Disturbance Decoupling for Gradient-Based Multi-Agent Learning With Quadratic Costs

Sarah H. Q. Li, *Graduate Student Member, IEEE*, Lillian Ratliff, *Member, IEEE*, and Behçet Açıkmeşe, *Senior Member, IEEE*

***Abstract*—Motivated by applications of multi-agent learning in noisy environments, this letter studies the robustness of gradient-based learning dynamics with respect to disturbances. While disturbances injected along a coordinate corresponding to any individual player's actions can always affect the overall learning dynamics, a subset of players can be *disturbance decoupled*—i.e., such players' actions are completely unaffected by the injected disturbance. We provide necessary and sufficient conditions to guarantee this property for games with quadratic cost functions, which encompass quadratic one-shot continuous games, finite-horizon linear quadratic (LQ) dynamic games, and bilinear games. Specifically, disturbance decoupling is characterized by both algebraic and graph-theoretic conditions on the learning dynamics, the latter is obtained by constructing a *game graph* based on gradients of players' costs. For LQ games, we show that disturbance decoupling imposes constraints on the controllable and unobservable subspaces of players. For two player bilinear games, we show that disturbance decoupling within a player's action coordinates imposes constraints on the payoff matrices. Illustrative numerical examples are provided.**

***Index Terms*—Game theory, machine learning, decentralized control.**

I. Introduction

AS THE application of learning in multi-agent settings gains traction, game theory has emerged as an informative abstraction for understanding the coupling between algorithms employed by individual players (see, e.g., [1]–[3]). Due to scalability, a commonly employed class of algorithms in both games and modern machine learning approaches to multi-agent learning is *gradient-based learning*, in which players update their individual actions using the gradient of their objective with respect to their action. In the gradient-based learning paradigm, continuous *quadratic games* stand out as a benchmark due to their simplicity and ability to exemplify state-of-the-art multi-agent learning methods such as policy gradient and alternating gradient-descent-ascent [4].

Despite the resurgence of interest in learning in games, a gap exists between algorithmic performance in simulation and physical application in part due to disturbances in measurements [5]. Robustness to environmental noise has been analyzed in a wide variety learning paradigms [6], [7]. Most analysis focuses on independent and identically distributed stochastic noise drawn from a stationary distribution.

In contrast, we study *adversarial disturbance* without any assumptions on its dynamics or bounds on its magnitude. Though some work exists on the effects of bounded adversarial disturbance in multi-agent learning [8], there is limited understanding of how gradient disturbance propagates through the network structure as determined by the coupling of the players' objectives. Does gradient-based learning fundamentally contribute to or reduce the propagation of disturbance through player actions? Our analysis aims to answer this question for gradient-based multi-agent learning dynamics. The insights we gain provide desiderata to support algorithm synthesis and incentive design, and will lead to improved robustness of multi-agent learning dynamics.

Contributions: The main contribution is providing a novel graph-theoretical perspective for analyzing disturbance decoupling in multi-agent learning settings. For quadratic games, we obtain a necessary and sufficient condition, which can be verified in polynomial time, that ensures complete decoupling between the corrupted gradient of one player and the learned actions of another player, stated in terms of algebraic and graph-theoretic conditions. The latter perspective leads to greater insight on the types of cost coupling structures that enjoy disturbance decoupling, and hence, provides a framework for designing agent interactions, e.g., via incentive design or algorithm synthesis. Applied to LQ games, a benchmark for multi-agent policy gradient algorithms, we show that disturbance decoupling enforces necessary constraints on the controllable subspace in relation to the unobservable subspace of individual players. Applied to bilinear games, we show that disturbance decoupling enforces necessary constraints on the players' payoff matrices.

Manuscript received March 17, 2020; revised May 23, 2020; accepted May 29, 2020. Date of publication June 9, 2020; date of current version June 24, 2020. This work was supported in part by NSF under Grant CNS-1736582 and Grant CMM1-1613235. Recommended by Senior Editor V. Ugrinovskii. *(Corresponding author: Sarah H. Q. Li.)*

Sarah H. Q. Li and Behçet Açıkmeşe are with the William E. Boeing Department of Aeronautics and Astronautics, University of Washington, Seattle, WA 98195 USA (e-mail: sarahli@uw.edu; behcet@uw.edu).

Lillian Ratliff is with the Department of Electrical and Computer Engineering, University of Washington, Seattle, WA 98195 USA (e-mail: ratliffl@uw.edu).

Digital Object Identifier 10.1109/LCSYS.2020.3001240

II. Related Work

We study gradient-based learning for N–player quadratic games with continuous cost functions and action sets. Convergence guarantees for gradient-based learning are studied from numerous perspectives including game theory [1], [3], [9], control [10], and machine learning [2], [11].

Convergence guarantees for gradient-based learning dynamics under stochastic noise are studied in [2], [3], [11]. Despite being an important property to understand for adversarial disturbance, how non-stochastic noise propagates through the player network has no guarantees.

Our analysis draws on geometric control [12]–[14]. In [12], algebraic conditions for disturbance decoupling within a single dynamical system is given. In [14], disturbance decoupling for a single structured dynamical system is studied with frequency-based techniques. In this letter, we provide both the algebraic and graph-theoretic conditions for disturbance decoupling of coupled dynamical systems in gradient-based multi-agent learning.

III. Continuous Games and the Game Graph Model

Let $[N] = \{1, 2, \ldots, N\}$ denote the index set where $N \in \mathbb{N}$. For a function $f \in C^r(\mathbb{R}^n, \mathbb{R})$ with $r \geq 2$, $D_i f = \partial f / \partial x_i$ is the partial derivative with respect to x_i.

