State Policy Couple Dynamics in Evolutionary Games*

Ilaria Brunetti, Yezekael Hayel and Eitan Altman†

Abstract— Standard Evolutionary Game framework is a useful tool to study large interacting systems and to understand the strategic behavior of individuals in such complex systems. Adding an individual state to model local feature of each player in this context, allows one to study a wider range of problems in various application areas as networking, biology, etc. In this paper, we introduce such an extension of evolutionary game framework and particularly, we focus on the dynamical aspects of this system. Precisely, we study the coupled dynamics of the strategies and the individual states inside a population of interacting individuals. We consider here a two strategies evolutionary game. We first obtain a system of combined dynamics and we show that the rest-points of this system are equilibria of our evolutionary game with individual state. Second, by assuming two different time scales between states and strategy dynamics, we can compute explicitly the equilibria. Then, by transforming our evolutionary game with individual states into a standard evolutionary game, we obtain an equilibrium which is equivalent, in terms of occupation measure, to the previous one. Finally, we show a generalization of the model. All our results are illustrated with numerical results.

I. Introduction

Evolutionary Game Theory (EGT) has been first introduced by J.Maynard Smith [1] to model the evolution of species in biology. It makes use of Game Theory tools to describe the dynamics of populations sizes as a result of a competition between them, where players are repeatedly and randomly matched through pairwise interactions. While in classical GT players are supposed to be rational individuals which interact and choose their strategies in order to maximize the individual utility function, in EGT there is no rationality assumption. All players in a population are supposed to use some strategy (or behavior type) and the utility is interpreted as a Darwinian fitness depending on the behaviors of others and thus on the population's profile, i.e. on the frequencies of the strategies in the whole population. Strategies with higher payoff are supposed to spread within the population.

A key notion in EGT is the *Evolutionarily Stable Strategy* (ESS), which is a strategy such that, if adopted by all the players, is robust against deviations of a (possibly small) fraction of the population. From a biological point of view it can be seen as a generalization of Darwin's idea of survival of the fittest, while from a GT perspective it is a refinement of the Nash Equilibrium, which satisfies a stability property.

In order to explain how a population get to a stable situation, one need to introduce another fundamental concept of EGT, the replicator dynamics, which serves to highlights the role of selection from a dynamic perspective. It is formalized by a system of ordinary differential equations and it establishes that the evolution of the size of the populations depends on the fitness they get in interactions. A strategy will spread if it performs better than the average strategy of the whole population. The *folk theorem of evolutionary games* allows to establish a strict connection between the stable points of the Replicator Dynamics and the Nash Equilibria.

In this work we want to extend the EGT models by introducing the concept of individual state. We analyze a particular simple case, in which we associate a state to each player, and we suppose that this state determines the set of available actions. We consider deterministic stationary policies and we suppose that the choice of a policy determines the fitness of the player and it impacts the evolution of the state. We define the interdependent dynamics of states and policies and we introduce the State Policy coupled Dynamics (SPcD) in order to study the evolution of the population profile and we prove the relation between the rest points of our system and the equilibria of the game. We then assume that the processes of states and policies move with different velocities: this assumption allows us to solve the system and then to find the equilibria of our game with two different methods: the singular perturbation method and a matrix approach.

The paper is organized as follows: in Section II we briefly present standard EGT main definitions and results. In Section III we extend EGT problems introducing individual states. We define the notion of equilibrium profile of the population in such context and the dynamics of states and policies. We show the relation between the rest points of the system and the equilibria of the game. We then provide, in section IV, the two different methods to solve this system, which can be applied when assuming a two-timescales behavior. We conclude with some numerical results.

II. Standard Evolutionary Game Theory

A. Evolutionary Stable Strategy

Consider an infinitely large population of players, where each player repeatedly meets a randomly selected individual within the population. Each individual disposes of a finite pure action space $\mathcal{A}$, $|\mathcal{A}| = K$. Let $\Delta(\mathcal{A}) = \{p \in \mathbb{R}_+^K \sum_{i \in \mathcal{A}} p_i = 1\}$ be the set of mixed strategies, that are probability measures over the action space. We define by $F(p, q)$ the expected payoff of an individual playing p against an opponent using q, where $p, q \in \Delta(\mathcal{A})$. If A is the

†The authors are with the Maestro team of INRIA, 2004, Route des Lucioles, 06902 Sophia Antipolis, France, and LIA/CERI, University of Avignon, 84911, Avignon, France. Email: {Ilaria.Brunetti,Eitan.Altman}@inria.fr, yezekael.hayel@univ-avignon.fr

payoff matrix associated to the pairwise interactions, then $F(p,q) = p^T A q$.

An ESS is a strategy that, if adopted by the whole population, it is resistant against mutations of a small fraction of individuals in the population. Suppose that the whole population adopts a strategy q, and that a fraction ϵ of *mutants* deviate to strategy p. Strategy q is an ESS if $\forall p \neq q$, there exists some $\epsilon_p > 0$ such that:

$$\forall \epsilon \in (0, \epsilon_p) \qquad F(q, \epsilon p + (1-\epsilon)q) > F(p, \epsilon p + (1-\epsilon)q) \quad (1)$$

The following proposition allows to characterize an ESS through its stability properties.

Proposition 1: $q \in \Delta(\mathcal{A})$ is an ESS if and only if it satisfies the following conditions:

- *Nash Condition:* $F(q,q) \geq F(p,q) \qquad \forall p$,
- *Stability Condition:* $F(q,q) = F(p,q) \Rightarrow F(q,p) \geq F(p,p) \qquad \forall p \neq q$.

It immediately follows that any strict Nash equilibrium is an ESS, while the converse is not true. When players are programmed to pure actions and mixed ones are not allowed, the notion of ESS is associated to the state of the population instead that to a mixed action, as introduced by Taylor and Jonker in [6]. The population state is given by the vector $q = (q_1, \ldots, q_K)$, where $q_i = \frac{N_i}{N}$ is the proportion of individuals in the population playing pure action i. We observe that $q \in \Delta(\mathcal{A})$, so it is formally equivalent to a mixed action in $\Delta(\mathcal{A})$. Let $F(q,p)$ denotes the immediate expected payoff of a group of individuals in state q playing against a population in state p, and let $F(i,p)$ denotes the immediate expected payoff of an individual playing pure strategy i against a population in state p. We have that: $F(q,p) = \sum_{i=1}^K q_i F(i,p)$.

