

# From Population Games to Payoff Dynamics Models: A Passivity-Based Approach

Shinkyu Park, Nuno C. Martins, and Jeff S. Shamma

**Abstract**—This tutorial article describes a dynamical systems framework rooted in evolutionary game principles to characterize non-cooperative strategic interactions among large populations of bounded rationality agents. It also overviews recent results that use passivity notions to characterize the stability of Nash-like equilibria. In our framework, each agent belongs to a population that prescribes to its members a strategy set and a strategy revision protocol. A so-called social state registers the proportions of agents in every population adopting each strategy and a pre-selected dynamic payoff mechanism, specified by a payoff dynamics model (PDM), determines the payoff as a causal map of the social state. According to the framework, each agent must take up a strategy at a time, which it can repeatedly revise over time based on its current strategy, and information about the payoff and social state available to it. The PDM class considered in our framework can model precisely or approximately prevalent dynamic behaviors such as inertia and delays that are inherent to learning and network effects, which cannot be captured using conventional memoryless payoff mechanisms (often referred to as population games).

We organize the article in two main parts. The first introduces basic concepts prevailing in existing approaches in which a population game determines the payoff, while the second considers rather general PDM classes, of which every population game is a particular case. The latter expounds a passivity-based methodology to characterize convergence of the social state to Nash-like equilibria.

## I. INTRODUCTION

Consider that  $m$  populations partition a possibly large number of agents. Each population is characterized by a finite strategy set and a bounded rationality strategy revision mechanism, or protocol. Each agent takes up a strategy at a time and a so-called social state registers the proportions of agents in every population adopting each strategy. Every agent of a population follows the assigned protocol to repeatedly revise its strategy based on its current strategy, and information

Park's work on this tutorial was supported in part by Princeton University's School of Engineering and Applied Science through the generosity of Lydia and William Addy '82. Martins' work on this tutorial was supported by AFOSR grant FA9550-19-1-0315. The work by Martins and Park reported on this tutorial was supported by AFOSR grant FA9550-15-1-0367, NSF ECCS 1408320 and NSF ECCS 1135726. Shamma's work on this tutorial was supported by funding from King Abdullah University of Science and Technology (KAUST).

Shinkyu Park is with the Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08544, USA. shinkyu@princeton.edu

Nuno Miguel Lara Cintra Martins is with the Department of Electrical and Computer Engineering and the Institute for Systems Research, University of Maryland, College Park, MD, 20742, USA. nmartins@umd.edu

Jeff S. Shamma is with King Abdullah University of Science and Technology (KAUST), Computer, Electrical and Mathematical Science and Engineering Division (CEMSE), Thuwal 23955-6900, Saudi Arabia. jeff.shamma@kaust.edu.sa

available to it about the social state and payoff vector, which quantifies the reward for every possible strategy selection and is determined by a payoff mechanism. Hence, each population is protocol-homogeneous, but is otherwise strategy-heterogeneous because its members predominantly adopt different strategies concurrently. The revision mechanism is, in general, stochastic due to two sources of randomness: (1) the timings at which each agent is given the opportunity to revise its strategy is typically governed by a Poisson process and (2) the strategy selection mechanism at the revision times often reflects the effects of noise or implements a randomized policy.

In our analysis, the payoff mechanism is specified by a payoff dynamics model (PDM) that determines the payoff in terms of the social state. Describing recent methods to analyze the strategy selection dynamics, and sufficient conditions on the PDM and revision mechanism under which certain Nash-like equilibria are stable is the central goal of this tutorial.

### A. Beyond Population Games: Payoff Dynamics Models

As can be inferred from [1] and references therein, most past related work focuses on a class of memoryless PDM often referred to as population games [2]. Such a PDM class includes intrinsic scenarios such as games of attrition [3] that model the competitive capture of resources and congestion games [4], [5] that model the effects of traffic assignment on delays. Population games can also model prescriptive scenarios [6] in which payoff mechanisms are designed to achieve a desired behavior, and the payoff is computed and broadcast to the agents by a coordinator.

In this article, we consider a rather general PDM class that includes population games and the dynamically modified payoffs proposed in [7]–[9] as particular cases. Population games cannot account for dynamical effects, such as anticipatory behavior, communication delays and inertia inherent to learning processes, all of which are relevant for multi-agent systems and can be modelled exactly or approximately by our PDM class.

### B. Mean Social State and Deterministic Payoff

Our technical approach hinges on a tractable deterministic approximation of the strategy selection dynamics that is arbitrarily accurate, in a probabilistic sense, as the number of agents in each population increases. More specifically, we focus on deterministic approximations for the social state and payoff that we call mean social state and deterministic payoff, respectively.

An evolutionary dynamics model (EDM) that accounts for the strategy revision protocol acts on the deterministic payoff as a causal map to yield the time evolution of the mean social state. Similarly, as is shown in [10], the PDM acts as a causal map of the mean social state to generate the deterministic payoff. Hence, the EDM and the PDM operate jointly as a feedback system.

### C. A Feedback Systems Approach

The feedback interconnection between the EDM and the PDM forms the so-called mean closed loop model, which can be expressed as a first order nonlinear ordinary differential equation. The mean closed loop model can also be viewed as an extension of the so-called mean dynamic [1] because in the latter the PDM is a population game. The concatenation of the mean social state and the internal state of the PDM constitutes the state of the mean closed loop model, which can be determined by solving the associated initial value problem.

### D. Stability Of The Mean Closed Loop Model

Establishing conditions under which the mean closed loop model is stable, defined as convergence of its state to a Nash-like equilibrium set, is a central theme of this tutorial. Recent work reported in [10]–[12] explores the passivity-based methodology pioneered in [7] to introduce new concepts, techniques, and results to characterize the stability of equilibria of the mean closed loop model when the PDM is a rather general finite-dimensional nonlinear system. We provide an overview of this work in the context of the framework described above, and we explain how the new results in [10]–[12] relate to and generalize their counterparts originally derived for conventional population games.

### E. Relevance To Applications and Finite Populations

The models and stability analysis results that we overview in this article are relevant in engineering applications which often comprise of a finite number of agents. As is shown in [13] for population games and established in [10] for a general PDM class, as the size of every population tends to infinity the social state and payoff converge uniformly (in a probabilistic sense), within any finite time interval, to the state of the mean closed loop model when its initial condition is appropriately matched. Also, as the population tends to infinity, the social state starting near stable states of the mean closed loop model remains in the vicinity of the states. These two facts support the applicability of the evolutionary game framework in assessing stability of multi-agent systems.

### F. Paper Organization

We organize the paper in the following two main parts and conclude it with a brief comparison with other multi-agent games and future directions.

- 1) The first part, which comprises §II–§IV, is an overview of the key concepts and stability analysis results for the conventional setting in which a population game determines the payoff. The following is a more detailed synopsis of the first part:

- In §II we introduce basic definitions and concepts for population games and EDMs.
  - In §III we define the stability notions used throughout the first part. We also describe key stability results for important classes of population games and EDMs.
  - In §IV we discuss engineering applications that have hitherto been investigated mostly in the context of population games, or similar payoff mechanisms that, other than delays, subsume no internal dynamics, but have the potential to be re-examined in the presence of dynamics in the payoff mechanism using the techniques presented in the second part.
- 2) In the second part, which consists of §V and §VI, we introduce dynamical payoff mechanisms specified by a PDM and discuss how passivity-based techniques are used to characterize the stability of the mean closed loop model. The following is a more detailed summary of the second part:
    - In §V we introduce the PDM class adopted throughout the paper, and discuss how it generalizes population games and dynamically modified payoffs.
    - In §VI we define the stability notions used throughout the second part and describe the passivity-based techniques used in our approach. In §VI-C we describe the most general stability results available for our framework.
  - 3) In §VII we compare our framework with other multi-agent games in the literature, and in §VIII we end the article with future directions.

## II. FOUNDATIONAL CONCEPTS AND MODELS

In this section, we present the definitions and key concepts of our deterministic approach, already alluded to in §I-B, for the simplest case in which the payoff is determined by a population game. Later on, starting in §V, we will extend this basic framework to allow for more general payoff mechanisms characterized by a PDM. We summarize the notation and acronyms adopted throughout the paper in Table I.

The deterministic approach will focus on the analysis of the dynamics governing the time evolution of the mean social state and deterministic payoff. As we will see in more detail in §IV-A.3 for population games and as discussed in [10] for the PDM class introduced in §V, under independence assumptions on the sources of randomness influencing how the agents revise their strategies, the mean social state and the deterministic payoff are accurate approximations of the social state and its associated payoff, respectively, for large but finite populations.

### A. Basic Definitions and Population Games

Consider that there are  $m$  populations with finite strategy sets  $\{\mathbb{S}^k\}_{k=1}^m$ , each defined as  $\mathbb{S}^k := \{s_1^k, \dots, s_{\bar{s}^k}^k\}$ , where  $\bar{s}^k$  represents the number of strategies available to population  $k$ . We adopt a continuous time framework in which the mean

$m$	number of populations.
$\bar{s}^k$	number of available strategies for population $k$ .
$\bar{s}$	number of available strategies for all populations defined by $\bar{s} = \sum_{k=1}^m \bar{s}^k$ .
$[r]_+$	given a real number $r$ , $[r]_+$ is defined as $[r]_+ = \max\{0, r\}$ .
$\mathbb{R}^n$	$n$ -dimensional real coordinate space.
$\mathbb{R}_+^n$	$n$ -dimensional nonnegative real coordinate space.
$\mathbb{R}_-^n$	$n$ -dimensional nonpositive real coordinate space.
$\mathbb{R}_*^n$	represents viable excess payoff vectors defined as $\mathbb{R}_*^n = \mathbb{R}^n - \text{int}(\mathbb{R}_-^n)$ .
$\Delta^k$	strategy profile set given to population $k$ .
$\text{int}(\Delta^k)$	is defined as $\{z^k \in \Delta^k \mid z_i^k > 0, 1 \leq i \leq \bar{s}^k\}$ .
$\text{bd}(\Delta^k)$	is given by $\Delta^k - \text{int}(\Delta^k)$ .
$\mathfrak{D}^n$	the set of all differentiable maps from $\mathbb{R}_+$ to $\mathbb{R}^n$ .
$\mathbf{P}(\cdot), \mathbf{E}(\cdot)$	a probability distribution and an expectation operator.
$\Phi$	a projection matrix defined by $\Phi_{ij} = \begin{cases} \frac{\bar{s}-1}{\bar{s}} & i = j \\ -\frac{1}{\bar{s}} & \text{otherwise} \end{cases}$
EDM	evolutionary dynamics model
PDM	payoff dynamics model

TABLE I

LIST OF NOTATION AND ACRONYMS.

social state of each population  $k$  at time  $t$  is represented with  $x^k(t)$  and is defined below.<sup>1</sup>

**Definition 1. (Mean social state)** *The set of choice profiles  $\Delta^k$  for population  $k$  is defined as follows:*

$$\Delta^k := \left\{ z \in \mathbb{R}_+^{\bar{s}^k} \mid \sum_{i=1}^{\bar{s}^k} z_i = 1 \right\} \quad (1)$$

At any given instant  $t$ ,  $x^k(t)$  is in  $\Delta^k$  and its entries are the proportions of the  $k$ -th population adopting each strategy. Namely,  $x_i^k(t)$  is the proportion of the  $k$ -th population adopting the strategy  $s_i^k$  at time  $t$ . For notational convenience, we adopt  $\Delta := (\Delta^1, \dots, \Delta^m)$ . Similarly, we write the mean social state at time  $t$  compactly as  $x(t) := (x^1(t), \dots, x^m(t))$ .

A population game assigns a deterministic payoff to each population  $k$  at every instant as follows.

**Definition 2. (Population games)** *A population game is specified by a Lipschitz continuous map  $\mathcal{F}^k : \Delta \rightarrow \mathbb{R}^{\bar{s}^k}$ . The associated payoff mechanism is governed by the memoryless map that determines the deterministic payoff as  $p^k(t) = \mathcal{F}^k(x(t))$  for  $t \geq 0$ . For notational convenience, we adopt  $\mathcal{F} := (\mathcal{F}^1, \dots, \mathcal{F}^m)$  and  $p(t) := (p^1(t), \dots, p^m(t))$ .*

The entries of the payoff vector  $p^k(t)$  are the rewards ascribed to each strategy of the population  $k$  at instant  $t$ . Namely,  $p_i^k(t)$  is the payoff associated with strategy  $s_i^k$  at time  $t$ .

## B. Recurring Examples

There are widely-known applications of this framework to biology and engineering. In order to illustrate how to employ

<sup>1</sup>We refer to  $x^k(t)$  as the ‘‘mean’’ social state to distinguish it from the social state used to describe the strategy profile of agents in large (but finite) population games, as described in §IV-A.

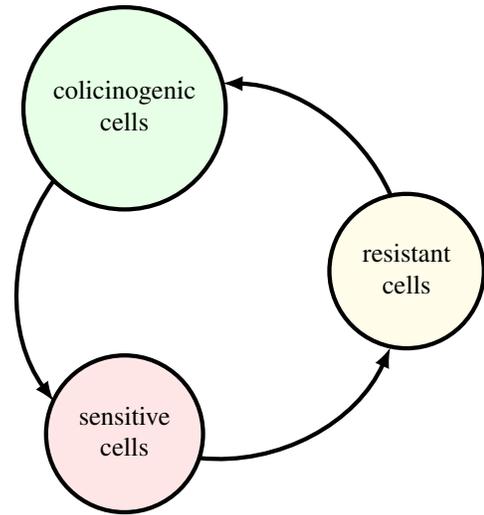


Fig. 1. A biodiversity example: RPS relationship among resistant cells (R), colicinogenic cells (C), and sensitive cells (S): S would displace R (because S has a higher growth rate), R would displace C (because R has a higher growth rate), and C would displace S (because C kills S).

our framework and key stability results, we use the following two examples of population games throughout the paper.

**Example 1. (Rock-paper-scissors (RPS) game)** *Consider a single-population RPS game whose payoff mechanism is defined by the following map:*

$$\mathcal{F}(x) = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad (2)$$

RPS games have been used in evolutionary biology to describe the interaction among different cells and predict how biodiversity depends on the spatial scale of the interactions [14]. In Fig. 1, we illustrate three different types of cells satisfying the RPS relationship. We note that the RPS game has its unique Nash equilibrium at  $(1/3, 1/3, 1/3)$ .

