2024 Tenth Indian Control Conference (ICC)
9 - 11 December, 2024
IISER Bhopal, India

Discounted Evolutionary Stochastic Games and Replicator Dynamics

Divya Murali[1*] and K. S. Mallikarjuna Rao[1†]

***Abstract*— In this paper, we study the discounted evolutionary stochastic game model. Using this model, we give an alternate definition for Nash equilibrium (NE) of stochastic games and also define evolutionarily stable strategy (ESS) for the same. Also, we provide a version of replicator dynamics for discounted stochastic games and prove that the defined ESS is asymptotically stable in it. Further, we explore finite stage stochastic game and study the relation between NE and replicator dynamics of the N-step stochastic game and the infinite horizon stochastic game.**

I. INTRODUCTION

Stochastic games were first introduced in Shapley (1953)[1], with discounted reward. In the same year, Gillette independently defined a version of stochastic game with undiscounted payoff[2]. Since then stochastic games have been well studied, mostly zero-sum games and the existence of value (see [3],[4],[5]). For non zero-sum stochastic games, the existence of equilibria has been studied in some special classes like irreducible games (see [6]). Meanwhile, evolutionary games came into existence through Maynard Smith and Price (1973)[7], where they defined an evolutionarily stable strategy (ESS).

In this paper, we study the discounted version of evolutionary stochastic games (see [8],[9],[10]), which may be considered as a multi-state version of classical evolutionary games (Section II). In the recent literature, there are a few approaches to integrate evolutionary game theory into stochastic games (see e.g. [11],[12],[13],[14],[15],[16]). The evolutionary stochastic game model and the modified discounted payoff function that we define in this paper, helps us to consider a stochastic game in the form of a population game. This allows us to work with a large population of players for the stochastic game, even though the interactions are pairwise. This model also helps us to define appropriate evolutionary dynamics and learning algorithms for stochastic games.

We establish the discounted game model with necessary definitions in Section II. In Section III, we look at the N-step game, where the stochastic game stops after N stages. Further in Section IV, we discuss the main results regarding the dynamic stability of the proposed ESS of the stochastic game w.r.t. the associated replicator dynamics and the relations between the N-step game and the infinite horizon game. Using these results, we may calculate the ESS of stochastic games, if we know it exists. Also, the linearised discounted reward $\tilde{F}_\beta$ defined here can be used to study further evolutionary properties of stochastic games due to similarity of the model to non-linear population games ([17]).

[1]School of Industrial Engineering and Operations Research, Indian Institute of Technology Bombay, Mumbai-400076, India.
*divyamurali963@gmail.com
†mallik.rao@iitb.ac.in

II. EVOLUTIONARY STOCHASTIC GAME MODEL

The evolutionary stochastic game model was introduced in Flesch et al(2013)[8]. In the paper, they considered two-person irreducible stochastic games with limiting average payoff. Here, we consider a discounted stochastic game and explore the static and dynamic stability conditions of different strategies.

A. The Model

In this model we consider a two-person, finite stochastic game. The state space S and the actions sets A and B for both the players are finite. At each state s, player i receives a payoff denoted by $r_i(s,a,b)$, where $a \in A$ and $b \in B$ are the actions chosen by the players. Let $s \in S$ be the initial state of the game. With probability $q(s'|s,a,b)$, the players move from state s to s'. We also assume that the stochastic game is irreducible, i.e., there is a positive probability to reach from state s to s', for all states $s, s' \in S$, and for all actions $a \in A$ and $b \in B$. Similar to the classical evolutionary games, we consider this stochastic game to be symmetric. That is,
1. The set of actions are same for both the players at all states, i.e., $A_s = B_s$ for all $s \in S$, where A_s and B_s are the action sets in state s for players 1 and 2 respectively;
2. $r_1(s,a,b) = r_2(s,b,a), \quad \forall s \in S,\ a \in A,\ b \in B$;
3. $q(s'|s,a,b) = q(s'|s,b,a), \quad \forall s, s', a, b.$
Therefore, in the rest of the paper, we look at the game through player 1.

