DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

INSTITUTE FOR SYSTEMS RESEARCH
A. JAMES CLARK SCHOOL OF ENGINEERING

Counterclockwise dissipativity, potential games and evolutionary Nash equilibria learning

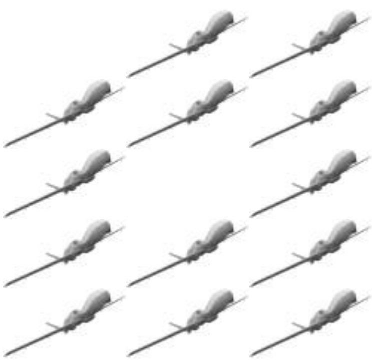
Nuno C. Martins
(*nmartins@umd.edu*)

Department of Electrical and Computer Engineering
Institute for Systems Research
University of Maryland at College Park

2024 AFOSR Dynamical Systems and Control Theory Review,
Arlington, VA, August 26, 2024.

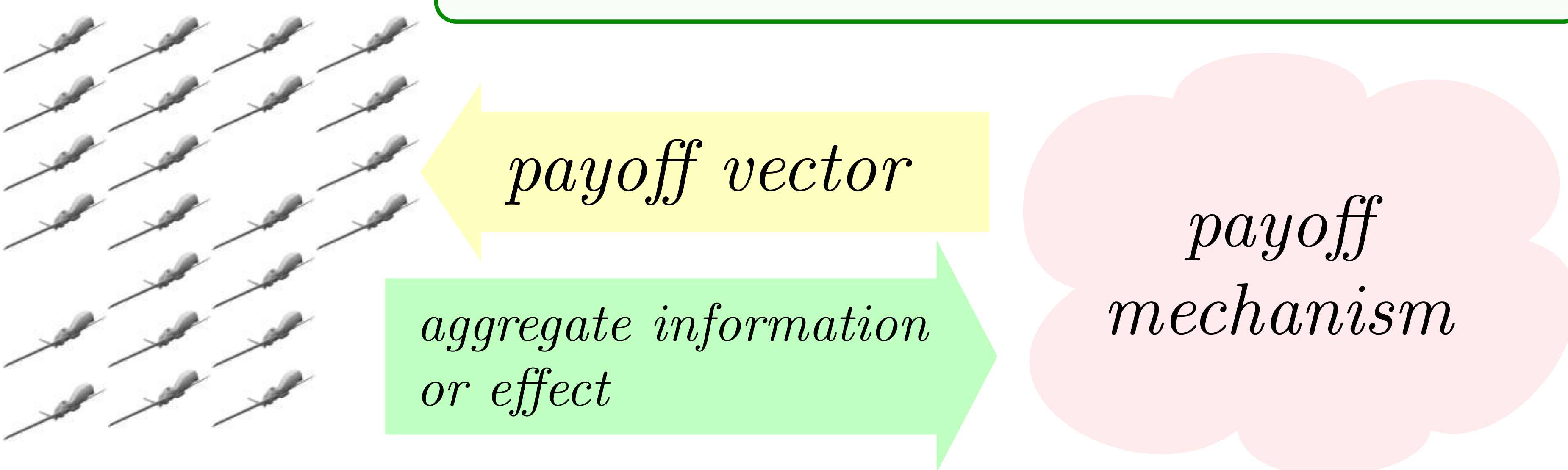
Population Games: Concept and Motivation

Framework for modeling noncooperative strategic interactions for large populations of agents.



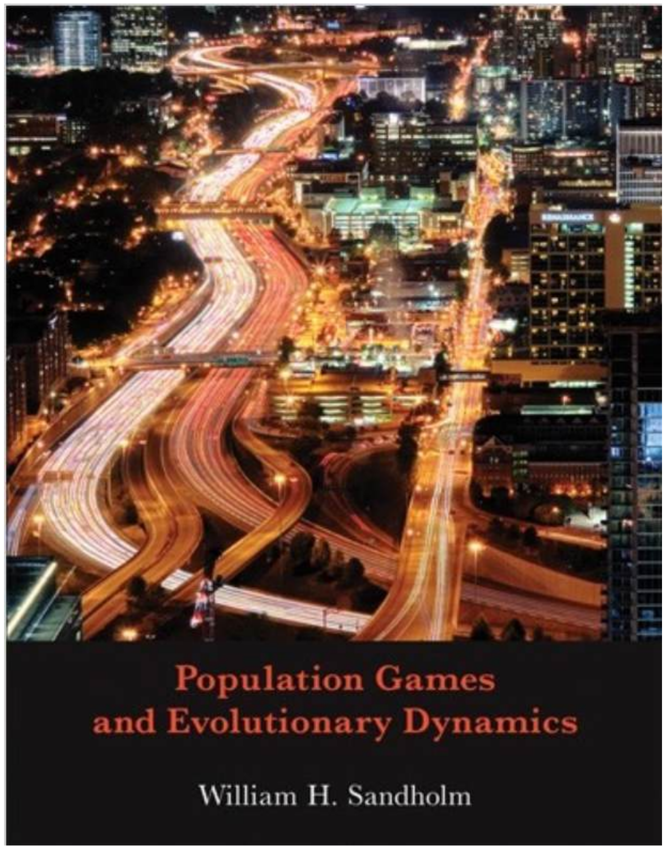
Basic tenets:

- Large number of nondescript agents grouped into populations.
- Agents select and repeatedly revise strategies.
- Each strategy has a payoff (reward or cost).
- Payoff mechanism ascribes vector of payoffs.

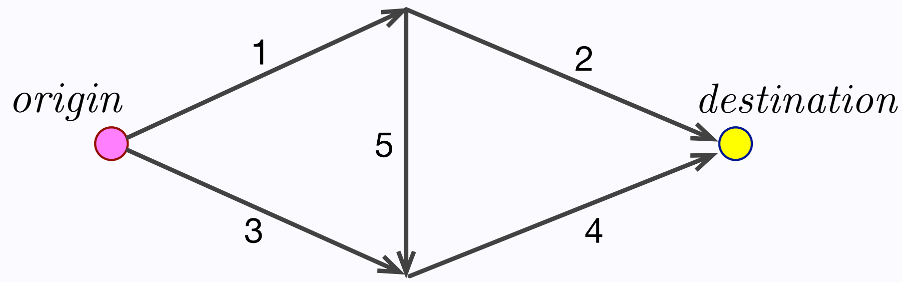


Population Games: Concept and Motivation

Textbook with many examples



Example: traffic assignment

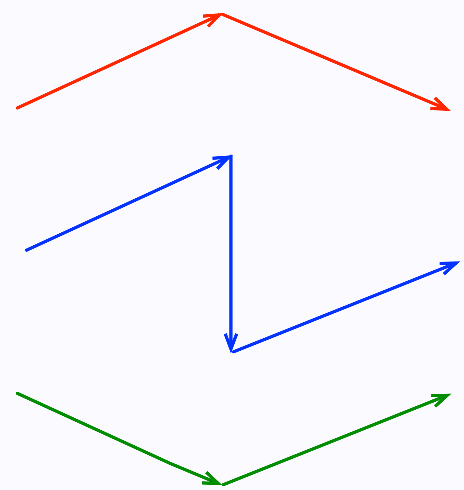


strategies available to each agent:

strategy 1

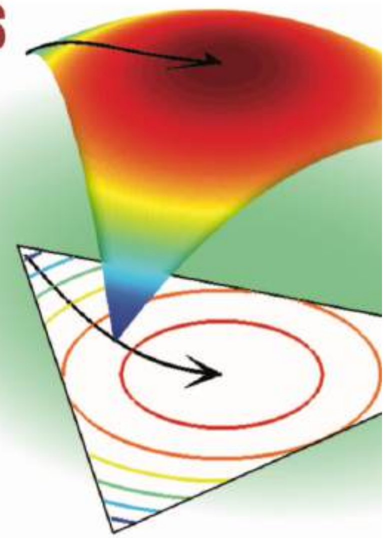
strategy 2

strategy 3



Population Games: Concept and Motivation

The Role of Population Games and Evolutionary Dynamics in Distributed Control Systems



THE ADVANTAGES OF EVOLUTIONARY GAME THEORY

NICANOR QUIJANO, CARLOS OCAMPO-MARTINEZ, JULIAN BARREIRO-GOMEZ, GERMÁN OBANDO, ANDRÉS PANTOJA, and EDUARDO MOJICA-NAVA

Recently, there has been an increasing interest in the control community in studying large-scale distributed systems. Several techniques have been developed to address the main challenges for these systems, such as the amount of information needed to guarantee the proper operation of the system, the economic costs associated with the required communication structure, and the high computational burden of solving for the control inputs for large-scale systems.

One way to overcome such problems is to use a multiagent systems framework, which may be cast in game-theoretical terms. Game theory studies interactions between self-interested agents and tackles the problem of interaction among agents using different strategies who wish to maximize their welfare. For instance in [1], the connections are provided among games, optimization, and learning for signal processing in networks. Other approaches, in terms of learning and games, can be found in [2]. In [3], distributed computation algorithms are developed based on generalized convex games that do not require full information and where there is a dynamic change in terms of network topologies. Applications of game theory in control of optical networks and game-theoretic methods for smart grids are described in [4]–[6]. Another approach in game-theoretical methods is to design protocols or mechanisms that possess some desirable properties [7]. This approach leads to a broad analysis of multiagent interactions, particularly those involving negotiation and coordination problems [8]. Other game-theoretical applications to engineering are reported in [9].