**Example 2. (Congestion game)** *Consider a three-strategy single-population congestion game whose payoff mechanism is defined by the following map:*

$$\mathcal{F}(x) = \begin{pmatrix} -1 & -1 & 0 \\ -1 & -2 & -1 \\ 0 & -1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix} \quad (3)$$

Congestion games have been used to model network traffic and to predict the pattern of the traffic at an equilibrium [15, Chapter 8]. Fig. 2 illustrates an example of a network traffic application, where the payoff map  $\mathcal{F}$  is defined as the negative of the total travel time of each route. We note that, in this particular example, the travel times of the segments  $A \rightarrow B$  and  $C \rightarrow D$  depend on the mean social state  $x = (x_1, x_2, x_3)$ , where each  $x_i$  represents the portion of the population using Route  $i$ . Whereas, all other segments,  $A \rightarrow C$ ,  $B \rightarrow C$ , and  $B \rightarrow D$ , have constant travel times. As was discussed in [15, Chapter 8], the game has a unique Nash equilibrium at which all the agents traverse

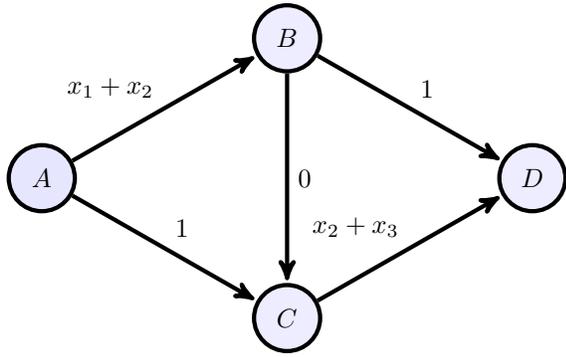


Fig. 2. A network traffic example: Agents are traversing from  $A$  to  $D$  using the following three routes:  $A \rightarrow B \rightarrow D$  (Route 1),  $A \rightarrow B \rightarrow C \rightarrow D$  (Route 2), and  $A \rightarrow C \rightarrow D$  (Route 3). Each  $x_i$  represents the portion of the population using Route  $i$  and the weight on each edge denotes the travel time.

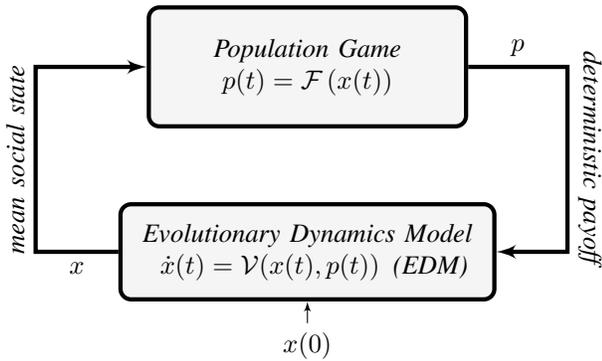


Fig. 3. A diagram representing the mean dynamics defined by a feedback interconnection between a population game and an EDM.

via  $A \rightarrow B \rightarrow C \rightarrow D$ , which underperforms the socially optimal strategy selection (known as Braess's Paradox).

### C. Protocols and Evolutionary Dynamics Model (EDM)

Evolutionary dynamics models (EDMs), as defined below, determine how  $x(t)$  evolves over time in response to initial conditions and a deterministic payoff  $p(t)$ .

**Definition 3. (Mean dynamics and EDM)** Given a population game  $\mathcal{F}$ , the mean social state  $x(t)$  and the associated deterministic payoff  $p(t)$  for  $t \geq 0$  are determined from the unique solution<sup>2</sup> of the initial value problem, with  $x(0) \in \Delta$ , of the following nonlinear ordinary differential equation referred to as mean dynamics:

$$\dot{x}_i^k(t) = \mathcal{V}_i^k(x^k(t), p^k(t)), \quad 1 \leq i \leq \bar{s}^k, \quad 1 \leq k \leq m \quad (4a)$$

$$p^k(t) = \mathcal{F}^k(x(t)) \quad (4b)$$

For each population  $k$ , the map  $\mathcal{V}^k : \Delta^k \times \mathbb{R}^{\bar{s}^k} \rightarrow \mathbb{R}^{\bar{s}^k}$  is Lipschitz continuous and guarantees that every solution of the mean dynamics remains in  $\Delta$ . The map  $\mathcal{V} := (\mathcal{V}^1, \dots, \mathcal{V}^m)$  is referred to as the evolutionary dynamics model (EDM).

In contrast with the ordinary differential equation  $\dot{x}_i^k(t) = \mathcal{V}_i^k(x^k(t), \mathcal{F}^k(x(t)))$  used in most existing work, we adopt

<sup>2</sup>See [1] for sufficient conditions for (4) to have a unique solution.

the interpretation in [7] according to which (4) is viewed as the feedback interconnection between the EDM and the population game (See Fig. 3 for an illustration).

We proceed to define what a revision protocol is and the way it affects the EDM. We do so by adapting to our framework the analogous definitions found in [1, Chapter 4].

**Definition 4. (Revision protocol)** All agents in population  $k$  follow the same revision protocol (or protocol for short) specified by Lipschitz continuous maps  $\mathcal{T}_{ij}^k : \Delta^k \times \mathbb{R}^{\bar{s}^k} \rightarrow \mathbb{R}_+$ ,  $1 \leq i, j \leq \bar{s}^k$ . We also adopt the convention that  $\mathcal{T} := (\mathcal{T}^1, \dots, \mathcal{T}^m)$  compactly represents the protocols for all populations. In particular, it determines the EDM as follows:

$$\mathcal{V}_i^k(z^k, r^k) = \sum_{j \in \mathcal{S}^k} [z_j^k \mathcal{T}_{ji}^k(z^k, r^k) - z_i^k \mathcal{T}_{ij}^k(z^k, r^k)], \quad 1 \leq i \leq \bar{s}^k, \quad z^k \in \Delta^k, \quad r^k \in \mathbb{R}^{\bar{s}^k} \quad (5)$$

An interpretation of  $\mathcal{T}_{ji}^k(z^k, r^k)$  is that it determines the rate at which agents from population  $k$  adopting strategy  $s_j^k$  switch to  $s_i^k$ , when the mean social state is  $z$  and the deterministic payoff is  $r$ . Accordingly, (5) quantifies the net rate at which agents adopt, or abandon, strategy  $s_i^k$ .

1) **Examples of Evolutionary Dynamics Models:** We proceed to review a few well-known EDM classes stemming from diverse research themes, including evolutionary games, resource allocation, and traffic assignment. We start by defining below the so-called excess payoff vector  $\hat{r}$  associated with a deterministic payoff  $r$  and mean social state  $z$ :

$$\hat{r}_i^k := r_i^k - \sum_{j=1}^{\bar{s}^k} z_j^k r_j^k \quad (6)$$

For notational convenience, we define  $\hat{r}^k := (\hat{r}_1^k, \dots, \hat{r}_{\bar{s}^k}^k)$ .

The following are well-known families of protocols carefully explained in [1, Chapters 5 and 6].

a) **Imitative:** These are protocols [16, Chapters 3 and 4] in which an agent is more likely to switch to strategies that not only offer a higher payoff but are also popular, in the sense that they are adopted by a significant portion of a population. This means that a strategy may go extinct when its popularity is overcome by that of other strategies. The imitative protocol is of the following form:

$$\mathcal{T}_{ij}^k(z^k, r^k) = z_j^k \rho_{ij}^k(z^k, r^k) \quad (7)$$

where  $\rho_{ij}^k(z^k, r^k)$  is called conditional imitation rates satisfying

$$r_j^k \geq r_i^k \iff \rho_{ij}^k(z^k, r^k) - \rho_{jl}^k(z^k, r^k) \geq \rho_{li}^k(z^k, r^k) - \rho_{il}^k(z^k, r^k), \quad 1 \leq i, j, l \leq \bar{s}^k \quad (8)$$

The well-known replicator EDM (see [17] and references therein) belongs to the class of imitative EDMs and has the

following conditional imitation rate:

$$\rho_{ij}^k(z^k, r^k) = [r_j^k - r_i^k]_+. \quad (9)$$

**b) Excess Payoff/Target (EPT):** When agents follow an excess payoff target (EPT) protocol, they are likely to switch to strategies whose payoff is higher than the average payoff for the population [18]. The higher the excess payoff for a given strategy, relative to the average, the more likely an agent will select it. The EPT protocol is of the following form:

$$\mathcal{T}_{ij}^k(z^k, r^k) = \mathcal{T}_j^{\text{EPT},k}(\hat{r}^k) \quad (10)$$

where  $\mathcal{T}_j^{\text{EPT},k}$  satisfies the acuteness condition given by

$$\sum_{i=1}^{\bar{s}^k} \hat{r}_i^k \mathcal{T}_i^{\text{EPT},k}(\hat{r}^k) > 0, \quad \hat{r}^k \in \text{int}\left(\mathbb{R}_*^{\bar{s}^k}\right)$$

*Brown-von Neumann-Nash (BNN)* [19] EDM is a widely used example of the EPT EDM, which was originally introduced in [20] to prove key properties of two-player zero-sum games, and its protocol is given by

$$\mathcal{T}_{ij}^{\text{BNN},k}(z^k, r^k) = [\hat{r}_j^k]_+. \quad (11)$$

**c) Impartial Pairwise Comparison (IPC):** According to an impartial pairwise comparison protocol, an agent is likely to switch to strategies offering a higher payoff. Typically, the likelihood of switching to a given strategy increases with its payoff. The protocol of IPC EDMs is of the following form:

$$\mathcal{T}_{ij}^k(z^k, r^k) = \mathcal{T}_j^{\text{IPC},k}(r_j^k - r_i^k) \quad (12)$$

where  $\mathcal{T}_j^{\text{IPC},k}$  satisfies the sign preservation condition given by

$$\begin{cases} \mathcal{T}_j^{\text{IPC},k}(r_j^k - r_i^k) > 0, & \text{if } r_j^k > r_i^k \\ \mathcal{T}_j^{\text{IPC},k}(r_j^k - r_i^k) = 0, & \text{if } r_j^k \leq r_i^k \end{cases} \quad (13)$$

A representative example is the so-called Smith EDM which was originally proposed in [21] to investigate a traffic assignment problem and is associated with the following protocol:

$$\mathcal{T}_j^{\text{Smith},k}(z^k, r^k) = [r_j^k - r_i^k]_+. \quad (14)$$

**d) Perturbed Best Response (PBR):** Agents following a perturbed best response (PBR) protocol transition to the strategy that they perceive to have the highest payoff. In this case, the agents would not have access to the actual payoff, but rather a version of it that is perturbed by noise. The following is the general form of the protocol for the PBR EDM, which was originally proposed in a slightly different but equivalent form in [22]:

$$\mathcal{T}_{ij}^k(z^k, r^k) = \mathbf{P} \left( r_j^k + \zeta_j^k = \max_{1 \leq i \leq \bar{s}^k} (r_i^k + \zeta_i^k) \right) \quad (15)$$

where the variable  $\zeta_j^k$  represents the noise in each agent's assessment of the deterministic payoff for strategy  $j$  of population  $k$ , and the probability distribution  $\mathbf{P}$  is computed with respect to the noise variables. More specifically,

$\zeta_i^k, \forall i \in \mathbb{S}^k, \forall k \in \{1, \dots, m\}$  are distributed according to i.i.d. additive noise processes that affect each agent's assessment of the deterministic payoff for each strategy. Hence, the PBR EDM drives the mean social state towards strategies that have the highest probability of yielding the largest payoff. [1, Chapter 6.2] provides a comprehensive discussion on the PBR EDM, with examples, most of which can also be found in [23], [24].

Interestingly, according to [23], the PBR EDM derived from (15) can be represented using an admissible deterministic perturbation  $\mathcal{Q}^k : \text{int}(\Delta^k) \rightarrow \mathbb{R}$  as follows:

$$\mathcal{V}^{\text{PBR},k}(z^k, r^k) = \arg \max_{y^k \in \text{int}(\Delta^k)} \left( \sum_{j=1}^{\bar{s}^k} y_j^k r_j^k - \mathcal{Q}^k(y^k) \right) - z^k \quad (16)$$

An important example of PBR EDMs is the logit EDM whose perturbation  $\mathcal{Q}^k$  is given by  $\mathcal{Q}^k(y^k) = \eta \sum_{j=1}^{\bar{s}^k} y_j^k \ln y_j^k$  with  $\eta > 0$  and the resulting EDM is represented by

$$\mathcal{V}_i^{\text{logit},k}(z^k, r^k) = \frac{\exp(\eta^{-1} r_i^k)}{\sum_{j=1}^{\bar{s}^k} \exp(\eta^{-1} r_j^k)} - z_i^k. \quad (17)$$

**Remark 1.** When the perturbation  $\mathcal{Q}^k$  is zero (no perturbation), the resulting EDM becomes what is known as the *Best Response (BR) EDM*. Although it belongs to an important family of EDMs, since its state-space representation is described by a differential inclusion rather than by a differential equation as in (4a), we do not discuss the BR EDM in this tutorial. We refer the interested reader to [1, Chapter 6].

### III. STABILITY ANALYSIS IN POPULATION GAMES

Characterizing stability of the mean dynamics is paramount in explaining how stable social states emerge from strategic interactions among multiple agents. The foundational work [25] in evolutionary biology explains how the notion of an evolutionarily stable strategy (ESS) is relevant in identifying an equilibrium strategy in population games that is resistant to invasion by mutant strategies and, hence, can be considered as a stable equilibrium. Notably, [26] shows that the mean social state trajectory governed by a replicator EDM converges to the ESS of a linear population game. Generalizations of this stability result are discussed in [16, Chapter 3]. Other related notions of evolutionary stability were proposed and analyzed in [27].

In spite of its importance, as not every population game has an ESS, the idea of using the ESS notion to establish stability is restricted to a certain class of population games, endowed with an ESS. To overcome this restriction, the authors of [28] introduced the concept of contractive population games<sup>3</sup>

<sup>3</sup>The concept of *contractive game* adopted here was initially referred to as *stable game* in the seminal article [28]. However, we follow recent work adopting the latter to avoid confusion with other notions of dynamical systems' stability used later on.

