DISTURBANCE DECOUPLING FOR GRADIENT-BASED MULTI-AGENT LEARNING

SARAH H.Q. LI[1]

1: AERONAUTICS AND ASTRONAUTICS ENGINEERING, UNIVERSITY OF WASHINGTON

Quadratic Continuous Games

In an N player quadratic continuous game (f_1, … f_N), where for each player i, f_i is its cost function and X_i is its action space, each player attempts to minimize its individual cost function f_i given the action of all other players.

$$\min_{x_i} f_i(x) = x_i^\top P_i x_i + x_i^\top \Big(\sum_{j\neq i} P_{ij} x_j + r_i\Big)$$

For such games, a stable outcome is a **Nash equilibrium**. Informally, a joint action is a Nash equilibrium when no player can improve their individual cost by adjusting their action. If this property only holds in a neighbourhood of x, it is a **local Nash equilibrium**.

Ex 1) Finite time horizon LQ games

$$\begin{aligned}&\min_{u_i\in\mathbb{R}^{n_i}} \frac{1}{2}(x^T)^\top Q_i x^T + \frac{1}{2}\sum_{t=0}^{T-1}(x^t)^\top Q_i x^t + \frac{1}{2}(u_i^t)^\top R_i u_i^t \\ &\text{s.t.}\quad x^{t+1} = Ax^t + \sum_{i=1}^{N} B_i u_i^t,\ t = 0,\dots,T-1,\end{aligned}$$

Ex 2) MARL Policy Evaluation

$$f_i(x) = -\frac{1}{2}\sum_{j\neq i}(x_i - x_j)^\top p_{ij}(x_i - x_j) + \frac{1}{2}\|\Phi x_i - J^\star\|_D^2,$$

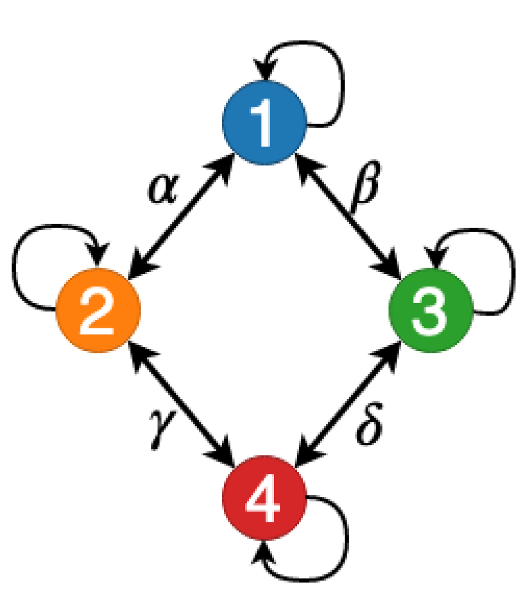
Adversarial Gradient Noise

We consider a class of gradient-based learning techniques, where each player myopically updates its own action based on its current gradient.

