A Coupling Approach to Analyzing Games with Dynamic Environments

Brandon C. Collins, Student Member, IEEE, Shouhuai Xu, Senior Member, IEEE and Philip N. Brown, Member, IEEE

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 real situations, the strategic environment varies as a result of past agent choices. Unfortunately, the analysis techniques that enabled a rich characterization of the emergent behavior in static environment games fail to cope with dynamic environment games. To address this, we develop a general framework using probabilistic couplings to extend the analysis of static environment games to dynamic ones. Using this approach, we obtain sufficient conditions under which traditional characterizations of Nash equilibria with best response dynamics and stochastic stability with log-linear learning can be extended to dynamic environment games. As a case study, we pose a model of cyber threat intelligence sharing between firms and a simple dynamic game-theoretic model of social precautions in an epidemic, both of which feature dynamic environments. For both examples, we obtain conditions under which the emergent behavior is characterized in the dynamic game by performing the traditional analysis on a reference static environment game.

Index Terms—Dynamic Environment, Game Theory, Multi-agent Systems, Learning in Games

I. INTRODUCTION

In social systems and distributed engineering 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 behaviors. For instance, game theory has long been used to study the spread of social conventions [2] using models such as the graphical coordination game [3] with the stochastic learning algorithm known as log-linear learning [4]. 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 analyses fail or become unwieldy when utility functions depend on the history of play.

Analysis techniques for history-dependent games have many 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 [5]–[8]. However, in a distributed control application, agents’ actions may directly modify the strategic environment; for 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 [9]–[11] and biology [12], [13].

Dynamic environments may be either random or deterministic functions of agent behaviors. The random case has been used to understand cell behavior and evolution [14], where the environment is a Markov model whose state represents the current conditions the cell inhabits. This approach is further studied in [15], which characterizes the importance of genetic mutations in dynamic environments. On the other hand, the deterministic case is often studied in the context of dynamical systems, where differential equations dictate the evolution of the environment. One such example is [16], where the authors characterize an oscillating tragedy of the commons effect under certain environmental feedback scenarios. Revisiting population games, [17] studies a zero-sum Rock Paper Scissors style game with replicator dynamics, where the environment responds negatively to the increase in frequency of any individual population. They show that the model is recurrent over time. This result is extended in [18] where the recurrence is generalized to a class of dynamics with environments and populations interacting in a networked fashion. Particularly, it is shown that the class of dynamics is equivalent to a zero-sum poly-matrix game under replicator dynamics and can be studied using traditional game theoretic techniques.

In this paper, we develop a general framework for analyzing binary-action games with dynamic environments, which we term history-dependent games. We develop a probabilistic coupling between a reference static game and a history-dependent game. We show that if the utility functions of the history-dependent game can be referenced appropriately to the static game, then traditional game theoretic results...
on the static game can be extended to the history-dependent

game. Specifically, the coupling provides a general inequality (Lemma 3) which compares the probability that the history-dependent
game is in a given state to that of the static game, for all time steps. Using this inequality, we develop sufficient
conditions under which traditional characterizations of Nash
equilibria with best response dynamics and stochastic stability
with log-linear learning can be extended to history-dependent
games.

To show the applicability of our results, we consider two
examples of history-dependent games in the contexts of cy-
ersecurity and epidemiology. The first example (given in
Section IV-A) is a Cyber Threat Intelligence (CTI) sharing
game where firms generate CTI over time and decide
whether to share CTI with each other. This problem has
been studied widely using game theory [19]–[24], but all
of these works consider a fixed value of CTI. This is an
unrealistic assumption, because different pieces of CTI may
have different intrinsic values or due to timeliness [25]. We
formulate a CTI sharing game which admits a broad class of
variable CTI functions (encoding the environment) and show
sufficient conditions under which the all-sharing action profile
is comparable to a Nash equilibrium in a fixed setting with best
response dynamics and is stochastically stable with log-linear
learning.

In the second example (developed in Section IV-B), we
study the feedback between the prevalence of an epidemic
and individuals’ adoption of preventative measures. It is known
that preventative measures like mask wearing or social distanc-
ing reduces the risk of infection spreading between individuals.
However, an individual’s willingness to practice such measures
depends on prevalence of the epidemic. That is, individuals are
more likely to adopt preventative measures when the epidemic
is widespread, which then in turn mitigates the epidemic,
creating a feedback loop. To capture this, we intertwine the
graphical coordination game model of social conventions and
the compartmental epidemic SIS model. In this new model,
the coordination game’s payoffs are impacted by the current
state of the SIS model, and the infectiousness parameter of
the SIS model depends on the current choices of agents.
We use the coupling to derive sufficient conditions under which
all individuals practice preventative conventions as the unique
stochastically stable state of log-linear learning.

II. MODEL

A. Game Formulation

In this work, we consider binary action games. Let \( N = \{1, 2, 3, \ldots, n \} \) denote the set of agents; agent \( i \in N \) 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 agent \( i \)'s action. Throughout, we use
state and action profile interchangeably to refer to a vector
of agent choices \( a \). We denote actions of all other agents by
\( a_{\bar{i}} = (a_1, a_2, \ldots, a_{i-1}, a_{i+1}, \ldots, a_N) \). We denote the all-
1 action profile as \( \bar{1} = (1)^N \) and similarly for the all-zero
profile \( \bar{0} \). Further, let \( \Delta(A) \) denote the standard probability
 simplex over \( A \). Let \( U_i : A \to \mathbb{R} \) be agent \( i \)'s utility function,
and let \( \mathcal{U} \) denote the space of all utility functions \( U_i : A \to \mathbb{R} \). We write \( U = \{U_i\}_{i \in N} \) to denote the collection of utility functions for all agents. Thus, we specify a static game using tuple \( g = (N, A, U) \).

We consider a general class of games which generalize the
preceding static game to enable the past behavior of agents to
influence the current behavior. To accomplish this, we allow
the utility function to be a function of previous actions taken
by agents. We begin by formalizing the past behavior or
history of agent actions. We use \( A_T \) to denote the set of joint
action histories of length \( T \in \mathbb{N} \), and denote the set of all
histories as \( A = \bigcup_{T \in \mathbb{N}} A_T \). We use \( \alpha \in A_T \) to refer to a history of action profiles (which we refer to as a \( \textit{path} \)) and use superscripts to denote time indices so that \( \alpha = (\alpha^1, \ldots, \alpha^T) \).

We use \( \alpha^T \) to denote the last action profile on a path \( \alpha \) of
length \( T \). We define \( A \) and \( A_T \) as partially ordered sets with
partial order \( \succeq \), where \( a' \succeq a \) when \( a_i' \in A \) and \( a_i' \geq a_i \)
for all \( i \in N \), recalling that \( a_i', a_i \in \{0, 1\} \). Using this, we
further define partial order \( A_T \succeq A_T \), such that \( \alpha \succeq \alpha' \) when
\( \alpha, \alpha' \in A_T \) and \( \alpha_i \succeq \alpha'_i \) for all \( i \in \{1, 2, \ldots, T\} \).

To model history-dependent utility functions, let \( U^\alpha \in \mathcal{U} \),
where this utility function is not only specific to agent \( i \)
but also to the history \( \alpha \). Let \( U^\alpha = (U^\alpha_0, U^\alpha_2, \ldots, U^\alpha_N) \) denote
each agent’s utility function given history \( \alpha \) and let \( U^A = \{U^\alpha \mid \alpha \in A \} \) be the set of utility functions across all paths.
We denote a \textit{history-dependent game} as tuple \( (N, A, U^A) \) and
let \( \mathcal{U}^A \) be the set of all such tuples.

In general, the analysis of history-dependent games is
difficult as the utility function can vary wildly between each
time step. Accordingly, we restrict our attention to a subset
of history-dependent games that have an important utility
function property, namely that there exists a reference static

game with certain properties relating the history-dependent
and static utility function. This is the basis for us to extend
the traditional game theoretic analysis of static games to
history-dependent games. We call this specific class \textit{aligned history-dependent games}, formally defined as follows.

\textit{Definition 1 (aligned history-dependent game):} We call a
tuple \( g = (N, A, U^A) \in \mathcal{U}^A \) an \textit{aligned history-dependent

game} if there exists a static game \( \bar{g} = (N, A, \bar{U}) \):

1) \( U^\alpha_i(1, \alpha^T_{-i}) \geq U_i(1, a_{-i}) \)

2) \( U_i(0, a_{-i}) \geq U^\alpha_i(0, \alpha^T_{-i}) \)

for any \( \alpha \in A, a, a' \in A, T \in \mathbb{N} \) such that \( \alpha^T \succeq \alpha_i \)
and \( a, a' \) vary by only a unilateral deviation. For convenience, we
denote ordering \( \succeq \) over \( A_{-i} \), \( \alpha_{-i} \) and \( a_{-i} \) equivalently to
\( \geq_{-i} \).