The definition of the ESS concerning states is then equivalent to that of the ESS concerning strategies: state q is an ESS if $\forall p \neq q$, there exists some $\epsilon_p > 0$ such that $F(q, \epsilon p + (1-\epsilon)q) > F(p, \epsilon p + (1-\epsilon)q)$, $\forall \epsilon < \epsilon_p$.

B. Replicator Dynamics

Let the population be of large but finite size N and let players be programmed to pure actions $\mathcal{A}$. Let N_i be the number of individuals adopting $i \in \mathcal{A}$. The population profile at time t is given by the vector $q(t) = (q_1(t), q_2(t), \ldots, q_K(t))$, $q(t) \in \Delta(\mathcal{A})$, where $q_i = \frac{N_i}{N}$ is the fraction of individuals playing pure action i. The replicator dynamics describes how the distribution of pure actions evolves in time depending on interactions between individuals. Replicator dynamics of action $i \in \mathcal{A}$ is expressed by the following equation:

$$\dot{q}_i(t) = q_i(t)(F_i(q(t)) - \bar{F}(q(t))), \qquad (2)$$

where $F_i(q(t))$ is the immediate fitness of an individual playing i and $\bar{F}(q(t))$ is the average immediate fitness of the population. In the two-actions case, with $\mathcal{A} = \{x,y\}$, if $q(t)$ indicates the share of the population playing action x at time t, the latter equation can be rewritten as follows: $\dot{q}(t) = q(t)(1 - q(t))(\bar{F}_x(q(t)) - \bar{F}_y(t))$.

The replicator equation has numerous properties and there is a close relationship between its rest points and the equilibria. The *folk theorem of evolutionary game theory* [7] states that:

1) any Nash equilibrium is a rest point of the replicator equation;
2) if a Nash equilibrium is strict then it's asymptotically stable;
3) if a rest point is the limit of an interior orbit for $t \to \infty$, then it is a Nash equilibrium;
4) any stable rest point of the replicator dynamics is a Nash equilibrium.

Any ESS is an asymptotically stable rest point and an interior ESS is globally stable, but the converse does not hold in general.

III. INDIVIDUAL STATE IN EGT FRAMEWORK

A. Individual State and its dynamics

In this section, we introduce the concept of *individual state* in EGT framework and we present the consequent dynamics of our model. We consider a population of N individuals, where each individual can be in one of two possible states, $\mathbb{S} = \{1, 0\}$; every individual goes through a cycle that starts at state 1, moves to states 0 after some random time at a rate that depends on its policy. After some exponentially distributed time it returns to state 1 and so on. At each pairwise interaction, the set of available actions of a player depends on its state: in state 1, $\mathcal{A}_1 = \{x, y\}$, whereas in state 0 an individual can only use y.

We consider the set of deterministic stationary policies, which are functions that associate to each state an action in $\mathcal{A}$ and do not depend on time. Let u_x (resp. u_y) be the deterministic stationary policy which consists in always playing action x (resp. y) in state 1. In state 0, an individual always plays y. Each player chooses one deterministic stationary policy and we denote by $q_x(t)$ the proportion of individuals in the population that play the deterministic stationary policy u_x at time t. In the same way, $q_y(t)$ denotes the proportion of individuals that play the deterministic stationary policy u_y at time t and we have $q_y(t) = 1 - q_x(t)$.

We suppose that the policy chosen impacts the utility of the player and also the time he spends in state 1. We define by μ_i the rate of decay from state 1 to state 0 when using policy u_i, $i \in \{x, y\}$, where $\mu_x > \mu_y$, and by μ the rate of change from state 0 to state 1. The rates depend on the action induced by the policy chosen by the player. Then, by abuse of description, we say that the rates are controlled by the policies, but in fact, they are controlled by the actions induced by the policies.

We define the proportion of individuals that are in state 1 at time t as $p_1 \equiv \frac{N_1}{N} = \frac{1}{N} \sum_{i=1}^N \mathbb{1}_{\{x_i=1\}}$, where x_i denotes the state of player $i \in \{1, \ldots, N\}$ and $\mathbb{1}_{\{x_i=1\}} = 1$ when $x_i = 1$ and it's zero otherwise. We denote by Y_i the random variable $Y_i \equiv \mathbb{1}\{x_i = 1\} \in \{0, 1\}$. Y_i has a Bernoulli distribution, such that $\mathbb{P}(Y_i = 1) = 1 - \mathbb{P}(Y_i = 0) = p_1^i$, and thus $\mathbb{E} = p_1^i$. The individual state dynamics can be described by the following differential equation:

$$\dot{p}_1^i(t) = -\mu_x p_1^i(t) q_x(t) + \mu(1 - p_1^i(t)) - \mu_y p_1^i(t)(1 - q_x(t)),$$

If each individual plays policy u_x with a probability q_x, $(Y_i)_i$ are i.i.d. and thus, for the strong law of large numbers, when $N \to \infty$, $p_1 \to p_1^i$ and consequently:

$$\dot{p}_1(t) = -\mu_x p_1(t) q_x(t) + \mu(1 - p_1(t)) - \mu_y p_1(t)(1 - q_x(t)), \tag{3}$$

where $q_x(t)$ is the proportion of individuals in the population that play the deterministic stationary policy u_x at time t.

B. Individual fitness

At each pairwise interaction, the immediate fitness obtained by an individual depends on his current action and the current action of its opponent. It is given by the following fitness matrix:

$$A := \begin{matrix} & \begin{matrix} x & \quad y \end{matrix} \\ \begin{matrix} x \\ y \end{matrix} & \begin{pmatrix} (a,a) & (b,c) \\ (c,b) & (d,d) \end{pmatrix} \end{matrix} \tag{4}$$

where x and y are the available actions and the matrix entry $A_{i,j}$ indicates the payoff respectively of the first (row) and the second (column) player. The immediate expected fitness of a player interacting at time t, depends on the population profile at that time. As we added a state component in our framework, the population profile at time t is now expressed by the couple $\xi(t) := (p_1(t), q_x(t))$.

