

# Hierarchical Decision-Making in Population Games

Yu-Wen Chen, Nuno C. Martins, Murat Arcak

**Abstract**—This paper introduces a hierarchical framework for population games, where individuals delegate decision-making to proxies that act within their own strategic interests. This framework extends classical population games, where individuals are assumed to make decisions directly, to capture various real-world scenarios involving multiple decision layers. We establish equilibrium properties and provide convergence results for the proposed hierarchical structure. Additionally, based on these results, we develop a systematic approach to analyze population games with general convex constraints, without requiring individuals to have full knowledge of the constraints as in existing methods. We present a navigation application with capacity constraints as a case study.

## I. INTRODUCTION

Population games study the aggregate behavior of strategic agents in large populations who adapt their actions via learning to maximize payoffs. A fundamental research question is determining conditions under which learning rules and payoff mechanisms ensure convergence to Nash equilibria. Early studies focused on memoryless payoff mechanisms, such as potential games [1]. Later, building on [2], the paper [3] introduced the concept of  $\delta$ -passivity to characterize stability properties and to allow for dynamic payoff mechanisms. A broader  $\delta$ -dissipativity approach was presented in [4]. Other passivity approaches were applied to imitation dynamics in [5]. Further advancing this line of work, [6] brought the notion of counterclockwise passivity (CCW) [7] to population games.

While the studies discussed above assume that individuals directly select their strategies, in many applications, decision-making is mediated by proxies with more information or computational power. For example, investors delegate portfolio management to financial managers, and travelers follow navigation applications in selecting routes. This leads to hierarchical, rather than direct, decision-making as illustrated in Figures 1a and 1b. Moreover, proxies may have strategy preferences that constrain the social state to a subset of the simplex in which it evolves. Limited work has addressed population games with such constraints. In [8], the authors considered box constraints which do not account for coupling restrictions between strategies. Moreover, the work requires a demanding information structure where the individuals have access to all system-wide constraints. In [9], a payoff penalty was used to ensure convergence to the feasible set, but the constraints were

not satisfied at all times. Recently, [10] studied constrained best-response dynamics for convex constraints; however, as in [8], these constraints are assumed to be common knowledge.

The main contributions of this paper are twofold. First, we introduce a hierarchical framework for population games, which captures practical scenarios involving multiple decision layers. With properly designed proxies and payoff structure, we provide convergence results akin to those in classical population games. Second, we address population games with general convex constraints within this framework. Unlike existing methods [8]–[10] requiring constraints known to individuals, we incorporate constraints through the hierarchical structure.

Section II gives an overview of population games. Section III introduces the hierarchical framework and formulates the main problem. Section IV presents the theoretical results. Section V illustrates the results on a navigation example.

## II. PRELIMINARIES

For a classical (single) population game, each individual in the population can choose from a set of strategies  $\mathcal{S} = \{1, \dots, d\}$ . Let  $\Delta^d = \{v \in \mathbb{R}_+^d : \mathbf{1}^T v = 1\}$  be the probability simplex in  $\mathbb{R}^d$ . Denote the state  $s(t) = [s_1(t), \dots, s_d(t)]^T \in \Delta^d$ , where  $s_i(t)$  describes the proportion of the population selecting strategy  $i$  at time  $t$ . A payoff function  $F : \Delta^d \rightarrow \mathbb{R}^d$  maps a state  $s(t)$  to a payoff vector  $F(s(t))$  with  $F_i(s(t))$  representing the payoff for choosing strategy  $i$ .

**Definition 1** (Best response to  $\pi$  within  $\mathcal{C}$ ). *The best response to a payoff vector  $\pi \in \mathbb{R}^d$  within a compact set  $\mathcal{C} \subseteq \Delta^d$  is denoted as  $BR_{\mathcal{C}}(\pi) = \arg \max_{s \in \mathcal{C}} s^T \pi$ .*

**Definition 2** (Nash equilibrium of  $F$  within  $\mathcal{C}$ ). *Let  $\mathcal{C} \subseteq \Delta^d$  be compact.  $s^* \in \mathcal{C}$  is a Nash equilibrium for the payoff function  $F$  within  $\mathcal{C}$  if  $(s - s^*)^T F(s^*) \leq 0$ , for  $s \in \mathcal{C}$ . The set of all Nash equilibria for  $F$  within  $\mathcal{C}$  is denoted as  $NE_{\mathcal{C}}(F)$ .*

Given payoff vectors  $\pi(t) \in \mathbb{R}^d$  for all  $t$ , the individuals switch between strategies to receive higher payoffs, which leads to a dynamics in  $s(t)$ . An evolutionary dynamics model (EDM) is used to describe this learning dynamics:

$$\dot{s}(t) = \mathcal{V}(s(t), \pi(t)), \quad t \geq 0, \quad (1)$$

where  $\mathcal{V} : \Delta^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  is defined based on the switching behaviors. Some common EDMs include Smith [11], Brown-von Neumann-Nash (BNN) [12], and best response dynamics [13, Chapter 6]. While the best response dynamics is defined over the simplex  $\Delta^d$ , a generalization to it is defined over a compact convex subset of  $\Delta^d$  as follows.

**Definition 3** (Constrained best response dynamics [10]).

$$\dot{s}(t) \in BR_{\mathcal{C}}(\pi(t)) - s(t), \quad \mathcal{C} \subseteq \Delta^d \text{ compact convex.} \quad (2)$$

This work was supported by NSF Grants 2135791 and 2135561 and AFOSR FA95502310467.

Yu-Wen Chen and Murat Arcak are with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA, USA. {yuwen\_chen, arcak}@berkeley.edu.

Nuno C. Martins is with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, USA. nmartins@umd.edu.

**Definition 4** (Positive Correlation). An EDM (1) is positively correlated if  $\mathcal{V}(s, \pi) \neq \mathbf{0} \implies \pi^T \mathcal{V}(s, \pi) > 0$ .

**Definition 5** (Nash Stationarity w.r.t.  $\mathcal{C}$ ). An EDM (1) is Nash stationary w.r.t.  $\mathcal{C}$  if  $\mathcal{V}(s, \pi) = \mathbf{0} \iff s \in BR_{\mathcal{C}}(\pi)$ .