In what follows, we overview key stability concepts and results for contractive population games [28]. We begin by describing the definition of equilibria of the mean dynamics (4) and stability concepts in population games. Then we describe the key stability results for contractive population games under EPT, IPC, and PBR EDMs.<sup>4</sup>

#### A. Equilibria of Mean Dynamics and Stability Definition

The set of equilibria of (4) for a given EDM  $\mathcal{V}$  and a pre-specified population game  $\mathcal{F}$  is defined as

$$\mathbb{O}^{\mathcal{F},\mathcal{V}} := \{z \in \Delta \mid \mathcal{V}^k(z^k, \mathcal{F}^k(z)) = 0, 1 \leq k \leq m\}. \quad (18)$$

In some cases, it is possible to establish a relationship between the set  $\mathbb{O}^{\mathcal{F},\mathcal{V}}$  and the set of Nash equilibria  $\mathbb{NE}(\mathcal{F})$  or perturbed versions  $\mathbb{PE}(\mathcal{F}, \mathcal{Q})$  of it, which are defined, respectively, as follows:

a) **Nash Equilibria:** Given a population game  $\mathcal{F}$ , we define the set of Nash equilibria of  $\mathcal{F}$  as

$$\mathbb{NE}(\mathcal{F}) := \left\{ z \in \Delta \mid z \in \arg \max_{y \in \Delta} y^T \mathcal{F}(z) \right\}. \quad (19)$$

b) **Perturbed Nash Equilibria:** Given a population game  $\mathcal{F}$  and a twice continuously differentiable map  $\mathcal{Q} : \text{int}(\Delta) \rightarrow \mathbb{R}$ , we define the set of perturbed Nash equilibria of  $\mathcal{F}$  as

$$\mathbb{PE}(\mathcal{F}, \mathcal{Q}) := \left\{ z \in \Delta \mid z = \arg \max_{y \in \text{int}(\Delta)} (y^T \mathcal{F}(z) - \mathcal{Q}(y)) \right\}. \quad (20)$$

It follows from the definitions that if  $\mathcal{V}$  satisfies the following property, called Nash stationarity, then  $\mathbb{O}^{\mathcal{F},\mathcal{V}}$  and  $\mathbb{NE}(\mathcal{F})$  are identical.

**Nash Stationarity:** A given EDM  $\mathcal{V}$  satisfies Nash stationarity when the following equivalence holds:

$$\mathcal{V}^k(z^k, r^k) = 0 \Leftrightarrow z^k \in \arg \max_{y^k \in \Delta^k} \sum_{j=1}^{\bar{s}^k} y_j^k r_j^k, \quad 1 \leq k \leq m \quad (21)$$

Note that both EPT and IPC EDMs satisfy Nash stationarity, and their equilibria in a pre-specified population game  $\mathcal{F}$  are  $\mathbb{NE}(\mathcal{F})$ . On the other hand, PBR EDMs do not satisfy Nash stationarity, however, its set of equilibria in  $\mathcal{F}$  coincides with the perturbed Nash equilibria  $\mathbb{PE}(\mathcal{F}, \mathcal{Q})$ .

We proceed to define the stability notions used to characterize the stability of the equilibria of the mean dynamic.<sup>5</sup>

**Definition 5.** [28] *Let an EDM  $\mathcal{V}$  and a population game  $\mathcal{F}$  be given. The stability of the set  $\mathbb{O}^{\mathcal{F},\mathcal{V}}$  is classified as follows:*

- **(Global Attractiveness)** *The set  $\mathbb{O}^{\mathcal{F},\mathcal{V}}$  is globally attractive if for every initial condition  $x(0)$  in  $\Delta$ , the solution trajectory  $x(t)$ ,  $t \geq 0$  of the*

<sup>4</sup>Notice that certain imitative EDMs, such as the replicator EDM, do not possess the same stability properties as other EDM classes. Analyzing stability for imitative EDMs is beyond the scope of this article. We refer the interested reader to [29] for further discussions on this topic.

<sup>5</sup>In this tutorial, we review global stability concepts and results that are relevant to the PDM framework that we discuss in §V. See, for instance, [30] for other local stability results.

mean dynamics (4) is such that the following limit holds:

$$\lim_{t \rightarrow \infty} \inf_{z \in \mathbb{O}^{\mathcal{F},\mathcal{V}}} \|x(t) - z\| = 0 \quad (22)$$

- **(Lyapunov Stability)** *The set  $\mathbb{O}^{\mathcal{F},\mathcal{V}}$  is Lyapunov stable if for every open set  $\mathbb{O}$  in  $\mathbb{R}^{\bar{s}}$  that contains  $\mathbb{O}^{\mathcal{F},\mathcal{V}}$ , there is another open set  $\mathbb{O}'$  for which the following holds:*

$$x(0) \in \mathbb{O}' \cap \Delta \implies x(t) \in \mathbb{O}, t \geq 0 \quad (23)$$

- **(Global Asymptotic Stability)** *The set  $\mathbb{O}^{\mathcal{F},\mathcal{V}}$  is globally asymptotically stable if it is globally attractive and Lyapunov stable.*

#### B. Lyapunov Function Approaches for Stability Analysis

Lyapunov function methods have been successfully employed to characterize the stability of the mean dynamics. More specifically, [1] and references therein provide a systematic methodology to construct Lyapunov functions for mean dynamics formed by a contractive game and an EPT, IPC, or PBR EDM. Subsequently, we summarize the main tenets of this methodology.

We begin by defining contractive population games.

**Definition 6. (Contractive population game)** *A given population game  $\mathcal{F}$  is qualified as contractive if it satisfies the following condition for every  $k$  in  $\{1, \dots, m\}$ :*

$$\sum_{k=1}^m (z^k - \bar{z}^k)^T (\mathcal{F}^k(z) - \mathcal{F}^k(\bar{z})) \leq 0, \quad z, \bar{z} \in \Delta \quad (24)$$

*The population game is said to be strictly contractive if the inequality above is strict for  $z \neq \bar{z}$ .*

An example of contractive population games is the concave potential game: A given population game  $\mathcal{F}$  is qualified as a potential game if there is a continuously differential map (potential)  $\mathcal{I}^{\mathcal{F}} : \mathbb{R}^{\bar{s}} \rightarrow \mathbb{R}$  satisfying:

$$\mathcal{F}^k(z) = \nabla_{z^k} \mathcal{I}^{\mathcal{F}}(z), \quad 1 \leq k \leq m, z \in \Delta \quad (25)$$

When  $\mathcal{I}^{\mathcal{F}}$  is (strictly) concave, we say that  $\mathcal{F}$  is a (strictly) concave potential game. Congestion games (see, for instance, [1, Example 13.16]) with concave payoff potentials are representative examples of concave potential games in which agents would compete for resources, e.g., communication link, traffic road, etc., and more agents compete less payoff they receive.

1) *Stability for EPT EDM:* When a contractive population game  $\mathcal{F}$  is played according to the EPT EDM satisfying the so-called integrability condition, described below, the resulting trajectory of the mean social state is guaranteed to have global convergence properties [28], which are obtained by applying the Lyapunov methods.

**Integrability:** There is a continuously differentiable function  $\mathcal{I}^{\text{EPT},k} : \mathbb{R}^{\bar{s}^k} \rightarrow \mathbb{R}$  such that the following holds:

$$\mathcal{I}^{\text{EPT},k}(\hat{r}^k) = \nabla \mathcal{I}^{\text{EPT},k}(\hat{r}^k), \quad \hat{r}^k \in \text{int}(\mathbb{R}_{*}^{\bar{s}^k}).$$

**Theorem 1.** [10], [28] *Consider mean dynamics formed by an integrable EPT EDM and a contractive population*

game  $\mathcal{F}$ . The set of Nash equilibria  $\mathbb{NE}(\mathcal{F})$  is globally asymptotically stable, where the Lyapunov function is given as

$$\mathcal{S}^{EPT,\mathcal{F}}(z) = \sum_{k=1}^m \mathcal{I}^{EPT,k}(\hat{\mathcal{F}}^k(z)) - \gamma, \quad z \in \Delta \quad (26)$$

with  $\hat{\mathcal{F}}_i^k(z) = \mathcal{F}_i^k(z) - \sum_{j=1}^{\bar{s}^k} z_j^k \mathcal{F}_j^k(z)$  and a constant  $\gamma$  such that  $\mathcal{S}^{EPT,\mathcal{F}}$  is positive.

**Remark 2.** The preliminary version of the stability result for the integrable EPT EDM in contractive population games was first reported in [28, Theorem 5.1] in which global asymptotic stability is only guaranteed when the EPT protocol is separable (as defined in [28, Section 4.5]) or  $\mathcal{F}$  has a unique Nash equilibrium. As an extension of it, the analysis presented in [10, Theorem 1] showed that any integrable EPT EDM in contractive population games achieves asymptotic stability and has the Lyapunov function of the form (26).

2) *Stability for IPC EDM:* For IPC EDMs, the following theorem states asymptotic stability of Nash equilibria of a contractive game  $\mathcal{F}$ .

**Theorem 2.** [28] Consider mean dynamics formed by an IPC EDM and a contractive population game  $\mathcal{F}$ . The set of Nash equilibria  $\mathbb{NE}(\mathcal{F})$  is globally asymptotically stable, where the Lyapunov function is given as

$$\mathcal{S}^{IPC,\mathcal{F}}(z) = \sum_{k=1}^m \sum_{i=1}^{\bar{s}^k} \sum_{j=1}^{\bar{s}^k} z_i^k \int_0^{\mathcal{F}_j^k(z) - \mathcal{F}_i^k(z)} \mathcal{T}_j^{IPC,k}(\tau) d\tau, \quad z \in \Delta \quad (27)$$

3) *Stability for PBR EDM:* For PBR EDMs, when the deterministic perturbation  $\mathcal{Q}$  is strictly convex, in a contractive game  $\mathcal{F}$ , the set  $\mathbb{PE}(\mathcal{F}, \mathcal{Q})$  has a unique element that is asymptotically stable. We summarize the stability result for PBR EDMs in contractive games in the following theorem.

**Theorem 3.** [28] Consider mean dynamics formed by a PBR EDM characterized by strictly convex perturbation  $\mathcal{Q}$  and a contractive population game  $\mathcal{F}$ . There is a unique state in  $\mathbb{PE}(\mathcal{F}, \mathcal{Q})$  which is asymptotically stable, where the Lyapunov function is given as

$$\mathcal{S}^{PBR,\mathcal{F}}(z) = \max_{y \in \text{int}(\Delta)} (y^T \mathcal{F}(z) - \mathcal{Q}(y)) - (z^T \mathcal{F}(z) - \mathcal{Q}(z)), \quad z \in \Delta \quad (28)$$

#### IV. APPLICATION OF POPULATION GAMES AND EVOLUTIONARY DYNAMICS

In this section, we explain how the framework consisting of population games and EDMs can be applied to analyze stability in engineering systems which often comprise of large but finitely many agents. We start with a discussion on how the finite population systems can be approximated by the framework we described in §II and how the stability results from §III would be relevant in establishing stability for the systems. Then, we present representative examples in

system control and optimization to which the approximation discussion is applicable.

##### A. Finite Population Approximation

In order to assess the predictive value of the (deterministic) mean social state trajectories obtained as solutions of the mean dynamics (4), we analyze how well they approximate the (stochastic) social state trajectories that would occur when the populations have a finite, albeit possibly large, number of agents. More specifically, in this section, we explain based on the analyses given in [13], [31] how the social state in the finite population case asymptotically approaches the solution of (4) as the number of agents of every population tends to infinity. To avoid heavy notation, we restrict to a single-population case ( $m = 1$ ) and omit the superscripts used to denote each population  $k$ .

For the sake of argument, consider that the population is formed by a finite number of agents, each adopting a single strategy at a time according to a mechanism that will be explained in detail below. The total number of agents is denoted by  $N$  and  $A_\alpha(t)$  taking values in  $\mathbb{S}$  represents the strategy adopted by agent  $\alpha$  at time  $t$ .

The strategy choice profile of the population is characterized by a continuous-time right-continuous jump process  $X$  taking values in  $\Delta^N$  defined by

$$\Delta^N := \left\{ z \in \mathbb{R}^{\bar{s}} \mid \sum_{i=1}^{\bar{s}} z_i = 1 \text{ and } z_j \in \{0, 1/N, \dots, 1\}, 1 \leq j \leq \bar{s} \right\}. \quad (29)$$

The portion of the population adopting strategy  $s_i$  at time  $t$  is represented by  $X_i(t)$ , or equivalently:

$$\begin{aligned} X_i(t) &= \frac{\# \text{ of agents adopting } s_i \text{ at time } t}{N} \\ &= \frac{\sum_{\alpha=1}^N \mathcal{I}(A_\alpha(t) = s_i)}{N} \end{aligned} \quad (30)$$

where  $\mathcal{I}$  is the indicator function. We use  $X(t) = (X_1(t), \dots, X_{\bar{s}}(t))$  to represent the social state in finite population game settings.

1) **Poisson Clock:** We define a Poisson process  $\Gamma_\alpha$  that specifies when each agent revises its strategy in which case  $A_\alpha$  is a right-continuous Markov process that is allowed to switch strategies at the jump times of  $\Gamma_\alpha$ . Based on the Poisson process, we define a revision protocol that specifies the rules determining how the agents revise their strategies.

The jump times of the Poisson process  $\Gamma_\alpha$ , with a positive rate  $\lambda$ , determines the times at which the  $\alpha$ -th agent may switch its strategy. Following [1], we call  $\Gamma_\alpha$  the Poisson clock of agent  $\alpha$ . Similarly,  $\Gamma$ , defined at each  $t$  as  $\Gamma(t) = \sum_{\alpha=1}^N \Gamma_\alpha(t)$ , is a Poisson process whose jump times indicate when some agents are allowed to switch their strategies. The rates of these processes are given by  $\lambda = N\lambda$ . In addition, these processes satisfy the following properties:

- The Poisson processes  $\Gamma_\alpha$ ,  $1 \leq \alpha \leq N$  are independent. This can be interpreted as each agent having its

own private Poisson clock that determines when it can revise its strategy.