Since we focus on analyzing history-dependent games corre-

csponding to relatively well-understood static ones, Definition 1
ensures that a suitable static game exists for comparison.

Intuitively, Definition 1 defines a subclass of history-dependent
games such that the 1 actions benefit from history relative to
a reference static game, but the 0 actions can only lose due to
history. More specifically, it ensures that playing the 1 action
is always more desirable for agents in the history-dependent
game relative to the static game whenever \( \alpha^T \succeq \alpha_i \), \( a_{-i} \)
(i.e., when a superset of agents are playing 1 in the history-
dependent game relative to the static game). Thus, an aligned
history-dependent game has the property that for all histories, having more agents playing 1 can only make playing 1 more desirable for other agents.

B. Properties of Learning Rules in Games

To establish maximum generality for our results, we provide conditions on learning rules under which the couplings can be applied, and show that both best response and log-linear learning satisfy these conditions. We begin by defining an individual learning rule as a function which gives the probability that action profile \( a \) will deviate to \( a' \) in a single time step given that only agent \( i \) can update its action.

**Definition 2 (individual learning rule):** Function \( P_i : A \times A \times U^n \rightarrow [0,1] \) is an individual learning rule if for any \( a, a' \in A, U \in U^n \) we have

1. \( \sum_{a \in A} P_i(a, a, U) = 1 \),
2. \( P_i(a, a', U) \geq 0 \), and
3. \( P_i(a, a', U) = 0 \) if \( a_i \neq a'_i \).

The first two conditions say that \( P_i \) is a valid probability measure over \( A \) given prior action \( a \) and utility vector \( U \). The third condition ensures that only agent \( i \) changes its action with a positive probability. We now give a condition on individual learning rules to derive our main results.

**Definition 3 (local individual learning rule):** We say an individual learning rule \( P_i \) is local if there exists \( P_i : A_i \times \mathbb{R}^{|A|} \rightarrow [0,1] \) such that for any \( a \in A, a'_i \in A_i \) we have:

\[
P_i(a, a'_i, a_{-i}, U) = \hat{P}_i(a'_i, \bar{U}_i(a_{-i}))
\]  

where \( \bar{U}_i(a_{-i}) = (U_i(a_i, a_{-i}))_{a_i \in A_i} \) is a vector of payoffs for each possible action of agent \( i \) given the actions of the other agents \( a_{-i} \).

Learning rule \( P_i \) being local has several implications. First, the probability that agent \( i \) selects a given action only depends on its utility function \( U_i \), which is a property known as uncoupled [26]. Additionally, the probability does not depend on agent \( i \)'s previous action \( a_i \) (i.e., agent \( i \) cannot be biased toward their previous action in any way). We now pose an additional monotonicity condition on learning rules to ensure that increases in payoffs do not decrease the probability that an action is played.

**Definition 4 (monotone individual learning rule):** An individual learning rule \( P_i : A \times A \times U^n \rightarrow [0,1] \) is monotone with respect to utility if for any utility function vector \( U \in U^n \), individual agent’s action \( a'_i \), action profile \( a \in A \), nonnegative constant \( l \geq 0 \), and defining \( U = (U_1, U_2, \ldots, U_{i-1}, U_i, U_{i+1}, \ldots, U_n) \) and

\[
\bar{U}_i(a) = \begin{cases} 
U_i(a) + l & a_i = a'_i \\
U_i(a) & \text{else,}
\end{cases}
\]

we have \( P_i(a, a'_i, U) \geq P_i(a, a', U) \) where \( a' = (a'_i, a_{-i}) \).

The preceding condition can be interpreted as follows. First, select some action of agent \( i \), action \( a'_i \), and increase its utility such that the increase does not depend on the actions of other agents. Then, if for any increase in utility the probability agent \( i \) selects \( a'_i \) does not decrease, the learning rule is monotone with respect to utility. We now use our definition of individual learning rule to define a learning rule where all agents can update their actions.

**Definition 5 (asynchronous learning rule):** Given a vector of individual learning rules \( \hat{P} = (P_1, P_2, \ldots, P_n) \), we define an asynchronous learning rule as

\[
P(a, a', U) = \frac{1}{n} \sum_{i \in N} P_i(a, a', U).
\]  

We call \( P \) asynchronous because it has the property that \( P(a, a', u) = 0 \) if \( a, a' \) vary by more than one action, which follows by Definition 2. That is, \( P \) only permits one agent to change its action at a time. More specifically, \( P \) selects a single agent \( i \) according to a uniform distribution and updates its action according to the distribution given by \( P_i \). We say asynchronous learning rule \( P \) is local and monotone if each individual learning rule in \( \hat{P} \) is local and monotone with respect to utility.

Throughout the paper, we couple a learning rule \( P \) with game \( g \). We adopt the convention that if game \( g \in \mathcal{G}^A \) is a history-dependent game, then the associated learning rules are given by

\[
P^\alpha(a') := P(\alpha^T, a', U^\alpha).
\]

Additionally, if \( \hat{g} = (N, A, U) \) is a static game, then we similarly define

\[
\hat{P}^\alpha(a') := P(a, a', \hat{U}).
\]

Correspondingly, the individual learning rules are given by

\[
P_i^\alpha(a'_i) := P(\alpha^T, (a'_i, a_{-i}^T), U^\alpha), \quad \text{and}
\]

\[
\hat{P}_i^\alpha(a'_i) := P(a, (a'_i, a_{-i}), \hat{U}).
\]

C. Example Learning Rules

In the previous section, we defined several properties of learning rules of the form \( P : A \times A \times U \rightarrow [0,1] \). In this section, we formulate a variety of well studied learning rules with respect to our learning rule definitions and show they satisfy all of the above properties.

The first example is called log-linear learning, where each individual learning rule \( P_i \) is defined as

\[
P_i(a, a', U; \tau) = \begin{cases} 
\frac{\exp\left(\frac{1}{\tau} U_i(a')\right)}{\sum_{a'_i \in A_i} \exp\left(\frac{1}{\tau} U_i(a_i, a_{-i})\right)} & a_i = a'_i \\
0 & a_i \neq a'_i,
\end{cases}
\]

where \( \exp(x) := e^x \), and \( \tau \) is the temperature parameter that 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. Traditional log-linear learning [4] can be implemented in the framework of (3) by selecting \( \tau \) and letting \( \hat{P} = (P_1(\cdot; \tau))_{i=1}^n \). Then, \( \hat{P}(\cdot; \hat{P}) \) is equivalent to the previously studied log-linear learning function.

One of the appeals of log-linear learning is that for a special class of static games known as potential games, log-linear learning has the desirable equilibrium selection properties. Formally, a static game \( g = (N, A, U) \) is an exact potential game if there exists a potential function \( \phi \) such that

\[
U_i(a'_i, a_{-i}) - U_i(a_i, a_{-i}) = \phi(a'_i, a_{-i}) - \phi(a_i, a_{-i})
\]
for any \( a \in A \), and \( a_i, a'_i \in A_i \). Under log-linear learning in potential games, maximizers of the potential function are stochastically stable [27]. We say \( a \in A \) is strictly stochastically stable if the following holds [28]: For any \( \epsilon > 0 \) there exists \( T > 0, T < \infty \) such that

\[
\Pr(s(t; P) = \hat{1}) > 1 - \epsilon \text{ whenever } t > T, \tau < T \tag{10}
\]

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 [2], [4], [27], [29] 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, and it is unclear how to apply resistance tree techniques generally on history-dependent games to show stochastic stability. (In Theorem 3 we will show how the aligned history-dependent framework may be used to apply stochastic stability to history-dependent games.)

The second example learning rule is the well-studied best response dynamics [30]. In this learning rule agents always best respond to the actions of other agents (breaking ties uniformly randomly). An individual agent’s best response learning rule can be given by

\[
P_i(a, a', U) = \begin{cases} \frac{\exp(\frac{1}{T} U_i'(a_i'))}{\sum_{\bar{a}_i \in A_i} \exp(\frac{1}{T} U_i(\bar{a}_i, a_i'))} & a_i = a'_i \\ 0 & a_i = a'_i \end{cases} \tag{11}
\]

where \( BR_i(a_i, U) \) is agent \( i \)’s best response set given utility function \( U \) to the other agent’s actions \( a_{-i} \). The best response is implemented using (3) by letting \( \hat{P} \) be a uniform vector of the above \( P_i \) in the form of \( P(\cdot; \hat{P}) \).

The best response learning rule has the property that once it selects a strict Nash equilibrium, the process stays there for all following time steps. Formally, suppose \( a \in A \) is a strict Nash equilibrium, then

\[
\Pr(s(t + 1, \hat{P}_x) = a) \geq \Pr(s(t, \hat{P}_x) = a) \tag{12}
\]

for all times \( t \) and initial distributions \( \pi \).