We define $\bar{J}_i^1(\xi(t))$ (resp. $\bar{J}_i^0(\xi(t))$) as the immediate expected fitness of an individual who is in state 1 (resp. 0) and plays deterministic stationary policy u_i against a population whose profile is $\xi(t)$. It follows that the value of $\bar{J}_x^1(\xi(t))$ corresponds to the immediate expected fitness of an individual playing action x, against a population whose profile is $\xi(t)$. We observe that $\bar{J}_y^1(\xi(t)) = \bar{J}_y^0(\xi(t)) = \bar{J}_x^0(\xi(t))$, being all equals to the value of the immediate expected fitness of an individual playing action y; we thus denote it simply as $\bar{J}_y(\xi(t))$. We obtain the following expressions:

$$\bar{J}_x^1(\xi(t)) := p_1(t)(q_x(t)a + (1 - q_x(t))b) + (1 - p_1(t))b,$$

$$\bar{J}_y(\xi(t)) := p_1(t)(q_x(t)c + (1 - q_x(t))d) + (1 - p_1(t))d.$$

The immediate expected fitnesses $\bar{F}_i(\xi(t))$ at time t of an individual playing deterministic stationary policy u_i, $i \in \{x, y\}$, are given by:

$$\bar{F}_x(\xi(t)) = p_1^i(t)\bar{J}_x^1(\xi(t)) + (1 - p_1^i(t))\bar{J}_y(\xi(t))$$

$$\bar{F}_y(\xi(t)) = \bar{J}_y(\xi(t)).$$

The average expected fitness of the whole population with profile $\xi(t) = (p_1(t), q_x(t))$ is defined as:

$$\bar{F}(\xi(t)) = q_x(t)\bar{F}_x(\xi(t)) + (1 - q_x(t))\bar{F}_y(\xi(t)).$$

C. Evolutionary Stable Strategies

We want to study the properties of stability of the population profile, supposing that individuals can play only deterministic policies.

We then define the equilibrium profile of the population as follows:

Definition 1 (Equilibrium profile): A population profile $\xi^* = (p_1^*, q_x^*)$ is an equilibrium profile if $\forall u_i \in supp(q_x^*)$ we have that:

$$\bar{F}_i(\xi^*) \geq \bar{F}_j(\xi^*) \qquad j \neq i,$$

where $supp(q_x^*) = \{u_i, \ i \in \mathcal{A} | q_i^* > 0 \text{ given the state } q_x^*\}$.

Proposition 2: If the population profile $\xi^* = (\tilde{p}_1^*, \tilde{q}_x^*)$ satisfies the indifference principle:

$$\bar{F}_x(\xi^*) = \bar{F}_y(\xi^*),$$

then it is an equilibrium profile.

Proof: Let's consider the population profile $\xi^* = (\tilde{p}_1^*, \tilde{q}_x^*)$. From the definition, if $supp(q_x^*) = \{x, y\}$ then ξ^* is an equilibrium profile if $\bar{F}_x(\xi^*) \geq \bar{F}_y(\xi^*)$ and $\bar{F}_y(\xi^*) \geq \bar{F}_x(\xi^*)$, which implies that $\bar{F}_x(\xi^*) = \bar{F}_y(\xi^*)$. In this case the indifference principle is thus equivalent to the equilibrium definition. If $supp(q_x^*) = \{i\}$, then if ξ^* satisfies the indifference principle $\bar{F}_x(\xi^*) = \bar{F}_y(\xi^*)$, then the equilibrium condition $\bar{F}_i(\xi^*) \geq \bar{F}_j(\xi^*) \quad j \neq i$, is satisfied with equality. ∎

We now look for the stability condition of the equilibrium state. This leads to the formal definition of the ESS: in our model it corresponds to an evolutionary stable population profile. Given $\tilde{q}_x \in [0, 1]$ and a population profile $\xi = (p_1, q_x)$, the average fitness of a group of individuals such that a proportion $\tilde{q}_x$ of the group play stationary deterministic policy u_x against a population whose profile is $\xi = (p_1, q_x)$ is given by:

$$\bar{F}_{\tilde{q}_x}(\xi) = \tilde{q}_x \bar{F}_x(\xi) + (1 - \tilde{q}_x)\bar{F}_y(\xi).$$

Definition 2 (ESS): A population profile $\xi^* = (p_1^*, q_x^*)$ is an ESS if $\forall \xi = (p_1, q_x)$, the two following conditions hold:
1) $\bar{F}_{q_x^*}(\xi^*) \geq \bar{F}_{q_x}(\xi^*)$,
2) $\bar{F}_{q_x^*}(\xi^*) = \bar{F}_{q_x}(\xi^*) \Rightarrow \bar{F}_{q_x^*}(\xi) \geq \bar{F}_{q_x}(\xi)$.

D. Policy Based Replicator Dynamics

As we focus here on policies instead that on strategies, we introduce a policy based replicator dynamics (PbRD), to study the evolution of the share of individuals $q_x(t)$ using pure policy u_x at time t. The PbRD is given by the following equation:

$$\dot{q}_x(t) := q_x(t)(\bar{F}_x(\xi(t)) - \bar{F}(\xi(t))). \tag{5}$$

Then, the growth rate of the population share using policy u_x is:

$$\frac{\dot{q}_x(t)}{q_x(t)} = \bar{F}_x(\xi(t)) - \bar{F}(\xi(t)), \tag{6}$$

The PbRD can be written as:

$$\begin{aligned} \dot{q}_x(t) &= q_x(t)(1 - q_x(t))(\bar{F}_x(\xi(t)) - \bar{F}_y(\xi(t))), \\ &:= g(p_1(t), q_x(t)). \end{aligned}$$

We now investigate the dynamics of actions, where the fitness is a function of the population profile depending on policies and states. If we pick one random individual in the population at time t, the probability that he plays pure action x, is given by the product $q(t) = q_x(t)p_1(t)$, which leads to:$\dot{q}(t) = \dot{q}_x(t)p_1(t) + q_x(t)\dot{p}_1(t)$. By carrying out the expression of $\dot{q}_x(t)$, after some basic algebra, we get the following equation for the growth rate of the proportion of individuals playing action x in the population at time t:

$$\frac{\dot{q}(t)}{q(t)} = (\bar{F}_x(\xi(t)) - \bar{F}(\xi(t))) + \frac{\dot{p}_1(t)}{p_1(t)} \qquad (7)$$

Equation (7) shows how the evolution of states impacts the dynamics of actions in our context. The growth rate of action x is increasing in the growth rate of state 1. We observe that a sufficiently high growth rate of state 1 can leads to a growing rate of action x even if policy u_x is non-optimal.

