

# Robust Stochastic Stability in Dynamic and Reactive Environments

Brandon C. Collins, Lisa Hines, Gia Barboza, and Philip N. Brown

**Abstract**—The theory of learning in games has extensively studied situations where agents respond dynamically to each other by optimizing a fixed utility function. However, in many settings of interest, agent utility functions themselves vary as a result of past agent choices. This paper develops a general framework using probabilistic coupling methods that can be used to derive the stochastically stable states of log-linear learning in certain games which feature such game-environment feedback. As a case study, we apply this framework to a simple dynamic game-theoretic model of social precautions in an epidemic and give conditions under which maximally-cautious social behavior in this model is stochastically stable.

## I. INTRODUCTION

In social systems and distributed engineered systems, collective behavior is the result of many individuals making intertwined self-interested choices. In many cases, the value of a particular choice depends not only on the current choices being made by others, but also on the history of past choices.

In principle, these socio-environmental feedback loops can be analyzed using techniques from *game theory*, which has a long history of analyzing the society-scale effects of self-interested behavior. For instance, game theory has long been used to study the spread of social conventions [1] using models such as the graphical coordination game [2] with the stochastic learning algorithm log-linear learning [3]. However, traditional analysis techniques almost uniformly assume that the game's utility functions are fixed for all time, so that the agents' choices over time can be described by a stationary Markov process. However, such analysis fails or becomes unwieldy when utility functions themselves depend on the history of play.

Analysis techniques for history-dependent games have broad potential applications. For example, in a global pandemic, the individual choice to adopt protective measures (e.g., wearing masks) may be made in response to the behavior of others and the prevalence of the disease. In turn, the prevalence of the disease is a function of the history of individual choices to adopt protective measures. As another example, game theoretic methods are frequently proposed in the area of distributed control of multiagent systems [4]–[7]. However, in a distributed control application, agents' actions may directly modify the strategic environment; for

This work was supported by the National Science Foundation under Grants #DEB-2032465 and #ECCS-2013779.

The authors are with the University of Colorado Colorado Springs, CO 80918, USA {bcollin3, lhines, gbarboza, philip.brown}@uccs.edu

instance if a search-and-rescue UAV identifies a disaster victim, that victim may be removed from the list of other UAVs' objectives. Other applications that can be modeled by history-dependent games are in machine learning [8]–[10] and biology [11], [12].

Owing in part to the challenges of modeling the complex game-environment feedback inherent to history-dependent games, general results on these games are elusive. Recent work has focused on specific learning algorithms and strategic environments, such as zero-sum games under replicator dynamics [13]. In [14] the authors characterize an oscillating tragedy of the commons effect under certain environmental feedback scenarios.

In this paper, we develop a general framework for analyzing the long-run behavior of binary-action history-dependent games. In particular, we study the stochastically stable states of the popular log-linear learning algorithm in such settings. We show that if the utility functions of the history-dependent game can be appropriately referenced to the utility functions of a corresponding exact potential game, then the history-dependent game of interest inherits the stochastically stable states of the reference potential game. To accomplish this we apply techniques from the theory of probabilistic couplings, and derive a monotone coupling that relates play in the history-dependent game with that in the reference potential game. To showcase an application of the framework, we present an epidemic model that intertwines the compartmental SIS disease model with a graphical coordination game. Using our analysis framework we provide conditions under which the stochastically stable states may be fully characterized, despite their history-dependence.

## II. MODEL

### A. Game Formulation

In this work we consider binary action games. Let  $N = \{1, 2, 3, \dots, |N|\}$  denote the player set; player  $i \in N$  has action set  $A_i = \{0, 1\}$ . The joint action space is then given by  $A = \{0, 1\}^{|N|}$ . We denote an action profile as  $a \in A$  and use  $a_i$  to denote player  $i$ 's action. The actions of all other players is then given by  $a_{-i} = (a_1, a_2, \dots, a_{i-1}, a_{i+1}, \dots, a_{|N|})$ . We refer to the all 1 action profile as  $\vec{1} = (1)_{i=1}^{|N|}$  and similarly for the all zero profile,  $\vec{0}$ . Further, let  $\Delta(A)$  denote the standard probability simplex over  $A$ .

Let  $U_i : A \rightarrow \mathbb{R}$  be player  $i$ 's utility function. We denote  $U = \{U\}_{i \in N}$  as the collection of all players' utility function.

A game  $g = (N, A, U)$  is an *exact potential game* if there exists a potential function  $\phi$  such that

$$U_i(a'_i, a_{-1}) - U_i(a_i, a_{-1}) = \phi(a'_i, a_{-1}) - \phi(a_i, a_{-1}) \quad (1)$$

for any  $a \in A$ , and  $a_i, a'_i \in A_i$ .

In this work we generalize the above by allowing each history of play to have a unique utility function. We write  $\mathcal{A}_T$  to denote the set of joint action histories of length  $T \in \mathbb{N}$ , and denote the set of all histories as  $\mathcal{A} = \cup_{T \in \mathbb{N}} \mathcal{A}_T$ . We write  $\alpha \in \mathcal{A}_T$  to refer to a history of action profiles (or *path*) and use superscripts to denote time indices so that  $\alpha = (\alpha^1, \dots, \alpha^T)$ . We abuse notation and write  $\alpha^T$  to denote the last action profile in any path  $\alpha$ . We also define  $A, \mathcal{A}_T$  as partially ordered sets by first defining partial order  $\geq_A$ , where  $a' \geq_A a$  whenever  $a, a' \in A$  and  $a'_i \geq a_i$  for all  $i \in N$ , recalling that  $a'_i, a_i \in \{0, 1\}$ . Using this we define partial order  $\geq_{\mathcal{A}_T}$  as  $\bar{\alpha} \geq_{\mathcal{A}_T} \alpha$  whenever  $\alpha, \bar{\alpha} \in \mathcal{A}_T$  and  $\bar{\alpha}^t \geq_A \alpha^t$  for all  $t \in \{1, 2, \dots, T\}$ .