A strategy specifies the actions (pure/mixed) taken by the players at each stage of the game, depending on the history (sequence of states and actions). In this paper we consider only stationary strategies, denoted by f, that depend only on the current state of the game and independent of the previous history, i.e., $f : S \to \wp(A)$. The set of all stationary strategies is denoted as Π_{SR}. If the actions at all states are pure (deterministic), i.e., the player chooses to play a particular action at all times with probability 1, then it is called a pure stationary strategy. Such a pure strategy may be defines as $f : S \to A$. In the evolutionary stochastic game model, we consider a large population, where the individuals are randomly matched in pairs. The population is characterised into different types, using pure stationary strategies. Let $\Pi_{SD} = \{e_l : 1 \le l \le L\}$ be the set of all pure stationary strategies. For example, in a stochastic game of three states, with two actions each, $e_1 = ((1,0),(1,0),(1,0))$ is a pure

stationary strategy. An individual is said to be of type l, if he/she uses strategy e_l. The fraction of the population using strategy e_l is given by d_l, and $d = (d_1, d_2, ..., d_L)$ is the population distribution, based on strategies. The simplex of all population distributions is denoted by Δ. Each population distribution induces a stationary strategy defined as

$$f_d(s,a) := \sum_{l=1}^{L} d_l \cdot e_l(s,a). \tag{1}$$

From the above equation, we can see that while each distribution induces a unique strategy, there may be infinitely many distributions inducing the same stationary strategy. For a stationary strategy f, we define the support of f as $C(f) = \{e_l : e_l(s,a) > 0 \implies f(s,a) > 0, \forall s \in S \text{ and } a \in A\}$. Similarly, for $d \in \Delta$, support of d is given by $C(d) = \{e_l : d_l > 0, l = 1, 2, 3..., L\}$.

The fitness of the individuals in this population is given by the discounted reward of the underlying stochastic game. Let $Q(f,g)$ be the transition probability matrix for strategy pair (f,g). The n-step transition probabilities are given by $Q^n(f,g)$ ([18]). The payoff vector $r(f,g)$, gives the payoff at each state, while using strategies f and g. For all $f, g \in \Pi_{SR}$, the discounted reward with discount factor $\beta \in (0,1)$, is defined as

$$F_\beta(f,g) := (1-\beta)\sum_{n=0}^{\infty} \beta^n Q^n(f,g) r(f,g) \tag{2}$$

$$= (1-\beta)(I - \beta Q(f,g))^{-1} r(f,g). \tag{3}$$

The s-th component of the vector $F_\beta(f,g)$ gives the payoff with initial state s. The discounted reward of an individual using pure stationary strategy e_l, against g, is given by

$$F_\beta(e_l,g) = (1-\beta)(I - \beta Q(e_l,g))^{-1} r(e_l,g). \tag{4}$$

When the population is in distribution d, inducing f, the average reward of the population against strategy g is given by

$$\tilde{F}_\beta(d,g) = \sum_{l=1}^{L} d_l F_\beta(e_l,g) \tag{5}$$

Now, let us fix the initial state s of the stochastic game. Using this payoff $\tilde{F}_\beta$, the Nash equilibrium and evolutionarily stable strategy (ESS) of the stochastic game can be defined as follows.

Definition 1. *A stationary strategy f^* is a Nash equilibrium, if for all stationary strategies f,*

$$\tilde{F}_\beta(s,d,f^*) \geq \tilde{F}_\beta(s,p,f^*), \tag{6}$$

whenever d and p are population distributions inducing f^* and f respectively.

This definition is equivalent to the existing standard definition of Nash equilibrium for a discounted stochastic game given below.

Definition 2. *A stationary strategy f^* is a Nash equilibrium, if for all stationary strategies f,*

$$F_\beta(s,f^*,f^*) \geq F_\beta(s,f,f^*). \tag{7}$$

Theorem 1. *A stationary strategy f^* is a Nash equilibrium as given in Definition* (1) *if and only if it is a Nash equilibrium as in Definition* (2)*.*