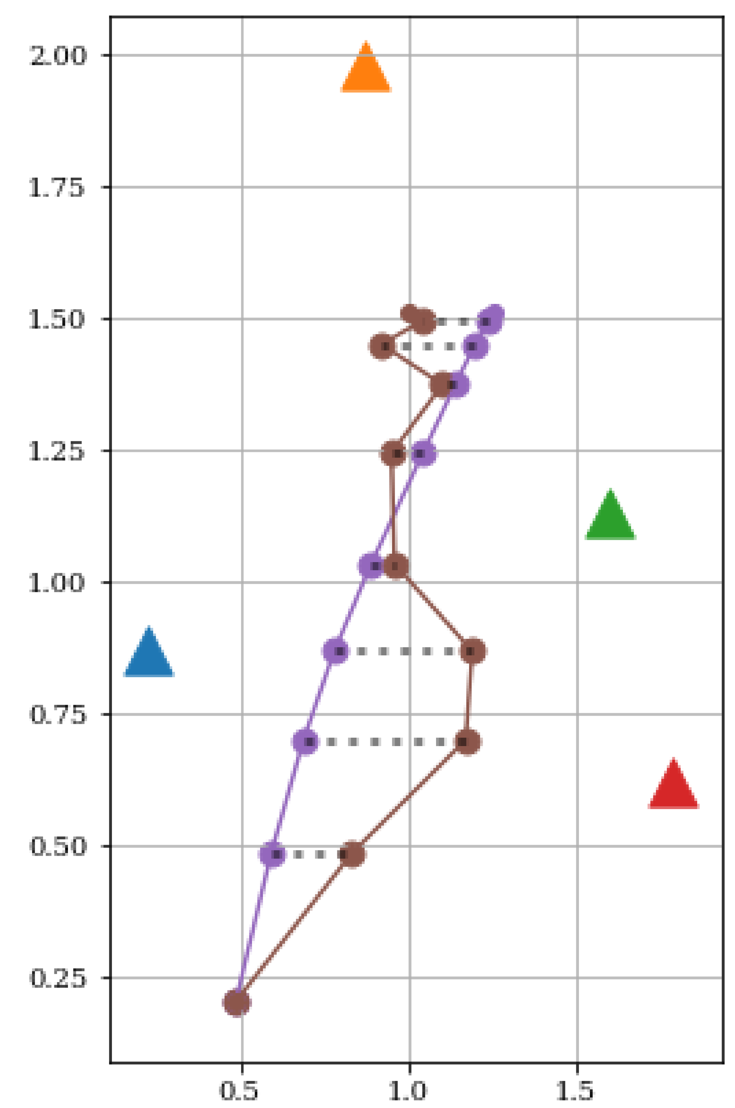
$$x_i^{k+1} = x_i^k - \gamma_i h_i(x_1^k,\dots,x_N^k),\quad h_i(x) = D_i f_i(x) + g_i,\quad g_i \in \mathbb{R}^{n_i}$$

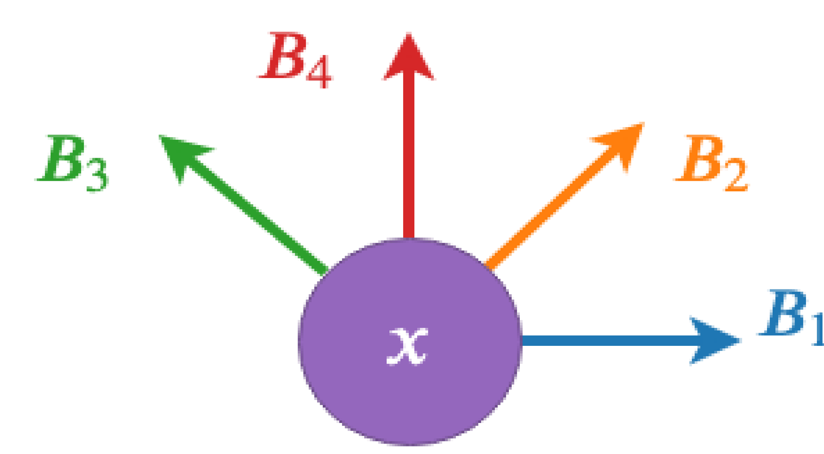
Assume an additive noise g_i corrupts the original gradient of player i. When we do not assume bounds or dynamics on the noise, such that it is arbitrary, the stability of the overall learning dynamics cannot be guaranteed. However, we show that in certain game graph structures, a subset of players' action updates remain completely unaffected by the disturbance. I.e., their learned action trajectories remain identical with and without gradient noise.

Game graph: we define a directed graph where the set of nodes correspond to the set of player actions, and the edges between each player exist if the gradient $D_{ij} f_j$ is non-zero.
We define a path of length k on a game graph as a sequence of k+1 nodes connected k edges.
Its path weight is given by the product of its edge weights