While BNN, Smith, and best response dynamics are Nash stationary w.r.t.  $\Delta^d$ , (2) is Nash stationary w.r.t.  $\mathcal{C}$ .<sup>1</sup>

### III. PROBLEM FORMULATION

We first present the core formulation and introduce key notation using a two-layer example in Section III-A. We extend this to the general hierarchical framework in Section III-B, and formulate several problems in this framework in Section III-C.

#### A. Illustrative example

Consider a group of investors who allocate capital across three investment targets through two managers. This leads to two layers as in Figure 1b: the first contains the investors; the second comprises two managers. We assume each investor has one unit of capital and must allocate it to either manager. Manager 1 distributes the funds across three investment targets, while Manager 2 invests only in the first and second.

Each decision group has a state variable representing its strategy distribution over its strategy set. Specifically, for each decision group  $j$  in layer  $i$ , denoted as the  $(i, j)$ -group, we define a state variable  $s^{i,j}(t) \in \Delta^{d^{i,j}}$ , where  $d^{i,j}$  is the number of available strategies to the  $(i, j)$ -group. The  $k$ -th entry of  $s^{i,j}(t)$ , denoted as  $s_k^{i,j}(t)$ , represents the proportion of the  $(i, j)$ -group selecting the  $k$ -th strategy at time  $t$ .

We define the social state  $x(t) \in \Delta^3$  as the final distribution over the three final strategies (investment targets) at time  $t$ :

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = \begin{bmatrix} s_1^{1,1}(t)s_1^{2,1}(t) + s_2^{1,1}(t)s_1^{2,2}(t) \\ s_1^{1,1}(t)s_2^{2,1}(t) + s_2^{1,1}(t)s_2^{2,2}(t) \\ s_1^{1,1}(t)s_3^{2,1}(t) \end{bmatrix} \quad (3a)$$

$$= \underbrace{\begin{bmatrix} 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}}_{:=W^2 \in \mathbb{R}^{3 \times 5}} \underbrace{\begin{bmatrix} s^{2,1}(t) & \mathbf{0}_{3 \times 1} \\ \mathbf{0}_{2 \times 1} & s^{2,2}(t) \end{bmatrix}}_{:=T^2(t) \in \mathbb{R}^{5 \times 2}} s^{1,1}(t). \quad (3b)$$

Equation (3a) follows from the paths in Figure 1b, while (3b) provides a compact representation. Here,  $T^2(t)$  models the redistribution of funds by the managers in layer 2 at time  $t$ , and  $W^2$  aggregates the resulting allocations. Further details on  $T^2(t)$  and  $W^2$  are provided in Section III-B.

With the defined state variables  $s^{i,j}(t)$  and social state  $x(t)$ , we now describe the system dynamics. A payoff function  $F: \Delta^3 \rightarrow \mathbb{R}^3$  maps the social state  $x(t)$  to a payoff vector  $F(x(t)) = [F_1(x(t)), F_2(x(t)), F_3(x(t))]^T$ , where  $F_i(x(t))$  represents the payoff for investing in the  $i$ -th investment target. Manager 1 updates portfolio allocations based on  $F(x(t))$ , while Manager 2 updates based on  $F_1(x(t))$  and  $F_2(x(t))$ . In addition, investors update their choices based on manager

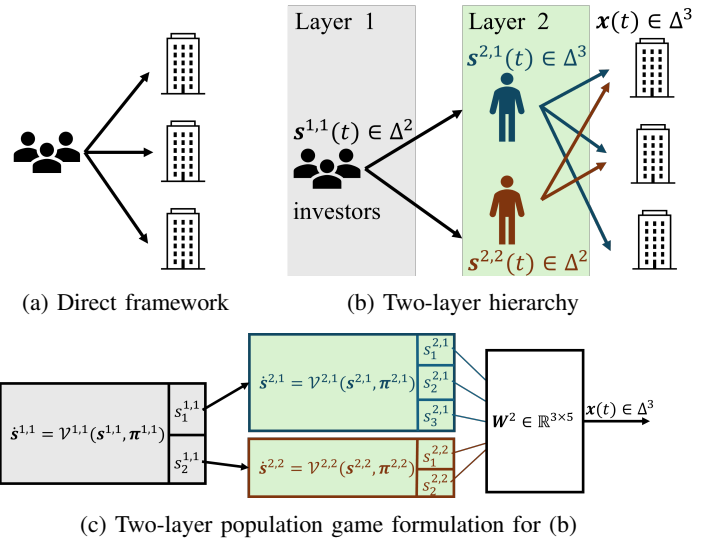


Fig. 1: In contrast to direct framework (a) in most literature, we provide an illustrative example with two-layer hierarchy.

performance. A natural performance metric for managers is their average payoff. Therefore, we can model the updates by:

$$\dot{s}^{i,j}(t) = \mathcal{V}^{i,j}(s^{i,j}(t), \pi^{i,j}(t)), \quad t \geq 0, \quad (4)$$

for  $(i, j) \in \{(1, 1), (2, 1), (2, 2)\}$ . Here,  $\mathcal{V}^{i,j}$  is the EDM defined in Section II and  $\pi^{i,j}(t)$  is the payoff at time  $t$ , for the  $(i, j)$ -group, while the payoffs are given as:

$$\pi^{2,1}(t) = F(x(t)) \quad (5a)$$

$$\pi^{2,2}(t) = [F_1(x(t)), F_2(x(t))]^T \quad (5b)$$

$$\pi^{1,1}(t) = \left[ \sum_{l=1}^3 s_l^{2,1}(t) F_l(x(t)), \sum_{l=1}^2 s_l^{2,2}(t) F_l(x(t)) \right]^T. \quad (5c)$$

The overall structure is visualized in Figures 1b and 1c.