Proof: (i) Let f^* be a NE of the stochastic game as in Definition 2; i.e.,

$$F_\beta(s,f^*,f^*) \geq F_\beta(s,f,f^*), \qquad \forall f \in \Pi_{SR}.$$

In particular

$$F_\beta(s,f^*,f^*) \geq F_\beta(s,e_l,f^*), \qquad l = 1,2,...,L. \tag{8}$$

From Corollary 4.6.6 in [6], we know if f^* is optimal against itself, then so is e_l for all $e_l \in C(f^*)$. From this it follows that whenever d is a population distribution inducing f^*,

$$F_\beta(s,f^*,f^*) = F_\beta(s,e_l,f^*),\ \forall e_l \in C(d) \qquad \text{and} \tag{9}$$

$$F_\beta(s,f^*,f^*) = \sum_{l=1}^{L} d_l F_\beta(s,e_l,f^*). \tag{10}$$

Now, from (8) and (9) we have

$$\tilde{F}_\beta(s,d,f^*) \geq F_\beta(s,e_l,f^*), \qquad 1 \leq l \leq L.$$

Therefore, for any stationary strategy f and distribution p inducing f,

$$\tilde{F}_\beta(s,d,f^*) \geq \tilde{F}_\beta(s,p,f^*),$$

since $\tilde{F}_\beta(s,p,f^*) = \sum_{l=1}^{L} p_l F_\beta(s,e_l,f^*)$.

Conversely, let f^* be a stationary strategy satisfying (6) in Definition 1, i.e.,

$$\tilde{F}_\beta(s,d,f^*) \geq \tilde{F}_\beta(s,p,f^*),$$

whenever d and p are population distributions inducing f^* and f respectively. This clearly yields,

$$F_\beta(s,e_l,f^*) = \max_k F_\beta(s,e_k,f^*), \quad \forall e_l \in C(d). \tag{11}$$

Again, from Corollary 4.6.6 in [6] we know that f^* optimal against itself since e_l optimal against f^* for all $e_l \in C(f^*)$. Using this fact and (11), we obtain

$$F_\beta(s,f^*,f^*) = \max_f F_\beta(s,f,f^*).$$

Hence

$$F_\beta(s,f^*,f^*) \geq F_\beta(s,f,f^*).$$

This shows that f^* is a NE as in Definition 2. □

From here onwards, we refer to Definition 1 of Nash equilibrium as just Nash equilibrium. Note that this definition is introduced using the modified payoff function $\tilde{F}_\beta$, but without changing the equilibrium set of the original game.

We now define an ESS of a stochastic game as a refinement of the Nash equilibrium.

Definition 3. *A stationary strategy $\bar{f}$ is an evolutionarily stable strategy, if the following two conditions are satisfied:*

- $\tilde{F}_\beta(s,d,\bar{f}) \geq \tilde{F}_\beta(s,p,\bar{f})$ *(NE as in Definition 1)*
- *For all $f \neq \bar{f}$,*

$$\tilde{F}_\beta(s,p,\bar{f}) = \tilde{F}_\beta(s,d,\bar{f}) \Rightarrow \tilde{F}_\beta(s,d,f) > \tilde{F}_\beta(s,p,f),$$

whenever d and p are population distributions inducing $\bar{f}$ and f respectively.

Remark 1. *Since the payoff $\tilde{F}_\beta$, for an evolutionary stochastic game is a non-linear, C^1 function on a compact set and hence, satisfying the Lipschitz continuity, we can prove that for an evolutionary stochastic game, $\bar{f}$ is an ESS as defined above, if and only if it is uninvadable (see Theorem 8.3.1 and Theorem 8.3.5, Sandholm 2010[17]); i.e., for every d inducing $\bar{f}$, there exists $\bar{\epsilon} \in (0,1)$, such that for all $\epsilon \in (0,\bar{\epsilon})$ and $f \neq \bar{f}$,*

$$\tilde{F}_\beta(s,d,\epsilon f + (1-\epsilon)\bar{f}) > \tilde{F}_\beta(s,p,\epsilon f + (1-\epsilon)\bar{f}), \quad (12)$$

whenever p is a population distribution inducing f.