- The Poisson clocks of all agents have identical rate  $\lambda$ . In particular,  $\Gamma_\alpha$  is described by the following complementary cumulative exponential distribution of the inter-arrival times:

$$\mathbf{P}(T_\alpha(\tau) - \tau > \delta) = e^{-\delta\lambda}, \quad \delta \geq 0 \quad (31)$$

where  $\tau$  is a jump time and  $T_\alpha$  is a random process according to which  $T_\alpha(\tau)$  is a random variable defined as the time when the first jump of  $\Gamma_\alpha$  occurs after time  $\tau$ , or equivalently,

$$T_\alpha(\tau) := \min \{t > \tau \mid \Gamma_\alpha(t) = \Gamma_\alpha(\tau) + 1\}$$

2) **Revision Protocol:** All agents follow the same revision protocol to switch their strategy selections. We specify the protocol by Lipschitz continuous maps  $\mathcal{T}_{ij} : \Delta \times \mathbb{R}^{\bar{s}} \rightarrow [0, 1]$ ,  $1 \leq i, j \leq \bar{s}$ . More specifically, the protocol governs the following transition probability:

$$\begin{aligned} \mathbf{P}(A_\alpha(t^+) = s_j \mid A_\alpha(t^-) = s_i, \Gamma_\alpha(t^+) = \Gamma_\alpha(t^-) + 1) \\ = \mathcal{T}_{ij}(X(t^-), P(t^-)), \quad 0 \leq t^- < t^+, \quad 1 \leq i, j \leq \bar{s} \end{aligned} \quad (32a)$$

$$\begin{aligned} \mathbf{P}(A_\alpha(t^+) = s_i \mid A_\alpha(t^-) = s_i, \Gamma_\alpha(t^+) = \Gamma_\alpha(t^-)) \\ = 1, \quad 0 \leq t^- < t^+, \quad 1 \leq i \leq \bar{s} \end{aligned} \quad (32b)$$

where  $P$  is the payoff assigned to the population according to a population game  $\mathcal{F}$ , e.g.,  $P = \mathcal{F}(X)$ .

From (32), the following quantifies the expected increments of the social state given that one agent is allowed to revise its strategy:

$$\begin{aligned} \mathbf{E}[X_i(t^+) - z_i \mid X(t^-) = z, \Gamma(t^+) = \Gamma(t^-) + 1] \\ = \frac{1}{N} \sum_{j=1}^{\bar{s}} [z_j \mathcal{T}_{ji}(z, \mathcal{F}(z)) - z_i \mathcal{T}_{ij}(z, \mathcal{F}(z))], \\ 0 \leq t^- < t^+, \quad 1 \leq i \leq \bar{s}, \quad z \in \Delta^N \end{aligned} \quad (33)$$

where we use  $z$  to represent a given social state at instance  $t^-$ . Using the fact that the intervals in-between jumps are exponential, at every jump time  $t$ , we further obtain:

$$\begin{aligned} \lim_{\delta \downarrow 0} \mathbf{E} \left[ \frac{X_i(t + \delta) - z_i}{\delta} \mid X(t) = z \right] = \mathcal{V}_i(z, \mathcal{F}(z)), \\ 1 \leq i \leq \bar{s}, \quad z \in \Delta^N \end{aligned} \quad (34)$$

where  $\mathcal{V} : \Delta \times \mathbb{R}^{\bar{s}} \rightarrow \mathbb{T}\Delta$ , with  $\mathbb{T}\Delta := \{z \in \mathbb{R}^{\bar{s}} \mid \sum_{i=1}^{\bar{s}} z_i = 0\}$ , is defined for each component as follows:

$$\begin{aligned} \mathcal{V}_i(z, r) = \lambda \sum_{j=1}^{\bar{s}} [z_j \mathcal{T}_{ji}(z, r) - z_i \mathcal{T}_{ij}(z, r)], \\ 1 \leq i \leq \bar{s}, \quad z \in \Delta^N, \quad r \in \mathbb{R}^{\bar{s}} \end{aligned} \quad (35)$$

Here  $r$  represents the possible deterministic payoff vectors that are, in the current context, determined by  $\mathcal{F}(z)$ , for a given mean social state  $z$ .

3) **Deterministic Approximation:** Since the above construction of  $X$  conforms with the framework in [32], we should expect from (34) that as the number  $N$  of agents tends to infinity, the trajectory of  $X$  would tend to the solution of the initial value problem for the mean dynamics (4) in which  $\mathcal{V}_i$  is given by (35).

In particular, consider a mean social state trajectory  $x(t)$ ,  $t \geq 0$  determined by (4). Notably, as is shown in [13, Theorem 4.1], if the population size tends to infinity and  $X(0)$  converges to  $x(0)$  then for every  $\bar{t} > 0$  the following holds:

$$\lim_{N \rightarrow \infty} \mathbf{P} \left( \sup_{t \in [0, \bar{t}]} \|X(t) - x(t)\| > \epsilon \right) = 0, \quad \forall \epsilon > 0 \quad (36)$$

Furthermore, [31, Lemma 1] also shows that for every  $\bar{t} > 0$ , the probability that  $X$  remains within an  $\epsilon$ -tube around  $x$  at any time in  $[0, \bar{t}]$  decays exponentially with  $\epsilon$  for  $N$  sufficiently large. According to [31, Proposition 2], the exponential decay result can be used to establish that if the equilibrium set  $\mathbb{O}^{\mathcal{F}, \mathcal{V}}$ , defined in (18), is asymptotically stable then, in the population limit, the stochastic process  $X(t)$  stays within an arbitrarily small neighborhood of  $\mathbb{O}^{\mathcal{F}, \mathcal{V}}$ . Specifically, let  $\mathbb{V}$  be any neighborhood of  $\mathbb{O}^{\mathcal{F}, \mathcal{V}}$ , then there is an open subset  $\mathbb{U}$  of  $\mathbb{V}$  containing  $\mathbb{O}^{\mathcal{F}, \mathcal{V}}$  for which the exit time  $\tau^N(\mathbb{U}) := \inf \{t \geq 0 \mid X(t) \notin \mathbb{U}\}$  of the process  $X(t)$  from  $\mathbb{U}$  satisfies

$$\mathbf{P} \left( \liminf_{N \rightarrow \infty} \tau^N(\mathbb{U}) = +\infty \right) = 1 \text{ if } X(0) \in \mathbb{U} \quad (37)$$

Clearly, the above two results indicate that the solutions of the mean dynamics (4) are accurate predictions of the behavior of the stochastic process  $X(t)$  over a finite time horizon and the stable equilibrium set of (4) is a correct predictor of the social state  $X(t)$  provided that  $X(t)$  starts near the equilibrium set.

### B. Applications in Engineering Systems and Optimization

Based on the arguments presented in §IV-A, we discuss examples illustrating applications of the population game framework to engineering problems, resource allocation and optimization. Although, in these examples, we consider payoff mechanisms to either be memoryless (population games) or have a fixed delay but no other internal dynamics, they describe important applications in which the results in §V would be relevant to allow internal dynamics in payoff mechanisms or revision protocols. As we discuss in §V-A, by including additional dynamics in the PDM, one can model important effects such as inertia and anticipation in the agents' perceptions of the payoff.

1) *Traffic assignment:* The seminal work reported in [21] investigates a traffic assignment problem in which the strategies are the possible routes each agent may traverse to travel from one location to another. The main result in [21] shows that certain simple protocols, which lead to what was latter called Smith dynamics [1, Example 5.6.1] and in this tutorial is referred to as Smith EDM (14), are guaranteed to steer the mean social state towards the Wardrop [33]

equilibrium of a given network cost-flow function. In the context of population games, such a cost is a congestion population game and the Wardrop equilibrium coincides with the befitting version of Nash's described in (19).

2) *Distributed control strategies for resource allocation and optimization:* The distributed control strategies discussed in [34] hinge on simple protocols to regulate flow distribution in urban drainage systems, voltage split for lighting systems and economic power dispatch in microgrids. The underlying population games determining the payoff for these applications have a strictly concave potential. In this setting, the same Smith EDM mentioned in §IV-B.1, or appropriate modifications thereof, guarantees that the mean social state will always converge to the maximum of the potential, which is also the unique Nash equilibrium of the population game. In fact, the potential can be used as a Lyapunov function to establish that its maximum is globally asymptotically stable. As is explained in [35] and [1, Section 3.1] the maximum of a strictly concave potential is globally asymptotically stable under any protocol satisfying positive correlation and Nash stationarity. More generally, the set of Nash equilibria of a potential game - as population games having a potential are called [36] - can be determined by the Karush-Kuhn-Tucker (KKT) conditions of local optimality of the potential. These results highlight the relevance of the population game formulation for distributed optimization and engineering applications in which a potential game is constructed and implemented by a coordinator, as is done in [34] and references therein, to promote distributed resource allocation. Notably, as explained in [1, Example 3.1.2], congestion population games in general admit a potential. In fact, the potential has been used since the early work in [5] to prove the existence of and also compute the Nash equilibria of congestion games.

3) *Distributed control strategies for wireless networks:* The approach in [37] uses the formalism of population games to design and analyze the performance of distributed algorithms that either trigger transmission or regulate the transmission power of a large number of agents in the scenarios of slotted ALOHA-based access network or W-CDMA power control, respectively. In these scenarios, the strategy sets, or equivalently the state of each agent, are  $\{T, N\}$  (transmit or not) and  $\{P_H, P_L\}$  (high or low power), respectively. The payoff for the ALOHA-based access network scenario captures the net reward of transmitting successfully subject to a cost quantifying the power spent. A population game captures the effect of collisions that may arise from simultaneous transmissions. In the W-CDMA power control scenario, the payoff captures the ability to broadcast to other nodes and the underlying population game also models the effect of interference that occurs when neighboring agents transmit concurrently.

## V. FROM POPULATION GAMES TO PAYOFF DYNAMICS MODELS

In this section, we introduce a new class of payoff dynamics models, which generalizes the conventional population games.

We begin by motivating the need to consider such models, and we illustrate how they can affect the stability of the closed loop in the context of the examples given in §II.

### A. Motivation and Definition of Payoff Dynamics Models

The mean dynamics (4) considered so far presume that the payoff mechanism is specified by a population game defined by a memoryless map  $\mathcal{F}^k : x(t) \mapsto p^k(t)$ ,  $1 \leq k \leq m$  for  $t \geq 0$ . According to such payoff mechanism, the payoff is instantaneously determined from the mean social state and immediately made available to the agents. To emphasize importance of considering dynamic models in defining the payoff mechanism, we list the following features which the memoryless framework cannot model (see Fig. 4):

(a) **Network effects and agents' learning dynamics:** More often than not, the payoff as determined by the game may be different from the perceived payoff that the agents use to make decisions, in which case they are related by a so-called learning dynamics. As a result, the learning dynamics may influence the behavior of the mean dynamics substantially. The following are examples of causes of learning dynamics:

- In a prescriptive scenario, when the central entity computes and broadcasts an estimated payoff, the learning dynamics may reflect inertia of the price-setting mechanism or communication delays.
- In an intrinsic scenario, the perceived payoff may be learned by each agent independently or as a networked group based on available information. In this case learning dynamics may be driven not only by communication and computation delays, but also by inertia in the beliefs of each agent.

(b) **Payoff mechanism learning dynamics:** There are also dynamics and possibly delays on the way in which the social state affects the payoff. It may stem from the following scenarios:

- In a prescriptive scenario, when the central entity computes and broadcasts the payoff, it may do so based on an estimate of the social state. The mechanism by which the estimate is determined may incur communication and data gathering delays, and be affected by the transient behavior caused by averaging or other estimation methods and algorithms.
- When the payoff results from mechanisms intrinsic to the game, causation dynamics may model the delays between when the social state attains a certain profile and the payoff reflects the associated value of the game. Congestion games are examples of this because the flow at the routes may not be an instantaneous function of the traffic density.

With the purpose of modeling dynamic effects in the payoff mechanism, such as learning and causation dynamics, we consider the following class of payoff dynamics models (PDMs).

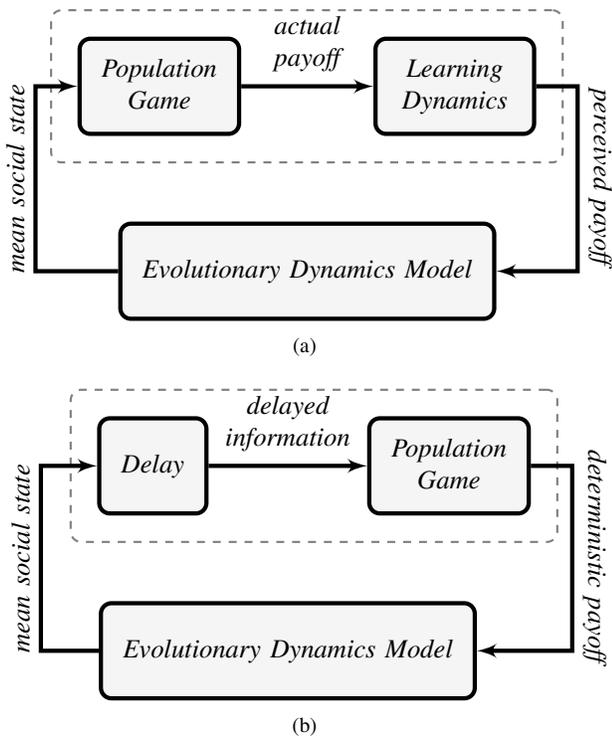


Fig. 4. Incorporating dynamic features in the payoff mechanism: (a) learning and (b) delays.

**Definition 7. (Payoff dynamics model)** [10] Consider that a payoff dynamics model (PDM) is defined as follows:

$$\dot{q}(t) = \mathcal{G}(q(t), x(t)) \quad (38a)$$

$$p^k(t) = \mathcal{H}^k(q(t), x(t)), \quad 1 \leq k \leq m \quad (38b)$$

where  $(q(0), x(0)) \in \mathbb{R}^n \times \Delta$ ,  $\mathcal{G} : \mathbb{R}^n \times \Delta \rightarrow \mathbb{R}^n$  is Lipschitz continuous, and  $\mathcal{H}^k : \mathbb{R}^n \times \Delta \rightarrow \mathbb{R}^{s^k}$  is continuously differentiable and Lipschitz continuous. The state-space representation (38) specifies a map from the mean social state trajectory  $x(t)$ ,  $t \geq 0$  and the initial condition  $q(0)$  to the state  $q(t)$  and the deterministic payoff  $p(t)$ . Henceforth, we specify a PDM by the pair  $(\mathcal{G}, \mathcal{H})$ , where  $\mathcal{H} = (\mathcal{H}^1, \dots, \mathcal{H}^m)$ .

The following technical condition will allow us to apply the concepts of Nash equilibrium and perturbed equilibrium to our PDM class. It requires that in the stationary limit every PDM should act as a population game.