#### B. Hierarchical framework

Generalizing the notation  $T^i(t)$  and  $W^i$  introduced in (3b), we define a transformation matrix  $T^i(t)$  by diagonally concatenating all state variables  $s^{i,j}(t) \in \Delta^{d^{i,j}}$  in layer  $i$ . Suppose there are  $n^i$  decision groups (state variables) in layer  $i$ , then for all  $t \geq 0$ ,

$$T^i(t) = \text{blkdiag}(s^{i,1}(t), \dots, s^{i,n^i}(t)) \in \mathbb{R}^{o^i \times n^i}, \quad (6)$$

where  $\text{blkdiag}$  denotes the block diagonal matrix composed by the arguments and  $o^i = \sum_j d^{i,j}$  is the total number of outputs of layer  $i$ . Since outputs from one layer may be grouped into decision groups in the next layer, we introduce an aggregation matrix to formalize this transition. These aggregation matrices define how outputs from layer  $i$  are assigned to decision groups in layer  $i+1$ , fully characterizing the hierarchical structure. An aggregation matrix is defined as:

$$W^i = [w_1^i, \dots, w_{o^i}^i] \in \mathbb{R}^{n^{i+1} \times o^i}, \quad (7)$$

where  $w_k^i \in \mathbb{R}^{n^{i+1}}$  represents how the  $k$ -th output of layer  $i$  distributes to the next layer. Let  $n^{L+1} = d$  represent the

<sup>1</sup>Since the best response dynamics and (2) are differential inclusions, a modified version of the Nash stationarity is used, as in [13, Theorem 6.1.4].

number of final strategies. The transformation matrix  $T^i(t)$  and aggregation matrix  $W^i$  satisfy the following properties:

**Property 1.**  $T^i(t)$  and  $W^i$  are elementwise nonnegative.

**Property 2.**  $1^T T^i(t) = 1^T$  and  $1^T W^i = 1^T$ .

We now formulate the general  $L$ -layer hierarchical framework using the notation introduced in (6) and (7). Suppose layer 1 consists of a single decision group representing the entire population, while each subsequent layer  $i$  has  $n^i \geq 1$  decision groups for  $i = 2, \dots, L$ . The decision group  $j$  in layer  $i$  is associated with a state variable  $s^{i,j}(t) \in \Delta^{d^{i,j}}$ , which describes the group's distribution over the group's strategy set.

Let the hierarchy be characterized by  $W^1$  to  $W^L$ . Note that since layer 1 has only one decision group,  $T^1(t) = s^{1,1}(t)$  and  $W^1$  is the identity matrix. Then, we can derive the distribution of the entire population input to layer  $i$  at time  $t$  by

$$M^i(t) = W^{i-1} T^{i-1}(t) \dots W^1 s^{1,1}(t) \in \Delta^{n^i}. \quad (8)$$

Further,  $W^L T^L(t) M^L(t)$  captures the distribution of the entire population over the  $d$  final strategies, which we call the social state. We denote the social state as  $x(t) \in \Delta^d$ :

$$\begin{aligned} x(t) &= W^L T^L(t) M^L(t) = \Pi_{i=1}^L W^i T^i(t) \\ &= W^L T^L(t) \dots W^2 T^2(t) W^1 s^{1,1}(t), \quad t \geq 0. \end{aligned} \quad (9)$$

Using Property 1 and Property 2, we can verify that  $x(t) \in \Delta^d$ , confirming that the population is properly distributed over the final  $d$  strategies. However, unlike in Figure 1a, where the individual selects a final strategy directly, here, decisions are made indirectly through multiple layers. Therefore, we refer to each layer  $i$ , for  $i = 2, \dots, L$ , as a proxy layer.

To complete the formulation, we describe the dynamics for  $s^{i,j}(t)$ . Let  $p(t) \in \mathbb{R}^d$  represent the  $d$  payoffs for selecting each final strategy. We temporarily assume  $p(t)$  is given by a payoff function  $F: \Delta^d \rightarrow \mathbb{R}^d$ , mapping the social state  $x(t)$  to the payoff vector  $F(x(t))$  for all  $t$ :

$$p(t) = F(x(t)), \quad t \geq 0. \quad (10)$$

Later, we will extend this to a dynamic model in Section IV-B. Denote  $\pi^{i,j}(t) \in \mathbb{R}^{d^{i,j}}$  as the  $d^{i,j}$  payoffs for selecting each strategy in  $(i, j)$ -group. Recall from Section III-A that  $\pi^{i,j}(t)$  is determined by the average payoff of the group it joins via back-propagating  $p(t)$ , as in (5). Let  $\pi^i(t)$  be the concatenation of payoffs  $\pi^{i,j}(t)$  in layer  $i$ :

$$\pi^i(t) = \begin{bmatrix} \pi^{i,1}(t) \\ \vdots \\ \pi^{i,n^i}(t) \end{bmatrix} \in \mathbb{R}^{o^i}, \quad t \geq 0. \quad (11)$$

Then, the payoffs are determined by

$$\pi^L(t) = W^L T^L p(t) \quad (12a)$$

$$\pi^i(t) = W^i T^{i+1}(t)^T \pi^{i+1}(t), \quad i = L-1, \dots, 1. \quad (12b)$$

With (12),  $s^{i,j}(t)$  evolves and follows, as in (4), the EDM:

$$\dot{s}^{i,j}(t) = \nu^{i,j} (s^{i,j}(t), \pi^{i,j}(t)), \quad t \geq 0. \quad (13)$$

The hierarchical framework is illustrated in Figure 2.

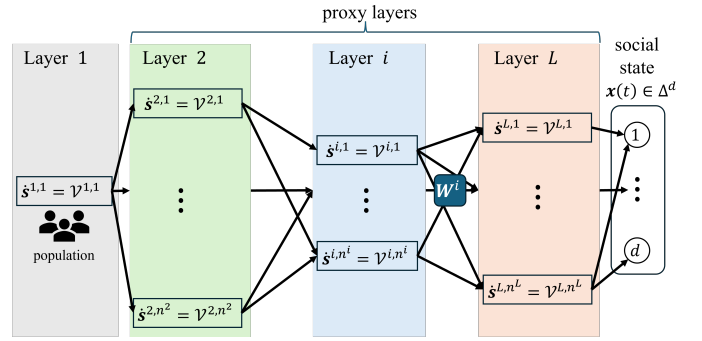


Fig. 2: Illustration of the hierarchical framework. The population contributes to the social state via multiple proxy layers.