B. Replicator Dynamics

The replicator dynamics is one of the most well-studied dynamics in classical evolutionary evolutionary game theory ([19],[17]). It was first introduced in Taylor and Jonker (1978)[20]. In the replicator dynamics there is no assumption of rationality in individuals. The fraction of individuals using a particular strategy increases (decreases) if their average payoff is greater (lesser) than the overall average payoff of the population. Based on this property, we propose the following replicator dynamics for a discounted evolutionary stochastic game.

$$\dot{d}_l = d_l\big(F_\beta(s,e_l,f_d) - \tilde{F}_\beta(s,d,f_d)\big), \quad 1 \leq l \leq L, \quad (13)$$

where f_d is the stationary strategy induced by d. This replicator dynamics is structurally similar to the classical evolutionary games version. Also, from the proof of Theorem 1 in the previous section, we know that $\tilde{F}_\beta(s,d,f_d) = F_\beta(s,e_l,f_d)$, whenever f_d is a Nash equilibrium and $e_l \in C(f_d)$. This along with the above equation shows that a Nash equilibrium strategy is always a rest point of the replicator dynamics 13.

III. The N-step Game

Consider an N-step discounted stochastic game, where the game stops and the payoff is calculated at the end of N-steps. The N-step discounted payoffs can be defined as follows.

$$F_\beta^N(f,g) = (1-\beta)\sum_{n=0}^{N} \beta^n Q^n(f,g) r(f,g), \quad (14)$$

where f and g are stationary strategies. Similar to the infinite horizon case, the s-th component of the vector gives the payoff with initial state s. The linearised discounted payoff is given by

$$\tilde{F}_\beta^N(d,g) = \sum_{l=1}^{L} d_l F_\beta^N(e_l,g), \quad (15)$$

where d is a population distribution and g is a stationary strategy.

Using this N-step game, we can define NE and ESS analogous to the infinite horizon game.

- **Nash Equilibrium:** A stationary strategy $f^N \in \Pi_{SR}$ is an N-step Nash equilibrium, with initial state s, if for all stationary strategy $f \in \Pi_{SR}$

$$\tilde{F}_\beta^N(s,d,f^N) \geq \tilde{F}_\beta^N(s,p,f^N), \quad (16)$$

whenever d and p are population distributions inducing f^N and f respectively.

- **Evolutionary stable strategy:** A stationary strategy $\bar{f}^N$ is an N-step evolutionarily stable strategy, if the following two conditions are satisfied:
  - $\tilde{F}_\beta^N(s,d,\bar{f}^N) \geq \tilde{F}_\beta^N(s,p,\bar{f}^N)$ (N-step NE as in the previous definition)
  - For all $f \neq \bar{f}^N$,

$$\tilde{F}_\beta^N(s,p,\bar{f}^N) = \tilde{F}_\beta^N(s,d,\bar{f}^N) \Rightarrow \tilde{F}_\beta^N(s,d,f) > \tilde{F}_\beta^N(s,p,f),$$

whenever d and p are population distributions inducing $\bar{f}^N$ and f respectively.

A. N-step Replicator Dynamics

The replicator dynamics associated with the N-step stochastic game can be defined as follows.

$$\dot{d}_l = d_l\big(F_\beta^N(s,e_l,f_d) - \tilde{F}_\beta^N(s,d,f_d)\big), \ 1 \leq l \leq L, \quad (17)$$

where f_d is the stationary strategy induced by d.

As $t \to \infty$, since the payoffs $F_\beta^N(s,e_l,f_d)$ and $\tilde{F}_\beta^N(s,d,f_d)$ converge to $F_\beta(s,e_l,f_d)$ and $\tilde{F}_\beta(s,d,f_d)$ respectively, the replicator equation also converges to the infinite horizon equation given in (13). Thus, for any $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that, (13) can be rewritten as

$$\dot{d}_l = d_l\big(F_\beta^N(s,e_l,f_d) - \tilde{F}_\beta^N(s,d,f_d)\big) \pm \epsilon, \ l = 1,2,...,L,$$

where f_d is the stationary strategy induced by d. This implies as $N \to \infty$ (and $\epsilon \to 0$), the trajectories of (17) converges to the corresponding trajectories of (13). Also, the stable states/sets of (17) converge to some stable state/set of (13).

Suppose $N \to \infty$. By definition, we know that the N-step game (payoff) converges to the infinite horizon stochastic game (payoff). In the next section, using this fact, we study the relations between the equilibria of the N-step and infinite horizon game and between the ESS and replicator dynamics of the infinite horizon game.