**Stationary population game of a PDM:** For every PDM, there is a continuous map  $\bar{\mathcal{F}}^k : \Delta \rightarrow \mathbb{R}^{s^k}$  for which the following implication holds for any initial condition  $q(0) \in \mathbb{R}^n$ :

$$\lim_{t \rightarrow \infty} \|\dot{x}(t)\| = 0 \implies \lim_{t \rightarrow \infty} \|p^k(t) - \bar{\mathcal{F}}^k(x(t))\| = 0, \quad 1 \leq k \leq m \quad (39)$$

In addition, we require that the set

$$\{(z, s) \in \Delta \times \mathbb{R}^n \mid \mathcal{H}^k(s, z) = \bar{\mathcal{F}}^k(z), \quad 1 \leq k \leq m\}$$

is either a compact subset of or the entire set  $\Delta \times \mathbb{R}^n$ . We refer to  $\bar{\mathcal{F}}$  as the stationary population game of the PDM.

The following example is adopted from [10].

**Example 3. (Smoothing-Anticipatory PDM)** Consider that  $\mathcal{F}^k : \Delta \rightarrow \mathbb{R}^{s^k}$ ,  $1 \leq k \leq m$  is a given continuously differentiable map defining a population game. Given a positive constant  $\alpha$  and non-negative parameters  $\mu_I$ ,  $\mu_S$ , and  $\mu_A$  satisfying  $\mu_I + \mu_S = 1$ , the associated smoothing-anticipatory PDM is defined as follows:

$$\dot{q}^k(t) = \alpha (\mathcal{F}^k(x(t)) - q^k(t)) \quad (40a)$$

$$p^k(t) = \mu_I \mathcal{F}^k(x(t)) + \mu_S q^k(t) + \mu_A \dot{q}^k(t) \quad (40b)$$

for  $t \geq 0$  and  $q^k(0) \in \mathbb{R}^{s^k}$ .

The smoothing-anticipatory PDM could model the learning dynamics of the agents. More specifically,  $\alpha^{-1}$  would represent the time constant quantifying how accurately the smoothing state  $q^k(t)$  can track rapid variations of  $\mathcal{F}^k(x(t))$ . In this context,  $\mathcal{F}^k(x(t))$  would represent the actual payoff, while  $p^k(t)$  is the payoff as perceived by the agents. The parameters  $\mu_I$ ,  $\mu_S$ , and  $\mu_A$  would determine the relative importance of the actual payoff  $\mathcal{F}^k(x(t))$  versus  $q^k(t)$  and its derivative  $\dot{q}^k(t)$ , which can be interpreted as an estimate of the direction towards which the actual payoff may be heading.

In order to show that (40) complies with (39), it suffices to notice that  $\dot{q}^k(t)$  vanishes when does  $\dot{x}^k(t)$ , in which case  $p^k(t)$  tends to  $\mathcal{F}^k(x(t))$  leading to the conclusion that  $\mathcal{F}$  is the stationary population game of the smoothing-anticipatory PDM.

**Remark 3.** The following are parameter choices leading to particular cases proposed in the literature:

- When  $\mu_I = 1$  and  $\mu_S = \mu_A = 0$  the PDM is a population game.
- When  $\mu_I = 1$ ,  $\mu_S = 0$ , and  $\mu_A > 0$  we obtain an anticipatory PDM, as considered in [8].
- The smoothing PDM considered in [7] is obtained when  $\mu_I = 0$ ,  $\mu_S = 1$ , and  $\mu_A = 0$ .

In addition to the smoothing-anticipatory PDM, another important class of higher-order dynamics models is investigated in [38] that are of integral-type payoff mechanisms. The models are used to generalize the standard first-order replicator EDM and to study the effect of *higher-order payoffs* on the mean social state trajectory derived by the replicator EDM.

## B. Examples Revisited

Recall the biodiversity and network traffic examples described in Examples 1 and 2, respectively. For each example, we demonstrate how the PDM framework can be used to model learning dynamics of agents and illustrate how the convergence properties of resulting mean dynamics differ from the population game case, which highlights the importance of devising a new approach to establish stability of the mean dynamics.

For both the examples, we use the following BNN EDM obtained from (11):

$$\dot{x}_i(t) = [\hat{p}_i(t)]_+ - x_i(t) \sum_{j=1}^3 [\hat{p}_j(t)]_+, \quad 1 \leq i \leq 3 \quad (41)$$

We consider two different types of PDMs defined below.

- **(Static PDM)** Consider that the payoff vector  $p$  is a static function of the mean social state  $x$ :

$$p_i(t) = \mathcal{F}_i(x(t)), \quad 1 \leq i \leq 3 \quad (42)$$

- **(Smoothing PDM)** Consider that agents are learning the deterministic payoff vector based on an exponentially-discounted reinforcement learning model [39] in which the perceived payoff vector  $p$  depends on  $x$  as follows:

$$\dot{p}_i(t) = \mathcal{F}_i(x(t)) - p_i(t), \quad 1 \leq i \leq 3 \quad (43)$$

Note that in both the cases the stationary population game  $\bar{\mathcal{F}}$  can be identified as  $\bar{\mathcal{F}} = \mathcal{F}$ .

Using the definition of contractive population games, as explained in (24), the static PDMs (42) for Examples 1 and 2 are both contractive. According to Theorem 1, since the BNN EDM (41) is an integrable EPT EDM, we can anticipate that all the trajectories derived by the mean dynamics consisting of (41) and (42) converge to Nash equilibrium. This is illustrated in Figs. 5(a) and 6(a).

On the other hand, when the mean dynamics are modified to include the smoothing PDM (43), the statements in Theorem 1 are no longer applicable to conclude the stability of the mean dynamics. Evidently, as illustrated in Figs. 5(b) and 6(b), resulting mean social state trajectories behave differently even if the stationary population games in both the examples are contractive. In particular, the reinforcement learning dynamics used to define (43) prevents the convergence of the trajectories to the Nash equilibrium in the RPS game (2). Motivated by these observations, in the next section, we discuss a new approach for establishing stability in the PDM framework.

## VI. EXTENSION OF STABILITY CONCEPTS AND PASSIVITY-BASED STABILITY ANALYSIS

In this section, we generalize the stability analysis of Section III-B for the case in which a PDM determines the payoff. We start by precisely defining the so-called mean closed loop model and associated stability concepts. Subsequently, we describe recent results that are stated in detail and proved in [10] using techniques that build on the passivity-based approach proposed in [7].

### A. Mean Closed Loop Model and Stability Concepts

The mean closed loop model is defined below (see Fig. 7 for an illustration).

**Definition 8. (Mean closed loop model)** Given a PDM and an EDM, the associated mean closed loop model is obtained by using  $x$  as an input to (38) and  $p$  as an input to (4a).

The following is the state space representation of the mean closed loop model:

$$\dot{q}(t) = \mathcal{G}(q(t), x(t)) \quad (44a)$$

$$\dot{x}^k(t) = \mathcal{V}_{\mathcal{H}}^k(q(t), x(t)), \quad 1 \leq k \leq m \quad (44b)$$

$$p^k(t) = \mathcal{H}^k(q(t), x(t)) \quad (44c)$$

where  $(q(0), x(0)) \in \mathbb{R}^n \times \Delta$  and  $\mathcal{V}_{\mathcal{H}}^k : \mathbb{R}^n \times \Delta \rightarrow \mathbb{R}^{s^k}$  is defined as:

$$\mathcal{V}_{\mathcal{H}}^k(s, z) := \mathcal{V}^k(z^k, \mathcal{H}^k(s, z)), \quad s \in \mathbb{R}^n, z \in \Delta$$

Notice that the state of the mean closed loop model is  $(q(t), x(t))$ , and the state trajectory  $(q(t), x(t))$ ,  $t \geq 0$  is the solution of the associated initial value problem in response to the pre-selected initial condition  $(q(0), x(0))$ .

**Remark 4. (Deterministic approximation for large but finite populations)** As is discussed in [10], an immediate extension of the analysis in §IV-A shows that for finite populations the social state and the payoff as determined by the PDM will converge, in a sense analogous to (36), to the mean social state and deterministic payoff, respectively, obtained from the solution of the initial value problem for the mean closed loop as the size of every population tends to infinity.

That investigating the stability of equilibria for the mean closed loop model is essential to characterize the long term behavior of large populations when a PDM specifies a dynamic payoff mechanism is an immediate conclusion of Remark 4 and the additional related discussion [10]. Hence, we proceed to describing key concepts and results to carry out such a stability analysis.

Note that when the state  $(q^*, x^*)$  of the mean closed loop model is at an equilibrium point, i.e.,  $\mathcal{G}(q^*, x^*) = \mathcal{V}_{\mathcal{H}}^k(q^*, x^*) = 0$ ,  $1 \leq k \leq m$ , according to the requirement that each PDM has a stationary game  $\bar{\mathcal{F}}$ , we can infer that  $\mathcal{V}^k(x^{k*}, \bar{\mathcal{F}}^k(x^*)) = 0$ ,  $1 \leq k \leq m$ . Note that in this case each equilibrium point of the mean closed loop model (44) is associated with an equilibrium point of the mean dynamic (4) defined by the population game  $\bar{\mathcal{F}}$ . This observation allows us to extend the stability concepts defined in Definition 5 as follows.

**Definition 9. (Stability concepts)** [10] Given an EDM and a PDM, consider the mean closed loop model (44). The stability of a set  $\mathbb{E}(\bar{\mathcal{F}})$  defined by

$$\mathbb{E}(\bar{\mathcal{F}}) := \{z \in \Delta \mid \mathcal{V}^k(z^k, \bar{\mathcal{F}}^k(z)) = 0, \quad 1 \leq k \leq m\}$$

which in our formulation is either  $\mathbb{NE}(\bar{\mathcal{F}})$  or  $\mathbb{PE}(\bar{\mathcal{F}}, \mathcal{Q})$ , is classified as follows:

- **(Global Attractiveness)** The set  $\mathbb{E}(\bar{\mathcal{F}})$  is globally attractive if for every initial condition  $(q(0), x(0))$  in  $\mathbb{R}^n \times \Delta$ , the solution trajectory  $(q(t), x(t))$ ,  $t \geq 0$  of the mean closed loop model (44)

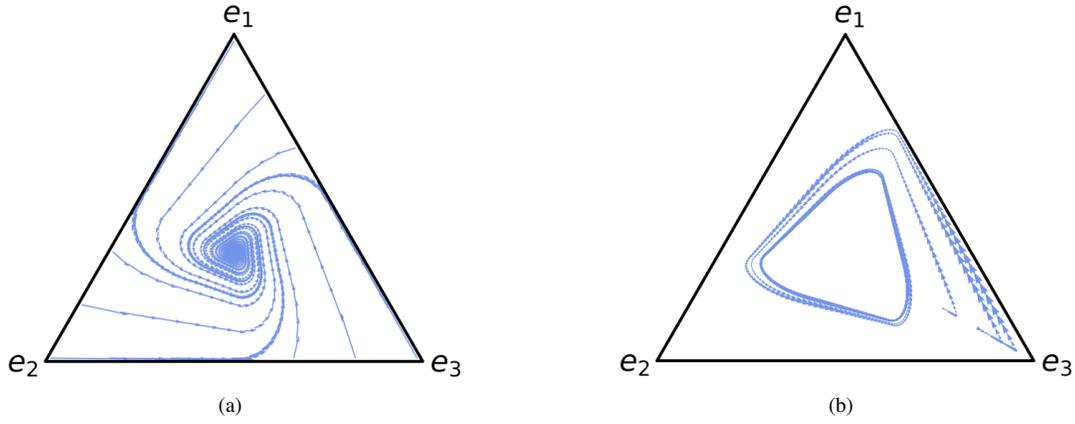


Fig. 5. Mean social state trajectories derived by the mean dynamics in the biodiversity example (Example 1): (a) the static PDM case and (b) the smoothing PDM case. For the static PDM case, the trajectories converge to the Nash equilibrium of the RPS payoff function (2), whereas for the smoothing PDM case, the trajectories form a limit cycle away from the Nash equilibrium.

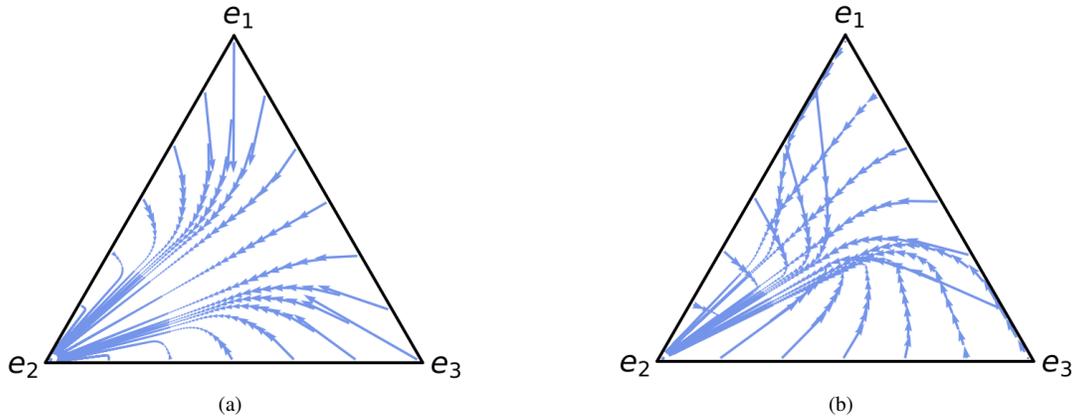


Fig. 6. Mean social state trajectories derived by the mean dynamics in the network traffic example (Example 2): (a) the static PDM case and (b) the smoothing PDM case. For both the cases, the trajectories converge to the Nash equilibrium of the congestion payoff function (3).