$$W_{j,v_{k-1}} \dots W_{v_1,i} = \prod_{l=0}^{k-1} W_{v_{l+1},v_l}$$

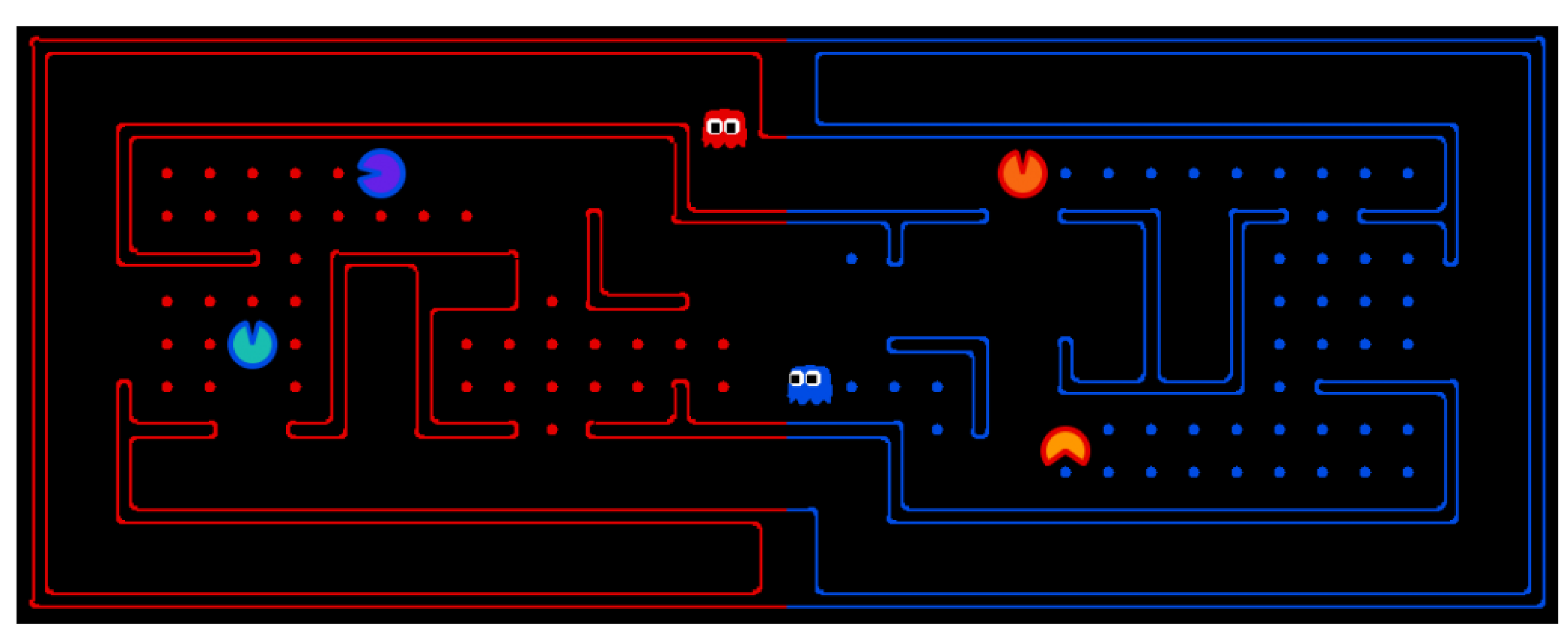




Top: A four player tug-of-war modelled as an LQ game. A target x is controlled by players with different control coordinates. Each pulls x towards a preset target in a finite time horizon.
Left: resulting trajectory of a four player tug-of-war game, preset player targets are given by triangles. Two trajectories are shown, the purple trajectory is a Nash equilibrium while the brown trajectory is with random player controls.

Problem Statement

Gradient-based learning proves to be an effective model for evaluating multi-agent learning dynamics found in games, reinforcement learning, etc. We utilize this framework to **study how noises in an individual player propagates to other players through coupled learning algorithms**. With a thorough understanding of the mechanism by which noises propagate within in a network of players, we can better design games and cost functions that maximizes **robustness of the overall learning dynamics to destabilizing adversarial noises**.



Disturbance Decoupling Networks

Theorem: Given a quadratic game (f_1, … f_N), where player i receives noisy gradient, player j is decoupled from player i's noisy gradient if and only if the path weights of length k satisfy

$$\sum_{p\in\mathcal{P}_k^{ij}} \prod_{l=1}^{k-1} W_{v_{l+1},v_l} = 0,\ \forall\, 0 \le k < N$$

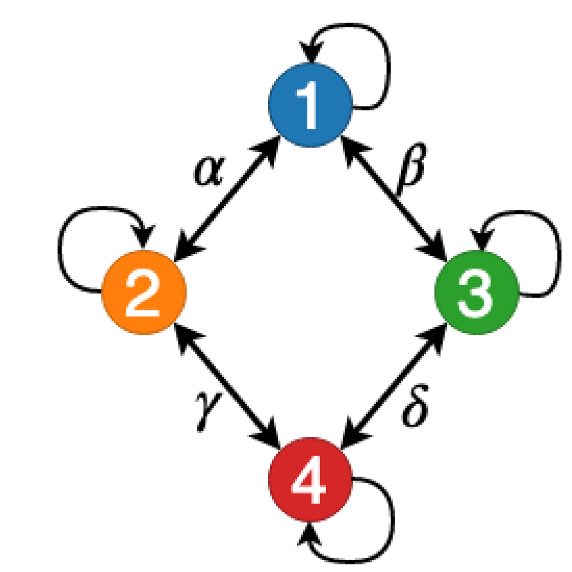
We consider this condition for player 1 and player 4 on the game graph given on the right, where each player has a scalar strategy and inter-player edges has weights as labelled. Each self loop has weight w_i.