IV. RESULTS

In this section, we first discuss the relation between the N-step Nash equilibria in (16) and the Nash equilibria in Definition 1, of the infinite horizon stochastic game.

Theorem 2. *Let f^N be an N-step equilibrium of the N-step game, for all $N \geq 1$. Then, for $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that f^N is an ϵ-equilibrium of the infinite horizon stochastic game.*

Proof: By the definition of $F_\beta(f,g)$ (2), we get that

$$\begin{aligned} F_\beta(f,g) =&(1-\beta)\sum_{n=0}^{\infty}\beta^n Q^n(f,g)r(f,g) \\ =&(1-\beta)\sum_{0}^{N}\beta^n Q^n(f,g)r(f,g)+ \\ &(1-\beta)\sum_{n=N+1}^{\infty}\beta^n Q^n(f,g)r(f,g) \\ =&F_\beta^N(f,g)+(1-\beta)\sum_{n=N+1}^{\infty}\beta^n Q^n(f,g)r(f,g). \end{aligned}$$

Since the second term on the right hand side is the tail end of a converging series, it can be arbitrarily small for large enough $N \in \mathbb{N}$.

Let f^N be a Nash equilibrium of the N-step game (16). Then, we know that $\big(\tilde{F}_\beta^N(s,d,f^N) - \tilde{F}_\beta^N(s,p,f^N)\big) \geq 0$. Thus,

$$\begin{aligned} \big(\tilde{F}_\beta(s,d,f^N) - \tilde{F}_\beta(s,p,f^N)\big) &= \big(\tilde{F}_\beta^N(s,d,f^N) - \tilde{F}_\beta^N(s,p,f^N)\big) \pm \epsilon \\ &\geq \pm\epsilon, \end{aligned}$$

for some $\epsilon > 0$. Therefore,

$$\tilde{F}_\beta(s,d,f^N) + \epsilon \geq \tilde{F}_\beta(s,p,f^N).$$

Hence f^N is an ϵ-equilibrium of the infinite horizon stochastic game. □

Now, we look a the dynamic stability of ESS in Definition 1, w.r.t. the associated replicator dynamics (13), for the infinite horizon game. As we know, ESS is a static stability concept. The next theorem extends the folk theorem of evolutionary game theory, which links the ESS and replicator dynamics in classical evolutionary games (see e.g., [21],[19], and [17]). This theorem connect strategies with static stability to dynamic stability. Before stating the theorem, we look at a lemma on ESS, that is useful in proving Theorem 3.

Lemma 1. *Consider an evolutionary stochastic game and a stationary strategy $\bar{f}$. Let D be the set of all population distributions inducing $\bar{f}$. Then, $\bar{f}$ is an ESS if and only if, for each $d \in D$, there exists a neighbourhood U_d in Δ such that $\tilde{F}_\beta(s,d,f_p) > \tilde{F}_\beta(s,p,f_p)$ for all $p \in U_d \setminus D$, where f_p is the stationary strategy induced by p.*

Proof: Let $\bar{f}$ be an ESS of the evolutionary stochastic game. By Remark 1, $\bar{f}$ is uninvadable and, for each $d \in D$, there is a uniform invasion barrier $\bar{\epsilon} \in (0,1)$. Let $Z_d \subset bd(\Delta)$ be the union of all boundary faces of Δ not containing d, i.e., $Z_d = \{q \in \Delta : q_l = 0 \text{ for some } l \in C(d)\}$. Let $U_d = \{p \in \Delta : p = \epsilon q + (1-\epsilon)d \text{ for some } q \in Z_d \text{ and } \epsilon \in [0,\bar{\epsilon})\}$. Clearly Z_d is a closed and non empty set which does not contain d. Suppose $p \in U_d \setminus D$. Since $\bar{f}$ is uninvadable, from Remark 1, we have

$$\tilde{F}_\beta(s,d,f_p) > \tilde{F}_\beta(s,q,f_p) = \tilde{F}_\beta(s,(\frac{p-(1-\epsilon)d}{\epsilon}),f_p).$$