From a game-theoretical perspective, there are three types of games: matrix games, continuous games, and

Digital Object Identifier 10.1109/MCS.2016.2621479
Date of publication: 19 January 2017

1066-033X/17©2017IEEE

Control systems magazine, 2017

voltage-source inverters, are connected to loads through an inverter. The magnitude and frequency of the output voltage are controlled by means of a droop-gain controller [64]. For more details about this approach, see [26] and references therein. An illustrative general scheme is presented in Figure 8 for a microgrid with seven DGs (DG 1, DG 2, ..., DG 7) and several loads connected in a electrical topology adapted from the IEEE 30-bus distribution system.

At the higher level of the microgrid control architecture, is a control strategy that dynamically dispatches active power setpoints. This controller has to generate the power setpoints based on economic criteria. In this model, the production costs of active power and load demands are considered as external inputs coming from the lower-level control to the microgrid central controller, where the dynamic dispatch based on replicator dynamics is being executed. This fact implies that costs and load demands could be time varying, allowing the inclusion of renewable-energy resources. The main focus here is on the higher level of the microgrid, where the dynamic dispatch is executed. The maximization of utility functions for the EDP is adopted, including active power generation by means of voltage-source inverters connected to DGs at a lower level. The EDP can be formulated as

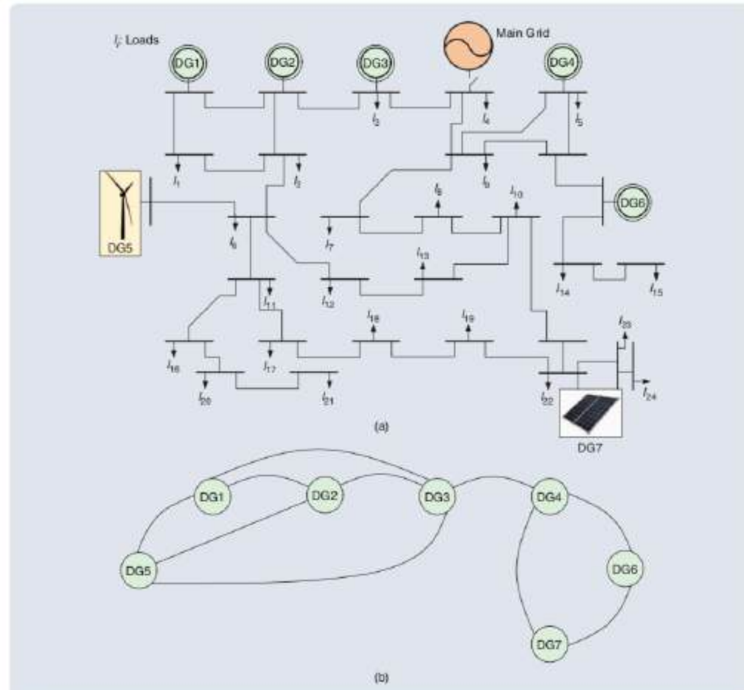


FIGURE 8 A microgrid test system with seven distributed generators and several loads. (a) An illustrative general scheme of a microgrid model adapted from an IEEE 30-bus distribution system and (b) a graph representing the topology of the communication network between distributed generators.

Dynamic Economic Dispatch for Microgrids

Basic Formulation

Without loss of generality we consider a single population

Strategy set for the population is $\{1, \ldots, n\}$

Population state

$$x(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}$$

$x_i(t)$ *is the portion of the population selecting strategy* i *at time* t.

We assume unit mass $\sum_{i=1}^{n} x_i(t) = 1$

Basic Formulation

Without loss of generality we consider a single population

Strategy set for the population is $\{1, \ldots, n\}$

Population state

$$x(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}$$

$x_i(t)$ *is the portion of the population selecting strategy* i *at time* t.

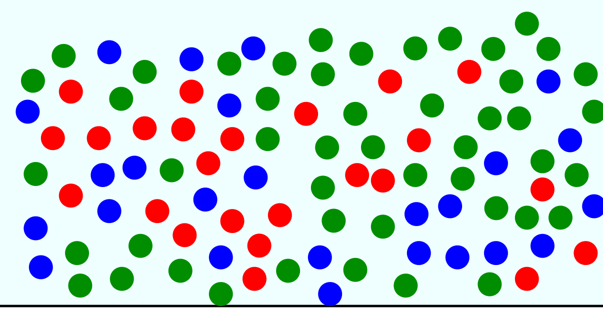
We assume unit mass $\sum_{i=1}^{n} x_i(t) = 1$

Example for $n = 3$ *and* $N = 100$

$$x(t^*) = \begin{bmatrix} 0.25 \\ 0.25 \\ 0.5 \end{bmatrix}$$

- *strategy 1*
- *strategy 2*
- *strategy 3*

population state at time t^*



Basic Formulation

$$x(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}$$

$x_i(t)$ is the portion of the population selecting strategy i at time t.

We assume unit mass $\sum_{i=1}^{n} x_i(t) = 1$

strategy profile set

$$\mathbb{X} \stackrel{\text{def}}{=} \left\{ x \in \mathbb{R}^n_+ \;\middle|\; \sum_{j=1}^{n} x_j = 1 \right\}$$

Basic Formulation

$$x(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}$$

$x_i(t)$ is the portion of the population selecting strategy i at time t.

We assume unit mass $\sum_{i=1}^{n} x_i(t) = 1$

strategy profile set $$\mathbb{X} \stackrel{\text{def}}{=} \left\{ x \in \mathbb{R}_+^n \;\middle|\; \sum_{j=1}^{n} x_j = 1 \right\}$$

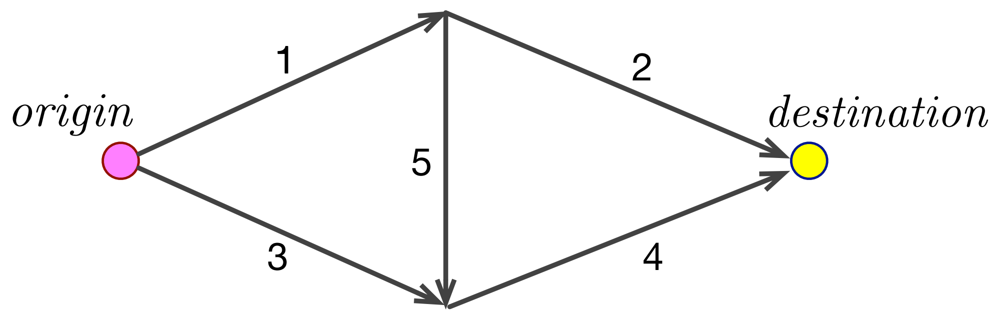
The payoff of a population game is specified as follows, where $\mathcal{F} : \mathbb{X} \to \mathbb{R}^n$ is a continuously differentiable map.

$$p(t) := \mathcal{F}\big(x(t)\big), \qquad p(t) = \begin{bmatrix} p_1(t) \\ \vdots \\ p_n(t) \end{bmatrix}$$

$p_i(t)$ is the payoff of strategy i at time t

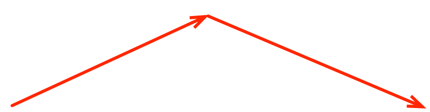
Basic Formulation

Example: 3-strategy congestion game



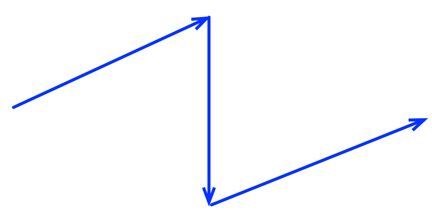
$\mathcal{D}_i : \mathbb{X} \to \mathbb{R}$ *is delay as a* C^1 *increasing function of utilization in link* i

strategy 1



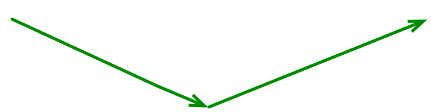
$$\mathcal{F}_1(x) = -\mathcal{D}_1(x_1 + x_2) - \mathcal{D}_2(x_1)$$

strategy 2



$$\mathcal{F}_2(x) = -\mathcal{D}_1(x_1 + x_2) - \mathcal{D}_4(x_2 + x_3) - \mathcal{D}_5(x_2)$$