To model history-dependent utility functions, let  $U_i^\alpha : A \rightarrow \mathbb{R}$ , where this utility function is not only specific to player  $i$  but also to the history  $\alpha$ . Let  $U^\alpha = (U_1^\alpha, U_2^\alpha, \dots, U_{|N|}^\alpha)$  denote each player's utility function given history  $\alpha$  and let  $U^{\mathcal{A}} = \{U^\alpha \mid \alpha \in \mathcal{A}\}$  be the set of utility functions across all paths. We denote a *history-dependent game* as tuple  $(N, A, U^{\mathcal{A}})$  and let  $\mathcal{G}^{\mathcal{A}}$  be the set of all such tuples. We now present a class of games that combines potential games and history dependence.

*Definition 1:* We call a tuple  $g = (N, A, U^{\mathcal{A}}) \in \mathcal{G}^{\mathcal{A}}$  an *aligned history-dependent game* if there exists an exact potential game  $\hat{g} = (N, A, \hat{U})$  with potential function  $\hat{\phi}$  such that:

- 1)  $\{\vec{1}\} = \arg \max_{z \in A} \hat{\phi}(z)$
- 2)  $U_i^\alpha(1, \alpha_{-i}^T) \geq \hat{U}_i(1, a_{-i})$
- 3)  $\hat{U}_i(0, a_{-i}) \geq U_i^\alpha(0, \alpha_{-i}^T)$

for any  $\alpha \in \mathcal{A}$ ,  $a, a' \in A$ ,  $T \in \mathbb{N}$  such that  $\alpha_{-i}^T \geq_{A_{-i}} a_{-i}$  and  $a, a'$  vary by only a unilateral deviation. For convenience, we denote ordering  $\geq_{A_{-i}}$  over  $A_{-i} = \{0, 1\}^{|N|-1}$  equivalently to  $\geq_A$ .

## B. Learning in Games

In this work we focus on the learning algorithm log-linear learning, which is a discrete time asynchronous learning algorithm [3], [15]. That is, for game  $g \in \mathcal{G}^{\mathcal{A}}$  at each time step log-linear learning selects a single agent uniformly at random to update their action.

$$\mathbb{P}_i^\alpha(a_i) = \frac{\exp(\frac{1}{\tau} U_i^\alpha(a_i, \alpha_{-i}^T))}{\sum_{a'_i \in A_i} \exp(\frac{1}{\tau} U_i^\alpha(a_i, \alpha_{-i}^T))}, \quad (2)$$

where  $\exp(x) := e^x$ , and  $\tau$  is the *temperature*, a parameter which governs the rationality of agents. As  $\tau \rightarrow 0$  agents best respond with high probability, and as  $\tau \rightarrow \infty$  agents choose actions uniformly at random. Note that we take the last action profile in the history  $\alpha^T$  as the behavior of the other agents. The probability that  $\alpha \in \mathcal{A}_T$  transitions to  $a' \in A$  under log-linear learning in a single transition is

$$P^\alpha(a') = \begin{cases} \frac{1}{|N|} \sum_{j \in N} \mathbb{P}_j^\alpha(a'_j) & \alpha^T = a' \\ \frac{1}{|N|} \mathbb{P}_i^\alpha(a'_i) & a_i^T \neq a'_i, \alpha_{-i}^T = a'_{-i} \\ 0 & \text{else.} \end{cases} \quad (3)$$

This can be interpreted as the probability that given history  $\alpha$  the next action profile  $\alpha^{T+1} = a'$ .

We say  $a \in A$  is *strictly* stochastically stable if the following definition holds, due to [16]. For any  $\epsilon > 0$  there exists  $\mathcal{T} > 0, T < \infty$  such that

$$\Pr(s(t; \tau, \pi, g) = \vec{1}) > 1 - \epsilon \text{ whenever } t > T, \tau < \mathcal{T} \quad (4)$$

where  $s(\cdot)$  is a random variable representing the action profile at time  $t$  under log-linear learning, given temperature  $\tau$ , initial distribution  $\pi \in \Delta(A)$  and game  $g$ .

Exact potential games under log-linear learning may be analyzed using a theory of *resistance trees* [1], [3], [15], [17] to relate potential function maximizers to stochastic stability. However, this analysis depends on the fact that log-linear learning induces an ergodic Markov process on the action profiles of any exact potential game, and it is unclear how to apply resistance tree techniques generally on history-dependent games to show stochastic stability.

## III. MAIN CONTRIBUTION

We now present our main result, giving that  $\vec{1}$  is stochastically stable in aligned history-dependent games

*Theorem 1:* If  $g \in \mathcal{G}^{\mathcal{A}}$  is an aligned history-dependent game then  $\vec{1}$  is strictly stochastically stable in  $g$  under log-linear learning.

The proof of Theorem 1 proceeds using Lemma 1, which we present here and prove in Section V. The interpretation of this lemma is that for an aligned history dependent game  $g$ , the probability at any time step that  $g$  is in the  $\vec{1}$  action profile is lower bounded by the probability its associated exact potential game  $\hat{g}$  is in the  $\vec{1}$  profile.

*Lemma 1:* If  $g \in \mathcal{G}^{\mathcal{A}}$  is an aligned history-dependent game with associated exact potential game  $\hat{g}$  then  $\Pr(s(T; \tau, \pi, g) = \vec{1}) \geq \Pr(s(T; \tau, \pi, \hat{g}) = \vec{1})$  for any temperature  $\tau > 0, \pi \in \Delta(A), T \in \mathbb{N}$ .

The proof of Lemma 1 is technically involved and depends on our novel monotone coupling framework which we present in Section V. Using this result we now present a straightforward proof of Theorem 1.