### C. Problem description

In this paper, we consider the general case where the *strategy distribution of interest* for each  $(i, j)$ -decision group is not necessarily the entire simplex but a subset  $\mathcal{K}^{i,j} \subseteq \Delta^{d^{i,j}}$ . This states that the  $(i, j)$ -group always makes a decision within  $\mathcal{K}^{i,j}$ , i.e.,  $s^{i,j}(t) \in \mathcal{K}^{i,j}$ . These subsets collectively determine an *admissible set*  $\mathcal{K} \subseteq \Delta^d$ , where the social state  $x(t)$  can reside. Denote  $[m]$  as the set  $\{1, \dots, m\}$ .

**Definition 6** (Admissible set  $\mathcal{K}$ ).

$$\mathcal{K} = \{ \Pi_{i=1}^L W^i K^i : k^{i,j} \in \mathcal{K}^{i,j}, i \in [L], j \in [n^i] \}, \quad (14)$$

where  $K^i = \text{blkdiag}(k^{i,1}, \dots, k^{i,n^i}) \in \mathbb{R}^{o^i \times n^i}$ .

The admissible set  $\mathcal{K}$  collects all possible social states generated from different combinations of group decisions. Consequently, (14) follows the same structures as (9) and (6). We provide a concrete example.

**Example 1** (Strategy distribution of interest). Consider the example in Section III-A. Suppose that Manager 1 favors the distribution  $s_{*}^{2,1} = [0.2, 0.2, 0.6]^T$  over the three investment targets but allows variations up to  $\varepsilon = 0.1$ , while Manager 2 prefers to invest at least  $1/3$  in both investment targets. Then, the strategy distributions of interest are

$$\mathcal{K}^{2,1} = \{ x \in \Delta^3 : \|x - s_{*}^{2,1}\|_2 \leq \varepsilon \} \quad \text{and} \quad (15a)$$

$$\mathcal{K}^{2,2} = \left\{ x \in \Delta^2 : x_1 \geq \frac{1}{3}, x_2 \geq \frac{1}{3} \right\}. \quad (15b)$$

Elements of  $\mathcal{K}^{2,2}$  are two-dimensional, corresponding to the first two final strategies. To visualize  $\mathcal{K}^{2,2}$  together with  $\mathcal{K}^{2,1}$  in  $\Delta^3$ , we lift elements in  $\mathcal{K}^{2,2}$  to 3D by appending a zero. If investors are free to choose between the two managers, i.e.,  $\mathcal{K}^{1,1} = \Delta^2$ , then the admissible set  $\mathcal{K}$  is the convex hull of  $\mathcal{K}^{2,1}$  and the lifted  $\mathcal{K}^{2,2}$ , as shown in Figure 3.

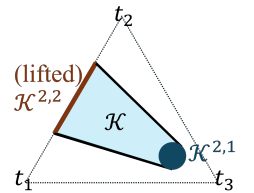


Fig. 3: Strategy distributions of interest over 3 investment targets.

**Main Problem:** Given the hierarchical framework (9)-(13), we aim to determine conditions such that the social state  $\mathbf{x}(t)$  converges to a Nash equilibrium characterized by  $F$  within the admissible set  $\mathcal{K}$  (14). This framework introduces new complexities due to multiple layers and additional challenges from the coupling dynamics among decision groups and their collective influence on  $\mathbf{x}(t)$  via (9).

As will be illustrated in Section V, our solution suggests a novel approach to constraining the social state within a desired set  $\mathcal{D} \subseteq \Delta^d$ , e.g. to avoid socially inefficient states. If we can design strategy distributions of interest  $\mathcal{K}^{i,j}$ s such that the resulting admissible set  $\mathcal{K}$  satisfies  $\mathcal{K} \subseteq \mathcal{D}$ , and the dynamics meets our conditions, then  $\mathbf{x}(t)$  evolves and converges within  $\mathcal{D}$ . Moreover, if we design  $\mathcal{K}^{1,1} = \Delta^{d^{1,1}}$ , then this approach does not impose direct constraints on individual decisions; instead, the constraint is embedded into the strategy distributions of the proxies.

We impose mild assumptions derived from the classical population games. Since not all results require all assumptions, we will state explicitly which assumptions apply in each case.

**Assumption 1.**  $\mathcal{K}^{i,j}, i \in [L], j \in [n^i]$ , are compactly convex.

**Assumption 2.**  $F$  is continuously differentiable.

**Assumption 3.** The learning dynamics (13) for each  $(i, j)$ -group ensures that  $\mathcal{K}^{i,j}$  is forward invariant.

**Assumption 4.**  $\pi^T \mathcal{V}^{i,j}(s, \pi), i \in [L], j \in [n^i]$ , are Lipschitz continuous w.r.t.  $s$  and  $\pi$ .

When  $L = 1$  and  $\mathcal{K}^{1,1} = \Delta^{d^{1,1}}$ , the assumptions reduce to the ones in classical population games. Here, Assumption 3 means that each  $(i, j)$ -group makes decisions consistently within its strategy distribution of interest  $\mathcal{K}^{i,j}$ . Generally, if  $\mathcal{K}^{i,j}$  is a convex set, the constrained best response dynamics (2) guarantees the forward invariance. Moreover, all learning dynamics introduced in Section II satisfy Assumption 4.

#### IV. ANALYSIS

We focus on static models for payoff (10) in Section IV-A and extend to dynamic models in Section IV-B.

##### A. Static Models for Payoff

We first characterize the system at its rest points, showing that the corresponding social state is a Nash equilibrium of  $F$  over the admissible set  $\mathcal{K}$ . We then establish conditions under which the system converges to these rest points.