E. State-Policy Coupled Dynamics

The replicator dynamics of our model are defined by the following system of State-Policy Coupled Dynamics (SPcD) which combines the dynamics of the individual state and the dynamics of the policies used in the population:

$$(S) \begin{cases} \dot{p}_1 = h(\xi(t)) \\ \dot{q}_x = g(\xi(t)) \end{cases}$$

where $\xi(t) = (p_1(t), q_x(t))$ corresponds to the population profile. The rest points of the SPcD is the couple $\xi^* = (p_1^*, q_x^*)$ satisfying:

$$\begin{cases} h(\xi^*) = 0 \\ g(\xi^*) = 0. \end{cases} \qquad (8)$$

Lemma 1: Any interior rest point of the SPcD (S) is a state equilibrium of the state-policy game.

Proof: Trivially, if ξ^* is internal, it satisfy the indifference principle $\bar{F}_x(\xi^*) = \bar{F}_y(\xi^*)$, so ξ^* is an equilibrium profile. ∎

Remark 1: Note that the converse does not necessarily hold. Any equilibrium state is a rest point of the PbRD in (5), but it's not necessarily a rest point of the individual state dynamics.

Lemma 2: Any stable rest point of the SPcD (S) is an equilibrium profile of the state-policy game.

Proof: We have already proved in lemma 1 that any interior rest point is an equilibrium profile; it only remains to prove that a boundary stable point is an equilibrium. Let suppose that the boundary point $\xi^* = (p_1*, 0)$ is stable but not an equilibrium. This implies that there exists some neighborhood U_{ξ^*} of ξ^* such that $\forall \xi \in U_{\xi^*}, \xi \neq \xi^*$, $\bar{F}_y(\xi) > \bar{F}_x(\xi)$. Then the PbRD in $U_{\xi^*} =$, for $\xi \neq \xi^*$ is such that $q_x = q_x(1-q_x)(\bar{F}_x(\xi) - \bar{F}_y(\xi)) < 0$, implying that there exists a time T such that $q(t) \notin U_{\xi^*}, \forall t > T$. It is a contradiction and hence ξ^* is an equilibrium profile. ∎

Finally, in order to guarantee that a rest point is an ESS, we need more properties on the rest point of the SPcD. A sufficient conditions to guaratee evolutionary stability of a rest point is the strong stability [7]. Another method to verify

that a rest point ξ^* is an ESS is to construct a suitable local Lyapunov function [8] for the raplicator dynamics in ξ^*.

IV. TWO TIME-SCALES BEHAVIOR

We assume here that the state and the policy dynamics move with different velocities. The individual state dynamics, given by equation (3), are supposed to move very fast compared to the slow updating strategy process modeled through the SPcD (3). This assumption allows us to find the equilibrium profile of the population with two different approaches. As we showed the relation between the equilibria of the game and the rest points of the SPcD, we can solve the system (S) to find the equilibria. Alternatively, we can consider the stationary distribution of states and rewrite our model as a matrix game.

A. Singular perturbations

If we consider the two-time-scales behavior of the system (S), we can approximate its solution using the standard Singular Perturbation Model [3] to find the rest points of the SPcD. We introduce the parameter $\epsilon > 0$, such that :

$$\epsilon \dot{p}_1 := h(p_1, q_x).$$

We rewrite the system of the two coupled differential equations:

$$(S_\epsilon) \begin{cases} \epsilon \dot{p}_1 = h(p_1, q_x), \\ \dot{q}_x = g(p_1, q_x). \end{cases}$$

The parameter ϵ is a small positive scalar which serves to represent the different timescales of the two processes, where the velocity of the state process, $\dot{p}_1 = h(p_1, q_x)/\epsilon$, is fast when ϵ is small.

The theory of singular perturbed differential equations gives an easy way to solve an approximation of the system when $\epsilon \to 0$. We can consider the *quasi-steady-state-model* [3] by first solving in p_1 the transcendental equation $0 = h(p_1, q_x)$ and then rewriting the differential equation $\dot{q}$ as a function of the obtained roots. As the latter equation has a unique real solution $\bar{p}_1 := \pi_1(q_x)$, our system is in *normal form* . This allows us to solve the second differential equation called the quasi-steady-state equation:

$$\dot{q}_x = g(\pi_1(q_x), q_x). \qquad (9)$$

As the assumption [3] $\frac{\partial h}{\partial p_1}(p_1, q_x) < 0$ is satisfied, the reduced model is a good approximation of the original system. The two-time-scale behavior of $p_1(t)$ and $q_x(t)$ has a geometric interpretation, as trajectories in $\mathbb{R}^2$. If we define the manifold sets $M_\epsilon := \{\phi$ s.t. $p_1 = \phi(q_x, \epsilon)$ and $\epsilon = h(q_x, \phi(q_x, \epsilon))\}$, it is possible to rewrite the problem in terms of invariant manifolds. When $\epsilon = 0$, the manifold M_0 corresponds to the expression of the quasi steady state model. When the condition $\frac{\partial h}{\partial p_1}(p_1, q_x) < 0$ is satisfied, we have that the equilibrium manifold M_0 is stable (attractive). Particularly, the important result is that the existence of a conditionally stable manifold M_0 for $\epsilon = 0$ implies the existence of an invariant manifold M_ϵ satisfying the following convergence for all $\epsilon \in [0, \epsilon^*]$:

$$\phi(\epsilon, q_x) \to \phi(0, q_x), \quad \text{and} \quad M_\epsilon \to M_0 \quad \text{as} \quad \epsilon \to 0.$$

The positive constant ϵ^* is determined such that the following manifold condition is satisfied:

$$\epsilon \frac{\partial \phi}{\partial x} g(\phi(q_x, \epsilon), q_x) = h(\phi(q_x, \epsilon), q_x),$$

for all q_x and $\epsilon \in [0, \epsilon^*]$. The attractiveness of the slow manifold M_0 is illustrated in the numerical illustrations section. Let us now compute the solution of the approximate system (S_0). We then consider the stationary regime of the individual state dynamics (expressed by Equation (3)). By imposing $\dot{p}_1 = 0$, we obtain the following slow manifold $M_0 := \{\phi \quad \text{s.t.} \quad p_1 = \phi(q_x, 0) \quad \text{and} \quad 0 = h(q_x, \phi(q_x, 0))\}$:

$$\phi(q_x, 0) = \frac{\mu}{\mu + \mu_x q_x + \mu_y (1 - q_x)} := \pi_1(q_x). \quad (10)$$