Consider an N-player continuous game $(f_1, \ldots, f_N)$ where for each $i \in [N]$, $f_i \in C^r(\mathbb{R}^n, \mathbb{R})$ with $r \geq 2$ is player i's cost function and $\mathbb{R}^n = \mathbb{R}^{n_1} \times \ldots \times \mathbb{R}^{n_N}$ is the joint action space, with $\mathbb{R}^{n_i}$ denoting player i's action space and $n = \sum_{i=1}^{N} n_i$. Each player's goal is to select an action $x_i \in \mathbb{R}^{n_i}$ to minimize its cost $f_i : \mathbb{R}^n \to \mathbb{R}$ given the actions of all other players. That is, player i seeks to solve the following optimization problem:

$$\min_{x_i \in \mathbb{R}^{n_i}} \; f_i(\underbrace{x_1, \ldots, x_i, \ldots, x_N}_{:=x}). \tag{1}$$

One of the most common characterizations of the outcome of a continuous game is a Nash equilibrium.

Definition 1 (Nash Equilibrium): For an N–player continuous game $(f_1, \ldots, f_N)$, a joint action $x^\star = (x_1^\star, \ldots, x_N^\star) \in \mathbb{R}^n$ is a Nash equilibrium if for each $i \in [N]$,

$$f_i(x^\star) \leq f_i(x_1^\star, \ldots, x_{i-1}^\star, x_i, x_{i+1}^\star, \ldots, x_N^\star), \;\; \forall \; x_i \in \mathbb{R}^{n_i}.$$

A. Gradient-Based Learning

We consider a class of simultaneous play, gradient-based multi-agent learning techniques such that at iteration k, player i receives $h_i(x^k)$ from an oracle to update its action as follows:

$$x_i^{k+1} = x_i^k - \gamma_i h_i(x_1^k, \ldots, x_N^k), \tag{2}$$

where $\gamma_i > 0$ is player i's step size,

$$h_i(x^k) = D_i f_i(x^k) + d_i^k \tag{3}$$

is player i's gradient evaluated at the current joint action x^k and affected by a player-specific, arbitrary additive disturbance $d_i^k \in \mathbb{R}^{n_i}$. In the setting we analyze, d_i^k can modify x_i^k to any other action within $\mathbb{R}^{n_i}$.

Under reasonable assumptions on step sizes—e.g., relative to the spectral radius of the Jacobian of h_i in a neighborhood of a critical point—it is known that the undisturbed dynamics converge [2], [3]. While such a guarantee cannot be given for arbitrary disturbances as considered in this letter, we provide conditions under which a subset of players still equilibriates and follows the undisturbed dynamics.

B. Quadratic Games

For an N–player continuous game $(f_1, \ldots, f_N)$, behavior of gradient-based learning around a local Nash equilibrium can be approximated by linearizing the learning dynamics, where the *linearization* corresponds to a quadratic game.

Definition 2 (Quadratic Game): For each $i \in [N]$, $f_i : \mathbb{R}^n \to \mathbb{R}$ is defined by

$$f_i(x) = x_i^\top P_i x_i + x_i^\top \Big(\sum\nolimits_{j \neq i} P_{ij} x_j + r_i\Big). \tag{4}$$

Quadratic games encompass potential games [15] with $P_{ij} = P_{ji}^\top$, and zero sum games [16] with $P_{ij} = -P_{ji}^\top$. We give further examples of quadratic games in Section III-D.

C. Game Graph

To highlight how an individual player's action updates depend on others' actions, we associate a directed graph to the gradient-based learning dynamics defined in (2).

We consider a directed graph $([N], \mathcal{E})$, where $[N]$ is the index set for the nodes in the graph, and $\mathcal{E}$ is the set of edges. Each node $i \in [N]$ is associated with action x_i of the i^{th} player. A directed edge (j, i) points from j to i and has weight matrix $W_{ij} \in \mathbb{R}^{n_i \times n_j}$, such that $(j, i) \in \mathcal{E}$ if $W_{ij} \neq 0$ element-wise. For each node i, we assume the self loop edge (i, i) always exists and has weight $W_{ii} \in \mathbb{R}^{n_i \times n_i}$. The composite matrix $W \in \mathbb{R}^{n \times n}$ with entries W_{ij} is the adjacency matrix of the *game graph*.

On a game graph, we define a path $p = (i, v_1, \ldots, v_{k-1}, j)$ as a sequence of nodes connected by edges. The set of paths $\mathcal{P}_{ij}^k$ includes all paths starting at i and ending at j, traversing $k + 1$ nodes in total. For a path $p = (i, v_1, \ldots, v_{k-1}, j)$, we define its *path weight* as the product of consecutive edges on the path, given by $W_{j,v_{k-1}} \ldots W_{v_1,i} = \prod_{l=0}^{k-1} W_{v_{l+1},v_l}$.

In the absence of disturbances d_i, the update in (2) for a quadratic game reduces to

$$x^{k+1} = W x^k - \Gamma \bar{r}, \tag{5}$$

where $\bar{r} = \begin{bmatrix} r_1^\top & \ldots & r_N^\top \end{bmatrix}^\top$, $W_{ii} = I_{n_i} - \gamma_i P_i$, $W_{ij} = -\gamma_i P_{ij}$, and $\Gamma = \text{blkdiag}(\gamma_1 I_{n_1}, \ldots, \gamma_N I_{n_N})$.

D. Subclasses of Games Within Quadratic Games

To both illustrate the breadth of quadratic games and provide exemplars of the game graph concept, we describe two important subclasses of games and their game graphs.