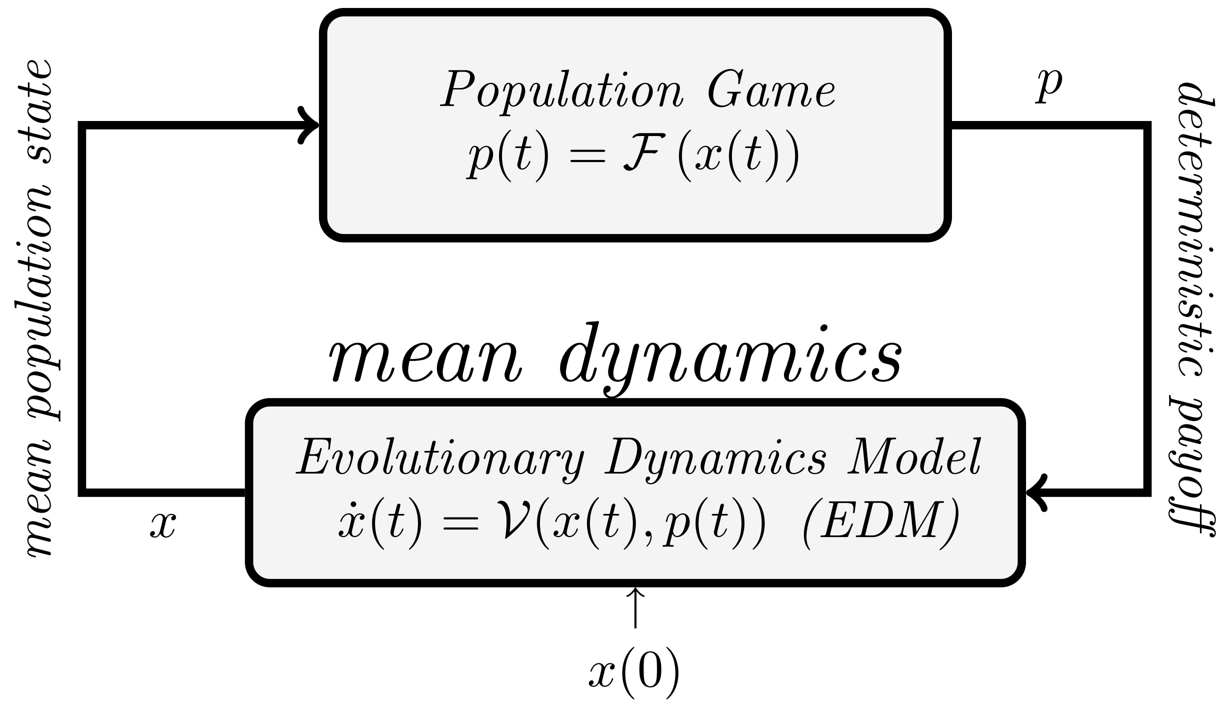
strategy 3



$$\mathcal{F}_3(x) = -\mathcal{D}_3(x_3) - \mathcal{D}_4(x_2 + x_3)$$

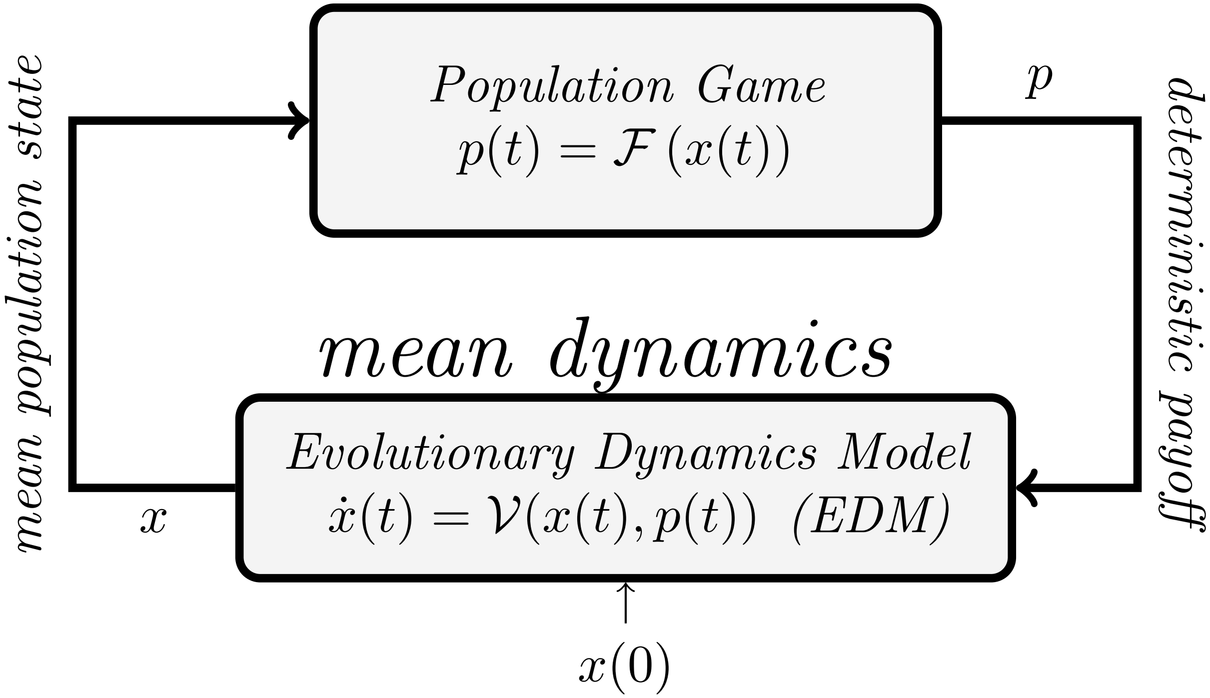
Mean Dynamics

Deterministic approximation for large population limit $(N \to \infty)$



Mean Dynamics

Deterministic approximation for large population limit ($N \to \infty$)



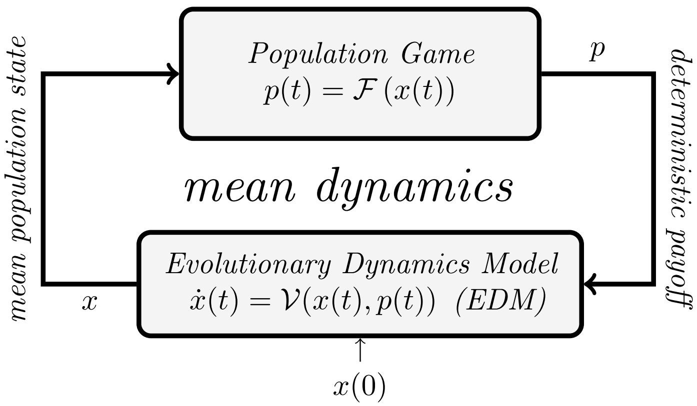
From learning rule to EDM

$$\dot{x}_i = \mathcal{V}_i(x,p) := \sum_{j=1}^{n} x_j \mathcal{T}_{ji}(x,p) - x_i \left(\sum_{j=1}^{n} \mathcal{T}_{ij}(x,p) \right),$$

$$x \in \mathbb{X},\ p \in \mathbb{R}^n,\ 1 \leq i \leq n$$

probabilistic model of learning rule

Mean Dynamics



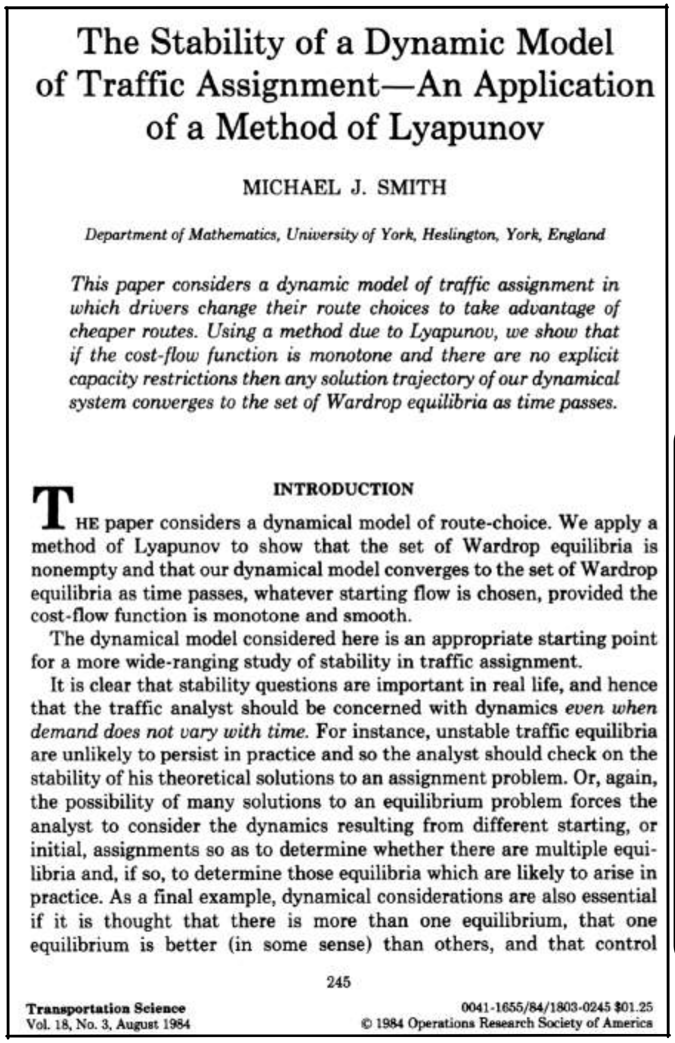
Flow interpretation

inflow switching to strategy i *outflow switching out of strategy i*

$$\dot{x}_i = \mathcal{V}_i(x,p) := \boxed{\sum_{j=1}^{n} x_j \mathcal{T}_{ji}(x,p)} - \boxed{x_i \left(\sum_{j=1}^{n} \mathcal{T}_{ij}(x,p) \right)},$$

$$x \in \mathbb{X}, \; p \in \mathbb{R}^n, \; 1 \leq i \leq n$$

(EDM) Evolutionary dynamics model



The Stability of a Dynamic Model of Traffic Assignment—An Application of a Method of Lyapunov

MICHAEL J. SMITH

Department of Mathematics, University of York, Heslington, York, England

This paper considers a dynamic model of traffic assignment in which drivers change their route choices to take advantage of cheaper routes. Using a method due to Lyapunov, we show that if the cost-flow function is monotone and there are no explicit capacity restrictions then any solution trajectory of our dynamical system converges to the set of Wardrop equilibria as time passes.

INTRODUCTION

THE paper considers a dynamical model of route-choice. We apply a method of Lyapunov to show that the set of Wardrop equilibria is nonempty and that our dynamical model converges to the set of Wardrop equilibria as time passes, whatever starting flow is chosen, provided the cost-flow function is monotone and smooth.

The dynamical model considered here is an appropriate starting point for a more wide-ranging study of stability in traffic assignment.