The PbRE (5) can now be rewritten as:

$$\begin{aligned} \dot{q}_x(t) = \quad & q_x(t)(1 - q_x(t)) \times \dots \\ & \left[\bar{F}_x(\pi_1(q_x(t)), q_x(t)) - \bar{F}_y(\pi_1(q_x(t)), q_x(t)) \right]. \end{aligned}$$

Proposition 3: Considering the singular perturbations method with $\epsilon \to 0$, the solutions of the coupled differential equations (S_0) is the population profile $\xi^* = (p_1^*, q_x^*)$, such that:

$$p_1^* = \frac{\mu - s^*(\mu_x - \mu_y)}{\mu + \mu_y} \quad \text{and} \quad q_x^* = \frac{s^*(\mu + \mu_y)}{\mu - s^*(\mu_x - \mu_y)}, \quad (11)$$

where s^* is the equilibrium of the standard replicator dynamics (2) when considering payoff matrix (4):

$$s^* = \frac{d - b}{\Delta} \quad \text{with} \quad \Delta = a - b - c + d.$$

Proof: Let us first study the equation $\dot{q}_x = 0$ before substituting the stationary equation of the state dynamics. This equation is equivalent to find the population profile $\xi = (\pi_1, q_x)$ such that:

$$\bar{F}_x(\pi_1, q_x) = \bar{F}_y(\pi_1, q_x).$$

By replacing and after some manipulations, we get the equivalent equation:

$$\begin{aligned} &\pi_1 a \pi_1 q_x + \pi_1 b (1 - \pi_1 q_x) + (1 - \pi_1) c \pi_1 q_x \\ &+ (1 - \pi_1) d (1 - \pi_1 q_x) = c \pi_1 q_x + d(1 - \pi_1 q_x). \end{aligned} \quad (12)$$

Then finally we get:

$$\pi_1 q_x = \frac{d - b}{\Delta} := s^*.$$

The stationary condition of the first differential equation (3) leads the following relation between p_1 and q_x:

$$p_1 = \pi_1(q_x) = \frac{\mu}{\mu + \mu_x q_x + \mu_y (1 - q_x)},$$

then we have to solve now: $\pi_1(q_x) q_x = s^*$. This last equation is equivalent to:

$$\frac{\mu q_x}{\mu + \mu_x q_x + \mu_y (1 - q_x)} = s^*.$$

After some simple manipulations we obtain:

$$q_x = \frac{s^*(\mu + \mu_y)}{\mu - s^*(\mu_x - \mu_y)} := q_x^*.$$

Finally, as we have $\pi_1(q_x) q_x = s^*$ then $p_1^* := \pi_1(q_x^*) = \frac{s^*}{q_x^*}$ which leads to: $p_1^* = \frac{\mu - s^*(\mu_x - \mu_y)}{\mu + \mu_y}$. ∎

Note that the rest point q_x^* of the PbRE (5) verifies:

$$q_x^* \pi_1(q_x^*) = s^*.$$

This result says that the equilibrium probability that any individual picked out randomly in the population, is playing action x is equal to s^*. This value is the mixed equilibrium of the standard matrix game. It means that, if we consider a state dependent action game, the equilibrium is obtained under conditional probability over the state.

We have the following necessary and sufficient condition under which the solution obtained is a strict interior point.

Lemma 3: The solution q_x^* obtained in proposition (3) is a strict interior point if and only if:

$$\mu > \mu_x \frac{s^*}{1 - s^*}.$$

Proof: The solution obtained in proposition (3) is:

$$q_x^* = \frac{s^*(\mu + \mu_y)}{\mu - s^*(\mu_x - \mu_y)}.$$

This solution is a strict interior point if and only if:

$$0 < q_x^* < 1.$$

First, let's look at the positivity condition $q_x^* > 0$. This is equivalent to:

$$0 < q_x^* \iff \mu > s^*(\mu_x - \mu_y).$$

After some basic algebras, the second condition is:

$$q_x^* < 1 \iff \mu > \mu_x \frac{s^*}{1 - s^*}.$$

We have clearly that for all $s^* \in]0, 1[$, μ_x and μ_y:

$$\frac{s^*}{1 - s^*} \mu_x > s^* \mu_x > s^*(\mu_x - \mu_y).$$

Then if $\mu > \mu_x \frac{s^*}{1 - s^*}$ the solution is a strict interior point, and the converse is true. This concludes the proof. ∎

An important remark is that this condition does not depend on the rate μ_y.

Singular perturbation theory has been widely used in many fields. In [2], the singularly perturbed o.d.e. is introduced to analyze the two timescales stochastic approximation. When considering two different stepsize schedules for the components of the iterations in stochastic approximation, under some regularity conditions, the iterates track the asymptotic behavior of the corresponding singularly perturbed o.d.e.. Stochastic approximation is a cornerstone in scientific computation, as it is easy to implement and it captures the average behavior of adaptive schemes.

In the next section, we present an alternative method based on rewriting our game problem into a matrix game considering only pure policies.