By rearrangement and linearity of the first argument of $\tilde{F}_\beta(s,\cdot,\cdot)$, we now obtain the required inequality,

$$\tilde{F}_\beta(s,d,f_p) > \tilde{F}_\beta(s,p,f_p).$$

Conversely, let $\tilde{F}_\beta(s,d,f_p) > \tilde{F}_\beta(s,p,f_p)$ for all $p \in U_d \setminus D$, where U_d is a neighbourhood of d, $d \in D$. Since U_d is an open neighbourhood of d, for each $d \in D$, there is an open ball centered at d and contained in U_d. This implies that there exists $\bar{\epsilon} \in (0,1)$ such that, for every $f \neq \bar{f}$ and $q \in \Delta$ inducing f, $p = \epsilon q + (1-\epsilon)d$ lies in this open ball and hence in U_d, whenever $0 \leq \epsilon < \bar{\epsilon}$. From this we have,

$$\tilde{F}_\beta(s,d,f_p) > \tilde{F}_\beta(s,p,f_p) = \tilde{F}_\beta(s,\epsilon q + (1-\epsilon)d,f_p).$$

After rearranging, we get $\tilde{F}_\beta(s,d,f_p) > \tilde{F}_\beta(s,q,f_p)$, i.e., $\bar{f}$ is uninvadable and, by Remark 1 it is an ESS. □

Theorem 3. *Consider an evolutionary stochastic game with an ESS $\bar{f}$. Let D be the set of all population distributions inducing $\bar{f}$. Then D is an asymptotically stable set of the replicator dynamics* (13)*.*

Proof: For $d \in D$, let $Q_d = \{p \in \Delta : C(d) \subset C(p)\}$ and $H_d : Q_d \to \mathbb{R}$ be a function defined by

$$H_d(p) := -\sum_{l \in C(d)} d_l \; log\Big(\frac{p_l}{d_l}\Big).$$

Here, $H_d(p)$ is the Kullback-Leibler divergence, and Q_d is the domain of $H_d(\cdot)$. The time derivative of H_d along the replicator trajectory $\xi(t,p)$ is given by

$$\begin{aligned} \dot{H}_d(p) = \frac{d}{dt}\big[H_d(\xi(t,p))\big]_{t=0} &= \sum_{l \in C(d)} \frac{\partial H_d(p)}{\partial p_l}[\dot{\xi}_l(t,p)]_{t=0} \\ &= -\big(\tilde{F}_\beta(s,d,f_p) - \tilde{F}_\beta(s,p,f_p)\big). \end{aligned} \tag{18}$$

Let U_d be as in Lemma 1 and $W_d = U_d \cap Q_d$. From (18) and the inequality in Lemma 1, it follows that $\dot{H}_d(p) \leq 0$ whenever $p \in W_d$, equality implying $p \in D$. Also, we know that $H_d(p) \geq 0$, and equality holds if and only if $p = d$.

Note that d is a (relative) interior point of W_d because it is so for both U_d and Q_d. Since H_d is continuous, there exists $\alpha_d > 0$ small enough so that the neighbourhood $P_d = \{p \in Q_d : H_d(p) < \alpha_d\}$ and its closure $\bar{P}_d$ are contained in W_d. This implies that the neighbourhood $P = \cup_{d \in D} P_d$, of D, is contained in $W = \cup_{d \in D} W_d$.

Let $p_0 \in P$ and $\xi(t,p_0)$, the replicator trajectory (13) starting from p_0. Then, there exists $d \in D$ such that $p_0 \in P_d \subset W_d$. Therefore $\dot{H}_d(p_0) \leq 0$ and $\xi(t,p_0) \in P_d$ for all $t > 0$. Let Ω be the set of all ω-limit points of $\xi(t,p_0)$. Then $H_d(p)$ is constant and $\dot{H}_d(p) = 0$ for all $p \in \Omega$. This together with (18) shows that, if $p \in \Omega$, $\tilde{F}_\beta(s,d,f_p) = \tilde{F}_\beta(s,p,f_p)$

and hence $p \in D$. Thus we get $\Omega \subseteq D$, that is, all ω-limit points of the trajectory $\xi(t, p_0)$ lie in D.