*Proof of Theorem 1:* Let  $g \in \mathcal{G}^{\mathcal{A}}$  be an aligned history-dependent game and  $\hat{g}$  be its associated exact potential game. It is well-known [15] that in an exact potential game  $a \in A$  is a stochastically stable state under log-linear learning if

$$a \in \arg \max_{a' \in A} \phi(a'). \quad (5)$$

Therefore, because  $\vec{1}$  is the lone maximizer of  $\hat{\phi}$ , it is strictly stochastically stable. We apply Lemma 1 directly to the definition of strict stochastic stability in (4). For any  $\epsilon > 0$  there exists  $\mathcal{T} > 0, T < \infty$  such that

$$\Pr(s(t; \tau, \pi, g) = \vec{1}) \geq \Pr(s(t; \tau, \pi, \hat{g}) = \vec{1}) > 1 - \epsilon \quad (6)$$

for all  $t > T, \tau < \mathcal{T}$ , yielding stochastic stability of  $\vec{1}$  in game  $g$ . ■

## IV. A SOCIAL DISTANCING EXAMPLE

To highlight Theorem 1's ability to analyze stochastic stability of history-dependent games, we exhibit a case study on a simple model of epidemics. One challenge of epidemic modeling is to account for the interplay between epidemic severity and the voluntary adoption of preventative social conventions. For example, in the absence of a epidemic

people prefer not to wear masks; however, in a widespread epidemic people may prefer to take preventative measures. To model this phenomenon we intertwine the SIS compartmental epidemic model and the graphical coordination game (GCG) which models the spread and adoption of the relevant preventative social conventions; we term this model SIS-GCG. The fraction of individuals in the society susceptible to infection is described by the nonlinear differential equation

$$\dot{s} = (1-s)(\gamma - \beta(t)s), \quad (7)$$

where  $\gamma > 0$  is the curing rate and  $\beta(t) > 0$  is a rate of infection which depends on agent actions. The action 1 represents a “safe convention” action in which a player is acting to reduce contagion; the action 0 represents conventions ignoring the pandemic. These actions are associated with infection coefficients  $0 < \beta_1 < \beta_0$ , respectively. Accordingly,  $\beta(t)$  is simply the average infection rate of all individuals, given their choices:

$$\beta(t) = \frac{1}{|N|} \sum_{i \in N} a_i^t \beta_1 + (1 - a_i^t) \beta_0, \quad (8)$$

where  $a_i^t$  is the action selected by player  $i$  at time  $t \in \mathbb{N}$ . Actions are selected by agents in  $N$  dynamically on undirected graph  $G = (N, E)$  according to log-linear learning (2). The utility of agent  $i$  at time  $t$  is given by

$$\bar{U}_i^\alpha(a_i^t, a_{-i}^t) = a_i^t |\mathcal{N}_i(1)| (q + I(t)) + (1 - a_i^t) |\mathcal{N}_i(0)|, \quad (9)$$

where  $\mathcal{N}_i(x) = \{j \in N \mid (i, j) \in E, a_j = x\}$  is the set of  $i$ ’s neighbors who play  $x \in \{0, 1\} = A_i$ , the fraction of infected individuals is given by  $I(t) := 1 - s(t)$ , and  $q \in (0, 1]$  represents the agent’s willingness to practice safe conventions in the absence of an epidemic.

*Proposition 1:* If  $s(0) \in [0, 1)$ , then if  $s(t)$  is a solution of (7) with  $\beta(t)$  given by (8), there exists a  $\bar{t}$  such that  $s(t) \leq \gamma/\beta_1$  for all  $t \geq \bar{t}$  almost surely.

This proof is omitted for space and appears online [18].

It can be seen from (9) that SISGCG can be represented by a history-dependent game, as the utility function depends on the history of play, so our Theorem 1 allows us to reference SISGCG to a related exact potential game and deduce conditions guaranteeing that  $\vec{1}$  is strictly stochastically stable.

*Proposition 2:* Let  $g^S$  be an instance of SISGCG. If  $\beta_1/\gamma > 1$ ,  $q + \gamma/\beta_1 > 1$  and  $I(0) > 0$  then  $\vec{1}$  is stochastically stable in  $g$ .

*Proof:* Let the SISGCG model be denoted by  $g^S$ , which played on graph  $G = (N, E)$  with  $q + \gamma/\beta_1 > 1$  and  $I(0) > 0$ , and we consider  $g^S$  as played after time  $\bar{t}$  as given by Proposition 1. Game  $g^S$  is a history-dependent game since (9) depends on  $I(t)$ , which is itself a function of the history  $\alpha$ . Thus we have  $g^S = (N, A, \bar{U}) \in \mathcal{G}^A$  where we let  $\bar{U} = \{\bar{U}^\alpha \mid \alpha \in \mathcal{A}\}$ .

Now we let  $\hat{g}^S = (N, A, \hat{U}^S)$  be a graphical coordination game played on graph  $G$ , where the utility function  $\hat{U}^S$  is given by (9) with  $I(t) = \gamma/\beta_1$ . Standard results give that  $\hat{g}^S$  is an exact potential game and that  $\vec{1}$  is its lone potential function maximizer [1].

We now use  $\hat{g}^S$  to show  $g^S$  is an aligned history-dependent game. Now we verify  $U_i^\alpha(1, \alpha_{-i}^T) \geq \hat{U}_i^S(1, a_{-1})$  anytime  $\alpha_{-i}^T \geq_{A_{-i}} a_{-i}$ ,  $t > \bar{t}$ . This can be rewritten for  $t > \bar{t}$  as

$$\sum_{j \in \mathcal{N}_i(1; \alpha_{-i}^T)} q + I(t) \geq \sum_{j \in \mathcal{N}_i(1; a_{-i})} q + \gamma/\beta_1 \quad (10)$$

where  $\mathcal{N}_i(1; a_{-i})$  denotes the neighbors of  $i$  who are playing 1 given profile  $a$ . This expression holds because  $\alpha_{-i}^T \geq_{A_{-i}} a_{-i} \Rightarrow |\mathcal{N}_i(1; \alpha_{-i}^T)| \geq |\mathcal{N}_i(1; a_{-i})|$  and by Proposition 1. An argument with the same structure holds for  $\bar{U}_i^\alpha(0, \alpha_{-i}^T) \leq \hat{U}_i^S(0, a_{-1})$ . Thus  $g^S$  is an aligned history-dependent game, and Theorem 1 gives  $\vec{1}$  is strictly stochastically stable. ■

## V. PROOF OF LEMMA 1

### A. A Primer on Monotone Couplings

We begin with the definition of a monotone coupling, the core analytical device for our paper.