To show that the preceding two example learning rules are applicable to the aligned history-dependent game framework we propose in this paper, we must show that (8) and (11) are individual, monotone with respect to utility, and local.

**Lemma 1**: Log-linear learning given by (8) and best response dynamics given by (11) are individual, monotone with respect to utility, and local.

**Proof**: We verify both log-linear learning and best response dynamics are individual, monotone with respect to utility, and local in order. Beginning with individual learning rule, we verify that log-linear learning, given by (8), satisfies the three conditions. Let \( U \in U^n \) be a vector of utility functions. The first condition of individuality is easy to verify algebraically because for some \( a \in A \) and some agent \( i \)

\[
\sum_{a' \in A} P_i(a, a', U; \tau) = \sum_{a_i' \in A_i} \frac{\exp(\frac{1}{T} U_i'(a_i', a_i))}{\sum_{\bar{a}_i' \in A_i} \exp(\frac{1}{T} U_i(\bar{a}_i, a_i'))} = 1 \tag{13}
\]

where the first equality follows as \( P(a, a', U) = 0 \) when \( a, a' \) vary by more than a single agent’s action. A similar argument holds for the best response dynamics so we omit it for space. It is easy to see the second condition of Definition 2 for both learning rules as both are strictly nonnegative by definition. The third condition holds since both learning rules explicitly have \( P_i(a, a') = 0 \) whenever \( a_{-i} \neq a_{-i}' \).

We now show both learning rules are monotone with respect to utility. Beginning with log-linear learning, define \( \hat{U}_i(a) \) according to (2) for some agent \( i \), \( l \geq 0 \), and action \( a_i' \in A_i \). Let \( U = (U_1, U_2, \ldots, U_i, \ldots, U_n) \) and let \( a' = (a_i', a_{-i}) \); we show \( P(a, a', U) \geq P(a, a', \bar{U}) \) directly:

\[
P(a, a', U) = \frac{\exp(\frac{1}{T} U_i(a_i'))}{\sum_{a_i \in A_i} \exp(\frac{1}{T} U_i(a_i, a_{-i}))} \geq \frac{\exp(\frac{1}{T} U_i(a_i'))}{\sum_{a_i \in A_i} \exp(\frac{1}{T} U_i(a_i, a_{-i}))} = P(a, a', \bar{U})
\]

for any action \( a \).

Monotonicity with respect to utility can be seen for best response dynamics in two cases. First, suppose \( a_i' \in BR_i(a_{-i}, U) \), then we have both \( a_i' \in BR_i(a_{-i}, \bar{U}) \) and \( |BR_i(a_{-i}, U)| \leq |BR_i(a_{-i}, \bar{U})| \). The first result follows intuitively as only the utility of \( a_i' \) increased from \( U_i \) to \( \bar{U}_i \) and thus \( a_i' \) must remain in the best response set. The second follows as if \( l > 0 \) then \( a_{-i}' \) becomes the unique best response and the best response set remains unchanged if \( l = 0 \). With both of these results together, we conclude \( P_i(a, a', U) \geq P_i(a, a', \bar{U}) \) from the definition of (11). The other case, \( a_i' \not\in BR_i(a_{-i}, U) \), is trivial as \( P_i(a, a', U) = 0 \) is a lower bound of the nonnegative function \( P(\cdot, U) \).

Locality for both log-linear learning and best response dynamics can be seen directly from their definitions, as they only depend on selected action \( a_i' \) and utilities of the form \( U(\cdot, a_{-i}) \) for prior action profile \( a \). Thus, log-linear learning and best response learning are individual learning rules, monotone with respect to utility, and local. ■

### III. Main Contribution

To characterize history-dependent game \( g \) using a reference game \( \hat{g} \) as per the definition of aligned history-dependent games, we develop a monotone coupling between \( P_\tau \) and \( \hat{P}_\tau \). Because developing the coupling is technically involved, we defer it to Section V; in this section we present the game theoretic significance of coupling. We begin by giving a broad result which follows immediately from the existence of the coupling to relate increasing metrics in the reference game to their counterparts in the history-dependent one.

**Theorem 1**: Let \( g \) be an aligned history-dependent game, \( P \) be a local and monotone asynchronous learning rule, and \( Z : A_T \rightarrow \mathbb{Z}^+ \) be an increasing function with respect to ordering \( \geq A_T \). Then, we have

\[
\mathbb{E}_{P_\tau}(Z) \geq \mathbb{E}_{\hat{P}_\tau}(Z) \tag{15}
\]

for any \( \pi \in \Delta(A) \), \( T \in \mathbb{N} \).

The proof of Theorem 1 will be given in Section V-E.
The above $Z$ functions can represent a wide range of metrics of interest. For instance, $Z$ could be the total time spent in the $\bar{1}$ state or the total time during which some proportion of agents were playing 1. Thus, Theorem 1 gives a lower bound on the expected value of a broad class of increasing metrics on the history-dependent game of interest. Figures 1-2 show this intuition by using two specific instances of aligned history-dependent games, whose details are deferred to Section IV-A and Section IV-B, respectively. In both plots, the faint red curves show trials of an aligned history-dependent game and the faint blue curves show the static reference game, both under log-linear learning. It can be seen in both models that trials of the history-dependent game frequently feature more agents playing the 1 action than in the static game, confirming Theorem 1 over a broad class of monotone metrics $Z$.

To develop further applications of Theorem 1, we consider a specific $Z$ function. In doing so we obtain a property known as stochastic dominance as shown in the following lemma.

**Lemma 2:** If $g \in \mathcal{G}^A$ is an aligned history-dependent game with corresponding static game $\bar{g}$, $P$ is a local and monotone asynchronous learning rule, and $\mathcal{I} \subseteq \mathcal{A}_T$ is an upper set (i.e., $\alpha \in \mathcal{I} \implies \alpha' \in \mathcal{I}$ if $\alpha' \geq \mathcal{A}_T\alpha$), then $P_\pi(\mathcal{I}) \geq \bar{P}_\pi(\mathcal{I})$.

**Proof:** Let $g \in \mathcal{G}^A$ be an aligned history-dependent game, $P$ be a local and monotone asynchronous learning rule, and $\mathcal{I} \subseteq \mathcal{A}_T$ be an upper set. Define $\mathbb{1}_\mathcal{I}(\alpha) := \mathbb{1}(\alpha \in \mathcal{I})$ as an indicator function. By definition of expectation we have

$$P_\pi(\mathcal{I}) = \mathbb{E}_{P_\pi}(\mathbb{1}_\mathcal{I}), \quad \bar{P}_\pi(\mathcal{I}) = \mathbb{E}_{\bar{P}_\pi}(\mathbb{1}_\mathcal{I}).$$

Because $\mathbb{1}_\mathcal{I}$ is increasing we apply Theorem 1 to the above to obtain

$$P_\pi(\mathcal{I}) \geq \bar{P}_\pi(\mathcal{I}).$$

This result can then be interpreted as a lower bound on the probability of any upper set $\mathcal{I}$ or collection of histories occurring in the aligned history-dependent game of interest. Although it can be difficult to interpret a bound of the form (16), it can be valuable to use two specific instances of aligned history-dependent games, whose details are deferred to Section IV-A and Section IV-B, respectively. In both plots, the faint red curves show trials of an aligned history-dependent game and the faint blue curves show the static reference game, both under log-linear learning. It can be seen in both models that trials of the history-dependent game frequently feature more agents playing the 1 action than in the static game, confirming Theorem 1 over a broad class of monotone metrics $Z$.

To develop further applications of Theorem 1, we consider

$$\Pr(s(t); \bar{P}_\pi) = \bar{1}$$

is the probability that at time $T$ game $g$ is in the $\bar{1}$ action profile given initial distribution $\pi \in \Delta(A)$, while noting that a parallel interpretation to (18) holds for $P_\pi$. We apply these to Lemma 2 to obtain

$$\Pr(s(t); \bar{P}_\pi) = \bar{1} \geq \Pr(s(t); P_\pi) = \bar{1}.$$
Further, if $\pi$ has full support over $A$, then
\[
\Pr(s(t; P_\pi) = \vec{1}) > 0 \text{ for all } t.
\]  
(21)

Proof: Let $g \in G^A$ be an aligned history-dependent game, $P$ be asynchronous best response as specified by (3) with individual learning rules given by (11), and the associated game $\hat{g}$ be a static game with $\vec{1}$ being a strict Nash equilibrium. Eq.(19) holds directly from Lemma 3 and the fact that Lemma 1 gives that $P$ is local and monotone. Further, supposing $\pi$ has full support over $A$, (21) follows directly from this, (12), and (19). Specifically, (12) can be applied because $\vec{1}$ is a strict Nash equilibrium. In combination with that $\pi$ have full support over $A$, we have
\[
\Pr(s(t; P_\pi) > 0
\]  
(22)

as desired.