It is clear that stability questions are important in real life, and hence that the traffic analyst should be concerned with dynamics *even when demand does not vary with time.* For instance, unstable traffic equilibria are unlikely to persist in practice and so the analyst should check on the stability of his theoretical solutions to an assignment problem. Or, again, the possibility of many solutions to an equilibrium problem forces the analyst to consider the dynamics resulting from different starting, or initial, assignments so as to determine whether there are multiple equilibria and, if so, to determine those equilibria which are likely to arise in practice. As a final example, dynamical considerations are also essential if it is thought that there is more than one equilibrium, that one equilibrium is better (in some sense) than others, and that control

245

Transportation Science
Vol. 18, No. 3, August 1984

0041-1655/84/1803-0245 $01.25
© 1984 Operations Research Society of America

$$\mathcal{V}_i(x,p) := \sum_{j=1}^{n} x_j \mathcal{T}_{ji}(x,p) - \left(\sum_{j=1}^{n} \mathcal{T}_{ij}(x,p)\right) x_i,$$

$$x \in \mathbb{X},\ p \in \mathbb{R}^n,\ 1 \le i \le n$$

(Smith EDM) *The Smith EDM is specified by the following Smith IPC learning rule:*

$$\mathcal{T}_{ij}^{Smith}(x,p) := [p_j - p_i]_+, \quad p \in \mathbb{R}^n, \quad x \in \mathbb{X}$$

$$\mathcal{V}_i^{Smith}(x,p) := \sum_{j=1}^{n} x_j [p_i - p_j]_+ - \left(\sum_{j=1}^{n} [p_j - p_i]_+\right) x_i$$

Nash Equilibria

John Nash

Average population payoff is $x^T \mathcal{F}(x)$.

Definition of NE for a game $\mathcal{F}$*:*

$$\mathbb{NE}(\mathcal{F}) := \left\{ x \in \mathbb{X} \mid x^T \mathcal{F}(x) \geq \tilde{x}^T \mathcal{F}(x),\ \tilde{x} \in \mathbb{X} \right\}$$

Nash Equilibria

John Nash

Average population payoff is $x^T \mathcal{F}(x)$.

Definition of NE for a game $\mathcal{F}$:

$$\mathbb{NE}(\mathcal{F}) := \left\{ x \in \mathbb{X} \mid x^T \mathcal{F}(x) \geq \tilde{x}^T \mathcal{F}(x), \ \tilde{x} \in \mathbb{X} \right\}$$

- From Kakutani's fixed point Theorem, $\mathbb{NE}(\mathcal{F}) \neq \emptyset$.
- Wardrop-type representation of Nash equilibria:

$$\mathbb{NE}(\mathcal{F}) = \left\{ x \in \mathbb{X} \,\middle|\, x_i > 0 \implies \mathcal{F}_i(x) = \max_{1 \leq j \leq n} \mathcal{F}_j(x) \right\}.$$

- Finite sub-pop. do not gain from deviating from $\mathbb{NE}(\mathcal{F})$.
- Admit refinements based on evolutionary concepts, such as:

$$\mathbb{GESS}(\mathcal{F}) := \{ x \in \mathbb{X} \mid (x - y)^T \mathcal{F}(y) > 0, y \in \mathbb{X} - \{x\} \}$$

Nash Equilibria

John Nash

Average population payoff is $x^T\mathcal{F}(x)$.

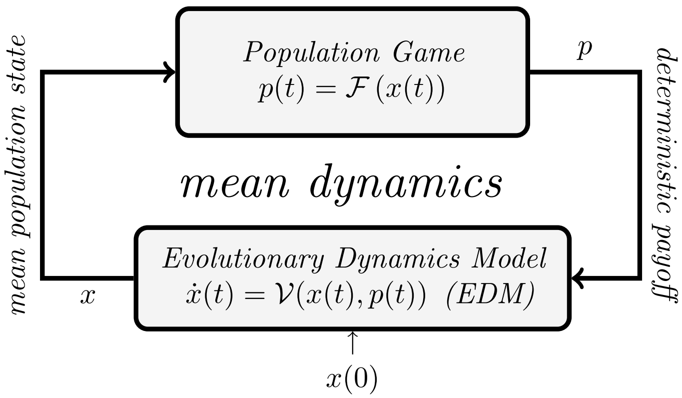
Definition of NE for a game $\mathcal{F}$:

$$\mathbb{NE}(\mathcal{F}) := \left\{x \in \mathbb{X} \mid x^T\mathcal{F}(x) \geq \tilde{x}^T\mathcal{F}(x),\ \tilde{x} \in \mathbb{X}\right\}$$

For potential games ($\mathcal{F} = \nabla f$) [Sandholm, J. Econ. Th., '01]:

- $\mathbb{NE}(\mathcal{F}) = \mathbb{KKT}(f) := \{x$ satisfies KKT for $\max_{x\in\mathbb{X}} f(x)\}$.
- If the game is homogeneous of degree $k \neq -1$ and f is concave then $\mathbb{NE}(\mathcal{F})$ is socially optimal.
- If $x^* \in \mathrm{relint}(\mathbb{GESS}(\mathcal{F}))$ then $\tilde{x}^T D\mathcal{F}(x^*)\tilde{x} < 0$ for $\tilde{x} \in \mathbb{TX} - \{0\}$, i.e., f is locally concave at x^* in $\mathbb{X}$.
  - x^* is local maximum of f.

Well-Behaved Learning Rules



Average population payoff is $x^T \mathcal{F}(x)$.

Definition of NE for a game $\mathcal{F}$*:*

$$\mathbb{NE}(\mathcal{F}) := \{x \in \mathbb{X} \mid x^T \mathcal{F}(x) \geq \tilde{x}^T \mathcal{F}(x),\ \tilde{x} \in \mathbb{X}\}$$

Well-behaved learning rules satisfy:

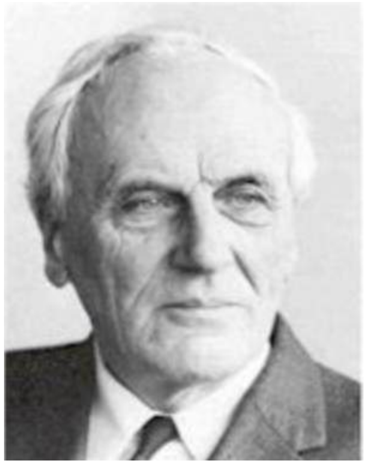
Nash stationarity

$$\mathcal{V}(x,p) = 0 \iff x \in \arg\max_{\tilde{x} \in \mathbb{X}} \tilde{x}' p$$

Positive correlation

$$\mathcal{V}(x,p) \neq 0 \iff p' \mathcal{V}(x,p) > 0$$

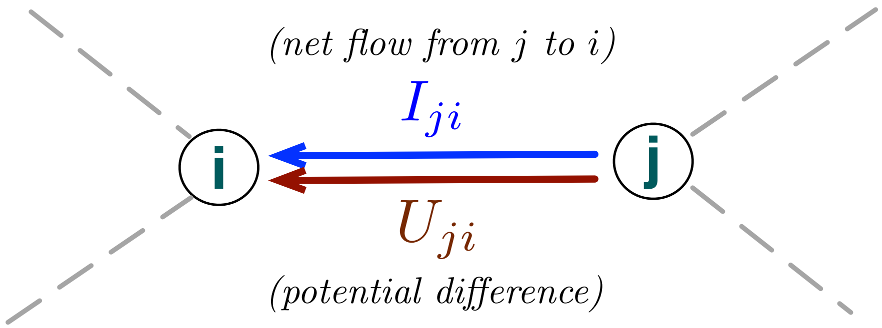
Well-Behaved Learning Rules



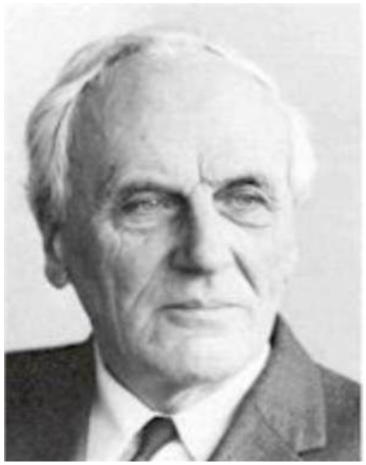
Bernard D. H. Tellegen

Tellegen-type interpretation of positive correlation

$$p'\mathcal{V}(x,p) = \sum_{i>j} \underbrace{(p_i - p_j)}_{U_{ji}} (\underbrace{x_j \mathcal{T}_{ji}(x,p) - x_i \mathcal{T}_{ij}(x,p)}_{I_{ji}})$$



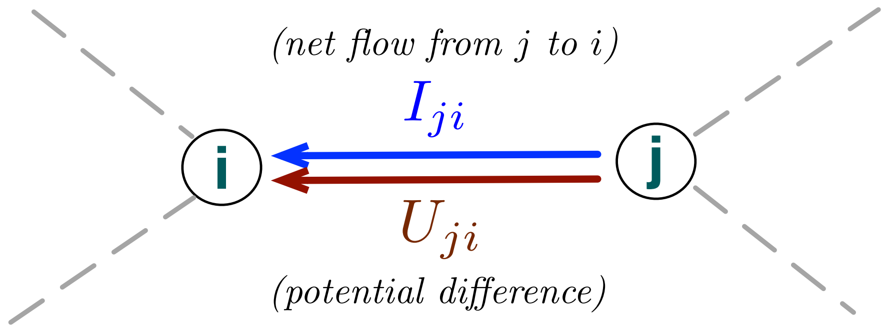
Well-Behaved Learning Rules



Bernard D. H. Tellegen

Tellegen-type interpretation of positive correlation

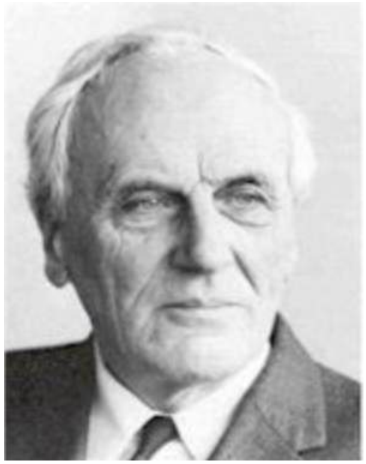
$$p'\mathcal{V}(x,p) = \sum_{i>j} \underbrace{(p_i - p_j)}_{U_{ji}} \underbrace{(x_j \mathcal{T}_{ji}(x,p) - x_i \mathcal{T}_{ij}(x,p))}_{I_{ji}}$$