*Definition 2:* Let  $X$  be a countable set with partial ordering  $\leq_X$  and  $p_1, p_2$  be probability measures on measure space  $(X, \mathcal{F})$ . Then a *monotone coupling* of  $p_1, p_2$  is a probability measure  $p$  on  $(X^2, \mathcal{F}^2)$  satisfying the following for all  $x, y \in X$

$$\sum_{x \leq_X y'} p(x, y') = p_2(y') \text{ and } \sum_{y \geq_X x'} p(x', y) = p_1(x'). \quad (11)$$

A monotone coupling is a useful tool for analyzing the component probability measures  $p_1$  and  $p_2$ . In particular the following property holds in general for monotone couplings.

*Proposition 3* (Paarporn et al., [19]): Let  $p_1, p_2$  be probability measures on  $(X, \mathcal{F})$ . If  $p$  is a monotone coupling of  $p_1, p_2$  then for any increasing random variable  $Z : X \rightarrow \mathbb{Z}_+$  we have

$$\mathbb{E}_{p_1}(Z) - \mathbb{E}_{p_2}(Z) = \sum_{\eta=0}^{\infty} p(Z_\eta^c, Z_\eta) \quad (12)$$

where  $Z_\eta = \{a \mid Z(a) > \eta\}$ .

Here we denote  $Z^c := X \setminus Z$  to be the complement set of  $Z \subseteq X$ . The proof is given in [19, Proposition 1].

### B. Notation Required for Proofs

Taking  $\hat{g} = (N, A, U)$ , we give equations analogous to (2), (3) that give the transition probabilities for  $\hat{g}$  under log-linear learning. In particular, if agent  $i$  is selected to update her action then she will do so with probabilities given by

$$\hat{P}_i^a(a_i) = \frac{\exp(\frac{1}{\tau} U_i(a_i, a_{-i}))}{\sum_{a'_i \in A_i} \exp(\frac{1}{\tau} U_i(a'_i, a_{-i}))} \quad (13)$$

Building on (13), we define the probability that action profile  $a$  transitions to  $a'$  under log-linear learning in a single transition as

$$\hat{P}^a(a') = \begin{cases} \frac{1}{|N|} \sum_{j \in N} \hat{P}_j^a(a_j) & a = a' \\ \frac{1}{|N|} \hat{P}^a(a'_i) & a_i \neq a'_i, a_{-i} = a'_{-i} \\ 0 & \text{else} \end{cases} \quad (14)$$

for some  $i \in N$  and  $a, a' \in A$ . Additionally, we define the probability that path  $\alpha \in \mathcal{A}_T$  occurs with initial distribution  $\pi \in \Delta(A)$  as

$$\hat{P}_\pi(\alpha) = \pi(\alpha^1) \prod_{t=1}^{T-1} \hat{P}^{\alpha^t}(\alpha^{t+1}) \quad (15)$$

noting that  $\pi(\alpha^1)$  denotes the probability of  $\alpha^1$  in initial distribution  $\pi$ .

Correspondingly, the probability that path  $\alpha \in \mathcal{A}_T$  occurs with initial distribution  $\pi \in \Delta(A)$  on  $g \in \mathcal{G}^A$  is

$$P_\pi(\alpha) = \pi(\alpha^1) \prod_{t=1}^{T-1} P^{\alpha^{\leq t}}(\alpha^{t+1}) \quad (16)$$

where we use  $\alpha^{\leq t} \in \mathcal{A}_t$  to mean history  $\alpha$  until time  $t \in \{1, 2, 3, \dots, T\}$ .

We now present a result connecting the utility conditions of aligned history-varying potential games to (13) and (2).

**Lemma 2:** Let  $g = (N, A, U^A) \in \mathcal{G}^A$ ,  $\hat{g} = (N, A, \hat{U})$  be an exact potential game and let  $i \in N, a \in A, \alpha \in \mathcal{A}$  such that  $\alpha_{-i}^T \geq_A a_{-i}$ . If  $U_i^\alpha(1, \alpha_{-i}^T) \geq \hat{U}_i(1, a_{-i})$  and  $\hat{U}_i(0, a_{-i}) \geq U_i^\alpha(0, \alpha_{-i}^T)$  then  $\mathbb{P}_i^\alpha(1) \geq \hat{\mathbb{P}}_i^a(1)$ .

This proof is omitted for space and appears online [18].

Our framework requires a careful partitioning of the action space corresponding to different types of agent action deviations. Let  $f: A \rightarrow 2^A$  be defined as  $f(a) = \{a' \in A \mid a_i \neq a'_i, a_{-i} = a'_{-i} \text{ for } i \in N\}$  be the set of action profiles reachable from  $a$  via exactly one unilateral deviation. For  $a, a' \in A$  let

$$g(a, a') = \begin{cases} i & a_i \neq a'_i \\ 0 & a = a' \end{cases} \quad (17)$$

indicate which agent unilaterally deviated their action between action profiles  $a, a'$ .

Now, let  $a, a' \in A$  where  $a' \geq_A a$ . We denote several disjoint subsets of  $f(a)$ :

- 1)  $r(a) = \{z \in f(a) \mid a_{g(a,z)} = 1\}$ ,
- 2)  $q(a, a') = \{z \in f(a) \mid z \leq_A a'\} \setminus r(a)$ , and
- 3)  $s(a, a') = f(a) \setminus (q(a, a') \cup r(a))$ .

These sets can be interpreted in the following way. The set  $r(a)$  is the set of action profiles which decreased with respect to  $\geq_A$  and  $q(\cdot), s(\cdot)$  both increased. Between  $q(\cdot)$  and  $s(\cdot)$ ,  $q(\cdot)$ 's action profiles remain less than  $a'$  and  $s(\cdot)$ 's profiles are greater than or incomparable to  $a'$ . We now present three more analogous sets that are disjoint subsets of  $f(a')$ :

- 1)  $R(a') = \{z \in f(a') \mid a'_{g(a',z)} = 0\}$ ,
- 2)  $Q(a, a') = \{z \in f(a') \mid z \geq_A a\} \setminus R(a')$ , and
- 3)  $S(a, a') = f(a') \setminus (Q(a, a') \cup R(a))$ .

The interpretation of these sets are reversed relative to  $r(\cdot)$ ,  $q(\cdot)$  and  $s(\cdot)$ .