We further develop results on potential games with log-linear learning, by applying the theory of potential games to a history-dependent game.

Theorem 3: Let $g \in G^A$ be an aligned history-dependent game and $P$ be log-linear learning as specified by (3) with individual learning rules given by (8). If the associated game $\hat{g}$ is an exact potential game with $\vec{1}$ being the sole maximizer of the potential function, then $\vec{1}$ is uniquely stochastically stable in $g$ under $P$.

Proof: Let $g \in G^A$ be an aligned history-dependent game and $P$ be a local and monotone asynchronous learning rule. Let the associated game $\hat{g}$ be an exact potential game with $\vec{1}$ being the sole maximizer of the potential function $\phi$. It is well-known [27] that in an exact potential game $a \in A$ is a stochastically stable state with log-linear learning if
\[
a \in \arg \max_{a' \in A} \phi(a').
\]  
(23)

Because $\vec{1}$ is the sole maximizer of $\phi$, it is strictly stochastically stable. Because $P$ is local and monotone, we apply Lemma 3 to the definition of strict stochastic stability given by (10). For any $\epsilon > 0$, there exists $T > 0, T < \infty$ such that
\[
\Pr(s(t; P_\pi) = \vec{1}) \geq \Pr(s(t; P_\pi) = \vec{1}) > 1 - \epsilon
\]  
(24)

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

IV. Sample Applications of the Aligned History-Dependent Framework

A. Cyber Threat Intelligence Sharing

Cyber Threat Intelligence (CTI) refers to any information pertaining to, for example, why and how a firm (or enterprise) has been attacked, including the techniques or tactics that are used by attackers. In principle, sharing CTI is mutually beneficial to firms as it allows firms to strengthen their cyber defense postures. In practice, firms frequently opt not to share CTI due to a myriad of reasons including competitor relationships with other firms and possible leakage of their internal cybersecurity posture.

The CTI sharing problem has been studied using game theory between two firms in [19]–[22], between many firms using a centralized repository in [23], and in a decentralized (i.e., networked) setting in [24]. However, one assumption common to all these studies is that the value of CTI is constant over time. This oversimplifies the problem because the value of CTI can vary wildly. For example, detecting a zero-day exploit is a much more valuable CTI than the detection of a known attack. In this example, we propose a networked model where the value of CTI may vary over time, which can reflect the history of sharing. We examine how the aligned history-dependent framework can be used to analyze such a model.

Let a group of firms be modeled by the set $N = \{1, 2, \ldots, n\}$. If two firms mutually value each other’s CTI (i.e. because they use similar technologies), then we consider them connected in the networked model. We denote this with an edge set $E$ such that firms $i, j \in N$ are connected if $(i, j) \in E$. Then, $G = (N, E)$ represents an undirected graph which we leverage to model how firms decide to share CTI. We use action space $A$ to represent a firm’s decision to share CTI; in particular, a firm chooses action $a_i \in A_i = \{0, 1\}$ where $a_i = 1$ means firm $i$ wishes to share CTI with its neighbors and $a_i = 0$ indicates firm $i$ does not share CTI.

We make the following assumptions on the mechanics of sharing CTI. First, CTI is only shared if two neighbor firms $i, j$ mutually wish to share CTI, meaning $a_i = a_j = 1$. This is because CTI is potentially sensitive and thus firms would guard their CTI closely and require reciprocity to share it with other firms. Second, CTI sharing incurs firms a fixed infrastructure cost $c_i \geq 0$. This cost represents all expenses and labor associated with creating, packaging and sharing CTI and is incurred whenever firm $i$ chooses $a_i = 1$. Third, the value of firm $i$’s CTI, given by $v_i : A \rightarrow \mathbb{R}_+$, is a function of the history of agent behavior. That is, when firms $i, j$ share CTI given history $\alpha \in A$, firm $i$ receives CTI from firm $j$ with value $v_j(\alpha)$ and firm $j$ receives CTI from $i$ with value $v_i(\alpha)$.

The cybersecurity meaning of $v_j(\alpha)$ can be interpreted in as follows. First, it is possible that the value of CTI is a function of time, say $v_i(\alpha) = v_i'(T)$, where $v_i'$ is a function of time $T$ and $T$ is the period of time corresponding to history $\alpha$. For example, it is possible that firm $i$ occasionally discovers and detects important pieces of CTI, which correspond to a high value $v_i'(T)$, but most often discovers and shares less valuable pieces of CTI, meaning a low $v_i'(T)$ for most $T$. Second, the value of CTI can vary based on the history of play as shown in the following two scenarios. (i) As firms gain experience sharing CTI, their ability to share and leverage other firms’ CTI matures. This can be modeled by $v_i(\alpha)$ which increases over time or as the history $\alpha$ evolves. (ii) Attackers often attempt to disrupt defenses, such as CTI sharing in this case. As a consequence, attackers may wage new attacks that are harder to detect and share after the employed attacks become easy to detect.

The preceding discussion leads to the following utility function:
\[
U_i^a(\alpha) = a_i \left( -c_i + \sum_{j \in N_i} a_j v_j(\alpha) \right),
\]  
(25)

where $N_i = \{j \in N \mid (i, j) \in E\}$ is the neighbor set of firm $i$ in the CTI sharing graph $G = (N, E)$. Since the value of CTI can vary in possibly infinitely many ways, to facilitate
analysis we consider a single family of them, by imposing a mild restriction on the lower bound of \( v_i \):

\[
y := \min_{i \in N} \inf_{a_i \in A} v_i(a). \tag{26}
\]

This is a lower bound of all firms in all histories \( \alpha \). Under this premise, we show that CTI sharing game \( g \) is an aligned history-dependent game as follows.

**Proposition 1:** If \( g \in \mathcal{G}^A \) be a CTI sharing game with utilities specified by (25), then \( g \) is an aligned history-dependent game.

**Proof:** Let \( g \in \mathcal{G}^A \) be a CTI sharing game. We begin by giving the static game defined by

\[
\hat{U}_i(a) = a_i \left( -c_i + \sum_{j \in N_i} a_j \Upsilon \right). \tag{27}
\]

We now show

1) \( U_i(1, \alpha^T_i; \alpha) \geq \hat{U}_i(1, \alpha_{-i}), \) and
2) \( U_i(0, a_{-i}) \geq \hat{U}_i(0, \alpha^T_i; \alpha) \)

for any \( \alpha \in A, a, a' \in A, T \in \mathbb{N} \) where \( \alpha^T_i \geq \alpha_{-i} \) and \( a, a' \) differ only by a unilateral deviation.

Note that condition 1 follows from (26) and the fact that if firm \( i \) plays 1, then firm \( i \)'s neighbors playing 1 can only increase firm \( i \)'s utility. Condition 2 is trivial as \( U_i(0, a_{-i}) = U_i(0, \alpha^T_i; \alpha) = 0 \).

Now that the aligned history-dependent game framework has been established on the CTI game, we show how it can be used to analyze best response dynamics and log-linear learning in a history-dependent game. Before developing these results, we show that the static game \( \hat{g} \) is a potential game.

**Proposition 2:** The static game \( \hat{g} \) is an exact potential game.

**Proof:** We first show that static game \( \hat{g} = (N, A, \hat{U}) \) with \( \hat{U} \) defined by (27) is an exact potential game by giving its potential function:

\[
\phi(a) = \sum_{i \in N} \left( (1 - a_i) c_i + a_i \frac{\Upsilon}{2} \sum_{j \in N_i} a_j \right). \tag{28}
\]

The potential function can be verified by checking (9). Letting \( N_i(1) \) denote the neighbors of agent \( i \) with play action 1. We compute the change in potential and utility if agent \( i \) switches from 0 to 1 given arbitrary action profile \( a_{-i} \):

\[
\phi(1, a_{-i}) - \phi(0, a_{-i}) = \Upsilon |N_i(1)| - c_i
\]

\[
= U_i(1, a_{-i}) - U_i(0, a_{-i}) \tag{29}
\]

and the situation where agent \( i \) switches from 1 to 0 is simply the above equality multiplied by \(-1\).

In the case of the best response dynamics, it is known that the process does not leave Nash equilibrium. By establishing these properties in the static game, we establish an analogous result in the history-dependent game by applying the framework. Particularly, we establish that the probability that \( \bar{\pi} \) is selected by the best response dynamics in the history-dependent game is lower bounded by the static game, as given below.

**Proposition 3:** Let \( g \in \mathcal{G}^A \) be a CTI sharing game, \( P \) be the best response dynamics defined by (11), distribution \( \pi \) have full support over \( A \), and \( \hat{g} \) be its associated static game defined by \( \Upsilon \). If

\[
y > \frac{c_i}{|N_i|} \text{ for all } i \in N. \tag{30}
\]

then

1) \( \bar{\pi} \) is a strict Nash equilibrium in \( \hat{g} \),
2) \( \Pr(s(t, P_z) = \bar{\pi}) \geq \Pr(s(t, P_x) = \bar{\pi}) \), and
3) \( \Pr(s(t, P_z) = 1) > 0 \) for all \( t \).