Sufficient condition for positive correlation

$$I_{ji}U_{ji} \geq 0 \text{ and } (I_{ji} \neq 0 \implies I_{ji}U_{ji} > 0)$$

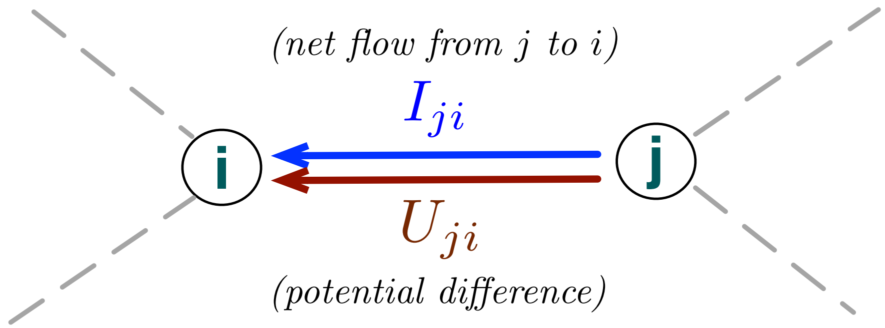
Well-Behaved Learning Rules



Bernard D. H. Tellegen

Tellegen-type interpretation of positive correlation

$$p'\mathcal{V}(x,p) = \sum_{i>j} \underbrace{(p_i - p_j)}_{U_{ji}} \big(\underbrace{x_j \mathcal{T}_{ji}(x,p) - x_i \mathcal{T}_{ij}(x,p)}_{I_{ji}}\big)$$



Tellegen's theorem when:
KCL $\Leftrightarrow \dot{x} = 0$

Sufficient condition for positive correlation

$$I_{ji}U_{ji} \geq 0 \text{ and } (I_{ji} \neq 0 \implies I_{ji}U_{ji} > 0)$$

Well-Behaved Learning Rules

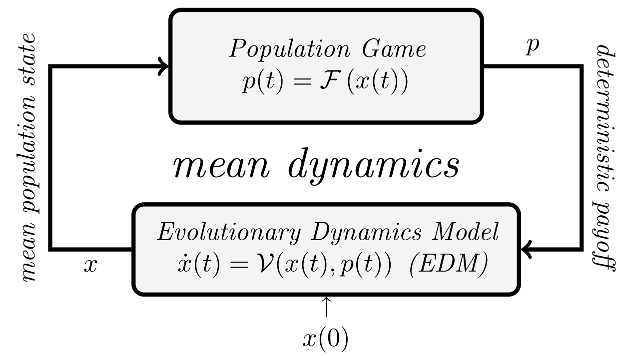
Well-behaved learning rules satisfy:

Nash stationarity

$$\mathcal{V}(x,p) = 0 \iff x \in \arg\max_{\tilde{x} \in \mathbb{X}} \tilde{x}' p$$

Positive correlation

$$\mathcal{V}(x,p) \neq 0 \iff p' \mathcal{V}(x,p) > 0$$



Examples include:

- Pairwise comparison rules (with $\phi_{ij}(\tau) = 0$ if $\tau \leq 0$, $\phi_{ij}(\tau) > 0$ otherwise):

$$\mathcal{T}_{ij}(x,p) = \phi_{ij}(p_j - p_i)$$

- Imitation protocols in the interior of $\mathbb{X}$ (includes replicator $\psi_{ij}(\tau) = [\tau]_+$):

$$\mathcal{T}_{ij}(x,p) = x_j \psi_{ij}(p_j - p_i)$$

- Excess payoff target protocol (includes best response approximations):

$$\mathcal{T}_{ij}(x,p) = \varphi_{ij}(\hat{p}), \quad \hat{p}_j := p_j - x^T p$$

- Hybrid rules formed as conic combinations.

Well-behaved rules & potential games

William H. Sandholm, "Potential Games with Continuous Player Sets," Journal of Economic Theory, Volume 97, Issue 1, 2001, Pages 81-108.

Nash stationarity

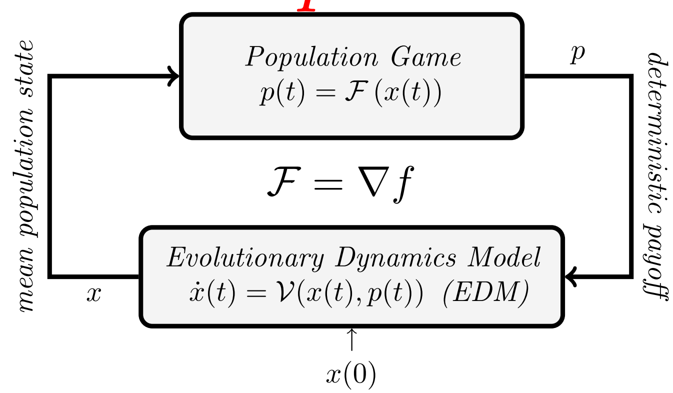
$$\mathcal{V}(x,p) = 0 \iff x \in \arg\max_{\tilde{x} \in \mathbb{X}} \tilde{x}' p$$

Positive correlation

$$\mathcal{V}(x,p) \neq 0 \iff p' \mathcal{V}(x,p) > 0$$

&

$\mathcal{F}$ *is potential*



THEOREM

$$x(t) \underset{t \to \infty}{\longrightarrow} \mathbb{NE}(\mathcal{F}) = \mathbb{KKT}(f)$$

evolutionary Nash equilibrium learning

Well-behaved protocols & potential games

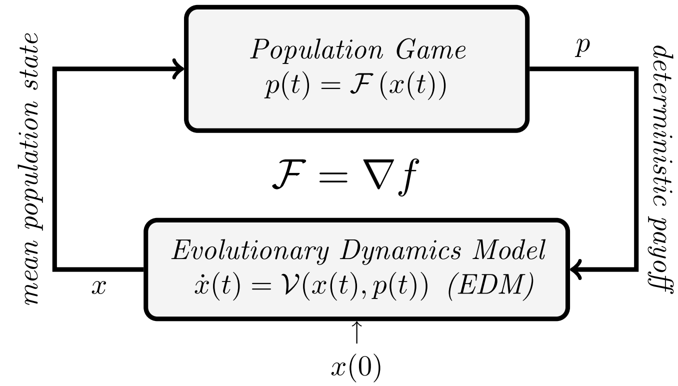
Nash stationarity

$$\mathcal{V}(x,p) = 0 \iff x \in \arg\max_{\tilde{x} \in \mathbb{X}} \tilde{x}' p$$

Positive correlation

$$\mathcal{V}(x,p) \neq 0 \iff p'\mathcal{V}(x,p) > 0$$

&



$$\boxed{x(t) \underset{t \to \infty}{\longrightarrow} \mathbb{NE}(\mathcal{F}) = \mathbb{KKT}(f)}$$

Proof:

$\frac{d}{dt} f(x(t)) = \nabla f(x(t))' \dot{x}(t) = p(t)' \mathcal{V}(x(t), p(t))$,
where $p(t) = \mathcal{F}(x(t))$. $\square$

Relevant observations:

- No need to precisely know $\mathcal{T}$.
- If f is concave then $\mathbb{NE}(\mathcal{F}) = \arg\max_{x \in \mathbb{X}} f(x)$.

Exploring δ passivity

$$\mathbf{u} : [0,\infty) \to \mathbb{U} \longrightarrow \boxed{\mathfrak{G} : \mathbf{u} \mapsto \mathbf{y}} \longrightarrow \mathbf{y} : [0,\infty) \to \mathbb{Y}$$

$\mathbf{u} \in \mathscr{U}$ and $\mathbf{y} \in \mathscr{Y}$, where $\mathscr{U}$ and $\mathscr{Y}$ are the input and output sets.

We assume all elements of $\mathscr{U}$ and $\mathscr{Y}$ are differentiable a.e.

We assume u(t) and y(t) have the same dimension, so we can write $u'(t)y(t)$

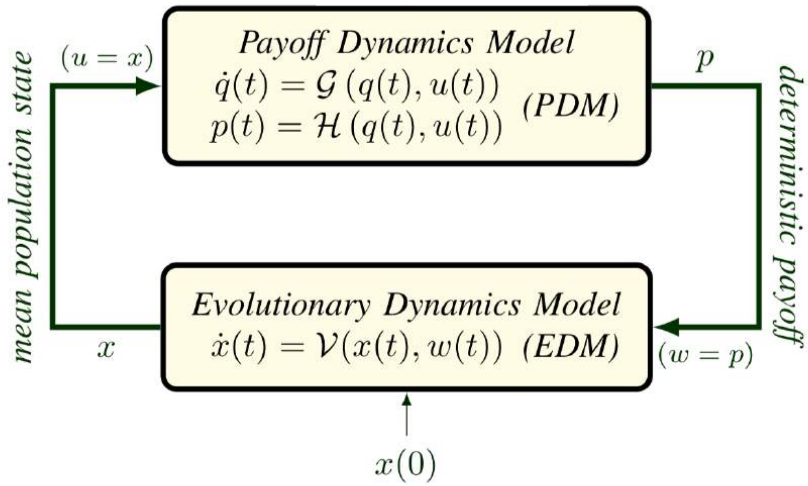
Fox MJ, Shamma JS. Population Games, Stable Games, and Passivity. Games. 2013; 4(4):561-583.