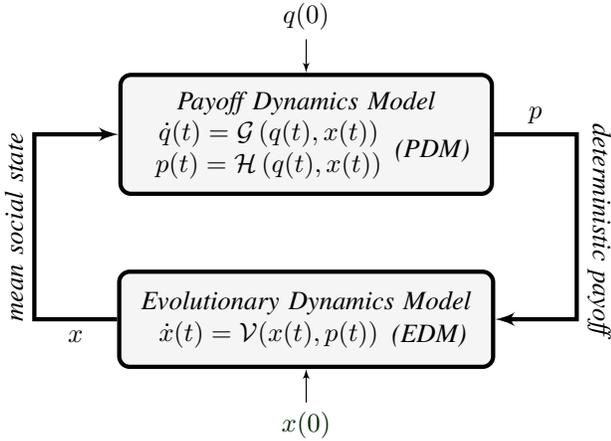


Fig. 7. A diagram representing the mean closed loop model defined by a feedback interconnection between a PDM and an EDM.

is such that the following limits hold:

$$\lim_{t \rightarrow \infty} \left( \inf_{z \in \mathbb{E}(\bar{\mathcal{F}})} \|x(t) - z\| \right) = 0 \quad (45a)$$

$$\lim_{t \rightarrow \infty} \|p^k(t) - \bar{\mathcal{F}}^k(x(t))\| = 0, \quad 1 \leq k \leq m \quad (45b)$$

- **(Lyapunov Stability)** The set  $\mathbb{E}(\bar{\mathcal{F}})$  is Lyapunov stable if for every open set  $\mathbb{O}$  in  $\mathbb{R}^n \times \mathbb{R}^s$  that contains

$$\mathbb{A} = \left\{ (q, z) \in \mathbb{R}^n \times \Delta \mid z \in \mathbb{E}(\bar{\mathcal{F}}) \text{ and } \mathcal{H}^k(q, z) = \bar{\mathcal{F}}^k(z), \quad 1 \leq k \leq m \right\},$$

there is another open set  $\mathbb{O}'$  that contains  $\mathbb{A}$  for which the following holds:

$$(q(0), x(0)) \in \mathbb{O}' \cap (\mathbb{R}^n \times \Delta) \implies (q(t), x(t)) \in \mathbb{O}, \quad t \geq 0 \quad (46)$$

- **(Global Asymptotic Stability)** The set  $\mathbb{E}(\bar{\mathcal{F}})$  is globally asymptotically stable if it is globally attractive and Lyapunov stable.

### B. Passivity-based Approach for Stability Analysis

We proceed to describe the concepts that we use to characterize the stability of the mean closed loop model, including the so-called  $\delta$ -passivity originally proposed in [7]. We refer the reader to [10, Section IV-D] for a comparative review of other notions of passivity such as equilibrium independent passivity (EIP) [40], which is adopted in [41] to characterize the stability of multi-agent finite games when reinforcement

learning is used. Notably, as is discussed in [10, Section IV-D], the EDM classes considered in our framework do not satisfy the technical assumptions required to adopt the EIP.

1) *EDM  $\delta$ -Passivity and Informative  $\delta$ -Storage Functions:* Given an EDM with input  $p$  and state  $x$ , the following inequality is central for the definition of  $\delta$ -passivity [7]:

$$\begin{aligned} \mathcal{S}(x(t), p(t)) - \mathcal{S}(x(t_0), p(t_0)) &\leq \\ \int_{t_0}^t [\dot{x}^T(\tau)\dot{p}(\tau) - \eta \dot{x}^T(\tau)\dot{x}(\tau)] d\tau, \\ t \geq t_0, x(t_0) \in \Delta, p \in \mathfrak{P} \end{aligned} \quad (47)$$

where  $\eta$  and  $\mathcal{S} : \Delta \times \mathbb{R}^{\bar{s}} \rightarrow \mathbb{R}_+$  are a nonnegative real constant and a map, respectively, and  $\mathfrak{P}$  is the set of differentiable maps defined as

$$\mathfrak{P} := \{p \in \mathfrak{X}^{\bar{s}} \mid \|p\| < \infty \text{ and } \|\dot{p}\| < \infty\}.$$

**Definition 10. (EDM  $\delta$ -passivity)** [10] *Given an EDM,  $\delta$ -passivity concepts are given as follows:*

- *The EDM is said to be  $\delta$ -passive if there is a continuously differentiable  $\mathcal{S}$  for which (47) is satisfied with  $\eta = 0$ .*
- *If the EDM is  $\delta$ -passive, let  $\eta^*$  be the supremum of all  $\eta$  for which there is a continuously differentiable  $\mathcal{S}$  satisfying (47). If  $\eta^*$  is positive then the EDM is qualified as  $\delta$ -passive with surplus  $\eta^*$ .*

*The map  $\mathcal{S}$  is referred to as a  $\delta$ -storage function. The EDM is referred to as strictly output  $\delta$ -passive when it is  $\delta$ -passive with some positive surplus  $\eta^*$ .*

Notice that the larger  $\eta^*$  the more stringent the requirement for strict output  $\delta$ -passivity. When it is positive, we view such  $\eta^*$  as a measure of  $\delta$ -passivity ‘‘surplus’’.

The concept of  $\delta$ -passivity stated in Definition 10 is closely related with the standard definition of passivity in feedback control theory. As is discussed in [7], an EDM is  $\delta$ -passive when the following augmented system with input  $\tilde{p}$  and output  $\tilde{x}$ , defined as  $\tilde{x} = (\tilde{x}^1, \dots, \tilde{x}^m)$ , is passive according to its standard definition [42]:

$$\dot{p}(t) = \tilde{p}(t), \quad p(0) \in \mathbb{R}^{\bar{s}}, \tilde{p} \in \tilde{\mathfrak{P}} \quad (48a)$$

$$\dot{x}^k(t) = \mathcal{V}^k(x^k(t), p^k(t)), \quad x^k(0) \in \Delta^k, 1 \leq k \leq m \quad (48b)$$

$$\tilde{x}^k(t) = \mathcal{V}^k(x^k(t), p^k(t)) \quad (48c)$$

where  $\tilde{\mathfrak{P}} := \{\tilde{p} \mid p \in \mathfrak{P}\}$ . Notably,  $(x(t), p(t))$  is the state of the augmented system and  $\mathcal{S}$  is a storage function for it.

Indeed the  $\delta$ -storage function  $\mathcal{S}$  of a  $\delta$ -passive EDM is used as a Lyapunov-like function in establishing stability for the mean closed loop model, in which the following definition of informativeness of the  $\delta$ -storage function plays a critical role.

**Definition 11. (Informative  $\mathcal{S}$ )** [10] *Let  $\mathcal{S} : \Delta \times \mathbb{R}^{\bar{s}} \rightarrow \mathbb{R}_+$  be a  $\delta$ -storage function for a given EDM specified by  $\mathcal{V}$ . We say that  $\mathcal{S}$  is informative if it satisfies the following two conditions:*

$$\mathcal{V}^k(z^k, r^k) = 0, 1 \leq k \leq m \implies \mathcal{S}(z, r) = 0 \quad (49a)$$

$$\begin{aligned} \nabla_{z^k}^T \mathcal{S}(z, r) \mathcal{V}^k(z^k, r^k) &= 0, 1 \leq k \leq m \\ \implies \mathcal{V}^k(z^k, r^k) &= 0, 1 \leq k \leq m \end{aligned} \quad (49b)$$

for every  $z$  and  $r$  in  $\Delta$  and  $\mathbb{R}^{\bar{s}}$ , respectively.

The implication in (49a) suggests that every equilibrium point of the EDM minimizes  $\mathcal{S}$ . In addition, we could conclude from (49b) and

$$\frac{d}{dt} \mathcal{S}(x(t), r) = \sum_{k=1}^m \nabla_{x^k}^T \mathcal{S}(x(t), r) \mathcal{V}^k(x^k(t), r^k)$$

that, for a constant deterministic payoff  $r$ ,  $\mathcal{S}(x(t), r)$  is constant only if  $x(t)$  remains at an equilibrium point of the EDM.

According to the definitions of the EDM classes described in §II-C.1, the EPT, IPC, and PBR EDMs are  $\delta$ -passive [7] and have informative  $\delta$ -storage functions [10], which we discuss in detail below. On the other hand, [11, Proposition III.5] shows that the imitative EDM is not  $\delta$ -passive.

- **$\delta$ -Passivity of EPT EDM:** If a given EPT EDM is integrable with revision potential  $\mathcal{I}^{\text{EPT}}$  then it is  $\delta$ -passive and there is a constant  $\gamma$  for which the following map  $\mathcal{S}^{\text{EPT}} : \Delta \times \mathbb{R}^{\bar{s}} \rightarrow \mathbb{R}_+$  is an informative  $\delta$ -storage function:

$$\mathcal{S}^{\text{EPT}}(z, r) = \mathcal{I}^{\text{EPT}}(\hat{r}) - \gamma, \quad z \in \Delta, r \in \mathbb{R}^{\bar{s}} \quad (50)$$

- **$\delta$ -Passivity of IPC EDM:** Let a Lipschitz continuous map  $\mathcal{T}_j^{\text{IPC}, k} : \mathbb{R} \rightarrow \mathbb{R}_+$  be the protocol of an IPC EDM. The IPC EDM is  $\delta$ -passive and the following map  $\mathcal{S}^{\text{IPC}} : \Delta \times \mathbb{R}^{\bar{s}} \rightarrow \mathbb{R}_+$  is an informative  $\delta$ -storage function:

$$\begin{aligned} \mathcal{S}^{\text{IPC}}(z, r) &= \sum_{k=1}^m \sum_{i=1}^{\bar{s}^k} \sum_{j=1}^{\bar{s}^k} z_i^k \int_0^{r_j^k - r_i^k} \mathcal{T}_j^{\text{IPC}, k}(\tau) d\tau, \\ z \in \Delta, r \in \mathbb{R}^{\bar{s}} \end{aligned} \quad (51)$$

- **$\delta$ -Passivity of PBR EDM:** Consider that an admissible deterministic perturbation  $\mathcal{Q}$  is given for a PBR EDM for which its  $\delta$ -storage function is defined as:

$$\begin{aligned} \mathcal{S}^{\text{PBR}}(z, r) &= \max_{\tilde{z} \in \text{int}(\Delta)} (\tilde{z}^T r - \mathcal{Q}(\tilde{z})) - (z^T r - \mathcal{Q}(z)), \\ z \in \Delta, r \in \mathbb{R}^{\bar{s}} \end{aligned} \quad (52)$$

Let  $\eta^*$  be the supremum of all nonnegative constants  $\eta$  for which the following holds:

$$\tilde{z}^T \nabla^2 \mathcal{Q}(z) \tilde{z} \geq \eta \tilde{z}^T \tilde{z}, \quad z \in \text{int}(\Delta), \tilde{z} \in \mathbb{T}\Delta \quad (53)$$

One of the two cases holds:

- **(Case I)** If  $\eta^* \geq 0$  then the PBR EDM is  $\delta$ -passive and  $\mathcal{S}^{\text{PBR}}$  is an informative  $\delta$ -storage function.
- **(Case II)** If  $\eta^* > 0$  then the PBR EDM is  $\delta$ -passive with surplus  $\eta^*$  and  $\mathcal{S}^{\text{PBR}}$  is an informative  $\delta$ -storage function.

Case II for  $\delta$ -passivity of the PBR EDM states that the EDM is strictly output  $\delta$ -passive when the perturbation  $\mathcal{Q}$  is strongly convex. We remark that according to [11, Corollary IV.3], the EPT and IPC EDMs cannot be strictly output  $\delta$ -passive.

2) *PDM  $\delta$ -antipassivity and weak  $\delta$ -antipassivity*: Similar to  $\delta$ -passivity for EDMs, in [10], a notion of  $\delta$ -antipassivity is defined for PDMs using the following conditions:

$$\begin{aligned} \mathcal{L}(z, s) &= 0 \\ \iff \mathcal{H}^k(s, z) &= \bar{\mathcal{F}}^k(z), \quad 1 \leq k \leq m, z \in \Delta, s \in \mathbb{R}^n \end{aligned} \quad (54a)$$

$$\begin{aligned} \mathcal{L}(x(t_0), q(t_0)) - \mathcal{L}(x(t), q(t)) &\geq \\ \int_{t_0}^t [\dot{p}^T(\tau)\dot{x}(\tau) - \nu\dot{x}^T(\tau)\dot{x}(\tau)] \, d\tau, \\ t \geq t_0, q(t_0) \in \mathbb{R}^n, x \in \mathfrak{X} \end{aligned} \quad (54b)$$

where  $\nu$  is a nonnegative constant,  $\bar{\mathcal{F}}$  is the stationary population game of the PDM,  $\mathcal{L} : \Delta \times \mathbb{R}^n \rightarrow \mathbb{R}_+$  is a map, and  $\mathfrak{X}$  is the set of differentiable maps defined as

$$\mathfrak{X} := \{x \in \mathfrak{R}^s \mid \|\dot{x}\| < \infty \text{ and } x(t) \in \Delta, t \geq 0\}$$

**Definition 12. (PDM  $\delta$ -antipassivity)** [10] *Given a PDM, consider the following cases:*

- The PDM is said to be  *$\delta$ -antipassive* if there is a continuously differentiable  $\bar{\mathcal{L}}$  for which (54) is satisfied for  $\nu = 0$ .
- The PDM is said to be  *$\delta$ -antipassive with deficit  $\nu^* > 0$*  if there is a continuously differentiable  $\bar{\mathcal{L}}$  for which (54) is satisfied for every  $\nu > \nu^*$ .

A map  $\mathcal{L}$  that satisfies either case is referred to as a  $\delta$ -antistorage function.

Notice that there is an antisymmetry between (47) and (54b) that is obtained from changing signs of certain terms and swapping the output with the input. This correspondence could be further strengthened by viewing  $\mathcal{L}$  as a  $\delta$ -antistorage function that would be the antisymmetric equivalent of  $\mathcal{S}$ . An analogy similar to (48) is done in [7] to compare  $\delta$ -antipassivity with the standard notion of passivity.

Given a PDM with input  $x$ , state  $q$ , and output  $p$ , the following inequality is central to characterizing a weaker notion of  $\delta$ -antipassivity:

$$\begin{aligned} \mathcal{A}(q(t_0), \|\dot{x}\|) &\geq \\ \int_{t_0}^t [\dot{p}^T(\tau)\dot{x}(\tau) - \nu\dot{x}^T(\tau)\dot{x}(\tau)] \, d\tau, \\ t \geq t_0, q(t_0) \in \mathbb{R}^n, x \in \mathfrak{X} \end{aligned} \quad (55)$$

where  $\nu$  and  $\mathcal{A} : \mathbb{R}^n \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$  are a nonnegative real constant and a map, respectively.

**Definition 13. (PDM Weak  $\delta$ -antipassivity)** [10] *Given a PDM, consider the following cases:*

- The PDM is said to be *weak  $\delta$ -antipassive* if there is  $\mathcal{A}$  for which (55) is satisfied for  $\nu = 0$ .
- The PDM is said to be *weak  $\delta$ -antipassive with deficit  $\nu^* > 0$*  if there is  $\mathcal{A}$  for which (55) is satisfied for every  $\nu > \nu^*$ .