**Proof:** It is easy to see that \( \bar{\pi} \) is a Nash equilibrium in \( \hat{g} \), as any deviation leads to a utility loss

\[
u_i(\bar{\pi}) - u_i(0, \bar{\pi}_{-i}) = |N_i| \Upsilon - c_i > 0. \tag{31}
\]

The second and third results hold from Theorem 2 as \( \bar{\pi} \) is a strict Nash equilibrium in \( \hat{g} \), \( g \) is an aligned history-dependent game by Proposition 1, and \( \pi \) has full support over \( A \).

Next, we examine the CTI sharing game under log-linear learning. As the long run behavior of potential games is well understood under log-linear learning, we can extend these results from the static game to the history-dependent one. By establishing a sufficient condition on \( y \) such that \( \bar{\pi} \) is a unique maximizer of \( \phi \), we can apply Theorem 3 to obtain stochastic stability in the history-dependent CTI game.

**Proposition 4:** Let \( g \in \mathcal{G}^A \) be a CTI sharing game, \( P \) be the log-linear learning dynamics defined by (8), and \( \hat{g} \) be its associated static game defined by \( \Upsilon \). If

\[
y > \frac{2c_i}{|N_i|} \text{ for all } i \in N, \tag{32}
\]

then we have

1) \( \{\bar{\pi}\} = \text{arg max}_{\pi} \phi(a) \), and
2) \( \bar{\pi} \) is strictly stochastically stable in \( g \) with log-linear learning.

**Proof:** Let \( g \in \mathcal{G}^A \) be a CTI sharing game, \( P \) be the log-linear learning dynamics defined by (8), and \( \hat{g} \) be its associated static game defined by \( \Upsilon \), and \( \phi \) be \( \hat{g} \)’s potential function defined in (28). To show that \( \bar{\pi} \) is the sole maximizer of \( \phi \), we show \( \phi(\bar{\pi}) > \phi(a) \) where \( a \) is any action profile \( a \neq \bar{\pi} \) as follows.

\[
\phi(\bar{\pi}) - \phi(a) = \frac{1}{2} \sum_{(i,j) \in E} \left( (1 - a_i) a_j \Upsilon - (1 - a_i) c_i \right) = \frac{1}{2} \sum_{(i,j) \in E} a_i a_j \Upsilon
\]

\[
\geq \frac{1}{2} \sum_{(i,j) \in E} (1 - a_i) c_i - \sum_{i \in N} (1 - a_i) c_i
\]

\[
= \frac{1}{2} \sum_{i \in N} (1 - a_i) |N_i| \Upsilon - \sum_{i \in N} (1 - a_i) c_i
\]

\[
> \sum_{i \in N} (1 - a_i) c_i - c_i = 0, \tag{33}
\]

where the first inequality holds because \( a_i \geq a_i a_j \) in domain \( \{0, 1\} \) and the second inequality is an application of (32) because there exists \( i \) such that \( a_i = 0 \) by the definition of \( a \).

A numerical comparison of the behavior of the CTI sharing game and the corresponding static game with log-linear
learning is shown in Figure 1. Both models consider a ring graph of 10 firms, and costs $c_i = 0.4$ for all $i \in N$. The history-dependent game model assumes a random model of CTI, where the value of CTI is a random value that is bounded from below, as given by

$$v(\alpha) = \max_{i \in N} \left( \frac{2c_i}{|N_i|} \right) + X + \epsilon = 0.4 + X + \epsilon,$$  \hspace{1cm} (34)

where $X$ is a uniform random variable in the interval $[0, 0.1]$ and $\epsilon$ is a positive constant. Thus, the value of CTI is random but sits in interval $[0.4 + \epsilon, 0.5 + \epsilon]$. This induces $\bar{X} = 0.4 + \epsilon$, defining the static game. Since $\bar{X}$ satisfies (32), Proposition 4 holds, meaning that $\bar{X}$ is stochastically stable in both the reference game and the history-dependent game. This is numerically confirmed as Figure 1 shows that the CTI sharing game rapidly finds the $\bar{X}$ state and stays there frequently. Although the static game also has $\bar{X}$ as the sole stochastically stable state, $\bar{X}$ only maximizes the potential function by an $\epsilon$ margin ($\epsilon = 0.001$ in this case) and thus the static game does not clearly frequently spend time in $\bar{X}$ state.

As per the definition of stochastic stability, if we take $t \rightarrow 0$ and $t \rightarrow \infty$ numerically, we can expect to see the static game tends toward $\bar{X}$ almost surely. Note that the key requirement for applying the aligned history-dependent framework is that $v(\alpha)$ is bounded from below, regardless of the exact form of (34), random or deterministic alike.

**B. Epidemics**

To show Theorem 3’s usefulness in analyzing the stochastic stability of history-dependent games, we consider another example, which is a simple model of epidemics. One challenge of epidemic modeling is to account for the interplay between epidemic severity and (in this example) the voluntary adoption of preventative social conventions. For instance, in the absence of an epidemic, people prefer not to wearing 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 SISGGC.

The fraction of individuals in the society susceptible to infection is described by the nonlinear differential equation

$$\dot{S} = (1 - S)(\gamma - \beta(t)S),$$ \hspace{1cm} (35)

where $\gamma > 0$ is the curing rate and $\beta(t) > 0$ is the rate of infection which depends on agent actions. The action 1 represents a “safe convention” action in which an agent acts to reduce contagion (e.g., wearing a mask); 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,

$$\beta(t) = \frac{1}{|N|} \sum_{i \in N} a_i^t \beta_1 + (1 - a_i^t)\beta_0,$$ \hspace{1cm} (36)

where $a_i^t$ is the action selected by agent $i$ at time $t \in N$. Actions are selected by agents in $N$ dynamically on undirected graph $G = (N, E)$ according to log-linear learning rule (6). The utility of agent $i$ at time $t$ is given by

$$U_i^\alpha(a_i^t, a_{-i}^t) = a_i|N_i|((\lambda + I(t)) + (1 - a_i)|N|),$$ \hspace{1cm} (37)

where $N_i(x) = \{ j \in N \mid (i, j) \in E, a_j = x \}$ is the set of $i$’s neighbors which play $x \in \{0, 1\} = A_i$. $I(t) = 1 - S(t)$ is the fraction of infected individuals, and $\lambda \in (0, 1)$ represents the agent’s willingness to practice safe conventions in the absence of epidemic.

The SISGGC model can be analyzed using the aligned history-dependent game framework. Specifically, a reference static game can be devised using utility function (37) by setting $I(t) = 1 - \gamma/\beta_1$, which is a lower bound of $I(t)$ after a sufficient time. The details of this lower bound are given in the following Proposition:

**Proposition 5:** If $S(0) \in [0, 1)$ and $S(t)$ is a solution to (35) with $\beta(t)$ being given by (36), then there exists a $\hat{t}$ such that $S(t) \leq \gamma/\beta_1$ for all $t \geq \hat{t}$ almost surely.

**Proof:** We write $S^*_t := \gamma/\beta_1$. Note that if $S(t) \geq S^*_t$, then because $\beta(t) \geq \beta_1$, we have that $\hat{S} \leq 0$ by (35) and that this inequality is strict when $S(t) > S^*_t$. Thus, the set $[0, S^*_t]$ is positively invariant for the hybrid nonlinear dynamics given in (35).

To see that $S(t)$ eventually enters $[0, S^*_t]$ almost surely, consider the event that $S(t) > S^*_t$ for all $t$. Note that $S^*_t$ is asymptotically stable when $\beta(t) \equiv \beta_1$. For any action profile $a \neq 1$ such that its associated $\beta > \beta_1$, the event that $\beta(t) \equiv \beta_1$ for all $t$ is the same event as $S(0) > S^*_t$ and $S(t) > S^*_t$ for all $t$. However, it can be seen that log-linear learning defines a stochastic process which visits every action profile in $A$ infinitely often. That is, the probability that $\beta(t) \equiv \beta_1$ is 0, and thus there must exist a $\hat{t}$ such that $S(t) \leq S^*_t$ for all $t \geq \hat{t}$ almost surely.

It can be seen from (37) that SISGGC can be represented by a history-dependent game, as the utility function depends on the history of play, so Theorem 3 allows us to reference SISGGC to a related exact potential game and deduce conditions guaranteeing that $\bar{X}$ is strictly stochastically stable under log-linear learning.