To satisfy Theorem 1, all sums of path weights must be zero for paths of lengths k = 0, 1, 2, 3. Since there are no paths of length zero or one, k = {0, 1} are automatically satisfied. For k = 2, we generate the condition

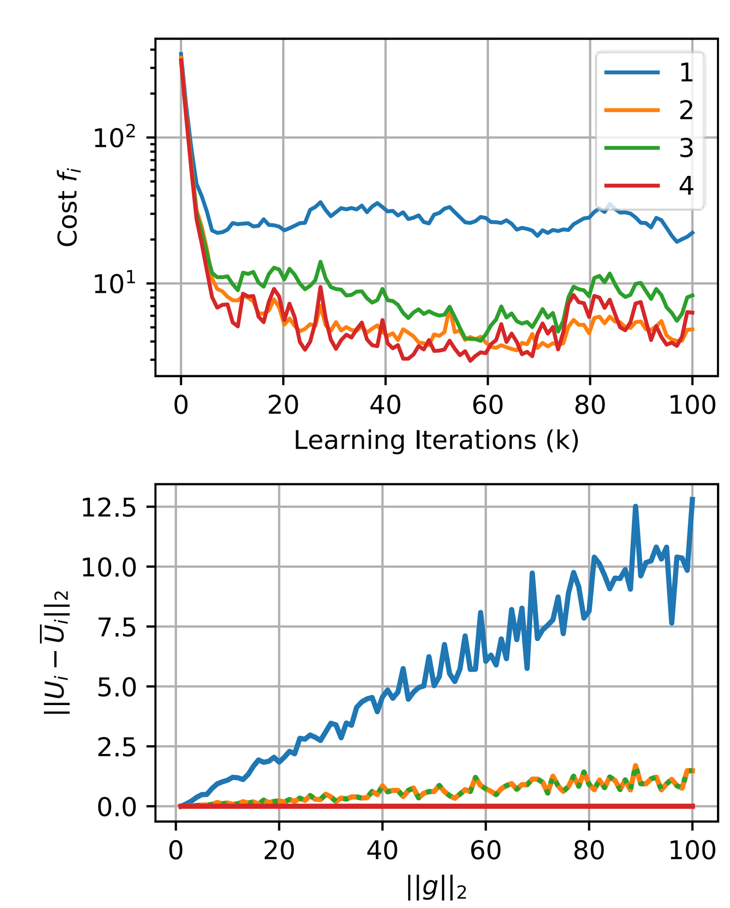
$$\alpha\gamma + \beta\delta = 0$$

For k = 3, we generate the condition

$$w_1\alpha\gamma + \alpha w_2\gamma + w_1\beta\delta + \beta w_3\delta + \alpha\gamma w_4 + \beta\delta w_4 = 0$$



Disturbance Decoupling in LQ Games



The tug-of-war LQ game as described in center right has game graph given by right top, where the weights satisfy disturbance decoupling. We simulate such this game with random initial conditions and random gradient noise in player one and plot the results as shown on the right. Top: we plot the cost of each player when player one's gradient is noise corrupted. Although player 4's action is disturbance decoupled from player 1. We note here that its cost is not. Bottom: we plot the error between individual player action from optimal action as a function of increasing noise magnitude. While player 1's action diverges as the noise increases, we note that player 4's action is completely unaffected.

ELECTRICAL & COMPUTER ENGINEERING
UNIVERSITY of WASHINGTON

ADVISORS: BEHÇET ACIKMEŞE[1], LILLIAN RATLIFF[2]

1: AERONAUTICS AND ASTRONAUTICS ENGINEERING, UNIVERSITY OF WASHINGTON

2: ELECTRICAL AND COMPUTER ENGINEERING , UNIVERSITY OF WASHINGTON