1) Finite Horizon LQ Game: Given initial state $z^0 \in \mathbb{R}^m$ and horizon T, each player i in an N-player, finite-horizon LQ game selects an action sequence $(u_i^0, \ldots, u_i^{T-1})$ with $u_i^t \in \mathbb{R}^{m_i}$ in order to minimize a cumulative state and control cost subjected to state dynamics:

$$\min_{u_i^t \in \mathbb{R}^{m_i}} \; \frac{1}{2}\left(\sum_{t=0}^{T} (z^t)^\top Q_i z^t + \sum_{t=0}^{T-1} (u_i^t)^\top R_i u_i^t\right)$$

$$\text{s.t.} \quad z^{t+1} = Az^t + \sum_{i=1}^{N} B_i u_i^t, \ \ t = 0, \ldots, T-1. \tag{6}$$

The LQ game defined by the collection of optimization problems (6) for each $i \in [N]$ is equivalent to a one-shot quadratic game in which each player selects $U_i = [(u_i^0)^\top, \ldots, (u_i^{T-1})^\top]^\top \in \mathbb{R}^{n_i}$ with $n_i = Tm_i$, in order to minimize their cost $f_i(U)$ defined by

$$\tfrac{1}{2}(\sum_{j=1}^{N} G_j U_j + Hz^0)^\top \bar{Q}_i (\sum_{j=1}^{N} G_j U_j + Hz^0) + \tfrac{1}{2} U_i^\top \bar{R}_i U_i,$$

where $U = (U_1, \ldots, U_N)$ is the joint action profile, and the cost matrices are given by $\bar{Q}_i = \text{blkdiag}\{Q_i, \ldots, Q_i\}$,

$$G_i = \begin{bmatrix} 0 & \ldots & 0 \\ B_i & \ldots & 0 \\ \vdots & \ddots & \vdots \\ A^{T-1}B_i & \ldots & B_i \end{bmatrix}, \ H = \begin{bmatrix} I \\ \vdots \\ A^T \end{bmatrix}, \tag{7}$$

and $\bar{R}_i = \text{blkdiag}\{R_i, \ldots, R_i\}$. This follows precisely from observing that the dynamics are equivalent to $Z = \sum_{i=1}^{N} G_i U_i + Hz^0$ where $Z = [(z^0)^\top, \ldots, (z^T)^\top]^\top$. From here, it is straight forward to rewrite the optimization problem in (6) as $\min_{U_i} f_i(U)$. The LQ game is a potential game if and only if $Q_i = Q_j$ and $R_i = R_j$ for all $i, j \in [N]$.

LQ Game Graph: Suppose each player uses step size γ_i. Since, $D_i f_i(U)$ is given by

$$(G_i^\top \bar{Q}_i G_i + \bar{R}_i) U_i + \sum_{j \neq i} G_i^\top \bar{Q}_i (G_j U_j + Hz^0), \tag{8}$$

the learning dynamics (5) are equivalent to

$$U^{k+1} = WU^k - \Gamma[\bar{Q}_1 G_1, \ldots, \bar{Q}_N G_N]^\top Hz^0, \tag{9}$$

where $W = I_n - M$, with $M \in \mathbb{R}^{n \times n}$ a blockwise matrix having entries $M_{ij} = \gamma_i G_i^\top \bar{Q}_i G_j$ if $i \neq j$ and $M_{ij} = \gamma_i (G_i^\top \bar{Q}_i G_i + \bar{R}_i)$ otherwise.

2) Bilinear Games: Bilinear games are an important class of games. For instance, a number of game formulations in adversarial learning have a hidden bilinear structure [17]. In evaluating and selecting hyper-parameter configurations in so-called *test suites*, pairwise comparisons between algorithms are formulated as bimatrix games [18], [19].

Formally, a *two player bilinear game*,[1] a subclass of continuous quadratic games, is defined by $f_1(x_1, x_2) = x_1^\top A x_2$ and $f_2(x_1, x_2) = x_1^\top B^\top x_2$ where $A \in \mathbb{R}^{n_1 \times n_2}$ and $B \in \mathbb{R}^{n_2 \times n_1}$ and $x_i \in \mathbb{R}^{n_i}$. Common approaches to learning in games [17], [20], simultaneous and alternating gradient descent both correspond to a linear system.

Game graph for simultaneous gradient play: Players update their strategies simultaneously by following the gradient of their own cost with respect to their choice variable:

$$x_1^{k+1} = x_1^k - \gamma_1 A x_2^k, \ \ x_2^{k+1} = x_2^k - \gamma_2 B x_1^k \tag{10}$$

[1]The bilinear game formulation and corresponding game graph for different gradient-based learning rules easily extend to an N-player setting, however the results in Section IV are presented for two player games.

The simultaneous gradient play game graph is given by

$$W_s = \begin{bmatrix} I & -\gamma_1 A \\ -\gamma_2 B & I \end{bmatrix}. \tag{11}$$

Game graph for alternating gradient play: In zero-sum bilinear games, it has been shown that alternating gradient play has better convergence properties [20]. Alternating gradient play is defined by

$$x_1^{k+1} = x_1^k - \gamma_1 A x_2^k, \ \ x_2^{k+1} = x_2^k - \gamma_2 B x_1^{k+1} \tag{12}$$

Examining the second player's update, we see that $x_2^{k+1} = (I + \gamma_1\gamma_2 BA)x_2^k - \gamma_2 B x_1^k$. The game graph in this case is defined by

$$W_a = \begin{bmatrix} I & -\gamma_1 A \\ -\gamma_2 B & I + \gamma_1\gamma_2 BA \end{bmatrix}. \tag{13}$$

Remark 1: Convergence of (10) and boundedness of (12) depend on choosing appropriate step sizes γ_1 and γ_2 [3], [20]. We consider disturbance decoupling for settings such as these where the undisturbed dynamics are convergent.

IV. Disturbance Decoupling on Game Graph

In this section, we derive the necessary and sufficient condition that ensures decoupling of gradient disturbance from the learning trajectory of a subset of players. We emphasize that the condition holds for disturbances with arbitrary magnitudes and functions. This is a useful result because it provides guarantees on both the equilibrium behavior and the learning trajectory under adversarial disturbance.