Unlike  $\delta$ -antipassivity, which requires the existence of a continuously differentiable  $\delta$ -antistorage function  $\mathcal{L}$ , weak  $\delta$ -antipassivity only requires the existence of a map  $\mathcal{A}$  satisfying (55). Although well-known results [43]–[45] indicate that a so-called “available” storage function can be constructed when (55) is satisfied, there are no guarantees that it will be continuously differentiable or satisfy (54a). The following remark outlines an argument to establish that  $\delta$ -antipassivity indeed implies weak  $\delta$ -antipassivity.

**Remark 5. ( $\delta$ -antipassivity implies weak  $\delta$ -antipassivity)** *Given a PDM, the following holds:*

- If the PDM is  $\delta$ -antipassive then it is weak  $\delta$ -antipassive and we can select  $\mathcal{A}$  as:

$$\mathcal{A}(q(0), \cdot) = \max_{z \in \Delta} \mathcal{L}(z, q(0)) \quad (56)$$

where  $\mathcal{L}$  is the  $\delta$ -antistorage function of the PDM.

- This choice for  $\mathcal{A}$  also allows us to conclude that if the PDM is  $\delta$ -antipassive with deficit  $\nu^*$  then it is also weak  $\delta$ -antipassive with deficit  $\nu^*$ .

As we will observe in §VI-C where we discuss convergence of the mean social state to  $\mathbb{N}\mathbb{E}(\bar{\mathcal{F}})$  or  $\mathbb{P}\mathbb{E}(\bar{\mathcal{F}}, \mathcal{Q})$ ,  $\delta$ -antipassivity and weak  $\delta$ -antipassivity are requirements for global asymptotic stability and global attractiveness, respectively. This confirms, as one should expect, that because the  $\delta$ -antipassivity condition is stricter than its weak version, it leads to stronger stability guarantees.

Based on the definition, we can examine  $\delta$ -antipassivity of certain classes of smoothing-anticipatory PDMs:

- According to Remark 3, population games are special cases of smoothing-anticipatory PDMs with the parameter selection of  $\mu_I = 1$  and  $\mu_S = \mu_A = 0$ . In [7], it is shown that contractive population games are  $\delta$ -antipassive.
- Other important cases of smoothing-anticipatory PDMs are defined by the following affine game: Let  $\mathcal{F}$  be an affine game specified by

$$\mathcal{F}(z) = Fz + \bar{r}, \quad z \in \Delta \quad (57)$$

where  $F \in \mathbb{R}^{s \times s}$  is such that  $\Phi F \Phi$  is symmetric, and  $\bar{r}$  is a constant vector in  $\mathbb{R}^s$ .

Consider that  $\mathcal{F}$ , a given positive  $\alpha$ , and non-negative  $\mu_I$ ,  $\mu_S$ , and  $\mu_A$  satisfying  $\mu_I + \mu_S = 1$  define a smoothing-anticipatory PDM. Let  $\bar{\lambda}$  be the largest eigenvalue of  $\Phi F \Phi$ . Then, according to [10, Proposition 7], the PDM satisfies the following:

- If  $\bar{\lambda} \leq 0$  then the PDM is weak  $\delta$ -antipassive.
- If  $\bar{\lambda} > 0$  and  $\mu_I + \alpha\mu_A \leq 1$  then the PDM is weak  $\delta$ -antipassive with deficit  $\bar{\lambda}$ .
- If  $\bar{\lambda} > 0$  and  $\mu_I + \alpha\mu_A > 1$  then the PDM is weak  $\delta$ -antipassive with deficit  $(\mu_I + \alpha\mu_A)\bar{\lambda}$ .

If, in addition to the conditions stated above, a PDM is specified by a symmetric and negative definite  $F$  then it is  $\delta$ -antipassive.

### C. Stability in Mean Closed Loop of an EDM and a PDM

We begin by stating the following main supporting lemma.

**Lemma 1. (Main supporting lemma)** [10] *Let a PDM and an EDM be given. We consider the following two cases.*

- **(Case I)** *The PDM is weak  $\delta$ -antipassive ( $\nu^* = 0$ ) and the EDM is  $\delta$ -passive ( $\eta^* = 0$ ) with respect to an informative  $\delta$ -storage function  $\mathcal{S}$ .*
- **(Case II)** *The PDM is weak  $\delta$ -antipassive with deficit  $\nu^* > 0$  and the EDM is  $\delta$ -passive with surplus  $\eta^* > \nu^*$  with respect to an informative  $\delta$ -storage function  $\mathcal{S}$ .*

*If either Case I or Case II is true then the following holds:*

$$\lim_{t \rightarrow \infty} \mathcal{S}(x(t), p(t)) = 0, \quad (x(0), q(0)) \in \Delta \times \mathbb{R}^n \quad (58)$$

*where the trajectory  $(x(t), p(t))$ ,  $t \geq 0$  is determined from the unique solution of initial value problem for (44).*

Notice that the lemma (Case II) asserts that if the PDM has a  $\delta$ -antipassivity deficit then (58) still holds provided that the EDM can make up for the deficit by having a large enough  $\delta$ -passivity surplus.

As is evident from the analysis in [10], showing that (58) holds is an important step in characterizing the stability of the mean closed loop model for several cases of interest. In fact, the way  $\delta$ -storage functions are employed in [10] is analogous to how Lyapunov functions were used to obtain the results in §III-B. In particular, when the  $\delta$ -storage function  $\mathcal{S}$  is informative, (58) implies that the mean social state  $x(t)$  converges to the set  $\{z \in \Delta \mid \mathcal{V}^k(z^k, \bar{\mathcal{F}}^k(z)) = 0, 1 \leq k \leq m\}$ , where  $\mathcal{V}$  is the EDM and  $\bar{\mathcal{F}}$  is the stationary game of the PDM. This fact, in turn, can be used to establish convergence of  $x(t)$  to  $\text{NE}(\bar{\mathcal{F}})$  when the EDM satisfies Nash stationarity.

In what follows, based on Lemma 1, we specialize the stability results for EPT, IPC, and PBR EDMs.

1) *Stability for EPT and IPC EDMs:* Recall that each PDM is associated with a stationary game  $\bar{\mathcal{F}}$ . Based on the fact that EPT and IPC EDMs satisfy Nash stationarity, the following theorem gives sufficient conditions for stability of the set of Nash equilibria of  $\bar{\mathcal{F}}$ .

**Theorem 4.** [10] *Consider a mean closed loop model formed by either an integrable EPT EDM or an IPC EDM and a PDM. If the PDM is weak  $\delta$ -antipassive then  $\text{NE}(\bar{\mathcal{F}})$  is globally attractive. If the PDM is  $\delta$ -antipassive then  $\text{NE}(\bar{\mathcal{F}})$  is globally asymptotically stable.*

We note that the theorem is stated only for  $\delta$ -antipassive PDMs. As we discussed in §VI-B.1, both integrable EPT and IPC EDMs are not strictly output  $\delta$ -passive and Case II of Lemma 1 does not apply here.

It is also important to notice that Theorem 4 combines [10, Theorem 1] and [10, Theorem 2], which are extensions of Theorems 1 and 2, respectively. In particular, Theorem 1 can be viewed as [10, Theorem 1] restricted to population games, which, as pointed out in Remark 2, gives stronger stability guarantees than [28, Theorem 5.1].

It is useful to observe that the notion of strict output  $\delta$ -passivity can be used to further specify a class of PDMs that attain stability under PBR EDMs but does not under EPT and IPC EDMs. The Hypnodisk game is one of such PDMs as explained in [12, Proposition IV.2].

2) *Stability for PBR EDM:* The following theorem from [10] gives sufficient conditions that characterize the stability of the set of perturbed Nash equilibria of  $\bar{\mathcal{F}}$  for a mean closed loop model consisting of a  $\delta$ -passive PBR EDM and a  $\delta$ -antipassive PDM.

**Theorem 5.** [10] *Consider a mean closed loop model formed by a  $\delta$ -passive PBR EDM characterized by an admissible payoff perturbation  $\mathcal{Q}$  and a PDM with a continuously differentiable  $\bar{\mathcal{F}}$ . One of the following two cases holds:*

- **(Case I)** *If the PDM is weak  $\delta$ -antipassive then  $\text{PE}(\bar{\mathcal{F}}, \mathcal{Q})$  is globally attractive. If the PDM is  $\delta$ -antipassive then  $\text{PE}(\bar{\mathcal{F}}, \mathcal{Q})$  is globally asymptotically stable.*
- **(Case II)** *If the PDM is weak  $\delta$ -antipassive with positive deficit  $\nu^*$  and the PBR EDM is  $\delta$ -passive with surplus  $\eta^* > \nu^*$  then  $\text{PE}(\bar{\mathcal{F}}, \mathcal{Q})$  is globally attractive. If the PDM is  $\delta$ -antipassive with positive deficit  $\nu^*$  and the PBR EDM is  $\delta$ -passive with surplus  $\eta^* > \nu^*$  then  $\text{PE}(\bar{\mathcal{F}}, \mathcal{Q})$  is globally asymptotically stable.*

Theorem 5 can be viewed as an extension of Theorem 3. Table II summarizes the stability results given both in §III-B and §VI-C.

3) *Application of Stability Results:* To explain application of the stability results, in particular Theorem 4, we revisit the numerical examples presented in §V-B. Other examples related with Theorem 5 can be found in [10].

According to the definition of  $\delta$ -antipassivity (Definition 12), in both the examples, the static PDM (42) is  $\delta$ -antipassive for which the mean closed loop model resulting from a feedback interconnection of the static PDM and the BNN EDM satisfies the conditions stated in Theorem 4. Figs. 5(a) and 6(a) depict the mean social state trajectories derived by the mean closed loop models from the examples. As claimed in Theorem 4, the trajectories converge to the Nash equilibrium of the respective stationary population games.

On the other hand, the smoothing PDM (43) defined by the congestion payoff function (3) is  $\delta$ -antipassive, whereas the smoothing PDM defined by the RPS payoff function (2) is not. In the latter case, Theorem 4 cannot be applied to the mean closed loop model consisting of the smoothing PDM and the BNN EDM. Figs. 5(b) and 6(b) illustrate the mean social state trajectories derived by the mean closed loop models. One remarkable observation is that the reinforcement learning dynamics used to define (43) prevents the convergence of the trajectories to the Nash equilibrium in the RPS game (2), and different from the static PDM case depicted in Fig. 5(a), the trajectories form a limit cycle away from the Nash equilibrium.

	contractive game	weak $\delta$ -antipassive PDM	weak $\delta$ -antipassive PDM with deficit $\eta^* > 0$	$\delta$ -antipassive PDM	$\delta$ -antipassive PDM with deficit $\eta^* > 0$
Integrable EPT EDM	asymptotically stable (Theorems 1 and 4)	attractive (Theorem 4)	not attractive in Hypnodisk game	asymptotically stable (Theorem 4)	not attractive in Hypnodisk game
IPC EDM	asymptotically stable (Theorems 2 and 4)	attractive (Theorem 4)	not attractive in Hypnodisk game	asymptotically stable (Theorem 4)	not attractive in Hypnodisk game
PBR EDM with surplus $\nu^* > 0$	asymptotically stable (Theorems 3 and 5)	attractive (Theorem 5)	attractive if $\nu^* > \eta^*$ (Theorem 5)	asymptotically stable (Theorem 5)	asymptotically stable if $\nu^* > \eta^*$ (Theorem 5)

TABLE II

A TABLE COMPARING THE STABILITY RESULTS FOR CONTRACTIVE GAMES AND  $\delta$ -ANTIPASSIVE PDMs

## VII. BRIEF COMPARISON WITH FINITE GAMES, AGGREGATIVE GAMES, AND MEAN FIELD GAMES

Apart from population games that we introduced in §II, in the literature of game theory, there are other important families of multi-agent games including finite games, aggregative games, and mean field games (MFG). The finite games [41] investigate mean deterministic dynamics of underlying learning rules adopted by a finite number of agents and the aggregative games [46] focus on how aggregation in games affects the existence of a pure Nash equilibrium and convergence to such an equilibrium under best-reply dynamics. On the other hand, similar to the population game framework, MFGs investigate finding control policies for a large population of agents in decision making problems. In this section, we review some of relevant works in the literature and compare how the population game framework (and hence its generalization, the PDM framework) is related to these works.

### A. Finite Games and Aggregative Games

Recent works [41], [47] propose a reinforcement learning framework in games with a finite number of agents and investigate conditions under which the learning rules adopted by the agents guarantee convergence of the strategy profile of the agents to a Nash equilibrium of a static payoff function. In particular, similar to the approach described in §V for the PDM framework, a passivity-based approach was employed to analyze stability of a Nash equilibrium.

In the aggregative game framework [46], payoffs each agent receive depend on its own individual strategy selection and an aggregate, such as average, of strategy selections of all other agents. Under such payoff mechanism, the aggregative game study investigates how the aggregation is relevant to the existence of a pure Nash equilibrium and seeks conditions under which best-reply strategy revision protocols ensure the convergence of the strategy profile to a Nash equilibrium.

The finite game and aggregative game studies are centered at analyzing stability in the strategic interaction among a finite number of agents whose strategy profile evolves based on strategy revision protocols that they adopt. Similar to the population game framework, the profile describes

strategies taken by agents and deterministic dynamic models are defined to describe how the profile evolves in response to payoffs assigned by respective payoff mechanisms.

Differences are that the population game framework considers that each agent in a population is elected to revise its strategy according to its own Poisson clock (independent of other's), whereas the finite game and aggregative game frameworks assume that the agents continuously or sequentially take turns to revise their strategy selections. Also, in terms of their mathematical models, as described in §IV-A, the mean dynamic models used in the population game framework approximate strategy revision dynamics of large populations. Whereas the finite game and aggregative game frameworks assume that underlying deterministic dynamic models approximate stochastic learning processes of individual agents based on a stochastic approximation principle [48]. For this reason, these two frameworks are more suitable to analyze stability of individual-level strategy revision processes.