The system $\mathfrak{G}$ is δ-passive when:

$$\inf_{T>0,\mathbf{u}\in\mathscr{U}} \int_0^T \dot{u}'(t)\dot{y}(t)dt \;>\; -\infty$$

Allowing Dynamics In Payoff Mechanism

$$\mathfrak{F} : \mathbf{x} \mapsto \mathbf{p}$$

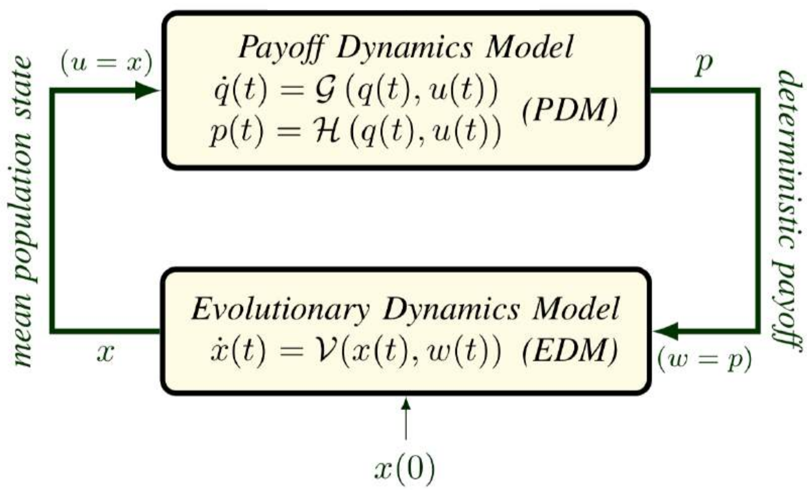


Dynamics in the payoff mechanism allows:

- higher-order learning dynamics.
- modeling information dissemination.
- more flexible design (coupled dynamics).

Allowing Dynamics In Payoff Mechanism

$$\mathfrak{F} : \mathbf{x} \mapsto \mathbf{p}$$



(Sandholm, 2001)

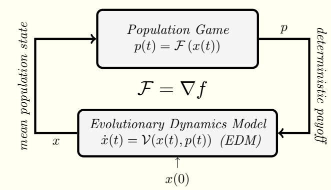
Nash stationarity

$\mathcal{V}(x,p) = 0 \iff x \in \arg\max_{\tilde{x}\in\mathbb{X}} \tilde{x}'p$

Positive correlation

$\mathcal{V}(x,p) \neq 0 \iff p'\mathcal{V}(x,p) > 0$

$+$



$$x(t) \underset{t\to\infty}{\longrightarrow} \mathrm{NE}(\mathcal{F}) = \mathbb{KKT}(f)$$

Contractive Games

(Hofbauer & Sandholm, 2009)

$$(x - \tilde{x})^T \big(\mathcal{F}(x) - \mathcal{F}(\tilde{x})\big) \leq 0$$

δ-passivity (Dynamic Payoffs)

(Fox & Shamma, 2013)

(Park, Martins, Shamma, tutorial 2019)

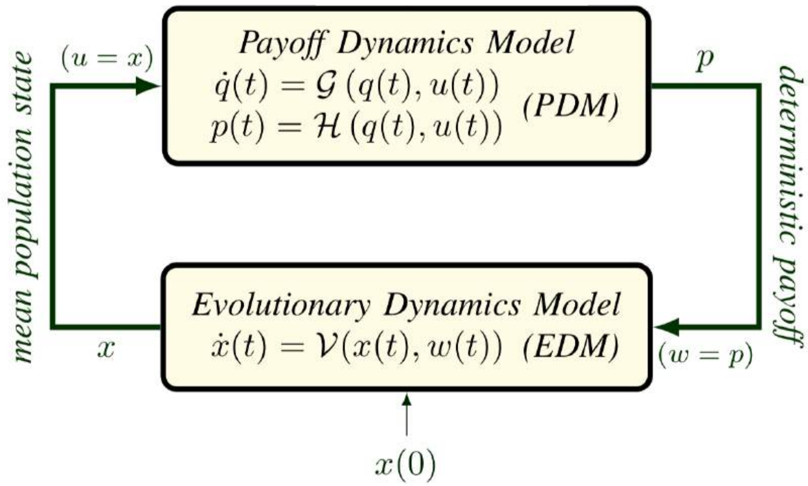
(Arcak & Martins, 2021)

Dynamics in the payoff mechanism allows:

- higher-order learning dynamics.
- modeling information dissemination.
- more flexible design (coupled dynamics).

Allowing Dynamics In Payoff Mechanism

$\mathfrak{F} : \mathbf{x} \mapsto \mathbf{p}$



Assuming well-defined $\mathcal{F} : \mathbb{X} \to \mathbb{R}^n$

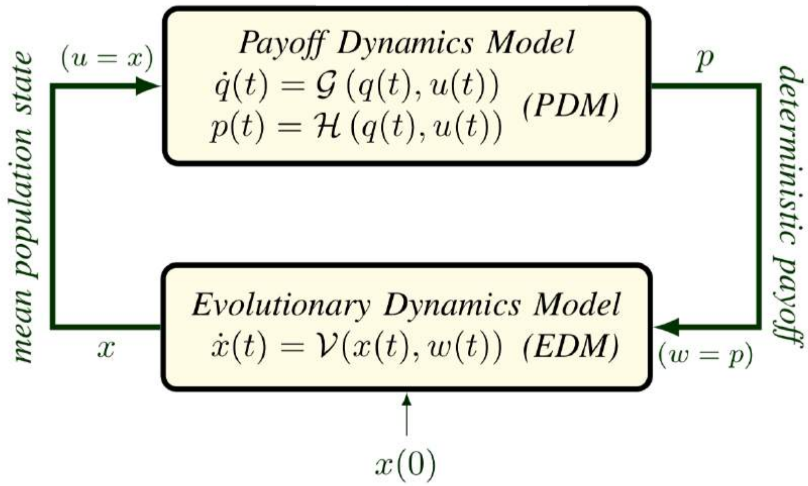
$$\mathcal{F}(x) := \lim_{t\to\infty} p(t),$$
$$\text{for } \lim_{t\to\infty} u(t) = x$$

Dynamics in the payoff mechanism allows:

- higher-order learning dynamics.
- modeling information dissemination.
- more flexible design (coupled dynamics).

Allowing Dynamics In Payoff Mechanism

$$\mathfrak{F} : \mathbf{x} \mapsto \mathbf{p}$$



Assuming well-defined $\mathcal{F} : \mathbb{X} \to \mathbb{R}^n$

$$\mathcal{F}(x) := \lim_{t\to\infty} p(t),$$
$$\text{for } \lim_{t\to\infty} u(t) = x$$

S. Park, N. C. Martins and J. S. Shamma, "From Population Games to Payoff Dynamics Models: A Passivity-Based Approac," 2019 IEEE 58th Conference on Decision and Control (CDC), Nice, France, 2019, pp. 6584-6601.

Nash stationarity

$$\mathcal{V}(x, p) = 0 \iff x \in \arg\max_{\tilde{x}\in\mathbb{X}} \tilde{x}'p$$

and

(EDM) is δ passive

& $-\mathfrak{F}$ is δ passive

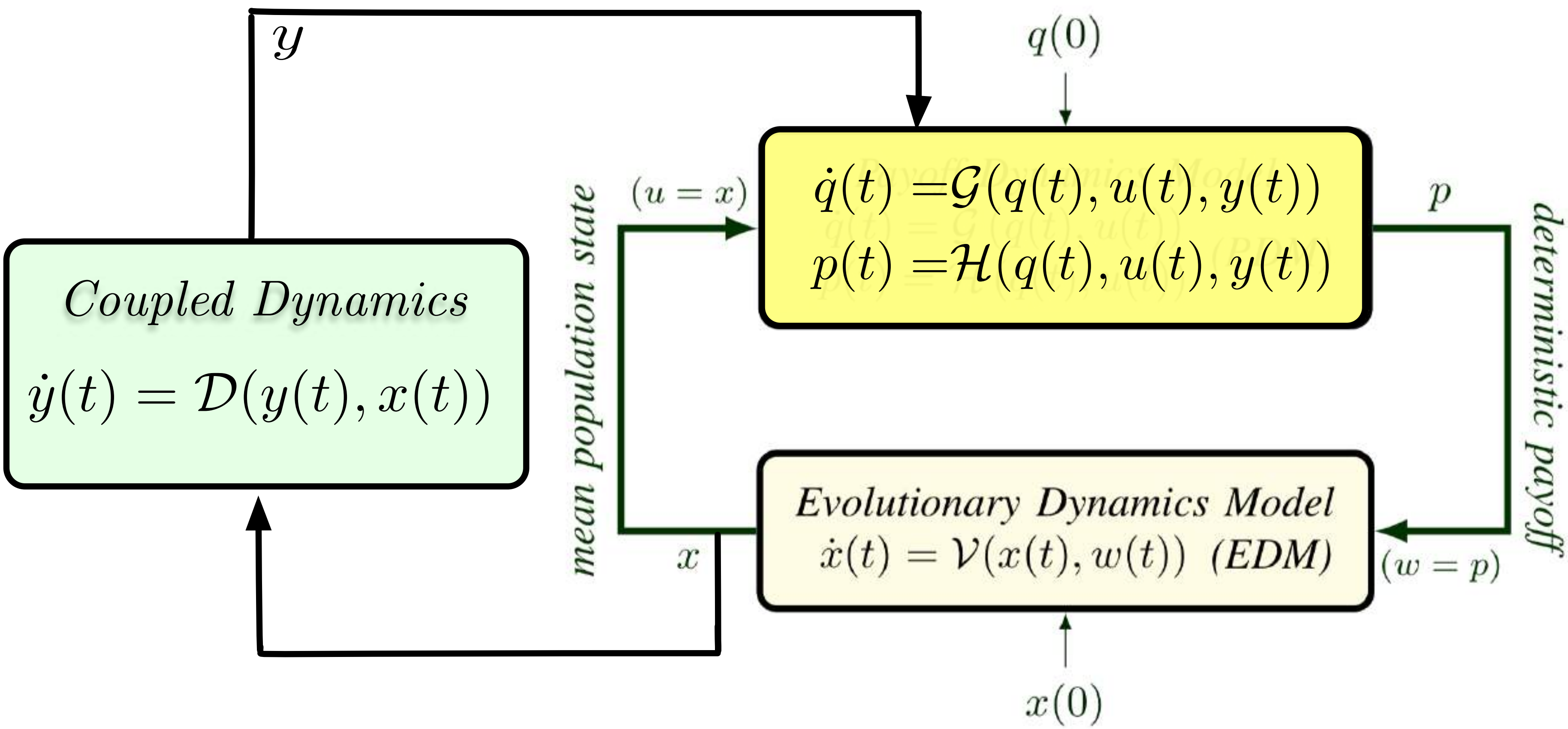
$$x(t) \underset{t\to\infty}{\longrightarrow} \mathbb{NE}(\mathcal{F})$$

evolutionary Nash equilibrium learning

Dynamics in the payoff mechanism allows:

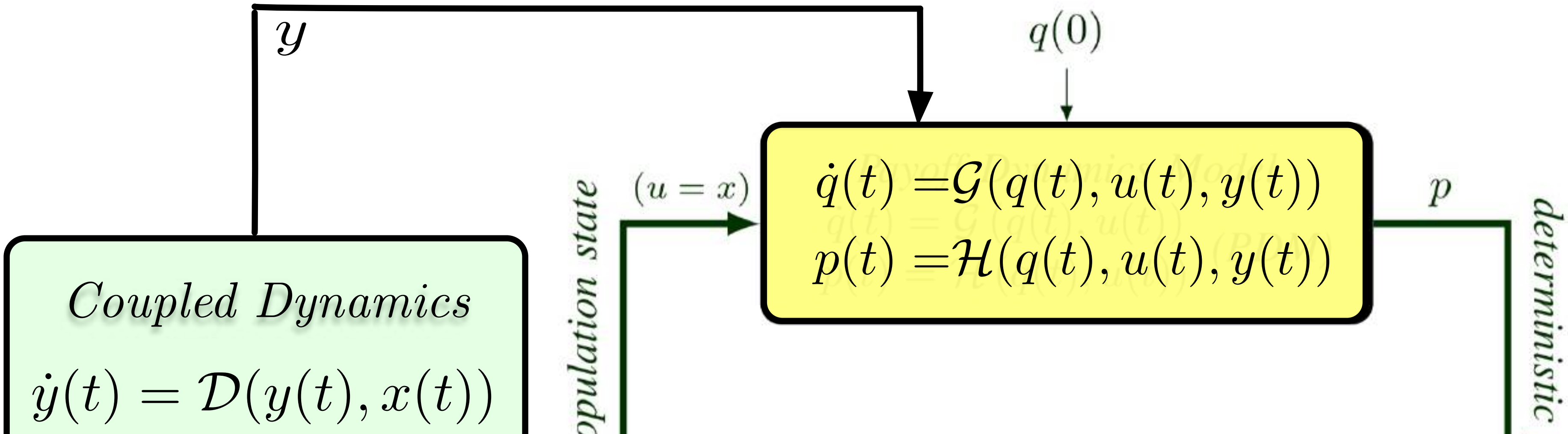
- higher-order learning dynamics.
- modeling information dissemination.
- more flexible design (coupled dynamics).

δ-Passivity And Coupling With Exogenous Dynamics



Goal: Given $\mathcal{D}$ and "some structural knowledge" of $\mathcal{V}$, design $\mathcal{G}$ and $\mathcal{H}$ for which (x, y, q) is guaranteed to converge to a desirable equilibrium set $\mathbb{E}^*$.

δ-Passivity And Coupling With Exogenous Dynamics



Progress made:

- **N. C. Martins, J. Certorio, R. J. La," Epidemic population games and evolutionary dynamics," Automatica, Volume 153, 2023. Discussed in last year's review.**
- **J. Certorio, N. C. Martins and R. J. La, "Epidemic Population Games With Nonnegligible Disease Death Rate," in IEEE Control Systems Letters, vol. 6, pp. 3229-3234, 2022.**
- **J. Certorio, N. C. Martins, R. J. La, M. Arcak, "Incentive Designs for Learning Agents to Stabilize Coupled Exogenous Systems," to appear at CDC'24.**

Generality of δ-passive learning rules

arXiv:2407.02083v2 [math.DS] 16 Jul 2024

Passivity Tools for Hybrid Learning Rules in Large Populations

Jair Certório [a], Nuno C. Martins [a], Kevin Chang [b], Pierluigi Nuzzo [b], Yasser Shoukry [c]

[a] *Department of ECE and ISR, University of Maryland, College Park, MD, 20742, USA.*

[b] *Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA, 90089, USA.*

[c] *Department of Electrical Engineering and Computer Science, University of California, Irvine, CA, 92697, USA.*

Abstract

Recent work has pioneered the use of system-theoretic passivity to study equilibrium stability for the dynamics of noncooperative strategic interactions in large populations of learning agents. In this and related works, the stability analysis leverages knowledge that certain "canonical" classes of learning rules used to model the agents' strategic behaviors satisfy a passivity condition known as δ-passivity. In this paper, we consider that agents exhibit learning behaviors that do not align with a canonical class. Specifically, we focus on characterizing δ-passivity for hybrid learning rules that combine elements from canonical classes. Our analysis also introduces and uses a more general version of δ-passivity, which, for the first time, can handle discontinuous learning rules, including those showing best-response behaviors. We state and prove theorems establishing δ-passivity for two broad convex cones of hybrid learning rules. These cones can merge into a larger one preserving δ-passivity in scenarios limited to two strategies. In our proofs, we establish intermediate facts that are significant on their own and could potentially be used to further generalize our work. We illustrate the applicability of our results through numerical examples.

Key words: Distributed learning, evolutionary dynamics, system-theoretic passivity, multi-agent systems, asymptotic stabilization.

1 Introduction

Developments in population games and evolutionary dynamics [1, 2] have contributed to systematic methods to model and analyze the dynamics of strategic noncooperative interactions among large populations of learning agents. Central to this approach is the use of learning rules [1] that model how agents update their strategies over time based on the payoffs of those strategies, which are in turn determined by a payoff mechanism. These learning rules can be explicitly programmed in artificial agents or represent innate preferences or bounded rationality in humans or other natural agents. This framework is well-suited to distributed optimization [3] and engineering systems [4, 5], where the payoff mechanisms might abstractly represent the specifics of the agents' interaction environments, such as in congestion games, or are purposefully designed and implemented by a coordinator to steer the population toward desired strategic outcomes [6, 7].

1.1 Studying Passivity for Hybrid Learning Rules

System-theoretic passivity tools have been employed to study how populations, adhering to certain learning rules, achieve and maintain Nash equilibria—a process often referred to as Nash equilibrium seeking. Pioneering work in [8] demonstrates that Nash equilibrium seeking and convergence are achieved under contractive payoff mechanisms if the learning rules are δ-passive, a concept inspired by classical notions of system-theoretic passivity that has been generalized in [9, 10]. The convergence results in these articles allow for dynamic payoff mechanisms and are not contingent upon the specific learning rule employed, as long as it satisfies δ-passivity, often confirmed through structural analysis.

Recent research has shown that δ-passivity can also be used to systematically design dynamic payoff mechanisms to guarantee, via Lyapunov analysis, global

* This work was supported in part by AFOSR Grant FA9550-23-1-0467 and NSF Grants 2139713, 2139781, 2139982, 2135561, and 1846524.

Email addresses: certorio@umd.edu (Jair Certório), nmartins@umd.edu (Nuno C. Martins), kcchang@usc.edu (Kevin Chang), nuzzo@usc.edu (Pierluigi Nuzzo), yshoukry@uci.edu (Yasser Shoukry).

[1] Often called revision protocols in the population game literature.

Good news:

- *δ-passive rules form large convex cones.*
  - *Include common hybrid rules.*
- *The cones include the best response rule.*

Generality of δ-passive learning rules

Passivity Tools for Hybrid Learning Rules in Large Populations

Jair Certório [a], Nuno C. Martins [a], Kevin Chang [b], Pierluigi Nuzzo [b], Yasser Shoukry [c]

[a] *Department of ECE and ISR, University of Maryland, College Park, MD, 20742, USA.*

[b] *Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, CA, 90089, USA.*

[c] *Department of Electrical Engineering and Computer Science, University of California, Irvine, CA, 92697, USA.*

Abstract

Recent work has pioneered the use of system-theoretic passivity to study equilibrium stability for the dynamics of noncooperative strategic interactions in large populations of learning agents. In this and related works, the stability analysis leverages knowledge that certain "canonical" classes of learning rules used to model the agents' strategic behaviors satisfy a passivity condition known as δ-passivity. In this paper, we consider that agents exhibit learning behaviors that do not align with a canonical class. Specifically, we focus on characterizing δ-passivity for hybrid learning rules that combine elements from canonical classes. Our analysis also introduces and uses a more general version of δ-passivity, which, for the first time, can handle discontinuous learning rules, including those showing best-response behaviors. We state and prove theorems establishing δ-passivity for two broad convex cones of hybrid learning rules. These cones can merge into a larger one preserving δ-passivity in scenarios limited to two strategies. In our proofs, we establish intermediate facts that are significant on their own and could potentially be used to further generalize our work. We illustrate the applicability of our results through numerical examples.

Key words: Distributed learning, evolutionary dynamics, system-theoretic passivity, multi-agent systems, asymptotic stabilization.