Definition 3 (Complete Disturbance Decoupling): Given initial joint action $x^0 \in \mathbb{R}^n$, game costs $(f_1, \ldots, f_N)$, step sizes $\Gamma \in \mathbb{R}^{n \times n}$, suppose that player i's gradient update is corrupted as in (3), then for player $j \neq i$, action x_j is decoupled from the disturbance in player i's gradient if the uncorrupted and corrupted dynamics, given respectively by

$$x^{k+1} = Wx^k - \Gamma \bar{r}, \ \ y^{k+1} = Wy^k - \Gamma \bar{r} - \Gamma d^k \tag{14}$$

result in identical trajectories for player j when $y^0 = x^0$. That is, $y_j^k = x_j^k$ holds for all $k \geq 0$, $d^k \in \mathcal{D}_i$, where

$$\mathcal{D}_i = \{d = [d_1, \ldots, d_N] \in \mathbb{R}^n \mid d_j = 0, \forall\, j \neq i\}.$$

A. Algebraic Condition

We first derive an algebraic condition on the joint action space for disturbance decoupling. Define $\mathcal{M}^\perp = \{x \in \mathbb{R}^n \mid x^\top \tilde{x} = 0, \ \forall\, \tilde{x} \in \mathcal{M}\}$ and let $\text{im}(A) = \{Ax \mid x \in \mathbb{R}^n\}$ denote the image of $A \in \mathbb{R}^{m \times n}$.

Proposition 1: Consider an N-player quadratic game $(f_1, \ldots, f_N)$ as in Definition 2 under learning dynamics as given by (2), where player i experiences gradient disturbance as given by (3). Let $\mathcal{S}(i) = \{x = [x_1, \ldots, x_N] \in \mathbb{R}^n \mid x_j = 0, \forall\, j \neq i\}$ be the joint action subset. For player $j \neq i$, the following statements are equivalent:

(i) Player j is disturbance decoupled from player i.
(ii) $W^k v \in \mathcal{S}(j)^\perp$, $\forall\, v \in \mathcal{S}(i)$, $\forall\, 0 \leq k < n$.
(iii) $\text{im}(W^k E) \subseteq \text{im}(Y)$, $\forall\, 0 \leq k < n$, where $E \in \mathbb{R}^{n \times n_i}$ and $Y \in \mathbb{R}^{n \times (n - n_j)}$ are matrices such that $\text{im}(E) = \mathcal{S}(i)$ and $\text{im}(Y) = \mathcal{S}(j)^\perp$.

Proof: For a quadratic game $(f_1, \ldots, f_N)$, the learning dynamics without and with disturbances reduce to the equations in (14). Given initial joint action x^0,

$$x^k = W^k x^0 - \begin{bmatrix} W^{k-1} & \ldots & W^0 \end{bmatrix} \Gamma \begin{bmatrix} \bar{r}^\top & \ldots, & \bar{r}^\top \end{bmatrix}^\top,$$
$$y^k = x^k - \begin{bmatrix} W^{k-1} & \ldots & W^0 \end{bmatrix} \Gamma \begin{bmatrix} (d^0)^\top & \ldots, & (d^{k-1})^\top \end{bmatrix}^\top.$$

Then, Definition 3 is equivalent to $\sum_{l=0}^{M-1} W^{M-l-1} d^l \in \mathcal{S}(j)^\perp$ satisfied for $M \geq 1$ and $d^l \in \mathcal{S}(i)$. Since the condition holds for all $M \geq 1$, it is equivalent to $W^k d^l \in \mathcal{S}(j)^\perp$ for all $k \geq 0$ and $d^l \in \mathcal{S}(i)$. This is then equivalent to $W^k d^l \in \mathcal{S}(j)^\perp$ for all $0 \leq k < n$ and $d^l \in \mathcal{S}(i)$. To see this equivalence, consider the following result from Cayley-Hamilton theorem, $W^k = \sum_{l=0}^{n-1} \alpha_l W^l$ for some $\alpha_l \in \mathbb{R}$. Thus, for $k \geq n$ and any $d \in \mathcal{S}(i)$, $W^k d = \sum_{l=0}^{n-1} W^l \alpha_l d = \sum_{l=0}^{n-1} W^l \hat{d}_l$ where $\hat{d}_l = \alpha_l d \in \mathcal{S}(i)$ for $l = 0, \ldots, n-1$, which implies that $W^k d \in \mathcal{S}(j)^\perp$. This concludes the equivalence.

Finally, we note that (*iii*) is a restatement of (*ii*). Furthermore, (*iii*) can be verified in polynomial time. ■

Remark 2: In connection to geometric control theory, condition (*iii*) of Proposition 1 is equivalent to the fact that $\text{im}([E, \ldots, W^{n-1}E])$, the smallest W-invariant subspace containing $\text{im}(E)$, must be a subset of $\mathcal{S}(j)^\perp$ [12, Th. 4.6].

B. Graph-Theoretic Condition

Next we derive the graph-theoretic condition on the joint action space for disturbance decoupling.

Theorem 1: Consider an N-player quadratic game $(f_1, \ldots, f_N)$ as in Definition 2 under learning dynamics as given by (2), where player i experiences gradient disturbance as given by (3). Player $j \neq i$ is disturbance decoupled if and only if the path weights of paths with length k satisfy

$$\sum_{p \in \mathcal{P}_{ij}^k} \prod_{l=0}^{k-1} W_{v_{l+1}, v_l} = 0, \ \forall\ 0 < k < n, \tag{15}$$

where (v_l, v_{l+1}) denotes consecutive nodes on path $p = (i, v_1, \ldots, v_{k-1}, j)$.

Proof: The result follows from equivalence between Proposition 1 condition (*ii*) and (15). Note that $x \in \mathcal{S}(i)$ is equivalent to $x_\ell = 0$ for all $\ell \neq i$, and $W^k x \in \mathcal{S}(j)^\perp$ is equivalent to $(W^k x)_j = 0$ for all $n > k \geq 0$. We prove the result by induction. For $k = 0$, $(W^0 x)_j = 0\ \forall\ x \in \mathcal{S}(i)$ holds if and only if $i \neq j$. For $k > 0$, $(W^k x)_j = 0\ \forall\ x \in \mathcal{S}(i)$ is equivalent to $i \neq j$ and $(W^k)_{ji} = 0$. Suppose that for $i, j \in [N]$, $(W^k)_{ji}$ is the sum of path weights over all paths of length k, originating at i and ending at j, then $(W^{k+1})_{ji}$ is the sum of path weights over all paths of length $k+1$, originating at i and ending at j. Let $W^k = M$, then $(W^{k+1})_{ji} = \sum_{q \in [N]} M_{jq} W_{qi}$, where $M_{jq} W_{qi} \neq 0$ if and only if the sum of path weights of length k from q to j is nonzero and there is an edge from i to q. Furthermore, $M_{jq} W_{qi}$ is the sum of path weights over all paths of length $k+1$ from i to j each of which contains $v_k = q$. Since we sum over $q \in [N]$, we conclude that $(W^{k+1})_{ji}$ is the sum of all paths weights of length $k+1$ from i to j, i.e., $(i, v_1, \ldots, v_k, j) \in \mathcal{P}_{ij}^{k+1}$. ■