**Theorem 1** (Characterization at rest points). *Suppose each  $\mathcal{V}^{i,j}$  in (13) is Nash stationary w.r.t.  $\mathcal{K}^{i,j}$ . Let  $\bar{\mathbf{x}} \in \mathcal{K}$  be an equilibrium social state when the system is at rest. Then,  $\bar{\mathbf{x}} \in NE_{\mathcal{K}}(F)$ .*

*Proof.* We use a bar on a variable to denote its value at a rest point of (13). Since  $\mathcal{V}^{i,j}$ s are Nash stationary w.r.t.  $\mathcal{K}^{i,j}$ s, by definition we have  $(\mathbf{s}^{i,j} - \bar{\mathbf{s}}^{i,j})^T \bar{\pi}^{i,j} \leq 0$ , for  $\mathbf{s}^{i,j} \in \mathcal{K}^{i,j}$ . Using (6), (11), and (14), we rewrite it compactly as

$$(\mathbf{K}^i - \bar{\mathbf{T}}^i)^T \bar{\pi}^i \preceq 0, \quad \mathbf{k}^{i,j} \in \mathcal{K}^{i,j}, i \in [L], j \in [n^i], \quad (16)$$

where  $\preceq$  denotes elementwise comparisons. For  $i = 1$ , since  $\mathbf{K}^1 = \mathbf{k}^{1,1}$  is a vector, we have

$$(\mathbf{K}^1 - \bar{\mathbf{T}}^1)^T \bar{\pi}^1 \leq 0 \quad (17a)$$

$$\iff (\bar{\mathbf{T}}^2 \mathbf{W}^1 \mathbf{K}^1 - \bar{\mathbf{T}}^2 \mathbf{W}^1 \bar{\mathbf{T}}^1)^T \bar{\pi}^2 \leq 0, \quad (17b)$$

where (17b) follows from (12b). For  $i = 2$  in (16), we obtain  $(\mathbf{K}^2 - \bar{\mathbf{T}}^2)^T \bar{\pi}^2 \preceq 0$ . By Property 1,  $\mathbf{W}^1 \mathbf{K}^1 \succeq 0$  and then

$$(\mathbf{W}^1 \mathbf{K}^1)^T (\mathbf{K}^2 - \bar{\mathbf{T}}^2)^T \bar{\pi}^2 \leq 0 \quad (18a)$$

$$\iff (\mathbf{K}^2 \mathbf{W}^1 \mathbf{K}^1 - \bar{\mathbf{T}}^2 \mathbf{W}^1 \mathbf{K}^1)^T \bar{\pi}^2 \leq 0. \quad (18b)$$

From (17b) and (18b), we get, for  $\mathbf{k}^{1,j} \in \mathcal{K}^{1,j}, \mathbf{k}^{2,j} \in \mathcal{K}^{2,j}$ ,

$$(\mathbf{K}^2 \mathbf{W}^1 \mathbf{K}^1 - \bar{\mathbf{T}}^2 \mathbf{W}^1 \bar{\mathbf{T}}^1)^T \bar{\pi}^2 \leq 0. \quad (19)$$

By recursively expanding  $\bar{\pi}^i$  using (12b) and combining with  $(\mathbf{W}^i \mathbf{K}^i \dots \mathbf{W}^1 \mathbf{K}^1)^T (\mathbf{K}^{i+1} - \bar{\mathbf{T}}^{i+1})^T \bar{\pi}^{i+1} \leq 0$ , we get

$$\begin{aligned} & (\mathbf{K}^L \mathbf{W}^{L-1} \mathbf{K}^{L-1} \dots \mathbf{W}^1 \mathbf{K}^1 \\ & - \bar{\mathbf{T}}^L \mathbf{W}^{L-1} \bar{\mathbf{T}}^{L-1} \dots \mathbf{W}^1 \bar{\mathbf{T}}^1)^T \mathbf{W}^{L,T} \bar{\mathbf{p}} \leq 0, \end{aligned} \quad (20)$$

for  $\mathbf{k}^{i,j} \in \mathcal{K}^{i,j}$ . By (9), (12a), and (14), this simplifies to

$$(\mathbf{x} - \bar{\mathbf{x}})^T \bar{\mathbf{p}} \leq 0, \quad \mathbf{x} \in \mathcal{K}. \quad (21)$$

Then, the theorem follows because  $\bar{\mathbf{p}} = F(\bar{\mathbf{x}})$  from (10). ■

Theorem 1 does not require Assumption 1 to hold. However, to show convergence results in the following, convexity is crucial for learning dynamics, e.g., (2), to satisfy Assumption 3.

**Lemma 1** (Set equivalence, [14, Theorem 3.2]). *If  $\mathcal{C}$  is a convex set and  $\alpha \geq 0, \beta \geq 0$ , then  $\alpha \mathcal{C} + \beta \mathcal{C} = (\alpha + \beta) \mathcal{C}$ .*

**Proposition 1** (Convexity of  $\mathcal{K}$ ). *If Assumption 1 holds, then  $\mathcal{K}$  is convex.*

*Proof.* Given two distinct points  $\mathbf{x}_p \in \mathcal{K}$ , for  $p = 1, 2$ , there exists  $\mathbf{K}_p^i$  for  $i \in [L]$ , such that  $\mathbf{x}_p = \mathbf{W}^L \mathbf{K}_p^L \dots \mathbf{W}^1 \mathbf{K}_p^1 \in \mathcal{K}$ . Note that  $\mathbf{K}_1^i$  and  $\mathbf{K}_2^i$  belong to the same convex set, denoted by  $\mathcal{K}^i$ . Then, for each  $\alpha, \beta \geq 0, \alpha + \beta = 1$ , we have  $\alpha \mathbf{x}_1 + \beta \mathbf{x}_2$

$$= (\alpha \mathbf{W}^L \mathbf{K}_1^L \dots \mathbf{W}^1) \mathbf{K}_1^1 + (\beta \mathbf{W}^L \mathbf{K}_2^L \dots \mathbf{W}^1) \mathbf{K}_2^1 \quad (22a)$$

$$= (\alpha \mathbf{W}^L \mathbf{K}_1^L \dots \mathbf{W}^1 + \beta \mathbf{W}^L \mathbf{K}_2^L \dots \mathbf{W}^1) \tilde{\mathbf{K}}^1 \quad (22b)$$

$$= (\alpha \mathbf{W}^L \mathbf{K}_1^L \dots \mathbf{K}_1^2 + \beta \mathbf{W}^L \mathbf{K}_2^L \dots \mathbf{K}_2^2) \mathbf{W}^1 \tilde{\mathbf{K}}^1 \quad (22c)$$

$$= \mathbf{W}^L \tilde{\mathbf{K}}^L \dots \mathbf{W}^1 \tilde{\mathbf{K}}^1 \in \mathcal{K}, \quad (22d)$$

where  $\tilde{\mathbf{K}}^i \in \mathcal{K}^i$ , which appears since we apply Lemma 1 to (22a) and recursively to (22c). ■

Theorem 1 states that the rest points lie in  $NE_{\mathcal{K}}(F)$ . In the following, we give conditions for convergence to these points.