### B. Mean Field Games

The description in [49, chapter 2] suggests that the class of mean field games (MFG) that is most similar to the population game framework is that of first order MFG. Starting with the pioneering work in [50]–[52], most work on first order mean field games in an engineering context seeks to control a collective of structurally dissociated dynamical systems (agents) in the limit when the number of agents tends to infinity. Each agent in such a setting is steered by a causal control policy that depends on its own state and the so-called (deterministic) mean state, which is analogous to the mean social state adopted in our approach in that both specify at every instant the portions of the population that attain every possible state or set of states. The control law is the same across an entire class (subgroup of identical agents), which is analogous to imposing the same protocol for all members of each population in the multi-population setting. In light of these similarities it is important that we list the following fundamental differences between our approach and that of first order MFG.

1) *Main differences between MFG and our approach:* In the first order MFG, the set of strategies available to each

agent class comprises control policies that govern the control action as a causal function of the mean state and the state of each agent. The so-called Nash Certainty Equivalence principle determines the control policy for each agent class from the solution of a Hamilton-Jacobi-Bellman equation. We proceed to describe the most significant differences between our approach and that of MFG.

a) *States and strategies*: In our approach, the state of each agent takes values in the strategy set. As a consequence, the strategy (or state) of each agent is revised repeatedly, which is in sharp contrast with the MFG approach in which the control policy of each agent class that is computed once off-line and used throughout the life of the MFG is a strategy.

b) *Cost and payoff mechanism*: While a cost is used to compute the control policies for each agent class of a MFG, in our approach a payoff mechanism is itself the feedback component providing the payoff vector used by the agents to revise their strategies. In the particular case in which the payoff mechanism is a population game and the revision protocols are of the *best response* type, our approach is somewhat *similar* to a repeated game [53] in which the stage game would be the population game. In this analogy, we emphasize the qualifier “similar” because unlike a standard repeated game, the techniques, models, concepts and objectives of our approach are tailored to deal with large populations.

c) *Equilibrium concepts*: Characterization of relevant mean social state equilibria for our approach is inspired by Nash’s and defined with respect to the payoff mechanism. When the payoff mechanism is a population game, the set of Nash equilibria, or perturbed versions of it, is functionally characterized in the usual way, and may be interpreted in the mass-action sense proposed in [54] and further analyzed in [55]. In contrast, the so-called decentralized  $\epsilon$ -Nash concept [50], which can be viewed as a relaxation of the person-to-person optimality notion [56] applied to the underlying cost, characterizes the relevant equilibrium set for a MFG.

d) *Convergence to equilibria*: In contrast to MFG, the main goal of our framework is to establish concepts and methods to determine whether the populations will react in a way that the mean social state converges towards equilibria of interest. Convergence to an equilibria set, such as Nash’s, gives credence to both the ability of the populations to self-organize using simple (bounded-rationality) protocols in a distributed fashion and the value of the equilibria concept to predict long-term behavior.

2) *Benefits of our approach*: The following are reasons to consider when deciding on whether to adopt our approach:

a) *Simplicity of protocols*: In our formulation, the protocols that the agents use to revise their strategies are simple to implement and amenable to analysis supported on conventional evolutionary principles. In addition, the agents need not be synchronized because independent Poisson clocks determine when each agent is allowed to revise its strategy.

b) *Information structure*: As is discussed in [57], the information structures required to implement the various types of revision protocols vary significantly. In general, the

strategy revision mechanism executed by each agent does not require access to the entire social state. Instead, it typically relies on either a portion or a specific function of the social state or payoff vector. As illustrated by imitative protocols that capture the tendency of an agent to follow a widely-adopted strategy, our framework can effectively model local interactions.

## VIII. FUTURE DIRECTIONS

An essential next step is to harness the potential of our framework and stability results for the design and analysis of engineered multi-agent systems.

The fact that our approach can tackle payoff mechanisms of the PDM type suggests that it can be used to generalize existing methods, like the ones employed in the applications listed in §IV-B, to account for the attributes expounded in §V-A, to name a few. Another worthy future direction is the investigation of ways to enhance the stability properties of algorithms based on replicator dynamics by modifying them to incorporate the more general classes of protocols analyzed here.

Compositional approaches, such as the one discussed in [58, Chapter 2], in which the stability of a networked system can be established by ascertaining certain passivity properties of each component separately have contributed to tractable analysis and design methods. Likewise, Theorems 4 and 5 involve separate conditions for the EDM and PDM that jointly are sufficient for the various stability guarantees provided (see also Table II). As a consequence of our results, designing the protocols, which may be randomized, and the PDM independently to guarantee the stability of a desirable subset of equilibria or cope with network limitations, information constraints, modelled and unmodelled learning dynamics, or other detrimental effects may be another viable future direction.

**Acknowledgment:** The authors would like to thank Semih Kara (UMD) for suggesting corrections and improvements to this article.

## REFERENCES

- [1] W. H. Sandholm, *Population games and evolutionary dynamics*. MIT Press, 2010.
- [2] L. E. Blume, *The economy as an evolving complex system II*, ch. Population games. CRC press, 2018.
- [3] D. T. Bishop and C. Cannings, “A generalized war of attrition,” *Journal of theoretical biology*, vol. 70, pp. 85–124, 1978.
- [4] Y. Sheffi, *Urban transportation networks*. Prentice Hall, 1985.
- [5] M. Beckmann, C. B. McGuire, and C. C. Koopmans, *Studies in the economics of transportation*. Research in economics at Yale University, Cowles Foundation, January 1956.
- [6] J. R. Marden and J. S. Shamma, *Handbook of Game Theory*, ch. Game Theory and Distributed Control. Elsevier Science, 2014.
- [7] M. J. Fox and J. S. Shamma, “Population games, stable games, and passivity,” *Games*, vol. 4, pp. 561–583, 2013.
- [8] G. Arslan and J. S. Shamma, “Anticipatory learning in general evolutionary games,” in *Proceedings of the IEEE conference on decision and control (CDC)*, pp. 6289–6294, 2006.
- [9] G. C. Chasparis and J. S. Shamma, “Distributed dynamic reinforcement of efficient outcomes in multiagent coordination and network formation,” *Dynamic games and applications*, vol. 2, no. 1, pp. 18–50, 2012.

- [10] S. Park, N. C. Martins, and J. S. Shamma, "Payoff dynamics model and evolutionary dynamics model: Feedback and convergence to equilibria (arxiv:1903.02018)." arXiv.org.
- [11] S. Park, J. S. Shamma, and N. C. Martins, "Passivity and evolutionary game dynamics," in *Proceedings of the IEEE conference on decision and control (CDC)*, 2018.
- [12] S. Park, J. S. Shamma, and N. C. Martins, "Passivity and evolutionary game dynamics (arxiv:1803.07744 )." arXiv.org, March 2018.
- [13] W. H. Sandholm, "Evolution and equilibrium under inexact information," *Games and economic behavior*, vol. 44, pp. 343–378, August 2003.
- [14] B. Kerr, M. A. Riley, M. W. Feldman, and B. J. M. Bohannan, "Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors," *Nature*, vol. 418, p. 171, 2002.
- [15] E. David and K. Jon, *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. New York, NY, USA: Cambridge University Press, 2010.
- [16] J. W. Weibull, *Evolutionary game theory*. MIT Press, 1995.
- [17] P. Schuster and K. Sigmund, "Replicator dynamics," *Journal of theoretical biology*, vol. 100, pp. 533–538, Feb. 1983.
- [18] W. H. Sandholm, "Excess payoff dynamics and other well-behaved evolutionary dynamics," *Journal of economic theory*, vol. 124, no. 2, pp. 149–170, 2005.
- [19] J. Hofbauer, "From Nash and Brown to Maynard Smith: equilibria, dynamics and ess," *Selection*, vol. 1, no. 81–88, 2000.
- [20] G. W. Brown and J. von Neumann, "Solutions of games by differential equations," tech. rep., RAND corporation, April 1950.
- [21] M. J. Smith, "The stability of a dynamic model of traffic assignment: an application of a method of Lyapunov," *Transportation science*, vol. 18, pp. 245–252, August 1984.
- [22] D. Fudenberg and D. M. Kreps, "Learning mixed equilibria," *Games and economic behavior*, vol. 5, pp. 320–367, July 1993.
- [23] J. Hofbauer and W. H. Sandholm, "Evolution in games with randomly disturbed payoffs," *Journal of economic theory*, vol. 132, no. 1, pp. 47–69, 2007.
- [24] J. Hofbauer and W. H. Sandholm, "On the global convergence of stochastic fictitious play," *Econometrica*, vol. 70, no. 6, pp. 2265–2294, 2002.
- [25] J. Maynard Smith and G. R. Price, "The logic of animal conflict," *Nature*, vol. 246, no. November, pp. 15–18, 1973.
- [26] P. D. Taylor and L. B. Jonker, "Evolutionarily stable strategies and game dynamics," *Mathematical Biosciences*, vol. 40, pp. 145–156, 1978.
- [27] T. Vincent and T. Vincent, "Evolution and control system design. The evolutionary game," *IEEE Control Systems*, vol. 20, pp. 20–35, 2000.
- [28] J. Hofbauer and W. H. Sandholm, "Stable games and their dynamics," *Journal of Economic Theory*, vol. 144, pp. 1665–1693.e4, July 2009.
- [29] J. Hofbauer and W. H. Sandholm, "Stable games," in *Proceedings of the IEEE conference on decision and control*, pp. 3416–3421, 2007.
- [30] W. H. Sandholm, "Local stability of strict equilibria under evolutionary game dynamics," *Journal of Dynamics and Games*, vol. 1, pp. 485–495, 2014.
- [31] M. Benaïm and J. Weibull, "Deterministic approximation of stochastic evolution in games," *Econometrica*, vol. 71, no. 3, pp. 873–903, 2003.
- [32] T. G. Kurtz, "Solutions of ordinary differential equations as limits of pure jump Markov processes," *Journal of applied probability*, vol. 7, pp. 49–58, April 1970.
- [33] J. G. Wardrop, "Some theoretical aspects of road traffic research," *Proceedings of the institution of civil engineers*, vol. 1, pp. 325–362, May 1952.
- [34] N. Quijano, C. Ocampo-Martinez, J. Barreiro-Gomez, G. Obando, A. Pantoja, and E. Mojica-Nava, "The role of population games and evolutionary dynamics in distributed control systems," *IEEE Control Systems Magazine*, vol. 37, no. 1, pp. 70–97, 2017.
- [35] W. H. Sandholm, "Potential games with continuous player sets," *Journal of economic theory*, vol. 97, pp. 81–108, 2001.
- [36] D. Monderer and L. S. Shapley, "Potential games," *Games and economic behavior*, vol. 14, pp. 124–143, 1996.
- [37] H. Tembine, E. Altman, R. El-Azouri, and Y. Hayel, "Evolutionary games in wireless networks," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 40, no. 3, pp. 634–646, 2010.
- [38] R. Laraki and P. Mertikopoulos, "Higher order game dynamics," *Journal of Economic Theory*, vol. 148, pp. 2666–2695, 2013.
- [39] P. Mertikopoulos and W. H. Sandholm, "Learning in games via reinforcement and regularization," *Math. Oper. Res.*, vol. 41, pp. 1297–1324, Nov. 2016.
- [40] G. H. Hines, M. Arcak, and A. K. Packard, "Equilibrium-independent passivity: A new definition and numerical certification," *Automatica*, vol. 47, pp. 1949–1956, 2011.
- [41] B. Gao and L. Pavel, "On passivity and reinforcement learning in finite games," in *2018 IEEE Conference on Decision and Control (CDC)*, pp. 340–345, Dec 2018.
- [42] R. Sepulchre, M. Janković, and P. Kokotović, *Constructive nonlinear control*. Springer, 1997.
- [43] P. J. Moylan, "Implications of passivity in a class of nonlinear systems," *IEEE Transactions on automatic control*, vol. 19, pp. 373–381, August 1974.
- [44] D. Hill and P. J. Moylan, "The stability of nonlinear dissipative systems," *IEEE transactions on automatic control*, vol. 21, pp. 708–711, October 1976.
- [45] J. C. Willems, "Dissipative dynamical systems part i: general theory," *Archive for rational mechanics and analysis*, vol. 45, pp. 321–351, January 1972.
- [46] M. K. Jensen, "Aggregative games and best-reply potentials," *Economic Theory*, vol. 43, pp. 45–66, 2010.
- [47] D. Gadjev and L. Pavel, "A passivity-based approach to Nash equilibrium seeking over networks," *IEEE Transactions on Automatic Control*, vol. 64, pp. 1077–1092, March 2019.
- [48] M. Benaïm, "Dynamics of stochastic approximation algorithms," in *Sminaire de Probabilits XXXIII. Lecture Notes in Mathematics*, vol. 1709, Springer, Berlin, Heidelberg.
- [49] D. A. Gomes, L. Nurbekyan, and E. A. Pimentel, "Economic models and mean-field games theory," in *30th Colóquio Brasileiro de matemática*, Publicações matemáticas, IMPA, 2015.
- [50] M. Huang, P. Caines, and R. Malhame, "Large-population cost-coupled LQG problems with nonuniform agents: individual-mass behavior and decentralized  $\epsilon$ -Nash equilibria," *IEEE Transactions on automatic control*, vol. 52, pp. 1560–1571, October 2007.
- [51] M. Huang, R. Malhame, and P. E. Caines, "The NCE (mean field) principle with locality dependent cost interactions," *IEEE transactions on automatic control*, vol. 55, no. 12, pp. 2799–2805, 2010.
- [52] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, pp. 229–260, March 2007.
- [53] D. Fudenberg and J. Tirole, *Game theory*. The MIT Press, 1991.
- [54] J. F. Nash Jr., "Non-cooperative games," *Annals of mathematics second series*, vol. 54, no. 2, pp. 286–295, 1951.
- [55] J. W. Weibull, "The mass-action interpretation of Nash equilibrium," No. 427 in Working paper series, 1995.
- [56] A. Mahajan, N. C. Martins, M. C. Rotkowitz, and S. Yüksel, "Information structures in optimal decentralized control," in *2012 IEEE 51st IEEE Conference on Decision and Control (CDC)*, pp. 1291–1306, Dec 2012.
- [57] W. H. Sandholm, "Pairwise comparison dynamics and evolutionary foundations for Nash equilibrium," *Games*, vol. 1, no. 1, pp. 3–17, 2010.
- [58] M. Arcak, C. Meissen, and A. Packard, "Networks of dissipative systems: compositional certification of stability, performance, and safety," in *Springer briefs in electrical and computer engineering : control, automation and robotics*, Springer, 2016.