We now highlight some useful features of these sets. It is evident that  $q(\cdot), r(\cdot), s(\cdot)$  are a disjoint partition of  $f(a)$ , and that  $Q(\cdot), R(\cdot), S(\cdot)$  are a disjoint partition of  $f(a')$ . For any  $a, a', a' \geq_A a$  we relate these sets by a function  $b^{a,a'}: f(a) \rightarrow f(a')$ . To evaluate  $b^{a,a'}(\bar{a})$ , identify the agent who deviated their action between  $a, \bar{a}$  and then deviate that agent's action in  $a'$ . Formally,  $b^{a,a'}(\bar{a}) = (\neg a'_{g(a,\bar{a})}, a'_{-g(a,\bar{a})})$  where we define  $\neg a_i \in \{0, 1\} \setminus \{a_i\}$  for  $a_i \in A_i = \{0, 1\}$ . In particular, this function relates the disjoint subsets of  $f(a), f(a')$  according to the following lemma.

**Lemma 3:** If  $a, a' \in A$  and  $a \leq_A a'$ , then the following statements hold:

- 1)  $b^{a,a'}: r(a) \rightarrow S(a, a')$  is a bijection,
- 2)  $b^{a,a'}: s(a, a') \rightarrow R(a')$  is a bijection, and
- 3)  $b^{a,a'}: q(a, a') \rightarrow Q(a, a')$  is a bijection.

This proof is omitted for space and appears online [18].

### C. The One-Step Couplings

To prove Lemma 1 and obtain Theorem 1, we construct a monotone coupling  $\nu_{\pi}^{\hat{g}}$  between measures  $P_\pi, \hat{P}_\pi$ . We first construct a family of monotone couplings for each one-step transition (Theorem 2), which we apply to show the coupling over histories (Theorem 3).

**Theorem 2:** Let  $g \in \mathcal{G}^A$  denote an aligned history-dependent game and  $\hat{g}$  be its associated exact potential game. Then a monotone coupling exists between  $\hat{P}^a$  and  $P^\alpha$  for any  $\alpha \in \mathcal{A}, a \in A$  whenever  $a \leq_A \alpha^T$ . This monotone coupling  $\nu^{a,\alpha}: A^2 \rightarrow [0, 1]$  is given in (25) in Figure 1.

*Proof:* Let  $a \in A, \alpha \in \mathcal{A}$  such that  $a \leq_A \alpha^T$  and let  $g \in \mathcal{G}^A$  be an aligned history-dependent game where  $\hat{g}$  is its associated exact potential game. To verify  $\nu^{a,\alpha}$  is a monotone coupling we must show the following conditions from Definition 2 for any  $\bar{a}, \bar{a}' \in A$ :

- 1)  $\nu^{a,\alpha}$  is a well-defined probability measure,
- 2)  $\sum_{z' \geq_A \bar{a}} \nu^{a,\alpha}(\bar{a}, z') = \hat{P}^a(\bar{a})$ , and
- 3)  $\sum_{z \leq_A \bar{a}'} \nu^{a,\alpha}(z, \bar{a}') = P^\alpha(\bar{a}')$ .

We begin by verifying Condition 2. We consider cases  $\bar{a} \notin (f(a) \cup \{a\})$ ,  $\bar{a} \in q, \bar{a} \in r, \bar{a} \in s$  and  $\bar{a} = a$  separately. We use the notational convention that  $q, s, Q, S$  are assumed to take arguments  $(a, \alpha^T)$  and  $r, R$  take the argument  $a, \alpha^T$  respectively. The first case represents any  $\bar{a}$  that cannot be achieved in a single unilateral deviation from  $a$ . Trivially, this gives that  $\hat{P}^a(\bar{a}) = 0$ , and thus all pairs of  $\bar{a}, z'$  must satisfy  $\nu^{a,\alpha}(\bar{a}, z') = 0$ . This holds as all parts of (25) require  $\bar{a} \in (f(a) \cup \{a\})$  except (25h), which has the desired property.

We now consider the second case that  $\bar{a} \in q$ . Note that only (25d) satisfies this condition, so

$$\begin{aligned} \sum_{z' \geq_A \bar{a}'} \nu^{a,\alpha}(\bar{a}, z') &= \nu^{a,\alpha}(\bar{a}, \alpha^T) \\ &= \hat{\mathbb{P}}_{g(a,\bar{a})}^a(\bar{a}_{g(a,\bar{a})})/|N| = \hat{P}^a(\bar{a}) \end{aligned} \quad (18)$$

as desired.

Next we consider  $\bar{a} \in r$  which satisfies (25c), (25f) uniquely since  $b^{a,\alpha^T}$  is a bijection by Lemma 3. Thus

$$\begin{aligned} \sum_{z' \geq_A \bar{a}'} \nu^{a,\alpha}(\bar{a}, z') &= \frac{1}{|N|} (\hat{\mathbb{P}}_{g(a,\bar{a})}^a(0) \\ &\quad - \hat{\mathbb{P}}_{g(a,\bar{a})}^a(0) + \hat{\mathbb{P}}_{g(\alpha^T,\bar{a}')}^a(0)) \\ &= \frac{1}{|N|} \hat{\mathbb{P}}_{g(a,\bar{a})}^a(0) = \hat{P}^a(\bar{a}) \end{aligned} \quad (19)$$

where the second equality follows as  $g(a, \bar{a}) = g(\alpha^T, \bar{a}')$  by definition of  $b^{a,\alpha^T}$ . The third equality follows as  $\bar{a} \in r \implies \bar{a}_{g(a,\bar{a})} = 0$ .

Considering  $\bar{a} \in s$ , we find only (25e) applies, thus for

$$\sum_{z' \geq_A \bar{a}'} \nu^{a,\alpha}(\bar{a}, z') = \frac{1}{|N|} \hat{\mathbb{P}}_{g(a,\bar{a})}^a(1) = \hat{P}^a(\bar{a}) \quad (20)$$

where  $\bar{a} \in s \implies \bar{a}_{g(a,\bar{a})} = 1$  or else  $\bar{a}$  would be in  $q$ .