Let $d^* \in \Omega$, and note that $\bar{P}_{d^*} = \{p \in Q_{d^*} : H_{d^*}(p) \leq \alpha_{d^*}\}$. Since $d^* \in \Omega$, the trajectory $\xi(t, p_0)$ will enter $\bar{P}_{d^*}$ at some time t and will remain in $\bar{P}_{d^*}$ as $\dot{H}_{d^*} \leq 0$. As a result, $\Omega \subseteq \bar{P}_{d^*}$. Since $\alpha_{d^*} > 0$ can be chosen arbitrarily small, we get $\Omega = \{d^*\}$. Thus, for each $p_0 \in P$, there exists a distribution $d^* \in D$ (which may depend on p_0) such that the trajectory $\xi(t, p_0)$ converges to d^*. Hence the set D is attracting.

It now remains to show that D is Lyapunov stable. To this end, let U be a neighbourhood of the set D. We can choose $\alpha_d > 0$, $d \in D$, such that $P = \cup_{d \in D} P_d$ is a subset of U. Let $p_0 \in P$ and $\xi(t, p_0)$, the trajectory with initial point p_0. This implies that $p_0 \in P_d$ for some $d \in D$ and hence $\xi(t, p_0) \in P_d \subseteq U$, for all $t > 0$. This shows that the set D is Lyapunov stable.

Since D is both attracting and Lyapunov stable, we conclude that the set D is asymptotically stable with respect to the replicator dynamics. □

V. EXAMPLE

First let us look at an example of a stochastic game that illustrates the dynamic stability result proved in the previous section.

Example 1. *Consider an stochastic game as given below.*

0 , (3/4,1/4)	*-2 , (1/2,1/2)*
2 , (1/2,1/2)	*0 , (3/4,1/4)*

state 1

-4 , (1/3,2/3)	*-6 , (1/4,3/4)*
6 , (1/4,3/4)	*4 , (1/5,4/5)*

state 2

The first entries in the cells give the payoff of player 1. Since we are considering symmetric games, the payoff matrix of the player 2 is the transpose of that of player 1. The ordered pair represents the probability with which the stochastic game moves from one state to the next.

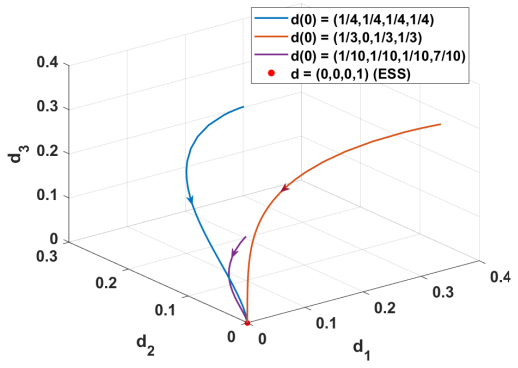


Fig. 1. Replicator trajectory for Ex.1, with $\beta = 0.3$.

It is easy to verify that the stationary strategy $f = ((0,1),(0,1))$, which is induced by a single population distribution $d = (0,0,0,1)$ is a Nash equilibrium of the above stochastic game. Moreover f is an ESS of the same. We plot the trajectories of the replicator dynamics (13)*, for this game with discount factor $\beta = 0.3$. Since d_4 is dependent on the other three components, we plot the trajectories in the (d_1, d_2, d_3) space. We find that the trajectories starting in a neighbourhood of d converge to d eventually (Fig 1).* □

VI. CONCLUSION

In this paper, we study the discounted reward version of the evolutionary stochastic game model introduced in Flesch et al(2013)[8]. We provide an alternate definition for Nash equilibrium of a stochastic game, which may be easier to work with (due to linearity of payoff in the first variable), and define an ESS for the same. This ESS is then found to asymptotically stable w.r.t. the associated replicator dynamics. We also discuss the N-step stochastic game and explore its relations with the infinite horizon game. Considering the N-step game may be useful in devising an actual play of the evolutionary stochastic game, that gives us the discounted reward $\tilde{F}_\beta$ and also in discretisation of the replicator dynamics. Even though the current setting considers only irreducible and symmetric stochastic settings, it may be possible to include any finite stochastic game, by making appropriate modifications.