The concept of disturbance decoupling is quite counter-intuitive: any change in player i's action does not affect player j's action, despite f_j being implicitly dependent on x_i through the network of player cost functions. As we see from the proof of Theorem 1, this situation arises when the dependencies 'cancel' each other out, i.e., the sum of path weights from i to j is always zero for equally lengthed paths.

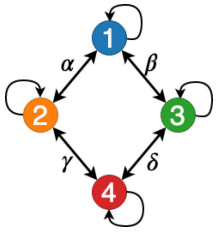

Fig. 1. A simple game graph between four players.

Example 1 (Disturbance Decoupled Players): Consider a 4 player quadratic game where $x_i \in \mathbb{R}$ and the game graph is given by Figure 1. Edge weights α, β, γ, and $\delta \in \mathbb{R}$, while each self loop has weight $w_i > 0$. Paths of length $k \leq 4$ from player 1 to player 4 are enumerated as $\mathcal{P}_{14}^1 = \{\emptyset\}$, $\mathcal{P}_{14}^2 = \{(1, 2, 4), (1, 3, 4)\}$, and $\mathcal{P}_{14}^3 = \{(1, 1, 2, 4), (1, 1, 3, 4), (1, 2, 2, 4), (1, 3, 3, 4), (1, 2, 4, 4), (1, 3, 4, 4)\}$. To satisfy Theorem 1, the sum of path weights for each $\mathcal{P}_{14}^k$ must be 0 for $0 < k < 4$. There are no paths of length one, summation for $k = 2$ implies the criteria $\alpha\gamma + \beta\delta = 0$, and summation for $k = 3$ implies the criteria $(w_1 + w_2 + w_4)\alpha\gamma + (w_1 + w_3 + w_4)\beta\delta = 0$. If $w_2 = w_3$, $\alpha\gamma + \beta\delta = 0$ is necessary and sufficient for disturbance decoupling between player 1 and player 4.

Remark 3: Disturbance decoupling is a structural property of the game in terms of disturbance propagation and attenuation. An open research problem is linking this structural property to robust decision making under uncertainties in cost parameters P_i, P_{ij} and step sizes γ_i.

The following corollary specializes to the class of potential games [15], which arise in many applications [21]–[23].

Corollary 1: Consider an N-player quadratic potential game under learning dynamics as given by (2), where player i experiences gradient disturbance as given by (3). Player i is disturbance decoupled from player $j \neq i$ if and only if player j is also disturbance decoupled from player i.

Proof: In a potential game graph, $W_{ij} = W_{ji}^\top$. Therefore, a path p with path weight $W_{j,v_{k-1}} \ldots W_{v_1,i}$ exists from i to j if and only if a path p' with path weight $W_{i,v_1}, \ldots, W_{v_{k-1},j}$ exists from j to i. Therefore, (15) holds from player i to player j if and only if it holds from player j to player i. ■

Corollary 2: Consider an N-player finite horizon LQ game as in (6) under learning dynamics as given by (9), where player i experiences gradient disturbance as given by (3), if disturbance decoupling holds between player j and gradient disturbance from player i, then

$$\begin{bmatrix} B_j^\top \\ \vdots \\ B_j^\top (A^\top)^{T-1} \end{bmatrix} Q_j \begin{bmatrix} B_i & \cdots & A^{T-1} B_i \end{bmatrix} = 0. \tag{16}$$

If Q_j is positive definite and $T \geq m$, the controllable subspace of $(\tilde{A}, \tilde{B}_i)$ must lie in the unobservable subspace of $(\tilde{B}_j^\top, \tilde{A}^\top)$ where $\tilde{A} = Q_j^{1/2} A Q_j^{-1/2}$, $\tilde{B}_i = Q_j^{1/2} B_i$, and $\tilde{B}_j = Q_j^{1/2} B_j$.

Proof: For player j to be disturbance decoupled from player i, edge (i, j) cannot exist, i.e., $-\gamma_j G_j^\top \bar{Q}_j G_i = 0$ from (7). Expanding $G_j^\top \bar{Q}_j G_i = M \in \mathbb{R}^{n_j \times n_i}$, $M_{pq} \in \mathbb{R}^{m_j \times m_i}$ is given by $\sum_{t=\min\{p,q\}}^{T-1} B_j^\top (A^\top)^{t-p} Q_j A^{t-q} B_i$. We unwrap these conditions starting from $p = T-1$, $q = T-1$; in this case $M_{pq} = B_j^\top Q_j B_i = 0$ is necessary. Then we consider $M_{T-2,T-2} = B_j^\top A^\top Q_j A B_i + B_j^\top Q_j B_i = 0$, which implies that $B_j^\top A^\top Q_j A B_i$ is necessary. Subsequently, this implies that all $B_j^\top (A^\top)^t Q_j A^t B_i = 0$ is necessary for $t \in [0, T)$. Similarly, we note that $M_{T-1,q} = B_j^\top Q_j A^q B_i = 0$ and $M_{p,T-1} = B_j^\top (A^\top)^p Q_j B_i = 0$. From these we can use the rest of M to conclude that $B_j^\top (A^\top)^p Q_j A^q B_i = 0$ for any $p, q \in [0, T)$. This condition is equivalent to (16). ∎

We apply Theorem 1 to two player bilinear games and prove a necessary condition for disturbance decoupling between different coordinates of each player's action space that is *independent* of players' step sizes.