The final case for Condition 2 is  $\bar{a} = a$ . we find cases (25a), (25b), and (25g) apply yielding:

$$\begin{aligned}
\sum_{z' \geq_A \bar{a}'} \nu^{a,\alpha}(\bar{a}, z') &= \frac{1}{|N|} \left( |N| - \sum_{z \in q \cup r} \hat{\mathbb{P}}_g^a(z_{g(a,z)}) \right. \\
&\quad \left. - \sum_{z' \in R} \hat{\mathbb{P}}_g^a(\alpha^T, z')(1) \right) \\
&= \frac{1}{|N|} \sum_{z \in f(a)} \left( 1 - \hat{\mathbb{P}}_g^a(z_{g(a,z)}) \right) = \hat{P}^a(\bar{a})
\end{aligned} \tag{21}$$

where the first equality follows as sums over  $Q \cup R$  are equivalent to the sums over  $Q$  and  $R$  as  $Q, R$  are disjoint, and that  $z' \in R \Leftrightarrow z'_{g(\alpha^T, z')} = 1$  by definition of  $R$ . The second equality follows as the  $R$  sum is equivalent to one over  $s$  by bijection  $b^{a,\alpha^T}$ , and then we may combine it with the sum over  $q \cup r$ , to a sum over  $f(a)$  and  $|f(a)| = |N|$ . We omit arguments for Condition 3 as they run parallel to Condition 2.

To verify Condition 1, we consider each case of (25) separately. Equations (25b), (25d), (25e), (25f), and (25h) are trivial as these probabilities are well defined by definition. Lemma 2 provides:

$$\mathbb{P}_i^\alpha(1) \geq \hat{\mathbb{P}}_i^a(1) \Leftrightarrow \hat{\mathbb{P}}_i^a(0) \geq \mathbb{P}_i^\alpha(0) \tag{22}$$

where the right hand side follows from  $\mathbb{P}_i(1, a', w) + \mathbb{P}_i(0, a', w) = 1 = \mathbb{P}_i(1, a', w_0) + \mathbb{P}_i(0, a', w_0)$ . Equation (25a) follows directly from the hypothesis and (25c) holds from the right side of the equivalence.

The lone remaining case is (25g), for which we define sets  $N_q = \{g(a, z) \mid z \in q\}$ ,  $N_Q = \{g(\alpha^T, z) \mid z \in Q\}$  and so on for  $r, s, R, S$ . We denote unions of these sets as  $N_{qr} := N_q \cup N_r$ ,  $N_{QR} := N_Q \cup N_R$  and so on for other combinations of  $q, r, s$  and  $Q, R, S$ . Recalling  $q, r, Q, R$  are partitions over states that  $a, \alpha^T$  may transition to, similarly,  $N_{qr}, N_{QR}$  are partitions of agents whose unilateral deviations result in such transitions. This enables us to expand (25g):

$$\begin{aligned}
\nu^{a,\alpha}(\bar{a}, \bar{a}') &= \frac{1}{|N|} \left( \sum_{i \in N_{qr} \cap N_{QR}} (1 - \hat{\mathbb{P}}_i^a(\neg a_i) \right. \\
&\quad \left. - \mathbb{P}_i^\alpha(\neg \alpha_i^T)) \right. \\
&\quad \left. + \sum_{i \in N_{qr} \setminus N_{QR}} (1 - \hat{\mathbb{P}}_i^a(\neg a_i)) \right. \\
&\quad \left. + \sum_{i \in N_{QR} \setminus N_{qr}} (1 - \mathbb{P}_i^\alpha(\neg \alpha_i^T)) \right).
\end{aligned} \tag{23}$$

This expansion takes advantage of  $|N| = |f(a)|$  which allows  $|N|$  to enter the sums as 1. It now suffices to show that the summand of each sum is a well defined probability, of which the last two terms clearly are.

We begin by investigating  $i \in N_{qr} \cap N_{QR}$ . In particular, we have  $N_q = N_Q, N_s = N_R, N_r = N_S$  due to  $b^{a,\alpha^T}$  and its bijectiveness due to Lemmas 3. By disjointness of  $q, r$  we have  $N_{qr} = N_{QS}$  which we apply to  $N_{qr} \cap N_{QR} = N_{QS} \cap N_{QR} = N_Q = N_q$ . Applying definitions of  $q, Q$  we find  $i \in N_q \implies \neg a_i = 1, \neg \alpha_i^T = 0$ . Thus the summand of the first sum for  $i \in N_q$  is given by

$$1 - \hat{\mathbb{P}}_i^a(1) - \mathbb{P}_i^\alpha(0) \geq 1 - \mathbb{P}_i^\alpha(1) - \mathbb{P}_i^\alpha(0) = 0 \tag{24}$$

wherein the inequality is by (22), giving that the summands in the first term of (23) are themselves well defined proba-

bilities. As all conditions have been met,  $\nu^{a,\alpha}$  is a monotone coupling as desired.  $\blacksquare$

#### D. A monotone coupling over histories

We now present coupling  $\nu_{\pi}^{\hat{g}}$  which is constructed using the one-step coupling. Using this coupling we then go on to prove Lemma 1. We define indicator function  $\mathbb{1}$  such that  $\mathbb{1}(P) = 1$  if  $P$  is a true logical proposition and  $\mathbb{1}(P) = 0$  else.

*Theorem 3:* Let  $g \in \mathcal{G}^A$  be an aligned history-dependent game and  $\hat{g}$  be its corresponding exact potential game. Then  $\nu_{\pi}^{\hat{g}} : \mathcal{A}_T^2 \rightarrow [0, 1]$  is a monotone coupling between  $\hat{P}_{\pi}, P_{\pi}$ . This coupling is given by

$$\nu_{\pi}^{\hat{g}}(\alpha, \bar{\alpha}) = \pi(\alpha^1) \mathbb{1}(\alpha^1 = \bar{\alpha}^1) \prod_{t=1}^{T-1} \nu^{\alpha^t, \bar{\alpha}^{\leq t}}(\alpha^{t+1}, \bar{\alpha}^{t+1}) \tag{26}$$

where  $\alpha, \bar{\alpha} \in \mathcal{A}_T, \pi \in \Delta(A)$ .