B. Matrix Approach

The two time-scales assumption implies that we can consider individuals in stationary state. We can thus rewrite our model as a matrix game in which individuals play stationary deterministic policies instead of actions. We get the following bimatrix game:

$$
\begin{array}{cc}
 & \begin{array}{cc} u_y & u_x \end{array} \\
\begin{array}{c} u_y \\ u_x \end{array} &
\left(
\begin{array}{cc}
(J(u_y,u_y),J(u_y,u_y)) & (J(u_y,u_x),J(u_x,u_y)) \\
(J(u_x,u_y),J(u_y,u_x)) & (J(u_x,u_x),J(u_x,u_x))
\end{array}
\right)
\end{array}
$$

where $J(u_i,u_j)$ is the expected average fitness of an individual playing pure policy u_i against an individual using u_j, $i,j \in \{x,y\}$. The stationary distributions in states 1 and 0 are given by the following time ratios:

$$
\bar{T}_1(i) = \frac{\frac{1}{\mu_i}}{\frac{1}{\mu} + \frac{1}{\mu_i}} = \frac{\mu}{\mu + \mu_i},
$$

$$
\bar{T}_0(i) = \frac{\frac{1}{\mu}}{\frac{1}{\mu} + \frac{1}{\mu_i}} = \frac{\mu_i}{\mu + \mu_i},
$$

where $i \in \mathcal{A}$ denotes the choice of policy u_i. The expected average fitness $J(u_i,u_j)$ can thus be expressed as a function of the time ratios:

$$
J(u_i,u_j) = \sum_{s,s' \in \mathcal{S}} \sum_{a,a' \in \mathcal{A}} T_s(a) T_{s'}(a') R(a,a'),
$$

where $R(a,a')$ is the immediate fitness of a player using action a against an opponent playing a'. We consider the payoffs bimatrix (4) and we thus obtain:

$$
J(u_y,u_y) = d, \quad J(u_x,u_y) = T_1(x)b + T_0(x)d,
$$
$$
J(u_y,u_x) = T_1(x)c + T_0(x)d,
$$
$$
J(u_x,u_x) = T_1(x)\left[T_1(x)a + T_0(x)b\right]
$$
$$
+ T_0(x)\left[T_1(x)c + T_0(x)d\right].
$$

If we consider this matrix game as a representation of a standard evolutionary game, we can write the replicator dynamics equation for this evolutionary game as:

$$
\dot{\delta}(t) = \delta(t)(1-\delta(t))(J(u_x,\delta(t)) - J(u_y,\delta(t))), \quad (13)
$$

where $\delta(t)$ is the proportion of individuals in the population who play pure policy u_x at time t. The standard replicator dynamics equation for this matrix game, if it converges, converges to the ESS of the evolutionary game. The mixed equilibrium δ^* for this matrix game is obtained by solving the indifference principle equation: $J(u_y,\delta^*) = J(u_x,\delta^*)$, where $J(u_i,q) = (1-q)J(u_i,u_y) + qJ(u_i,u_x)$ with $i \in \mathcal{A}$. We know from standard evolutionary game theory that the mixed equilibrium is expressed by:

$$
\delta^* = \frac{J(u_y,u_y) - J(u_x,u_y)}{J(u_x,u_x) - J(u_y,u_x) + J(u_y,u_y) - J(u_x,u_y)}.
$$

Thus, after some algebras, the ESS obtained by rewriting our game by considering stationary deterministic policies and average fitnesses is:

$$
\delta^* = \frac{s^*}{\bar{T}_F(x)}. \quad (14)
$$

C. Relations between equilibria

We Now compare the two equilibria we obtained by considering two different point of view of the problem. In section IV-A, we supposed that each individual plays a deterministic policy u_x which consists in always choosing action x in state F and, by applying the singular perturbation method, we have been able to determine the equilibrium of such a game. In section IV-B, we assumed that individuals are in their stationary states and we rewrite the game as a standard evolutionary game.

Proposition 4: The relation between the equilibrium δ^* and the equilibrium q_x^* is the following:

$$
q_x^* < \delta^*.
$$

Proof: We compare the value of the equilibria δ^*, obtained by solving our model as a matrix game in (14) with the value of the equilibrium obtained through the singular perturbation method, q_x^* in (3):

$$
\delta^* - q_x^* = \frac{s^*(\mu + \mu_x)}{\mu} - \frac{s^*(\mu + \mu_y)}{\mu - s^*(\mu_x - \mu_y)}
$$
$$
= s^* \frac{(\mu_x - \mu_y)(\mu - s^*(\mu + \mu_x))}{\mu(\mu - s^*(\mu_x - \mu_y))}.
$$

From lemma 3, the denominator is always positive. We observe that $s^* > 0$, $\mu_x > \mu_y$ for definition, and, from lemma 3, $(1 - s^*)\mu > s^*\mu_x$, so the nominator is also positive. This implies that the strict inequality $q_x^* < \delta^*$ always hold. ∎

We are able also to compare the two equilibria in terms of average fitness obtained by the population, i.e. $J(\delta^*,\delta^*)$ and $\bar{F}(q_x^*,p_1^*)$.

Proposition 5: The average fitnesses of the population at the two equilibria points obtained with the two approaches are equals, i.e. $\bar{F}(q_x^*,p_1^*) = J(\delta^*,\delta^*)$.

Proof: Considering the first approach based on the singular perturbations method, we have:

$$
\bar{F}(q_x^*,p_1^*) = q_x^* \bar{F}_x(q_x^*,p_1^*) + (1-q_x^*)\bar{F}_y(q_x^*,p_1^*)
$$
$$
= \bar{F}_y(q_x^*,p_1^*) + q_x^*(\bar{F}_x(q_x^*,p_1^*) - \bar{F}_y(q_x^*,p_1^*))
$$

At the equilibrium state, we have $\bar{F}_x(q_x^*,p_1^*) = \bar{F}_y(q_x^*,p_1^*)$ and thus $\bar{F}(q_x^*,p_1^*) = \bar{F}_x(q_x^*,p_1^*) = \bar{F}_y(q_x^*,p_1^*) = \bar{J}_y(q_x^*,p_1^*)$. Then, we get the average fitness of the population at the equilibrium with the first approach is:

$$
\bar{F}(q_x^*,p_1^*) = p_1^*(q_x^*c + (1-q_x^*)d) + (1-p_1^*)d.
$$

But we have the following relation $q_x^* p_1^* = s^*$ and then:

$$
\bar{F}(q_x^*,p_1^*) = s^*c + (1-s^*)d.
$$

Considering the second method of rewriting the game into a matrix game, we obtain the following ESS: $\delta^* = \frac{s^*}{\bar{T}_F(x)}$. The average fitness of the population in this case is:

$$
J(\delta^*,\delta^*) = \delta^* J(u_x,\delta^*) + (1-\delta^*)J(u_y,\delta^*)
$$
$$
= J(u_y,\delta^*) + \delta^*(J(u_x,\delta^*) - J(u_y,\delta^*)).
$$

At the ESS, we have the following equality $J(u_x,\delta^*) = J(u_y,\delta^*)$ and then the average fitness of the population becomes simply:

$$
J(\delta^*,\delta^*) = J(u_y,\delta^*) = \delta^* J(u_y,u_x) + (1-\delta^*)J(u_y,u_y).
$$