**Proposition 6:** Let $g^S$ be an instance of SISGGC, and $P$ be log-linear learning as defined by (8) and (3). If $\beta_1/\gamma > 1$, $\lambda + \gamma/\beta_1 > 1$, and $I(0) > 0$, then $\bar{X}$ is stochastically stable in $g$ with $P$.

**Proof:** Denote the SISGGC model by $g^S$, which is played on graph $G = (N, E)$ with $g + \gamma/\beta_1 > 1$ and $I(0) > 0$. Consider $g^S$ as played after time $t$ as given by Proposition 5. Game $g^S$ is a history-dependent game since (37) depends on $I(t)$, which is a function of the history $\alpha$. Thus, we have $g^S = (N, A, U) \in G^A$ where $U = \{U^\alpha \mid \alpha \in A\}$.

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 (37) with $\hat{I}(t) = 1 - \gamma/\beta_1$. Standard results give that $\hat{g}^S$ is an exact potential game and that $\bar{X}$ is its sole potential function maximizer [2].

We now use $\hat{g}^S$ to show that $g^S$ is an aligned history-dependent game. We verify $U_i^\alpha(1, a^T_{-i}) \geq \hat{U}^S_i(1, a_{-i})$ holds
for $\alpha^{T^i}_{a_i} \geq A_{a_i}$ and $t > \tilde{t}$. It can be rewritten for $t > \tilde{t}$ as
\[ \sum_{j \in N_i(1; a_{\tilde{t}})} \lambda + I(t) \geq \sum_{j \in N_i(1; a_{\tilde{t}})} \lambda + \gamma / \beta_1, \] (38)
where $N_i(1; a_{\tilde{t}})$ denotes the neighbors of $i$ which play 1 given profile $a$. This expression holds because $\alpha^{T^i}_{a_i} \geq A_{a_i}, a_{\tilde{t}} \Rightarrow |N_i(1; a^{T^i}_{a_i})| \geq |N_i(1; a_{a_{\tilde{t}}})|$ and by Proposition 5. An argument with the same structure holds for $U^{\alpha}_{a_i} \leq \tilde{U}^i_{\alpha}(0, a_{\tilde{t}})$. Thus, $g^T$ is an aligned history-dependent game, and Theorem 3 gives that $\tilde{I}$ is strictly stochastically stable.

Figure 2 plots a numerical example of the SISCGC and its associated static game, where $\gamma = 0.25, \beta_0 = 0.9, \beta_1 = 0.45$, and $\lambda = \gamma / \beta_1 + \delta$. It can be seen that the average $I(t)$ across SISCGC rapidly shoots above $1 - \gamma / \beta_1$, which is the fixed disease $I(t)$ for the GCG game. This confirms the intuition of Proposition 5 that $1 - \gamma / \beta_1$ can be treated as an effective lower bound of $I(t)$ which we use to define the static game. Figure 2 shows that the SISCGGC model rapidly finds the all $\tilde{I}$ state faster and more consistently than the reference game, which supports Proposition 6 experimentally. Similar to the above CTI sharing example, $\tilde{I}$ is only stochastically stable by an $\epsilon$ margin, explaining the phenomenon it does not appear to find the $\tilde{I}$ state frequently. Critically, the history-dependent game framework can be applied to characterize agent behaviors in a stochastically complex epidemic model.

V. PROVING THE MONOTONE COUPLING

A. A Primer on Monotone Couplings

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

**Definition 6:** Let $X$ be a countable set with partial ordering $\leq_X$ and $p_1, p_2$ be probability measures on $(X, \mathcal{F})$, where $\mathcal{F}$ is a $\sigma$-algebra of $X$. 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 $x, y \in X$
\[ \sum_{x \leq_X x'} p(x, y') = p_2(y') \quad \text{and} \quad \sum_{y \geq_X y'} p(x', y) = p_1(x'). \] (39)

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

**Proposition 7 (Paarpporn et al., [31]):** 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 $Z : X \to \mathbb{Z}_+$ we have
\[ \mathbb{E}_{p_1}(Z) - \mathbb{E}_{p_2}(Z) = \sum_{\eta=0}^{\infty} p(Z^\eta, \{a \mid Z(a) > \eta\}), \] (40)
where $Z_0 = \{a \mid Z(a) > \eta\}$ and $Z^\eta := X \setminus Z$ is the complement set of $Z \subseteq X$.

B. Notations Required for Proofs

We write the probability that path $\alpha \in A_T$ occurs with initial distribution $\pi \in \Delta(A)$ as
\[ \tilde{P}_\pi(\alpha) = \pi(\alpha^1) \prod_{t=1}^{T-1} \tilde{P}_{\alpha^t}((\alpha^{t+1}), \] (41)
where $\pi(\alpha^1)$ is the probability of $\alpha^1$ in initial distribution $\pi$. Correspondingly, the probability that path $\alpha \in A_T$ occurs with initial distribution $\pi \in \Delta(A)$ on $g \in \mathcal{G}^A$ is
\[ P_\pi(\alpha) = \pi(\alpha^1) \prod_{t=1}^{T-1} P_{\alpha^t}((\alpha^{t+1}), \] (42)
where $\alpha \leq t \in A_t$ is the history $\alpha$ until time $t \in \{1, 2, 3, \ldots, T\}$. We now present a result to connect the utility conditions of aligned history-dependent games to (6) and (7).

**Lemma 4:** Let $P_t : A \times A \times U^n \to [0, 1]$ be local and increasing with respect to utility. Consider utility functions $U, \bar{U} \in U^n$. If $U_i(1, a^-_{\tilde{t}}) \geq \bar{U}_i(1, a^-_{\tilde{t}})$ and $\bar{U}_i(0, a^-_{\tilde{t}}) \geq U_i(0, a^-_{\tilde{t}})$, then $P_1(a', 1, a^-_{\bar{t}}) \geq \bar{P}_1(a, 1, a^-_{\tilde{t}})$ for some $a, a' \in A$.

**Proof:** Assume $U_i(1, a^-_{\tilde{t}}) \geq \bar{U}_i(1, a^-_{\tilde{t}})$ and $\bar{U}_i(0, a^-_{\tilde{t}}) \geq U_i(0, a^-_{\tilde{t}})$ for some $a, a' \in A$. Define
\[ \bar{U}_i(x) = \begin{cases} U_i(1, a^-_{\tilde{t}}) & x_1 = 1 \\ \bar{U}_i(0, a^-_{\tilde{t}}) & x_i = 0. \end{cases} \] (43)

Let $\bar{U} = (\bar{U}_1, \ldots, \bar{U}_{i}, \ldots, \bar{U}_n)$. Because $U(1, a^-_{\tilde{t}}) = \bar{U}(1, a^-_{\tilde{t}}) + l$ for some $l \geq 0$, we have $P_1(a, (1, a^-_{\bar{t}}), \bar{U}) \geq P_1(a, (1, a^-_{\tilde{t}}), U)$ by definition of increasing with respect to utility. Similarly, because $\bar{U}(0, a^-_{\tilde{t}}) = U(0, a^-_{\bar{t}}) + l$ for some $l \geq 0$, we have $P_1(a', (0, a^-_{\bar{t}}), \bar{U}) \geq P_1(a', (0, a^-_{\tilde{t}}), U)$ again by definition of increasing with respect to utility. By adding $P_1(a', (0, a^-_{\bar{t}}), \bar{U})$ to both sides of this expression, we derive $P_1(a', (1, a^-_{\bar{t}}), \bar{U}) \geq P_1(a', (1, a^-_{\tilde{t}}), \bar{U})$. Using these inequalities, we find
\[ P_1(a', (1, a^-_{\tilde{t}}), \bar{U}) \geq P_1(a', (1, a^-_{\bar{t}}), \bar{U}) \]
\[ = \bar{P}_1(1, \bar{U}_i(a^-_{\bar{t}})) \]
\[ = \bar{P}_1(1, \bar{U}_i(a^-_{\tilde{t}})) \]
\[ = \bar{P}_1(a, (1, a^-_{\bar{t}}), \bar{U}) \]
\[ \geq \bar{P}_1(a, (1, a^-_{\tilde{t}}), \bar{U}), \]
where $\bar{P}_1$ is defined according to (1) as $P_t$ is local by hypothesis.

