & \frac{c-x_2^*}{x_0^*} \\ \frac{c-y_1^*}{y_0^*} & 1 & 0 \\ \frac{c-y_2^*}{y_0^*} & 0 & 1 \end{pmatrix} \quad a(c) = \frac{c(x_0^* + y_0^* - 1) + \sum_{i=1}^{2} x_i^* y_i^*}{x_0^* y_0^*}$$

Each player can run **FTRL with the regularizer h_i** and this satisfies the conservation law. This can be extended for an arbitrary network constant sum game with charges.

Future Work

- Investigate more network configurations where competition leads to cooperation.

- To understand how these results can carry over in discrete time dynamical systems.

- Taking the results on conservation laws and dimensionality reduction to multi-agent reinforcement learning.

Summary

- We introduced **NGC framework using FTRL dynamics**. To analyze the complex systems in NGC we did the following:

- Made a **coordinate transform** to the **payoff space**.

- Identify **conservation laws** in NGC.

- Exploit the **inherent symmetries** w.r.t to the row/column agents in the **payoff space**.

- We provided special cases of our results which lead to **simple**, **non-chaotic** behavior and where **cooperation arose from competition**.

- We answer the **inverse question** of implementing a base game with the **required mixed NE profile** in the NGC framework that satisfies the given **conservation laws**.

Thank You

Please email with questions or for a copy of the paper:

sai_nagarajan@mymail.sutd.edu.sg

References

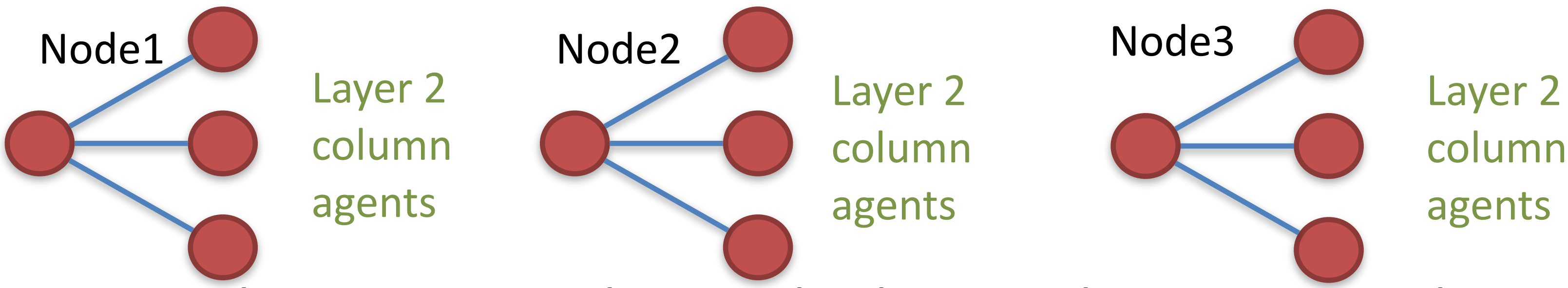
- Szabó, G., & Fath, G. (2007). Evolutionary games on graphs. *Physics reports*, *446*(4-6), 97-216.
- Wang, Z., Wang, L., Szolnoki, A., & Perc, M. (2015). Evolutionary games on multilayer networks: a colloquium. *The European physical journal B*, *88*(5), 124.
- Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative adversarial nets. In Advances in neural information processing systems. 2672–2680.
- Elad Hazan et al. 2016. Introduction to online convex optimization. Foundations and Trends® in Optimization 2, 3-4 (2016), 157–325.
- M. J. Kearns, M. L. Littman, and S. P. Singh. 2001. Graphical Models for Game Theory. In UAI.
- Nagarajan, Sai Ganesh, Sameh Mohamed, and Georgios Piliouras. "Three body problems in evolutionary game dynamics: Convergence, periodicity and limit cycles." AAMAS, 2018.

- Oriol Vinyals, Igor Babuschkin, Junyoung Chung, Michael Mathieu, Max Jaderberg, Wojciech M. Czarnecki, Andrew Dudzik, Aja Huang, Petko Georgiev, Richard Powell, Timo Ewalds, Dan Horgan, Manuel Kroiss, Ivo Danihelka, John Agapiou, Junhyuk Oh, Valentin Dalibard, David Choi, Laurent Sifre, Yury Sulsky, Sasha Vezhnevets, James Molloy, Trevor Cai, David Budden, Tom Paine, Caglar Gulcehre, Ziyu Wang, Tobias Pfaff, Toby Pohlen, Yuhuai Wu, Dani Yogatama, Julia Cohen, Katrina McKinney, Oliver Smith, Tom Schaul, Timothy Lillicrap, Chris Apps, Koray Kavukcuoglu, Demis Hassabis, and David Silver. 2019. AlphaStar: Mastering the Real-Time Strategy Game StarCraft II. https://deepmind.com/blog/ alphastar-mastering-real-time-strategy-game-starcraft-ii/. (2019).
- David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, Yutian Chen, Timothy Lillicrap, Fan Hui, Laurent Sifre, George van den Driessche, Thore Graepel, and Demis Hassabis. 2017. Mastering the game of Go without human knowledge. Nature 550 (2017), 354–359.

- Panayotis Mertikopoulos, Christos Papadimitriou, and Georgios Piliouras. 2018. Cycles in adversarial regularized learning. In Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms. SIAM, 2703–2717.
- Gerasimos Palaiopanos, Ioannis Panageas, and Georgios Piliouras. 2017. Multiplicative weights update with constant step-size in congestion games: Convergence, limit cycles and chaos. In Advances in Neural Information Processing Systems. 5872–5882.
- G. Piliouras and J. S. Shamma. 2014. Optimization Despite Chaos: Convex Relaxations to Complex Limit Sets via Poincaré Recurrence. In SODA.
- Yuzuru Sato, Eizo Akiyama, and J. Doyne Farmer. 2002. Chaos in learning a simple two-person game. Proceedings of the National Academy of Sciences 99, 7 (2002), 4748–4751. https://doi.org/10.1073/pnas.032086299 arXiv:http://www.pnas.org/content/99/7/4748.full.pdf+html
- Mazumdar, Eric, et al. "Policy-Gradient Algorithms Have No Guarantees of Convergence in Linear Quadratic Games." *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*. 2020.
- Bailey, James P., and Georgios Piliouras. "Multiplicative weights update in zero-sum games." Proceedings of the 2018 ACM Conference on Economics and Computation. 2018.
- Balduzzi, David, et al. "Smooth markets: A basic mechanism for organizing gradient-based learners." arXiv preprint arXiv:2001.04678 (2020).

Proof Sketch

Consider the first layer (unravelled) as follows:



- Main observation is that each player in layer 1, sees the **same payoff** coming from the column agents in layer 2 (up to a **scaling factor** which is their **charge**).
- Then we can obtain the **invariant equations** for each **strategy** in the **payoff space** between the players of the layer 1.
- Inductively applying this idea **layer by layer** we obtain required dimensionality reduction.