Corollary 3: Consider a two player bilinear game under learning dynamics (10) and (12), where coordinates $x_{1,i}$ and $x_{2,i}$ experience gradient disturbance as given by (3). If $j \neq i$ and coordinate $x_{1,j}$ is disturbance decoupled from coordinate $x_{1,i}$, (A, B) must satisfy $\sum_{\ell=1}^{n_2} b_{\ell i} a_{j\ell} = 0$, where a_{pq} and b_{pq} denote the $(p, q)^{th}$ elements of A and B, respectively. Similarly, if $j \neq i$ and coordinate $x_{2,j}$ is disturbance decoupled from coordinate $x_{2,i}$, (A, B) must satisfy $\sum_{\ell=1}^{n_1} b_{j\ell} a_{\ell i} = 0$.

Proof: We construct games played by $n_1 + n_2$ players with actions $\{x_{1,1}, \ldots, x_{1,n_1}, x_{2,1}, \ldots, x_{2,n_2}\}$ and whose game graphs are identical to W_s (11) and W_a (13). First consider disturbance decoupling of $x_{1,j}$ from $x_{1,i}$. In both learning dynamics, $\{x_{1,1}, \ldots, x_{1,n_1}\}$ do not have any edges between players. Therefore, paths between $x_{1,i}$ and $x_{1,j}$ with length 2 is given by $\mathcal{P} = \{(x_{1,i}, x_{2,\ell}, x_{1,j}) \mid \ell \in [n_2]\}$. We sum path weights over $\mathcal{P}$ to obtain $\sum_{\ell=1}^{n_2} b_{\ell i} a_{j\ell} = 0$ for disturbance decoupling of $x_{1,j}$ from $x_{1,i}$ in (10) and (12). A similar argument follows for disturbance decoupling of $x_{2,j}$ from $x_{2,i}$ in (10). For disturbance decoupling of $x_{2,j}$ from $x_{2,i}$ in (12), we note that a edge from $x_{2,i}$ to $x_{2,j}$ exists with weight $\gamma_1\gamma_2(BA)_{ji}$ when $j \neq i$. Disturbance decoupling requires $\gamma_1\gamma_2(BA)_{ji} = 0$, therefore $\sum_{\ell=1}^{n_1} b_{j\ell} a_{\ell i} = 0$. ∎

Corollary 4: Consider a two player bilinear game under learning dynamics (10) and (12), where coordinates $x_{1,i}$ and $x_{2,i}$ experience gradient disturbance as given by (3). If coordinate $x_{2,j}$ is disturbance decoupled from coordinate $x_{1,i}$, (A, B) must satisfy $b_{ji} = 0$ and $\sum_{q=1}^{n_2} b_{qi} \sum_{\ell=1}^{n_1} a_{\ell q} b_{j\ell} = 0$, where a_{pq} and b_{pq} denote the $(p, q)^{th}$ elements of A and B, respectively. If coordinate $x_{1,j}$ is disturbance decoupled from coordinate $x_{2,i}$, (A, B) must satisfy $a_{ji} = 0$ and $\sum_{q=1}^{n_1} a_{qi} \sum_{\ell=1}^{n_2} b_{\ell q} a_{j\ell} = 0$.

Proof: We construct games played by $n_1 + n_2$ players with actions $\{x_{1,1}, \ldots, x_{1,n_1}, x_{2,1}, \ldots, x_{2,n_2}\}$ and whose game graphs are identical to W_s (11) and W_a (13). In both learning dynamics, disturbance decoupling requires no direct path between the decoupled players. Therefore $a_{ji} = 0$ or $b_{ji} = 0$.

Consider disturbance decoupling of $x_{1,j}$ from $x_{2,i}$ in (10), paths of length 3 from $x_{2,i}$ to $x_{1,j}$ without self loops is given by $\mathcal{P} = \{(x_{2,i}, x_{1,q}, x_{2,\ell}, x_{1,j}) \mid q \in [n_1],\ \ell \in [n_2]\}$. A path of length 3 with self loops must also include $(x_{2,i}, x_{1,j})$, whose weight is 0. We sum path weights over $p \in \mathcal{P}$ to obtain $\sum_{q=1}^{n_1} a_{qi} \sum_{\ell=1}^{n_2} b_{\ell q} a_{j\ell} = 0$. A similar argument is made for disturbance decoupling of $x_{2,j}$ from $x_{1,i}$ in (10).

Consider disturbance decoupling of $x_{2,j}$ from $x_{1,i}$ in (12), paths of length 2 from $x_{1,i}$ to $x_{2,j}$ without self loops is given by $\mathcal{Q} = \{(x_{1,i}, x_{2,q}, x_{2,j}) | q \in [n_2]\}$. A path of length 2 with self loops must also include $(x_{1,i}, x_{2,j})$, whose weight is 0. Weight of $(x_{2,q}, x_{2,j})$ is given by $\gamma_1\gamma_2(BA)_{jq} = \gamma_1\gamma_2 \sum_{\ell=1}^{n_1} b_{j\ell} a_{\ell q}$. We sum path weights over $p \in \mathcal{Q}$ to obtain $\sum_{q=1}^{n_2} b_{qi} \sum_{\ell=1}^{n_1} a_{\ell q} b_{j\ell} = 0$. A similar argument is made for disturbance decoupling of $x_{1,j}$ from $x_{2,i}$ in (12). ∎

V. Numerical Example

We provide an example of disturbance decoupling in a LQ game. Consider a tug-of-war game in which a single target $z \in \mathbb{R}^2$ is controlled by four players. We assume that player i can move z along vector $B_i \in \mathbb{R}^2$ by $u_i \in \mathbb{R}$, and that z is stationary without any player input, i.e., $A = I$. Starting with a randomized initial condition z^0, at each step t, the target moves according to the dynamics $z^{t+1} = z^t + \sum_{i=1}^4 B_i u_i^t$ where $B_1 = [1, 0]^\top$, $B_2 = [\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]^\top$, $B_3 = [\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]^\top$, $B_4 = [0, 1]^\top$. Each player i's cost function is given by

$$\tfrac{1}{2}\left\|z^9 - c_i\right\|_2^2 + \sum_{t=0}^{8} \tfrac{1}{2}\left\|z^t - c_i\right\|_2^2 + 10\left\|u_i^t\right\|_2^2$$

which describes player i's objective to move target z towards $c_i \in \mathbb{R}^2$ in a finite time $T = 10$ by using minimal amount of control. By designing the game dynamics to satisfy Theorem 1, we ensure that player 4's action is disturbance decoupled from player 1's.