Then, the average fitness of the population is: $J(\delta^*, \delta^*) = \delta^*(T_1(x)c + T_0(x)d) + (1-\delta^*)d$. We have that $\delta^* T_1(x) = s^*$ which leads to: $J(\delta^*, \delta^*) = s^* c + (1 - s^*)d$. ∎

Finally, we prove that the two mixed strategies obtained with the two approaches are not equal, but are in the same equivalent class in terms of occupation measures. We denote by $\bar{T}_1(q)$ the average sojourn time in state F for an individual who plays mixed strategy q. This mixed strategy has two possibilities. First, it can be a mixed policy between the pure policies u_y and u_x, like proposed in the section IV-B. Second, a mixed policy characterized by a probability q to play action x in state F, at each time an individual is in state F. This second point of view is proposed in section IV-A. The two equilibria obtained by the singular perturbations method and the matrix game reformulation are in the same equivalent class in terms of the occupation measures. It means that they should satisfy: $\bar{T}_1(\delta^*) = \bar{T}_1(q_x^*)$.

We have for the case of pure policies:

$$\bar{T}_1(\delta^*) = \delta^* \frac{\mu}{\mu + \mu_x} + (1-\delta^*)\frac{\mu}{\mu + \mu_y}. \quad (15)$$

For the case of mixed policies, we have:

$$\bar{T}_1(q_x^*) = \pi_1(q_x^*) = \frac{\mu}{\mu + \mu_x q_x^* + \mu_y(1 - q_x^*)}. \quad (16)$$

This important result is proved in the following proposition.

Proposition 6: The mixed equilibrium δ^* over the pure policies and the equilibrium obtained by the singular perturbation approach yield to the same occupation measures, i.e.
$$\bar{T}_1(\delta^*) = \bar{T}_1(q_x^*).$$

Proof: We first rewrite δ^* as a function of the immediate payoffs $\{a, b, c, d\}$:

$$\delta^* = \frac{(\mu_x + \mu)(d - b)}{\mu \Delta}.$$

where $\Delta := a - b - c + d$. We substitute it in (15), and we get:

$$\bar{T}_1(\delta^*) = \frac{\mu \Delta + (\mu_x + \mu_y)}{\mu \Delta (\mu + \mu_y)}.$$

Analogously, we substitute the expression of s^* in q_x^* in proposition(3), and we rewrite it in (16), which leads to:

$$\bar{T}_1(q_x^*) = \frac{\mu \Delta + (\mu_x + \mu_y)}{\mu \Delta (\mu + \mu_y)}.$$

which proves that $\bar{T}_1(q_x^*) = \bar{T}_1(\delta^*)$. ∎

This two previous results show that we can define two equivalent classes for deterministic stationary policies that yield same average fitness and occupation measures. This also leads us to generalize to several states and actions.

D. Generalization of the Model

When considering the population state to be in its stationary regime, we can easily extend our model to a general context with several states and actions. Let $\mathcal{S}$ be the finite state space of a player and $\mathcal{A}$ the finite set of actions. Let $P_{s,a}(s')$ be the transition probability for a player to move

from s to s' if he uses action a. Define the finite set $\mathcal{U}_D$ of deterministic policies of a player and $\mathcal{U}_S$ the set of stationary policies. Let $\bar{F}(u, v)$ denote the immediate expected utility of a player that uses policy $u \in \mathcal{U}_D$ against a population playing $v \in \mathcal{U}_D$. We define the expected state-action frequency at time t of the pair $(s, a) \in \mathcal{S} \times \mathcal{A}$ under policy $u \in \mathcal{U}_D$ for an initial state distribution of η:

$$\phi_\eta^t(u; s, a) = \frac{1}{t} \sum_{n=1}^t \mathbb{P}_{\eta, t}^u(s, a).$$

Let $\phi_\eta^t(u)$ denote the set $\{\phi_\eta^t(u; s, a)\}_{s,a}$ and let $\Phi_\eta(u)$ be the set of accumulation points of the sequence $\phi_\eta^t(u)$ when $t \to \infty$. Denote by V^* the set of policies u such that $\Phi_\eta(u)$ is a singleton.

Let $s_q \in \mathcal{U}_S$ be a stationary (mixed) policy defined as follows:

$$s_q(a|s) = \sum_{u \in \mathcal{U}_D} q(u) 1\{g_u(s) = a\}$$

where $g_u : \mathcal{S} \to \mathcal{A}$ is the function that associates to each state $s \in \mathcal{S}$ the corresponding action $a \in \mathcal{A}$ chosen by the pure policy $u \in \mathcal{U}_D$, and $q = (q_1, \ldots, q_n)$ is a distribution over $\mathcal{U}_D$. We have that:

$$s_q \in V^*, \quad \text{and} \quad \phi_\eta(s_q) = \sum_{u \in \mathcal{U}_D} q(u)\phi(\eta, u) \qquad \forall \eta \in \mathcal{S}.$$

Let $J(s, a; s', a')$ be the immediate reward that a player gets when it is at state s and plays action a against an individual in state s' using action a'. If $u, v \in \mathcal{U}_S$, then

$$\bar{F}(u, v) = \sum_{s,s' \in \mathcal{S} \, a, a' \in \mathcal{A}} \phi(u; s, a) J(s, a; s', a')\phi(v; s', a')$$
$$(17)$$

Based on the previous expression, we define the dynamics of the distribution $q(u)$ over time, where u is a stationary policy in $\mathcal{U}_S$:

$$\dot{q}(u, t) = q(u, t)\left(\sum_{v \in \mathcal{U}_D} \bar{F}(u, v)q(v, t) - \sum_{v, v' \in \mathcal{U}_D} q(v, t)\bar{F}(v, v')q(v', t)\right).$$

If the trajectory of q converges to a stable rest point q^*, then the policy used at time t converges to the stationary one, i.e. $s_{q(t)} \to_{t \to \infty} s_{q^*}$ and s_{q^*} is a Nash equilibrium policy.