**Definition 7** (Potential game). *A payoff function  $F : \Delta^d \rightarrow \mathbb{R}^d$  is a potential game if there exists a  $C^1$  potential function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ , such that  $\nabla_{\mathbf{x}} f(\mathbf{x}) = F(\mathbf{x})$ , for  $\mathbf{x} \in \Delta^d$ .*

While Theorem 1 describes the analog of the Nash stationarity property for the hierarchical structure, the following

lemma states the analog of the positive correlation property and is used to prove the convergence results in Theorem 2.

**Lemma 2.** *If each  $\mathcal{V}^{i,j}$  is positively correlated, then  $\mathbf{p}(t)^T \dot{\mathbf{x}}(t) \geq 0$  for all  $t$ .*

*Proof.*

$$\mathbf{p}(t)^T \dot{\mathbf{x}}(t) = \mathbf{p}(t)^T \left( \mathbf{W}^L \dot{\mathbf{T}}^L(t) \mathbf{M}^L(t) + \dots + \mathbf{W}^L \mathbf{T}^L(t) \dots \mathbf{W}^1 \dot{\mathbf{T}}^1(t) \right) \quad (23a)$$

$$= \boldsymbol{\pi}^L(t)^T \dot{\mathbf{T}}^L(t) \mathbf{M}^L(t) + \dots + \boldsymbol{\pi}^1(t)^T \dot{\mathbf{T}}^1(t) \quad (23b)$$

$$= \sum_{i=1}^L \sum_{j=1}^{n^i} M_j^i(t) \mathcal{V}^{i,j}(\mathbf{s}^{i,j}(t), \boldsymbol{\pi}^{i,j}(t))^T \boldsymbol{\pi}^{i,j}(t) \geq 0. \quad (23c)$$

Here, (23b) follows from (12), while the inequality in (23c) holds by positive correlation and Property 1. ■

**Theorem 2** (Convergence for potential game  $F$ ). *Let  $F$  be a potential game, and let Assumptions 1 to 4 hold. Suppose that each  $\mathcal{V}^{i,j}$  in (13) is Nash stationary w.r.t.  $\mathcal{K}^{i,j}$  and positively correlated. Then,  $\mathbf{x}(t)$  asymptotically approaches  $NE_{\mathcal{K}}(F)$ .*

*Proof.* Since  $F$  is a potential game, there exists a potential function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ . Define  $f^* = \max_{\mathbf{x} \in \mathcal{K}} f(\mathbf{x})$ , which exists since  $\mathcal{K}$  is compact. We introduce the Lyapunov function  $V : \mathcal{K} \rightarrow \mathbb{R}_+$ , given by  $V(\mathbf{x}) = f^* - f(\mathbf{x})$ . By Lemma 2,

$$\frac{d}{dt} V(\mathbf{x}(t)) = -\nabla_{\mathbf{x}} f(\mathbf{x}(t))^T \dot{\mathbf{x}}(t) = -\mathbf{p}(t)^T \dot{\mathbf{x}}(t) \leq 0. \quad (24)$$

By (23c) and Assumption 4, we have  $\mathbf{p}(t)^T \dot{\mathbf{x}}(t) \rightarrow 0$ . Then, by continuity, we have either  $\mathcal{V}^{i,j}(\mathbf{s}^{i,j}(t), \boldsymbol{\pi}^{i,j}(t))^T \boldsymbol{\pi}^{i,j}(t) \rightarrow 0$  or  $M_j^i(t) \rightarrow 0$ . For the former, by positive correlation, we have  $\mathbf{s}^{i,j}(t) \rightarrow BR_{\mathcal{K}^{i,j}}(\boldsymbol{\pi}^{i,j}(t))$ . For the latter, it means no one selects the  $(i, j)$ -group, rendering  $\mathbf{x}(t)$  unaffected by  $\mathbf{s}^{i,j}(t)$ . Therefore, by Theorem 1, we conclude  $\mathbf{x}(t) \rightarrow NE_{\mathcal{K}}(F)$ . ■

### B. Dynamic Models for Payoff

Next, we consider scenarios where the payoff, instead of being given by (10), follows the dynamical model:

$$\dot{\mathbf{q}}(t) = g(\mathbf{q}(t), \mathbf{x}(t)), \quad \mathbf{p}(t) = h(\mathbf{q}(t), \mathbf{x}(t)), \quad t \geq 0, \quad (25)$$

where  $\mathbf{q}(t) \in \mathbb{R}^a$  and the functions  $g : \mathbb{R}^a \times \Delta^d \rightarrow \mathbb{R}^a$  and  $h : \mathbb{R}^a \times \Delta^d \rightarrow \mathbb{R}^d$  are Lipschitz continuous.

**Definition 8** (Payoff Dynamics Model, PDM [15]). *A dynamic model (25) is called a Payoff Dynamics Model (PDM) if it is bounded-input bounded-output stable and recovers the static model (10) in steady state, i.e.,*

$$\lim_{t \rightarrow \infty} \|\dot{\mathbf{x}}(t)\| = 0 \implies \lim_{t \rightarrow \infty} \mathbf{p}(t) = F(\mathbf{x}(t)). \quad (26)$$

Motivated by [6], we extend our results from Theorem 2 to support PDMs that satisfy the following input-output property.

**Definition 9** (Counterclockwise dissipativity, CCW [7]). *A dynamic model (25) is Counterclockwise dissipative (CCW) if its input-output relation satisfies*

$$\alpha_c := \liminf_{T \rightarrow +\infty} \int_0^T \dot{\mathbf{p}}(t)^T \mathbf{x}(t) dt > -\infty. \quad (27)$$

A potential game  $F$  is equivalent to a *memoryless* CCW PDM [6]. Thus, the following theorem generalizes Theorem 2 in the sense that the convergence results hold for any CCW PDM. More examples of CCW dynamics can be found in [6].