Our framework requires a careful partitioning of the action space corresponding to different types of agent action deviations. Let $f : A \to \mathcal{A}$ be defined as
\[ f(a) = \{ a' \in A \mid a_i \neq a'_i, a_{-i} = a'_{-i} \text{ for } i \in N \}, \]
which is the set of action profiles reachable from $a$ via exactly one unilateral deviation. For $a, a' \in A$, let
\[ g(a, a') = \begin{cases} i & a_i \neq a'_i \\ 0 & a = a' \end{cases} \] (45)
indicate which agent unilaterally deviated its action between action profiles $a, a'$. Consider $a, a' \in A$ where $a' \geq A a$. We consider several disjoint subsets of $f(a)$:

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

These sets can be interpreted as follows. The set $r(a)$ is the set of action profiles that decrease with respect to $\geq A$; both $g(\cdot, \cdot)$
and $s(\cdot, \cdot)$ increase with respect to $\geq_A$. Between $g(\cdot, \cdot)$ and $s(\cdot, \cdot)$, $g(\cdot, \cdot)$’s action profiles remain less than $a'$ and $s(\cdot, \cdot)$’s profiles are greater than or incomparable to $a'$. We now present three more analogous sets that are disjoint subsets of $f(a')$:

1. $R(a') = \{x \in f(a') | a' \notin Q(a', x)\}$,
2. $Q(a, a') = \{x \in f(a') | x \geq_A a \} \cap R(a')$, and
3. $S(a, a') = f(a') \setminus (Q(a, a') \cup R(a'))$.

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

We now highlight some useful features of these sets. It is evident that $r(\cdot), q(\cdot, \cdot), s(\cdot, \cdot)$ are decreasing or disjoint on $f(a)$, and that $R(\cdot), Q(\cdot), S(\cdot)$ are disjoint or partition on $f(a')$. For any $a, a'$ such that $a' \geq_A a$, we relate these sets by a function $b^{a, a'}: f(a) \mapsto f(a')$. To evaluate $b^{a, a'}(\bar{a})$, we identify the agent that deviates its action between $a$ and $\bar{a}$ and then devise the agent’s action in $a'$. Formally, we have $b^{a, a'}(\bar{a}) = (\bar{a} - a, \bar{a} - a, P_g(a, a, \bar{a}))$ where we define $\bar{a}_i \in \{0, 1\}$ for $a_i \in A_i, \{0, 1\}$. In particular, this function relates the disjoint subsets of $f(a)$ and $f(a')$ according to the following lemma.

Lemma 5: If $a, a' \in A$ and $a' \geq_A a'$, then the following statements hold:

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

Proof: Let $a, a' \in A$ such that $a' \geq_A a$. We proceed by proving $b^{a, a'}: r(a) \mapsto S(a, a')$ is a bijection; the other bijection statements can be proved similarly.

We begin by proving injectiveness, namely $b^{a, a'}(z) = b^{a, a'}(z') \implies z = z'$ for $z, z' \in r(a)$. Observe that

$g(a, z) = g(a', b^{a, a'}(z)) = g(a', b^{a, a'}(z')) = g(a, z'),$

where the first and third equalities follow the definition of $b^{a, a'}$ and the middle one follows the hypothesis. Injectiveness follows from $g(a, z) = g(a, z')$, meaning that $a, z$ and $a, z'$ differ by the same agent’s unilateral deviation. Thus, the possible actions of agent $g(a, z)$ is given by $A_g(a, z) \setminus \{a_g(a, z)\}$, which is a singleton by the binary action property, leaving only one possible state $a$ that could transition to $r(a)$ via a unilateral deviation. Thus $z = z'$ as desired.

Next we show surjection, namely that for any $z' \in S(a, a')$, there exists a $z \in r(a)$ such that $b^{a, a'}(z) = z'$ for $a, a' \in A$ and $a \leq_A a'$. By definition of $S(a, a')$, we have $z' \notin a'$. Since $z' \in f(a')$, we observe that $z'$ and $a'$ differ only by a single unilateral deviation corresponding to some agent $i$. By partial ordering $\leq_A$, we infer $a_i = 1$ and $z_i = 0$; otherwise, $z' \notin a'$ would be violated. Further, we may infer $a_i = 1$; otherwise, $a_i = 0$ leads to $z' \notin Q(a, a')$, which gives a contradiction to the definition of $z'$. It is easy to see by definition of $r(a)$ that $a_i = 1 \implies z \in r(a)$ satisfying $g(a, z) = g(a', z')$ as $z \neq a_i$, but $z_i = a_i$ as $z \in f(a)$. Note that $g(a, z) = g(a', z')$ is always satisfied when $b^{a, a'}(z) = z'$, by the definition of the function.

C. The One-Step Couplings

To prove Theorem 1, we construct a monotone coupling $\mathbb{P}_\pi$ between measures $P_\pi$ and $\hat{P}_\pi$. We first construct a family of monotone couplings for each one-step transition (Theorem 4), and apply it to show the coupling over histories (Theorem 5).

Theorem 4: Let $g \in G^A$ denote an aligned history-dependent game, $\hat{g}$ be its reference static game, and $P$ be a local and monotone asynchronous learning rule. Then, a monotone coupling exists between $P^a$ and $P^\alpha$ for any $\alpha \in A$ and $a \in A$ when $a \leq_A a\alpha$. This monotone coupling $\mathbb{P}_{\pi, a, a\alpha} : A^2 \mapsto [0, 1]$ is given in (46) as shown in Figure 3.

Proof: Let $a \in A$ and $\alpha \in A$ such that $a \leq_A a\alpha$. Let $g \in G^A$ be an aligned history-dependent game where $g$ is its reference static game, and let $P$ be a local and monotone asynchronous learning rule. To verify that $\mathbb{P}_{\pi, a, a\alpha}$ is a monotone coupling, we must show the following conditions given by Definition 6. For any $\bar{a}, a' \in A$, it holds that

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

We begin by verifying Condition 2. We consider cases $\bar{a} \notin (f(a) \cup \{a\})$, $\bar{a} \in \bar{a} \in r, a \in s$ and $\bar{a} = a$ separately. We use the notational convention that $q, s, Q, S$ are assumed to take arguments $(a, a\alpha)$, and that $r, R$ take arguments $a, a\alpha$ respectively. The first case represents any $\bar{a}$ which cannot be achieved in a single unilateral deviation from $a$. This gives that $P^\alpha(\bar{a}) = 0$, and thus all pairs of $\bar{a}, z'$ must satisfy $\mathbb{P}_{\pi, a, a\alpha}(\bar{a}, z') = 0$. This holds as all parts of (46) require $\bar{a} \in (f(a) \cup \{a\})$ except (46h), which has the desired property.

We now consider the second case that $\bar{a} \in q$. Note that only (46d) satisfies this condition, so we have

$$\sum_{a' \geq_A \bar{a}} \mathbb{P}_{\pi, a, a\alpha}(\bar{a}, a') = \hat{P}_{\pi}^{a\alpha}(\bar{a}, a') = \hat{P}_{\pi}^{a\alpha}(\bar{a})$$

Next we consider $\bar{a} \in r$, which satisfies (46c), (46f) because $b^{a, a\alpha}$ is a bijection by Lemma 5. Thus

$$\sum_{a' \leq_A \bar{a}} \mathbb{P}_{\pi, a, a\alpha}(\bar{a}, a') = \frac{1}{|N|} \left( \hat{P}_{\pi}^{a\alpha}(0) - P_{\pi}(0) + P_{\pi}(a', 0) \right)$$

where the second equality follows as $g(\bar{a}, a) = g(a^\alpha, a')$ by definition of $b^{a, a\alpha}$. The third equality follows as $\bar{a} \in r \implies \bar{a}_g(a, a) = 0$.

Considering $\bar{a} \in s$, we find only (46e) applies, and thus we have

$$\sum_{a' \geq_A \bar{a}} \mathbb{P}_{\pi, a, a\alpha}(\bar{a}, a') = \frac{1}{|N|} \hat{P}_{\pi}^{a\alpha}(1) = \hat{P}_{\pi}^{a\alpha}(\bar{a})$$

where $\bar{a} \in s \implies \bar{a}_g(a, a) = 1$ (otherwise, $\bar{a}$ would be in $q$). The final case for Condition 2 is $\bar{a} = a$. We find cases (46a),
\[ P^{\alpha,\alpha}(\tilde{a}, \tilde{a}') = \begin{cases} \frac{1}{|N|} \left( P_{g^T}^\alpha(\tilde{a}, \tilde{a}') - P_{g^T}^\alpha(\tilde{a}', \tilde{a}) \right)(1) & \tilde{a} = a, \tilde{a}' \in R \\
\frac{1}{|N|} P_{g^T}^\alpha(\tilde{a}, \tilde{a}') & \tilde{a} = a, \tilde{a}' \in Q \\
\frac{1}{|N|} \left( P_{g^T}^\alpha(\tilde{a}, \tilde{a}) - P_{g^T}^\alpha(\tilde{a}, \tilde{a}') \right)(0) & \tilde{a} \in r, \tilde{a}' = \alpha^T \\
\frac{1}{|N|} P_{g^T}^\alpha(\tilde{a}, \tilde{a}) & \tilde{a} \in q, \alpha^T = \tilde{a}' \\
\frac{1}{|N|} \left( P_{g^T}^\alpha(\tilde{a}, \tilde{a}') - P_{g^T}^\alpha(\tilde{a}, \tilde{a}) \right)(0) & \tilde{a} \in r, \tilde{a}' = b^{\alpha,\alpha^T}(\tilde{a}) \\
\frac{1}{|N|} \left( |N| - \sum_{z \in g^T} P_{g^T}(\tilde{a}, z)(z(g(a), z)) - \sum_{z' \in g^T \cup T} P_{g^T}(\alpha, z')(z(g(a), z')) \right) & a = \tilde{a}, \tilde{a}' = \tilde{a}' \\
0 & \text{otherwise}. \end{cases} \]

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 \(b^{\alpha,\alpha^T}\) and its bijectiveness due to Lemmas 5. By disjointness of \(q, r\) we have \(N_{qr} = N_{QS}\) which we apply to \(N_{qr} \cap N_{QR} = N_{NQ} \cap N_{QR} = N_Q = N_q\). Applying definitions of \(q, Q\), we find \(i \in N_q \iff -a_i = 1, -\alpha_i^T = 0\). Thus, the summand of the first term for \(i \in N_q\) is given by

\[ 1 - P_{i}\left(1 - P_{i}\right)(1) - P_{i}(1) - P_{i}(0) = 0, \]

where the inequality follows (51), giving that the summands in the first term of (52) are themselves well-defined probabilities. As all conditions have been met, \(P^{\alpha,\alpha}\) is a monotone coupling.