*Proof:* Let  $\alpha, \bar{\alpha} \in \mathcal{A}_T$  and let  $g \in \mathcal{G}^A$ , and let  $\hat{g}$  be the corresponding exact potential game. We begin by showing that if  $\alpha \not\leq_{\mathcal{A}_T} \bar{\alpha}$ , then  $\nu_{\pi}^{\hat{g}}(\alpha, \bar{\alpha}) = 0$ . Immediately, we have  $\nu_{\pi}^{\hat{g}}(\alpha, \bar{\alpha}) = 0$  if  $\alpha^1 \neq \bar{\alpha}^1$ , so we need only consider cases where  $\alpha^1 = \bar{\alpha}^1$ . Inductively we find that if  $\alpha \not\leq_{\mathcal{A}_T} \bar{\alpha}$  there must exist some  $t \in \{1, 2, 3, \dots, T-1\}$  such that  $\alpha^t \leq_A \bar{\alpha}^t$  but  $\alpha^{t+1} \not\leq_A \bar{\alpha}^{t+1}$ , and let  $t$  be the minimal such value. In this case we have  $\nu^{\alpha^t, \bar{\alpha}^{\leq t}}(\alpha^{t+1}, \bar{\alpha}^{t+1}) = 0$  because  $\nu^{\alpha^t, \bar{\alpha}^{\leq t}}$  is a well defined monotone coupling by Theorem 2, yielding  $\nu_{\pi}^{\hat{g}}(\alpha, \bar{\alpha}) = 0$  as desired. It also follows that  $\nu_{\pi}^{\hat{g}}$  will always yield a well defined probability as it is either 0 or a product of well defined probabilities. Thus we only need to show that the marginal probabilities are preserved given by (11). We begin by showing the left equation of (11), that is:

$$\sum_{\alpha \leq_{\mathcal{A}_T} z} \nu_{\pi}^{\hat{g}}(\alpha, z) = \hat{P}_{\pi}(\alpha) \text{ for each } z \in \mathcal{A}_T \tag{27}$$

and omit the proof for the right hand equation as it proceeds identically. By inspecting (26), we only need to consider  $z$  such that  $z^1 = \alpha^1$  and  $z$  features at most a single unilateral deviation between any  $t, t+1$ . With these two conditions we rewrite

$$\begin{aligned}
\sum_{\alpha \leq_{\mathcal{A}_T} z} \nu_{\pi}^{\hat{g}}(\alpha, z) &= \sum_{\alpha \leq_{\mathcal{A}_T} z} \pi(\alpha^1) \prod_{t=1}^{T-1} \nu^{\alpha^t, z^{\leq t}}(\alpha^{t+1}, z^{t+1}) \\
&= \pi(\alpha^1) \sum_{\alpha^2 \leq_A z^2} \nu^{\alpha^1, z^{\leq 1}}(\alpha^2, z^2) \dots \\
&\quad \sum_{\alpha^T \leq_A z^T} \nu^{\alpha^{T-1}, z^{\leq T-1}}(\alpha^T, z^T).
\end{aligned} \tag{28}$$

as the combinatorial form. Critically, this allows us to apply the marginal sum properties of  $\nu^{\alpha^t, z^{\leq t}}$  from Theorem 2 for each  $t \in \{1, 2, \dots, T\}$ . First, considering the rightmost sum in (28), it holds that

$$\sum_{\alpha^T \leq_A z^T} \nu^{\alpha^{T-1}, z^{\leq T-1}}(\alpha^T, z^T) = \hat{P}^{\alpha^{T-1}}(\alpha^T). \tag{29}$$

Because this has no dependence on  $z$  we may factor out  $\hat{P}^{\alpha^{T-1}}(\alpha^T)$  and repeat the process on the new rightmost sum. After performing this process recursively on all sums, we have

$$\nu^{a,\alpha}(\bar{a}, \bar{a}') = \begin{cases} \frac{1}{|N|} \left( \mathbb{P}_{g(\alpha^T, \bar{a}')}^\alpha(1) - \hat{\mathbb{P}}_{g(a', \bar{a}')}^a(1) \right) & \bar{a} = a, \bar{a}' \in R \\ \frac{1}{|N|} \mathbb{P}_{g(\alpha^T, \bar{a}')}^\alpha(\bar{a}'_{g(\alpha^T, \bar{a}')} ) & \bar{a} = a, \bar{a}' \in Q \\ \frac{1}{|N|} \left( \hat{\mathbb{P}}_{g(a, \bar{a})}^a(0) - \mathbb{P}_{g(a, \bar{a})}^\alpha(0) \right) & \bar{a} \in r, \bar{a}' = \alpha^T \\ \frac{1}{|N|} \hat{\mathbb{P}}_{g(a, \bar{a})}^a(\bar{a}_{g(a, \bar{a})}) & \bar{a} \in q, \alpha^T = \bar{a}' \\ \frac{1}{|N|} \hat{\mathbb{P}}_{g(a, \bar{a})}^a(1) & \bar{a} = b^{a,\alpha^T}(\bar{a}'), \bar{a}' \in R \\ \frac{1}{|N|} \mathbb{P}_{g(\alpha^T, \bar{a}')}^\alpha(0) & \bar{a} \in r, \bar{a}' = b^{a,\alpha^T}(\bar{a}) \\ \frac{1}{|N|} \left( |N| - \sum_{z \in q \cup r} \hat{\mathbb{P}}_{g(a, z)}^a(z_{g(a, z)}) - \sum_{z' \in Q \cup R} \mathbb{P}_{g(\alpha^T, z')}^\alpha(z'_{g(\alpha^T, z')}) \right) & a = \bar{a}, \alpha^T = \bar{a}' \\ 0 & \text{otherwise.} \end{cases}$$

Fig. 1. The full specification of the one-step monotone coupling for Theorem 2. We adopt the notational convention that  $q, s, Q, S$  are assumed to take arguments  $a, a'$  and  $r, R$  take the argument  $a, a'$  respectively.

$$\sum_{\alpha \leq \mathcal{A}_T z} \nu_{\pi}^{\hat{g}}(\alpha, z) = \pi(\alpha^1) \prod_{t=1}^{T-1} \hat{P}^{\alpha^t}(\alpha^{t+1}) = \hat{P}_{\pi}(\alpha) \quad (30)$$

as desired, noting we accounted for the indicator functions in  $\nu_{\pi}^{\hat{g}}$ . This concludes the proof of Theorem 3.  $\blacksquare$

Now that the necessary results have been developed we proceed with the proof of Lemma 1.