This model introduces a new approach to look at stochastic games through the lens of evolutionary game theory. It presents an analogy between population games ([17]) and stochastic games, thus enabling us to define an ESS an replicator dynamics for stochastic games. Since the basic model is established it can used to further explore the evolutionary properties of stochastic game like best-response dynamics. Also, from an evolutionary game theoretic viewpoint, in this model we are able to concentrate on more than one payoff matrices (i.e. more than one type of interaction) between the same individuals. In a biological setting this may mean that we can include fight for resources/territory and mating preferences in the game and look at it simultaneously as it often happens in a real life setting.

ACKNOWLEDGEMENT

The authors thank Prof. A.J. Shaiju for useful discussions during the preparation of this paper.

REFERENCES

[1] L. S. Shapley, "Stochastic games," *Proc. Natl. Acad. Sci. USA*, vol. 39, no. 10, pp. 1095–1100, 1953.

[2] D. Gillette, *Representable infinite games*. University of California, Berkeley: PhD.Thesis, 1953.

[3] A. M. Fink, "Equilibrium in a stochastic n-person game," *J. Sci. Hiroshima Univ. Ser. A-I Math.*, vol. 28, no. 1, p. 89–93, 1964.

[4] M. Takahashi, "Equilibrium points of stochastic non-cooperative n-person games," *J. Sci. Hiroshima Univ. Ser. A-I Math.*, vol. 28, no. 1, pp. 95–99, 1964.

[5] J. Mertens and A. Neymen, "Stochastic games," *Internat. J. Game Theory*, vol. 10, p. 53–66, 1981.

[6] J. Filar and K. Vrieze, *Competitive Markov Decision Processes*. New York: Springer-Verlag, 1997.

[7] J. Maynard Smith and G. R. Price, "The logic of animal conflict," *Nature*, vol. 246, pp. 15–18, 1973.

[8] J. Flesch, T. Parthasarathy, F. Thuijsman, and P. Uyttendaele, "Evolutionary stochastic games," *Dyn. Games Appl.*, vol. 3, pp. 207–219, 2013.

[9] D. Murali and A. J. Shaiju, "Dynamic stability of the set of nash equilibria in stable stochastic games," *Journal of Dynamics and Games*, vol. 10, no. 3, pp. 270–286, 2023.

[10] ——, "Best-response dynamics for evolutionary stochastic games," *International Game Theory Review, Early Access*, vol. 25, no. 4, p. 2350010, 2023.

[11] V. Borkar, "Reinforcement learning in markovian evolutionary games," *Advances in Complex Systems*, vol. 5, no. 1, p. 55–72, 2002.

[12] D. Hennes, K. Tuyls, and M. Rauterberg, "State-coupled replicator dynamics," in *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 2009, pp. 789–796.

[13] E. Altman, Y. Hayel, H. Tembine, and R. ElAzouzi, "Markov decision evolutionary games with time average expected fitness criterion," in *3rd International ICST Conference on Performance Evaluation Methodologies and Tools*, vol. 5. ICST, 2010.

[14] I. Brunetti, Y. Hayel, and E. Altman, "State policy couple dynamics in evolutionary games," in *2015 American Control Conference (ACC)*, 2015, pp. 1758–1763.

[15] C. Hilbe, Š Šimsa, K. Chatterjee, and M. A. Nowak, "Evolution of cooperation in stochastic games," *Nature*, vol. 559, p. 246–249, 2018.

[16] D. S. Leslie, S. Perkins, and Z. Xu, "Best-response dynamics in zero-sum stochastic games," *Journal of Economiv Theory*, vol. 189, p. 105095, 2020. [Online]. Available: https://doi.org/10.1016/j.jet.2020.105095

[17] W. Sandholm, *Population Games and Evolutionary Dynamics*. Cambridge: MIT Press, 2010.

[18] O. J. Vrieze, "Stochastic games and stationary strategies," *In: Neyman A., Sorin S. (eds) Stochastic Games and Applications. NATO Science Series*, vol. 570, pp. 37–50, 2003.

[19] R. Cressman, *Evolutionary Dynamics and Extensive Form Games*. Cambridge: MIT Press, 2003.

[20] P. D. Taylor and L. B. Jonker, "Evolutionarily stable strategies and game dynamics," *Math. Biosci.*, vol. 40, pp. 145–156, 1978.

[21] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics*. Cambridge: Cambridge University Press, 1998.