V. APPLICATIONS

A. Energy Control in Wireless Network

The two-states two-actions model can be applied to describe a particular case of a problem that arises in dynamic power control in mobile networks, which has been presented in [12]. A large number of mobiles transmit packets occasionally. Each transmitter can be in Full (F) or Almost empty (A) battery state. When a mobile is in F state it can choose to transmit packets using high (h) or low (l) power, whereas if it is in state A, it can only transmit packets using l power. In general, several mobiles try to join a common receiver at the same time and interferences occur between the received signals. We suppose that transmissions are sparse so that the probability that more than two mobiles transmit simultaneously is negligible. We assume also that a

transmission is successful either if the mobile is the only one transmitting during a slot or if it transmits at higher power than all the others. The time spent in state F depends on the action chosen by the mobile. Then the state of the mobile changes to the other battery state A. After an exponentially distributed time, its battery state becomes empty. We assume that the battery is immediately recharged, so that the mobile goes back to state F. When transmitting at high power, the mobile's battery is consumed faster, and thus the transition rate from F to A is faster.

B. Network Formation Games

Another application of the proposed model can be found in network formation games [13]. We consider a large number of nodes where each node is in one of two possible states: Infected or Susceptible, so that $\mathcal{S} = \{I, S\}$. Nodes interact through pairwise interactions, during which, both nodes exchange contents. If a node is in state S he determines the type of unidirectional link to the node he is interacting with. The type of link can be charged at a price (p) or free (f); if a node is in the infected state (state I), it can only create free links. Pay connection is safer, so that when a link is not a free one, the probability for a node to be infected is lower, independently of the choice of the other node to pay or not. After some random time in I state, a node becomes susceptible again.

VI. NUMERICAL ILLUSTRATIONS

We illustrate here the theoretical results obtained in previous sections with numerical solutions. We consider a first numerical example with the following transition rates: $\mu = 10$, $\mu_x = 1.5$ and $\mu_y = 1$. The fitnesses of the matrix game are: $a = -0.3$, $c = 0$, $b = 1$ and $d = 0.5$. Those values yield to the following equilibrium of the standard evolutionary game $s^* = \frac{5}{8} = 0.625$.

We plot on figure 1 the trajectories of the system (S_ϵ) of the coupled differential equations for different initial conditions and for $\epsilon = 0.01$. We simulate a discrete time version of the differential equations. We plot also the invariant manifold M_0 and we observe that it is an attractor of the trajectories.

More, based on proposition (3), we have the following solution of the system, by considering the singular perturbation method based on the steady-state model:

$$q_x^* = 0.7097, \quad \text{and} \quad p_1^* = 0.8807.$$

This couple corresponds exactly to the attractor of the trajectories on figure 1.

VII. CONCLUSIONS AND PERSPECTIVES

In this paper, we have considered a particular type of evolutionary game in which the action of the individual not only determines its immediate fitness but it also impacts his state. The aim of this paper is to describe the coupled dynamics of the individual states which is due to the direct control and the dynamics of policies which is determined by the replicator dynamics mechanism. Once we have introduced these combined dynamics, we have proved that any stable
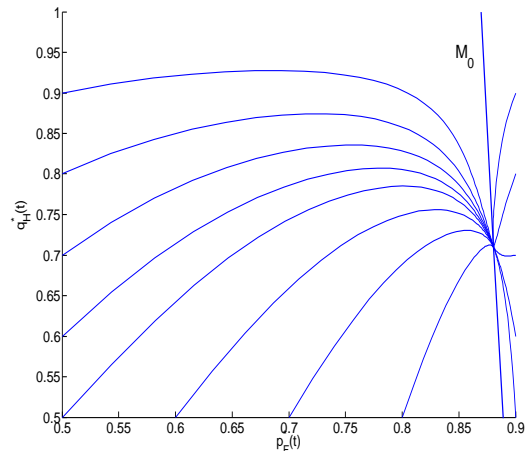

Fig. 1. Trajectories of the system (S_ϵ) from different starting points and the slow manifold M_0 with $\epsilon = 0.01$.

rest point corresponds to an equilibrium of the game. We have proposed two methods to obtain the rest points under the assumption that the two dynamics have different time scales. Finally, we discuss how to generalize our framework to any finite number of states and actions, which can lead to several applications of our framework in networking, social networks and complex systems. In perspective, we propose to further investigate the general case to extend our SPcD to the MDEG framework. We are also interested in studying the stability conditions for a rest point of the SPcD to be an ESS profile of our game.

REFERENCES

[1] John Maynard Smith and George R. Price, (1973). "The logic of animal conflict". Nature 246 (5427): 15-18.
[2] V. Borkar, *Stochastic Approximation: A Dynamical Systems Viewpoint*, Cambridge University Press, 2008.
[3] P. Kokotovic, H. Khalil, J. O'reilly, *Singular perturbation methods in Control*, SIAM, 1986.
[4] E. Altman, Y. Hayel, *Markov Decision Evolutionary Games*, in IEEE Transactions on Automatic Control, vol.55, no.7, 2010.
[5] J. Weibull, *Evolutionary Game Theory*, the MIT Press, 1997.
[6] P. Taylor, L. Jonker *Evolutionary Stable Strategies and Game Dynamics*, in Mathematical Biosciences, vol. 40, pp. 145-156, 1978.
[7] J. Hofbauer, K. Sigmund *Evolutionary Game Dynamics*, in Bulletin of the American Mathematical Society, vol. 40, no. 4, pp. 479-519, 2003.
[8] J. Hofbauer, P. Schuster, and K. Sigmund, A Note on Evolutionarily Stable Strategies and Game Dynamics, Journal of Theoretical Biology, 81:609,612, 1979.
[9] S. Hafner, *On the evolution of beliefs*, technical report, 2014.
[10] J. Nachbar, *Evolutionary Selection Dynamics in Games: Convergence and Limit Properties*, in International Journal of Game Theory, vol. 19, pp. 59-89, 1990.
[11] J. Flesch, T. Parthasarathy, F. Thuijsman, P. Uyttendaele *Evolutionary stochastic games* in Dynamic Games and Applications 3, 207-219, 2013.
[12] E. Altman et Y. Hayel *A Stochastic Evolutionary Game of Energy Management in a Distributed aloha Network*, in Proceedings of IEEE INFOCOM, 2008.
[13] M. Jackson, *A Survey of Network Formation Models: Stability and Efficiency*, in Group Formation in Economics: Networks, Clubs, and Coalitions book, Cambridge University Press, 2005.