*Proof of Lemma 1:* Let  $g \in \mathcal{G}^A$  be an aligned history-dependent game and  $\mathcal{I} \subset \mathcal{A}_T$  be an upper set. Define  $\mathbb{1}_{\mathcal{I}}(\alpha) := \mathbb{1}(\alpha \in \mathcal{I})$  as an indicator function. Consider probability measures  $P_{\pi}, \hat{P}_{\pi}$  coupled by  $\nu_{\pi}^{\hat{g}}$  in Theorem 3, we have

$$\begin{aligned} P_{\pi}(\mathcal{I}) - \hat{P}_{\pi}(\mathcal{I}) &= \mathbb{E}_{P_{\pi}}(\mathbb{1}_{\mathcal{I}}) - \mathbb{E}_{\hat{P}_{\pi}}(\mathbb{1}_{\mathcal{I}}) \\ &= \nu_{\pi}^{\hat{g}}(\mathcal{I}^C, \mathcal{I}) \geq 0. \end{aligned} \quad (31)$$

where the second equality follows by Proposition 3 as  $\mathbb{1}_{\mathcal{I}}$  is increasing in  $\mathcal{A}_T$ . Note (31) runs parallel to the proof of [19, Corollary 3]. That is, for any upper set  $\mathcal{I} \subset \mathcal{A}_T$  we have

$$P_{\pi}(\mathcal{I}) \geq \hat{P}_{\pi}(\mathcal{I}). \quad (32)$$

Let  $((\vec{0})_{t=1}^{T-1}, \vec{1}) \in \mathcal{I}$ . This induces  $\mathcal{I}$  such that it includes every path such that at time  $T$  the  $\vec{1}$  state is played. This yields the following interpretation

$$P_{\pi}(\mathcal{I}) = \Pr(s(T; \tau, \pi, g) = \vec{1}) \quad (33)$$

representing the probability that at time  $T$  game  $g$  is in the  $\vec{1}$  action profile given initial distribution  $\pi \in \Delta(A)$  and learning temperature parameter  $\tau$ . Noting a parallel interpretation to (33) holds for  $\hat{P}_{\pi}, \hat{g}$ , we apply these to (32) to obtain

$$\Pr(s(T; \tau, \pi, g) = \vec{1}) \geq \Pr(s(T; \tau, \pi, \hat{g}) = \vec{1}) \quad (34)$$

as desired.  $\blacksquare$

## REFERENCES

- [1] H. P. Young, "The evolution of conventions," *Econometrica: Journal of the Econometric Society*, pp. 57–84, 1993.
- [2] M. Kearns, M. L. Littman, and S. Singh, "Graphical models for game theory," in *UAI'01*, (San Francisco, CA, USA), p. 253–260, Morgan Kaufmann Publishers Inc., 2001.
- [3] J. R. Marden and J. S. Shamma, "Revisiting log-linear learning: Asynchrony, completeness and payoff-based implementation," *Games and Economic Behavior*, vol. 75, no. 2, pp. 788–808, 2012.
- [4] R. Chandan, D. Paccagnan, and J. R. Marden, "When Smoothness is Not Enough: Toward Exact Quantification and Optimization of the Price-of-Anarchy," in *58th IEEE CDC*, pp. 4041–4046, 2019.
- [5] B. C. Collins and P. N. Brown, "Exploiting an Adversary's Intentions in Graphical Coordination Games," in *ACC2020*, pp. 4638–4643, 2020.
- [6] J. R. Marden and A. Wierman, "Distributed welfare games," *Operations Research*, vol. 61, no. 1, pp. 155–168, 2013.
- [7] A. Kanakia, B. Touri, and N. Correll, "Modeling multi-robot task allocation with limited information as global game," *Swarm Intelligence*, vol. 10, no. 2, pp. 147–160, 2016.
- [8] C. Wang, C. Xu, X. Yao, and D. Tao, "Evolutionary generative adversarial networks," *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 6, pp. 921–934, 2019.
- [9] U. Garciaarena, R. Santana, and A. Mendiburu, "Evolved gans for generating pareto set approximations," in *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 434–441, 2018.
- [10] V. Costa, N. Lourenço, J. Correia, and P. Machado, "Coegan: evaluating the coevolution effect in generative adversarial networks," in *Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 374–382, 2019.
- [11] A. R. Tilman, J. R. Watson, and S. Levin, "Maintaining cooperation in social-ecological systems," *Theoretical Ecology*, vol. 10, no. 2, pp. 155–165, 2017.
- [12] A. R. Tilman, J. B. Plotkin, and E. Akçay, "Evolutionary games with environmental feedbacks," *Nature communications*, vol. 11, no. 1, pp. 1–11, 2020.
- [13] S. Skoulakis, T. Fiez, R. Sim, G. Piliouras, and L. Ratliff, "Evolutionary game theory squared: Evolving agents in endogenously evolving zero-sum games," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, pp. 11343–11351, 2021.
- [14] J. S. Weitz, C. Eksin, K. Paarporn, S. P. Brown, and W. C. Ratcliff, "An oscillating tragedy of the commons in replicator dynamics with game-environment feedback," *Proceedings of the National Academy of Sciences*, vol. 113, no. 47, pp. E7518–E7525, 2016.
- [15] C. Alós-Ferrer and N. Netzer, "The logit-response dynamics," *Games and Economic Behavior*, vol. 68, no. 2, pp. 413–427, 2010.
- [16] P. N. Brown, H. P. Borowski, and J. R. Marden, "Security against impersonation attacks in distributed systems," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 1, pp. 440–450, 2019.
- [17] B. S. Pradelski and H. P. Young, "Learning efficient nash equilibria in distributed systems," *Games and Economic behavior*, vol. 75, no. 2, pp. 882–897, 2012.
- [18] B. C. Collins, L. Hines, G. Barboza, and P. N. Brown, "Robust stochastic stability in dynamic and reactive environments," *arXiv preprint arXiv:2103.13475*, 2021.
- [19] K. Paarporn, C. Eksin, J. S. Weitz, and J. S. Shamma, "Networked SIS Epidemics with Awareness," *IEEE Transactions on Computational Social Systems*, vol. 4, no. 3, pp. 93–103, 2017.