**D. A Monotone Coupling Over Histories**

We now present coupling \(\hat{P}_T^g\), which is constructed using the one-step coupling. We define the indicator function \(1\) such that \(1(P) = 1\) if \(P\) is a true logical proposition and \(1(P) = 0\) otherwise.

**Theorem 5:** Let \(g \in G^A\) be an aligned history-dependent game, \(\hat{g}\) be its corresponding static game, \(P\) be a local and
monotone asynchronous learning rule, and \( \pi \in \Delta(A) \) be a distribution over all action profiles. Then, \( \mathbb{P}_\pi^\alpha : A_\pi^\alpha \rightarrow [0, 1] \) is a monotone coupling between \( P_\pi \), \( \bar{P}_\pi \). This coupling is given by

\[
\mathbb{P}_\pi^\alpha(\alpha, \bar{\alpha}) = \pi(\alpha^1) \mathbb{I}(\alpha^1 = \bar{\alpha}^1) \prod_{t=1}^{T-1} \mathbb{P}^\alpha_t, \alpha^{t+1} = (\alpha^{t+1}, \bar{\alpha}^{t+1}),
\]

(54)

where \( \alpha, \bar{\alpha} \in A_T \) and \( \pi \in \Delta(A) \).

**Proof:** Let \( \alpha, \bar{\alpha} \in A_T \) and \( g \in G^A \). Let \( \hat{g} \) be the corresponding static game and \( P \) be a local and monotone asynchronous learning rule. We begin by showing that if \( \alpha \not\in A_T \) \( \alpha \), then \( \mathbb{P}_\pi^\alpha(\alpha, \bar{\alpha}) = 0 \). Since \( \mathbb{P}_\pi^\alpha(\alpha, \bar{\alpha}) = 0 \) if \( \alpha^1 \neq \bar{\alpha}^1 \), we only need to consider cases where \( \alpha^1 = \bar{\alpha}^1 \). Inductively, we find that if \( \alpha \not\in A_T \) \( \bar{\alpha} \), there must exist some \( t \in \{1, 2, 3, \ldots, T-1\} \) such that \( \alpha^t \leq_A \bar{\alpha}^t \) but \( \alpha^{t+1} \not\in A_T \bar{\alpha}^{t+1} \); let \( t \) be the minimum of these values. We have \( \mathbb{P}^\alpha_t, \alpha^{t+1} = (\alpha^{t+1}, \bar{\alpha}^{t+1}) = 0 \) because \( \mathbb{P}^\alpha_t, \alpha^{t+1} \) is a well-defined monotone coupling by Theorem 4, yielding \( \mathbb{P}_\pi^\alpha(\alpha, \bar{\alpha}) = 0 \). It also follows that \( \mathbb{P}_\pi^\alpha \) 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 (39). We begin by showing the left-hand equation of (39), namely

\[
\sum_{\alpha \leq A_T} \mathbb{P}_\pi^\alpha(\alpha, z) = \hat{P}_\pi(\alpha) \text{ for each } z \in A_T
\]

(55)

and omit the proof for the right-hand equation because it is similar. By inspecting (54), we observe that we only need to consider \( z \) such that \( z^1 = \alpha^1 \) and \( z \) differ at most one unilateral deviation between any \( t, t + 1 \). With these two conditions, we rewrite

\[
\sum_{\alpha \leq A_T} \mathbb{P}_\pi^\alpha(\alpha, z) = \sum_{\alpha \leq A_T} \pi(\alpha^1) \prod_{t=1}^{T-1} \mathbb{P}^\alpha_t, z^{t+1} = (\alpha^{t+1}, z^{t+1}) = \pi(\alpha^1) \sum_{\alpha^2 \leq A_T^2} \prod_{t=1}^{T-1} \mathbb{P}^\alpha_t, z^{t+1} = (\alpha^{t+1}, z^{t+1}) \sum_{\alpha^3 \leq A_T^3} \prod_{t=1}^{T-1} \mathbb{P}^\alpha_t, z^{t+1} = (\alpha^{t+1}, z^{t+1}) \sum_{\alpha^4 \leq A_T^4} \prod_{t=1}^{T-1} \mathbb{P}^\alpha_t, z^{t+1} = (\alpha^{t+1}, z^{t+1}) = \hat{P}_\pi(\alpha),
\]

(56)

as the combinatorial form. This allows us to apply the marginal sum properties of \( \mathbb{P}^\alpha_t, z^{t+1} \) from Theorem 4, where \( t \in \{1, 2, \ldots, T\} \), as follows. First, consider the rightmost sum in (56). It holds that

\[
\sum_{\alpha^T \leq A_T} \mathbb{P}^\alpha_{T-1}, z^{T-1} = (\alpha^T, z^T) = \hat{P}^\alpha_{T-1} = (\alpha^T).
\]

(57)

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

\[
\sum_{\alpha \leq A_T} \mathbb{P}_\pi^\alpha(\alpha, z) = \pi(\alpha^1) \prod_{t=1}^{T-1} \hat{P}^\alpha_{t+1} = \hat{P}_\pi(\alpha)
\]

(58)

as desired, while accounting for the indicator functions in \( \mathbb{P}_\pi^\alpha \). This concludes the proof of Theorem 5.

**E. Proof of Theorem 1**

**Proof:** Let \( g \) be an aligned history-dependent game, \( \hat{g} \) be the associated static game, \( P \) be a local and monotone asynchronous learning rule, and \( \pi \in \Delta(A) \) be a distribution over all action profiles. Theorem 5 gives that there exists a monotone coupling between measures \( P_\pi \) and \( \bar{P}_\pi \), which is given by \( \mathbb{P}_\pi^\alpha \). This enables us to apply Proposition 7 to obtain

\[
E_{P_\pi}(z) - E_{\bar{P}_\pi}(z) \geq 0
\]

(59)

for any function that is monotone with respect to ordering \( \geq_A T \). This holds because the coupling \( \mathbb{P}_\pi^\alpha \) is a nonnegative-valued function, which concludes the proof.

**VI. CONCLUSION**

We have developed new analytical tools to relate behaviors of history-dependent games to some related, tractable static games. This represents a significant step towards making theoretical studies more faithful to reality. There are many directions for future research. One is to extend the present study to more complex types of history-dependent interaction, such as games with richer action spaces and settings where the relationship between the history-dependent game and the static game holds in expectation rather than for all time. Another direction is to explore the usefulness of the concepts and techniques in other settings, such as the cybersecurity dynamics framework [32], or preventive and reactive defense dynamics [33]. Although games have been investigated in the cybersecurity dynamics framework [34], much research needs to be done. Moreover, it would be interesting to investigate whether chaotic behaviors can be exhibited by history-dependence as such behaviors are relevant to proactive defense dynamics [35].

**REFERENCES**

[1] B. C. Collins, L. Hines, G. Barboza, and P. N. Brown, “Robust stochastic stability in dynamic and reactive environments,” in 2021 60th IEEE Conference on Decision and Control (CDC), pp. 1892–1897, 2021.

[2] H. P. Young, “The evolution of conventions,” Econometrica: Journal of the Econometric Society, pp. 57–84, 1993.

[3] 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.

[4] 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.

[5] 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.

[6] B. C. Collins and P. N. Brown, “Exploiting an Adversary’s Intentions in Graphical Coordination Games,” in ACC 2020, pp. 4638–4643, 2020.

[7] J. R. Marden and A. Wierman, “Distributed welfare games,” Operations Research, vol. 61, no. 1, pp. 155–168, 2013.

[8] 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.

[9] 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.

[10] 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.

[11] 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.