Using the equivalent formulation as described in Section III-D1, $D_i f_i(U) = (G_i^\top \bar{Q}_i G_i + \bar{R}_i)U_i + \sum_{j \neq i} G_i^\top \bar{Q}_i (G_j U_j + Hz^0 - C_i)$ where $C_i = [c_i^\top, \ldots, c_i^\top]^\top$. Hence, the learning dynamics are $U^{k+1} = WU^k + \Gamma \bar{Q}_i [G_1, \ldots, G_N]^\top [(Hz^0 - C_1)^\top, \ldots, (Hz^0 - C_N)^\top]^\top$, where $W_{ij} = G_i^\top \bar{Q}_i G_j = E \otimes B_i^\top B_j$ with $B_1^\top B_2 = B_1^\top B_3 = B_2^\top B_4 = \frac{1}{\sqrt{2}}$, $B_2^\top B_3 = B_1^\top B_4 = 0$, $B_3^\top B_4 = -\frac{1}{\sqrt{2}}$, $B_1^\top B_1 = B_2^\top B_2 = B_3^\top B_3 = B_4^\top B_4 = 1$, and

$$E = \begin{bmatrix} 9 & 8 & 7 & \ldots & 1 \\ 8 & 8 & 7 & \ldots & 1 \\ 7 & 7 & 7 & \ddots & 1 \\ \vdots & & \ddots & & 1 \\ 1 & & \ldots & & 1 \end{bmatrix} \in \mathbb{R}^{9 \times 9}.$$

To ensure convergence of the undisturbed learning dynamics [3], we use uniform step sizes such that $\Gamma = \text{blkdiag}(\gamma_1 I, \ldots, \gamma_4 I)$ with $\gamma_i = \frac{\sqrt{\alpha}}{\beta}$, where $\alpha = \lambda_{\min}(\frac{1}{4}(W + W^\top)^\top (W + W^\top))$ and $\beta = \lambda_{\max}(W^\top W)$ with $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$ denoting the maximum and minimum eigenvalues of their arguments, respectively. The associated game graph is given in Figure 1, where $\alpha = \beta = \gamma = \frac{1}{\sqrt{2}}E$ and $\delta = -\frac{1}{\sqrt{2}}E$. A path $p = (1, v_1, \ldots, v_{k-1}, 4)$ of length k must have path weight $(\frac{-1}{\sqrt{2}})^{m_\delta}(\frac{1}{\sqrt{2}})^{m_\gamma} E^k$, where m_δ (m_γ) denotes the number of times the edge with weight δ (γ) is traversed in p.

Disturbance decoupling between players 1 and 4 is guaranteed if all paths of length $k \in (0, 36)$ satisfy (15). We can

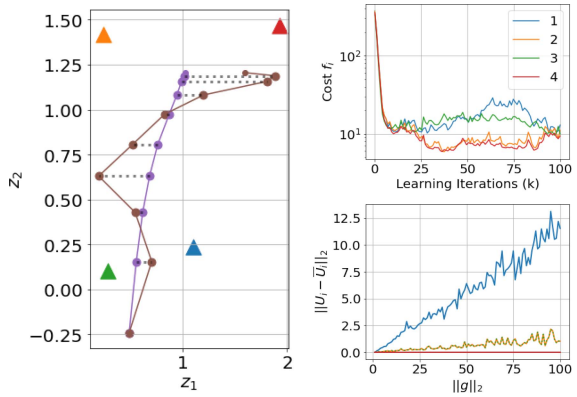

Fig. 2. Left: Trajectory of z with and without disturbances. Players' preferred destinations are given by triangles. Top right: Players' game costs during learning. Bottom right: Players' control error as a function of disturbance magnitude.

numerically verify that Proposition 1 is satisfied or make the following graph-theoretic observations based on Theorem 1. First, due to the symmetry within the game graph, the existence of path $p = (1, v_1, \ldots, v_{k-1}, 4)$ with path weight $L = (\frac{-1}{\sqrt{2}})^{m_\delta}(\frac{1}{\sqrt{2}})^{m_\gamma}E^k$ implies the existence of path $\hat{p} = (1, \hat{v}_1, \ldots, \hat{v}_{k-1}, 4)$ with path weight $\hat{L} = (\frac{-1}{\sqrt{2}})^{\hat{m}_\delta}(\frac{1}{\sqrt{2}})^{\hat{m}_\gamma}E^k$, where $m_\gamma = \hat{m}_\delta$ and $m_\delta = \hat{m}_\gamma$. Second, since edges $(3, 4)$ and $(2, 4)$ form a cut between player 1 and player 4 in the game graph, any path between them has the property that $m_\gamma + m_\delta$ is odd. From these observations, we can conclude that $L = -\hat{L}$. Since each path p of length k and weight L can be paired with path $\hat{p}$ of equivalent length k and weight $\hat{L} = -L$, we conclude that all path sets $\mathcal{P}_{14}^k$ where $k > 0$ must satisfy Theorem 1.