**Theorem 3** (Convergence for CCW PDM). *Let the payoff  $\mathbf{p}(t)$  be given by a CCW PDM, and let Assumptions 1 to 4 hold. Suppose that each  $\mathcal{V}^{i,j}$  in (13) is Nash stationary w.r.t.  $\mathcal{K}^{i,j}$  and positively correlated. Then,  $\mathbf{x}(t) \rightarrow BR_{\mathcal{K}}(\mathbf{p}(t))$ .*

*Proof.* We first prove that  $\mathbf{p}(t)^T \dot{\mathbf{x}}(t) \rightarrow 0$ . By integration by parts and (27), we obtain

$$\begin{aligned} \int_0^\infty \mathbf{p}(t)^T \dot{\mathbf{x}}(t) dt &= [\mathbf{p}(t)^T \mathbf{x}(t)]_0^\infty - \int_0^\infty \dot{\mathbf{p}}(t)^T \mathbf{x}(t) dt \\ &\leq 2\|\mathbf{p}\|_\infty - \alpha_c < \infty. \end{aligned} \quad (28)$$

Since  $\mathbf{p}(t)^T \dot{\mathbf{x}}(t) \geq 0$  by (23), the integral in (28) exists. By (23c) and Assumption 4,  $\mathbf{p}(t)^T \dot{\mathbf{x}}(t)$  is Lipschitz continuous. Therefore, by Barbalat's lemma, we obtain that  $\mathbf{p}(t)^T \dot{\mathbf{x}}(t) \rightarrow 0$ . Then, by (23c) and positive correlation, we conclude that  $\mathbf{x}(t)$  tends to be the best response to  $\mathbf{p}(t)$  within  $\mathcal{K}$ . ■

Theorem 3 implies, when  $\dot{\mathbf{x}}(t) \rightarrow 0$ ,  $\mathbf{x}(t) \rightarrow NE_{\mathcal{K}}(F)$ , which follows from (26). Even when  $\dot{\mathbf{x}}(t) \not\rightarrow 0$ , we can nevertheless conclude that  $\mathbf{x}(t)$  tends to be the best response to  $\mathbf{p}(t)$  within  $\mathcal{K}$ . This weaker convergence result compared to Theorem 2 is since  $\mathbf{p}(t)$  is given by the dynamic payoff model instead of the static one.

## V. APPLICATIONS TO SCENARIOS WITH CONSTRAINTS

We demonstrate the results on a routing scenario where travelers use a navigation app that provides three route selection channels, each displaying an estimated travel time based on the average travel time of current users in that channel. Travelers are only informed of the estimated travel times and are unaware of differences between the channels, e.g., variations in routing sets or dispatching rules. Further, travelers select a channel based on the estimated times displayed. Each channel then assigns its users to routes within its routing set according to its dispatching strategy. For simplicity, we assume that all channels share the same routing set, comprising all possible routes. An illustration is provided in Figure 4, where  $\mathbf{s}^Y(t) \in \Delta^3$  represents how the channel  $Y = A, B, C$ , dispatches its users among the three routes at time  $t$ .

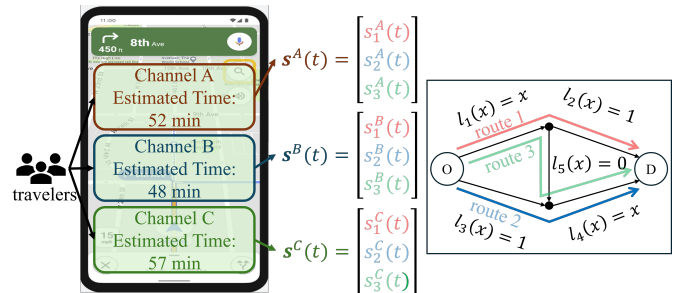


Fig. 4: Travelers select channels based on the estimated times and each channel assigns its users to routes.



Let  $\mathbf{x}(t) \in \Delta^3$  be the proportion of travelers on each route, then given the delay functions for links displayed in Figure 4, we define the payoff by (negative) travel times:

$$F(\mathbf{x}(t)) = - \begin{bmatrix} 1 + x_1(t) + x_3(t) \\ 1 + x_2(t) + x_3(t) \\ x_1(t) + x_2(t) + 2x_3(t) \end{bmatrix}. \quad (29)$$

Then, at time  $t$ , each channel  $Y$  displays the estimated travel time by  $-s^Y(t)^T F(\mathbf{x}(t))$  and updates its dispatching strategy  $s^Y(t)$  based on  $F(\mathbf{x}(t))$ , for  $Y = A, B, C$ . This forms a two-layer structure with  $\mathbf{W}^1 = \mathbf{I}_3$  and  $\mathbf{W}^2 = [\mathbf{I}_3, \mathbf{I}_3, \mathbf{I}_3] \in \mathbb{R}^{3 \times 9}$ , where  $\mathbf{I}_3$  is the 3-by-3 identity matrix. Note that  $s^A(t)$ ,  $s^B(t)$ ,  $s^C(t)$  in this example are aliases of  $s^{2,1}(t)$ ,  $s^{2,2}(t)$ ,  $s^{2,3}(t)$ .

We next demonstrate how to leverage this two-layer framework to ensure constraint satisfaction. Consider two scenarios: (i) the proportion of travelers on route 2,  $x_2(t)$ , must remain below 15%, and (ii) the proportion of travelers on route 3,  $x_3(t)$ , must not exceed 90%. The feasible sets, denoted as  $\mathcal{D}$ , are shown in Figure 5, where the vertex  $R_3$  is the unique Nash equilibrium. In case (i), the constraint set includes the original Nash equilibrium, whereas in case (ii), it does not.