To numerically verify disturbance decoupling, we simulate the uncorrupted learning trajectory of z, shown in the left plot of Figure 2 in purple. We then inject a random disturbance into player 1's gradient updates as given by (3) with increasing magnitude, and observe its effects on each player's action. A sample corrupted trajectory is shown in the left plot of Figure 2 in brown. In the bottom right plot of Figure 2, we show the total error in each player's action from to the uncorrupted optimal action. We observe that player 4 does not deviate from the optimal action, while player 1's action error increases as the disturbance magnitude increases. We note that these results hold despite the fact that gradient-based learning no longer converges. In the top right plot of Figure 2, individual player costs are compared in one round of gradient-based learning where $\|d_i\| \leq 50$ is injected. Interestingly, despite action remaining uncorrupted, player 4's cost *is* disturbance affected. Note that the disturbance decoupling in actions does not necessarily imply disturbance decoupling in costs.

VI. Conclusion

In this letter, we investigated and characterized the effects of gradient disturbances on an N–player gradient-based learning dynamics. For quadratic games, we defined disturbance decoupling for arbitrary disturbances, and showed the cost coupling structure is crucial in facilitating decoupling individual player's action from input disturbance. Our future work aims to leverage these analysis results to design incentives for players to ensure disturbance decoupling.

References

[1] D. Fudenberg, F. Drew, D. K. Levine, and D. K. Levine, *The Theory of Learning in Games*, vol. 2. Cambridge, MA, USA: MIT Press, 1998,

[2] E. Mazumdar, L. J. Ratliff, and S. S. Sastry, "On gradient-based learning in continuous games," *SIAM J. Math. Data Sci.*, vol. 2, no. 1, pp. 103–131, 2020.

[3] B. Chasnov, L. J. Ratliff, E. Mazumdar, and S. Burden, "Convergence guarantees for gradient-based learning in continuous games," in *Proc. 35th Conf. Uncertainty Artif. Intell. (UAI)*, Tel Aviv-Yafo, Israel, 2019.

[4] E. Mazumdar, L. J. Ratliff, M. I. Jordan, and S. S. Sastry, "Policy-gradient algorithms have no guarantees of convergence in linear quadratic games," in *Proc. Int. Conf. Auton. Agents Multi Agent Syst. (AAMAS)*, 2020, pp. 860–868.

[5] S. Shalev-Shwartz, O. Shamir, and S. Shammah, "Failures of gradient-based deep learning," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 3067–3075.

[6] S. Li, Y. Wu, X. Cui, H. Dong, F. Fang, and S. Russell, "Robust multi-agent reinforcement learning via minimax deep deterministic policy gradient," in *Proc. AAAI Conf. Artif. Intell.*, 2019, pp. 4213–4220.

[7] L. Bottou, "Large-scale machine learning with stochastic gradient descent," in *Proc. Comput. Stat.*, 2010, pp. 177–186.

[8] Q. Jiao, H. Modares, S. Xu, F. L. Lewis, and K. G. Vamvoudakis, "Multi-agent zero-sum differential graphical games for disturbance rejection in distributed control," *Automatica*, vol. 69, pp. 24–34, Jul. 2016.

[9] L. J. Ratliff, S. A. Burden, and S. S. Sastry, "On the characterization of local nash equilibria in continuous games," *IEEE Trans. Autom. Control*, vol. 61, no. 8, pp. 2301–2307, Aug. 2016.

[10] J. S. Shamma and G. Arslan, "Dynamic fictitious play, dynamic gradient play, and distributed convergence to nash equilibria," *IEEE Trans. Autom. Control*, vol. 50, no. 3, pp. 312–327, Mar. 2005.

[11] Z. Zhou, P. Mertikopoulos, A. L. Moustakas, N. Bambos, and P. Glynn, "Mirror descent learning in continuous games," in *Proc. 56th IEEE Conf. Decis. Control*, Melbourne, VIC, Australia, Dec. 2017, pp. 5776–5783.

[12] H. L. Trentelman, A. A. Stoorvogel, and M. Hautus, *Control Theory for Linear Systems*. London, U.K.: Springer, 2012.

[13] W. M. Wonham, "Linear multivariable control," in *Optimal Control Theory and Its Applications*. New York, NY, USA: Springer, 1974, pp. 392–424.

[14] J.-M. Dion, C. Commault, and J. Van Der Woude, "Generic properties and control of linear structured systems: A survey," *Automatica*, vol. 39, no. 7, pp. 1125–1144, 2003.

[15] D. Monderer and L. S. Shapley, "Potential games," *Games Econ. Behav.*, vol. 14, no. 1, pp. 124–143, 1996.

[16] D. B. Gillies, "Solutions to general non-zero-sum games," *Contributions to the Theory of Games*, vol. 4. Princeton, NJ, USA: Princeton Univ. Press, 1959, pp. 47–86.

[17] E.-V. Vlatakis-Gkaragkounis, L. Flokas, and G. Piliouras, "Poincaré recurrence, cycles and spurious equilibria in gradient-descent-ascent for non-convex non-concave zero-sum games," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2019, pp. 10450–10461.

[18] D. Balduzzi, K. Tuyls, J. Perolat, and T. Graepel, "Re-evaluating evaluation," in *Proc. 32nd Conf. Adv. Neural Inf. Process. Syst.*, 2018, pp. 3268–3279.

[19] D. Balduzzi *et al.*, "Smooth markets: A basic mechanism for organizing gradient-based learners," 2020. [Online]. Available: arXiv:2001.04678.

[20] J. P. Bailey, G. Gidel, and G. Piliouras, "Finite regret and cycles with fixed step-size via alternating gradient descent-ascent," 2019. [Online]. Available: arXiv:1907.04392.

[21] D. Paccagnan, B. Gentile, F. Parise, M. Kamgarpour, and J. Lygeros, "Distributed computation of generalized Nash equilibria in quadratic aggregative games with affine coupling constraints," in *Proc. IEEE 55th Conf. Decis. Control*, Las Vegas, NV, USA, 2016, pp. 6123–6128.

[22] B. Lutati, V. Levit, T. Grinshpoun, and A. Meisels, "Congestion games for V2G-enabled EV charging," in *AAAI Conf. Artif. Intell.*, 2014, pp. 1440–1446.

[23] T. Alpcan and T. Başar, *Network Security: A Decision and Game-Theoretic Approach*. New York, NY, USA: Cambridge Univ. Press, 2010.