We assume decisions made by travelers are unconstrained, i.e.,  $\mathcal{K}^{1,1} = \Delta^3$ . If we design  $\mathcal{K}^A, \mathcal{K}^B, \mathcal{K}^C$  such that  $\mathcal{K} \subseteq \mathcal{D}$ , then  $\mathbf{x}(t) \in \mathcal{D}$  is ensured for all  $t$ . Since  $\mathbf{W}^1 = \mathbf{I}_3$ ,  $\mathcal{K}$  is a convex combination of the sets  $\mathcal{K}^A, \mathcal{K}^B$ , and  $\mathcal{K}^C$ . We present design examples in Figure 5 where  $\mathcal{K} = \mathcal{D}$ . Let each channel  $Y$ , update  $s^Y(t)$  by constrained best response dynamics (2):

$$\dot{s}^Y(t) \in BR_{\mathcal{K}^Y}(F(\mathbf{x}(t))) - s^Y(t), \quad t \geq 0. \quad (30)$$

Since congestion games are potential games, we conclude by Theorem 2 that  $\mathbf{x}(t)$  satisfies the constraint and converges to a Nash equilibrium within  $\mathcal{D}$ , although travelers are unaware of the constraints and react only to the estimated times.

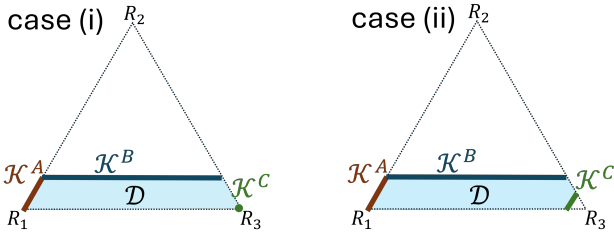


Fig. 5: Constructions of  $\mathcal{K}^A, \mathcal{K}^B, \mathcal{K}^C$  such that  $\mathcal{K} = \mathcal{D}$ .

Figure 6 demonstrates the simulation results. Thin yellow lines represent the channels' dispatching strategy trajectories, and green dots and red crosses indicate the initial and final dispatching strategies, respectively. The dotted pink line is the route distribution trajectory when travelers select routes directly. Despite convergence to the Nash equilibrium  $R_3$ , the constraints are violated during evolution. In contrast, the trajectory for the two-layer case, shown by the thick blue line, converges within  $\mathcal{D}$ . In case (ii), it converges to a new Nash equilibrium. Compared to the original Nash equilibrium  $R_3$ , flows are dispersed, showcasing the ability to steer the system toward an equilibrium with greater social efficacy.

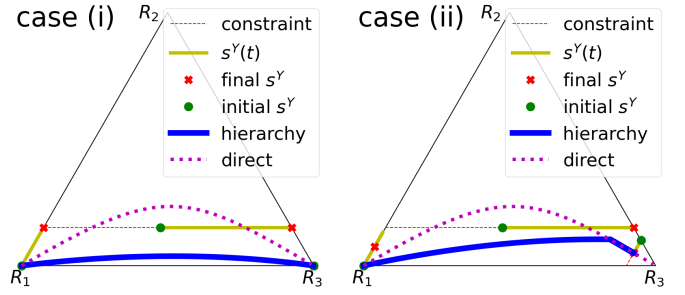


Fig. 6: Trajectory plots projected from the simplex to 2D.

## VI. CONCLUSIONS

We introduced a hierarchical decision-making framework for population games, extending the classical formulation to multiple decision layers. We characterized equilibrium properties and established convergence under payoff conditions, such as potential games and CCW payoff dynamics. Using these results, we proposed a novel approach for population games with constraints that individuals are unaware of. Some interesting avenues for future work include investigating broader classes of learning dynamics that satisfy Assumption 3 over general convex sets, as well as exploring broader classes of games that converge within the hierarchical framework.

## REFERENCES

- [1] W. H. Sandholm, "Potential games with continuous player sets," *Journal of Economic Theory*, vol. 97, no. 1, pp. 81–108, 2001.
- [2] J. Hofbauer and W. H. Sandholm, "Stable games and their dynamics," *Journal of Economic Theory*, vol. 144, no. 4, pp. 1665–1693, 2009.
- [3] M. J. Fox and J. S. Shamma, "Population games, stable games, and passivity," *Games*, vol. 4, no. 4, pp. 561–583, 2013.
- [4] M. Arcak and N. C. Martins, "Dissipativity tools for convergence to Nash equilibria in population games," *IEEE Transactions on Control of Network Systems*, vol. 8, pp. 39–50, 2020.
- [5] M. A. Mabrok, "Passivity analysis of replicator dynamics and its variations," *IEEE Transactions on Automatic Control*, vol. 66, no. 8, pp. 3879–3884, 2021.
- [6] N. C. Martins, J. Certório, and M. S. Hankins, "Counterclockwise dissipativity, potential games and evolutionary Nash equilibrium learning," 2024. [Online]. Available: <https://arxiv.org/abs/2408.00647>
- [7] D. Angeli, "Systems with counterclockwise input-output dynamics," *IEEE Transactions on Automatic Control*, vol. 51, no. 7, pp. 1130–1143, 2006.
- [8] J. Barreiro-Gomez and H. Tembine, "Constrained evolutionary games by using a mixture of imitation dynamics," *Automatica*, vol. 97, pp. 254–262, 2018.
- [9] J. Martinez-Piazuelo, N. Quijano, and C. Ocampo-Martinez, "A payoff dynamics model for equality-constrained population games," *IEEE Control Systems Letters*, vol. 6, pp. 530–535, 2022.
- [10] Y.-W. Chen, C. Kizilkale, and M. Arcak, "Solving monotone variational inequalities with best response dynamics," in *IEEE 63rd Conference on Decision and Control (CDC)*, 2024, pp. 1751–1756.
- [11] M. J. Smith, "The stability of a dynamic model of traffic assignment—an application of a method of Lyapunov," *Transportation Science*, vol. 18, no. 3, pp. 245–252, 1984.
- [12] G. W. Brown and J. Von Neumann, *Solutions of games by differential equations*. Rand Corporation, 1950.
- [13] W. H. Sandholm, *Population games and evolutionary dynamics*. MIT press, 2010.
- [14] R. T. Rockafellar, *Convex analysis*. Princeton University Press, 1970.
- [15] S. Park, N. C. Martins, and J. S. Shamma, "From population games to payoff dynamics models: A passivity-based approach," in *IEEE 58th Conference on Decision and Control (CDC)*, 2019, pp. 6584–6601.