1 Introduction

Developments in population games and evolutionary dynamics [1,2] have contributed to systematic methods to model and analyze the dynamics of strategic noncooperative interactions among large populations of learning agents. Central to this approach is the use of learning rules [1] that model how agents update their strategies over time based on the payoffs of those strategies, which are in turn determined by a payoff mechanism. These learning rules can be explicitly programmed in artificial agents or represent innate preferences or bounded rationality in humans or other natural agents. This framework is well-suited to distributed optimization [3] and engineering systems [4,5], where the payoff mechanisms might abstractly represent the specifics of the agents' interaction environments, such as in congestion games, or are purposefully designed and implemented by a coordinator to steer the population toward desired strategic outcomes [6,7].

1.1 Studying Passivity for Hybrid Learning Rules

System-theoretic passivity tools have been employed to study how populations, adhering to certain learning rules, achieve and maintain Nash equilibria—a process often referred to as Nash equilibrium seeking. Pioneering work in [8] demonstrates that Nash equilibrium seeking and convergence are achieved under contractive payoff mechanisms if the learning rules are δ-passive, a concept inspired by classical notions of system-theoretic passivity that has been generalized in [9,10]. The convergence results in these articles allow for dynamic payoff mechanisms and are not contingent upon the specific learning rule employed, as long as it satisfies δ-passivity, often confirmed through structural analysis.

Recent research has shown that δ-passivity can also be used to systematically design dynamic payoff mechanisms to guarantee, via Lyapunov analysis, global

* This work was supported in part by AFOSR Grant FA9550-23-1-0467 and NSF Grants 2139713, 2139781, 2139982, 2135561, and 1846524.

Email addresses: `certorio@umd.edu` (Jair Certório), `nmartins@umd.edu` (Nuno C. Martins), `kcchang@usc.edu` (Kevin Chang), `nuzzo@usc.edu` (Pierluigi Nuzzo), `yshoukry@uci.edu` (Yasser Shoukry).

[1] Often called revision protocols in the population game literature.

Good news:

- *δ-passive rules form large convex cones.*
  - *Include common hybrid rules.*
- *The cones include the best response rule.*

Remaining challenge:

Positive correlation

$$\mathcal{V}(x,p) \neq 0 \iff p'\mathcal{V}(x,p) > 0$$

$$\Uparrow \qquad \not\Downarrow$$

δ-passivity

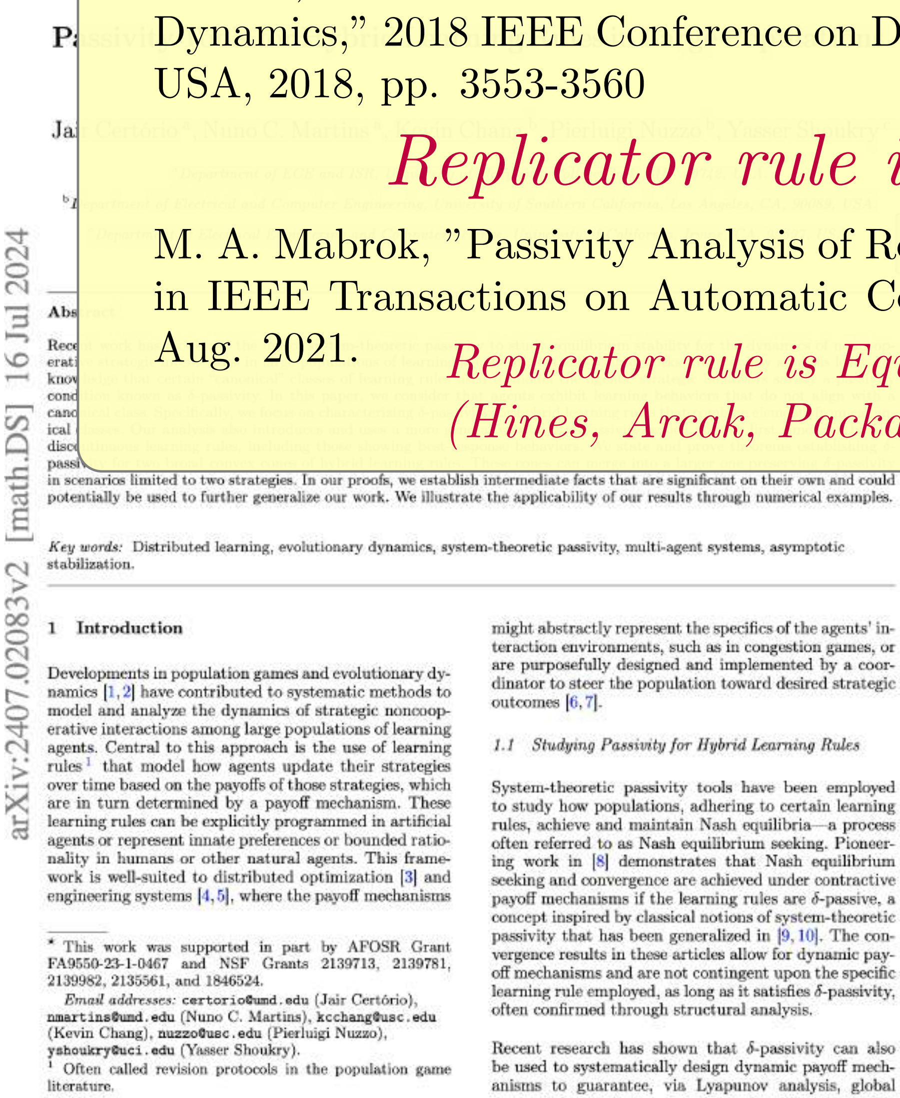
Generality of δ-passive learning rules

S. Park, J. S. Shamma and N. C. Martins, "Passivity and Evolutionary Game Dynamics," 2018 IEEE Conference on Decision and Control (CDC), Miami, FL, USA, 2018, pp. 3553-3560

Replicator rule is not δ-passive

M. A. Mabrok, "Passivity Analysis of Replicator Dynamics and Its Variations," in IEEE Transactions on Automatic Control, vol. 66, no. 8, pp. 3879-3884, Aug. 2021.

Replicator rule is Equilibrium Independent Passive (Hines, Arcak, Packard, Automatica'11).



arXiv:2407.02083v2 [math.DS] 16 Jul 2024

in scenarios limited to two strategies. In our proofs, we establish intermediate facts that are significant on their own and could potentially be used to further generalize our work. We illustrate the applicability of our results through numerical examples.

Key words: Distributed learning, evolutionary dynamics, system-theoretic passivity, multi-agent systems, asymptotic stabilization.

1 Introduction

Developments in population games and evolutionary dynamics [1, 2] have contributed to systematic methods to model and analyze the dynamics of strategic noncooperative interactions among large populations of learning agents. Central to this approach is the use of learning rules[1] that model how agents update their strategies over time based on the payoffs of those strategies, which are in turn determined by a payoff mechanism. These learning rules can be explicitly programmed in artificial agents or represent innate preferences or bounded rationality in humans or other natural agents. This framework is well-suited to distributed optimization [3] and engineering systems [4, 5], where the payoff mechanisms might abstractly represent the specifics of the agents' interaction environments, such as in congestion games, or are purposefully designed and implemented by a coordinator to steer the population toward desired strategic outcomes [6, 7].

1.1 Studying Passivity for Hybrid Learning Rules

System-theoretic passivity tools have been employed to study how populations, adhering to certain learning rules, achieve and maintain Nash equilibria—a process often referred to as Nash equilibrium seeking. Pioneering work in [8] demonstrates that Nash equilibrium seeking and convergence are achieved under contractive payoff mechanisms if the learning rules are δ-passive, a concept inspired by classical notions of system-theoretic passivity that has been generalized in [9, 10]. The convergence results in these articles allow for dynamic payoff mechanisms and are not contingent upon the specific learning rule employed, as long as it satisfies δ-passivity, often confirmed through structural analysis.

Recent research has shown that δ-passivity can also be used to systematically design dynamic payoff mechanisms to guarantee, via Lyapunov analysis, global

* This work was supported in part by AFOSR Grant FA9550-23-1-0467 and NSF Grants 2139713, 2139781, 2139982, 2135561, and 1846524.

Email addresses: certorio@umd.edu (Jair Certório), nmartins@umd.edu (Nuno C. Martins), kcchang@usc.edu (Kevin Chang), nuzzo@usc.edu (Pierluigi Nuzzo), yshoukry@uci.edu (Yasser Shoukry).

[1] Often called revision protocols in the population game literature.

...ex cones.

...ules.

...onse rule.

Remaining challenge:

Positive correlation

$$\mathcal{V}(x,p) \neq 0 \iff p'\mathcal{V}(x,p) > 0$$

$$\Uparrow \qquad \not\Downarrow$$

δ-passivity

Generality of δ-passive learning rules

S. Park, J. S. Shamma and N. C. Martins, "Passivity and Evolutionary Game Dynamics," 2018 IEEE Conference on Decision and Control (CDC), Miami, FL, USA, 2018, pp. 3553-3560

Replicator rule is not δ-passive

M. A. Mabrok, "Passivity Analysis of Replicator Dynamics and Its Variations," in IEEE Transactions on Automatic Control, vol. 66, no. 8, pp. 3879-3884, Aug. 2021.

Replicator rule is Equilibrium Independent Passive (Hines, Arcak, Packard, Automatica'11).

Remaining challenge:

Well-behaved learning rules satisfy:

Nash stationarity

$$\mathcal{V}(x,p) = 0 \iff x \in \arg\max_{\tilde{x}\in\mathbb{X}} \tilde{x}'p$$

Positive correlation

$$\mathcal{V}(x,p) \neq 0 \iff p'\mathcal{V}(x,p) > 0$$

δ-passivity does not hold for all well-behaved rules.

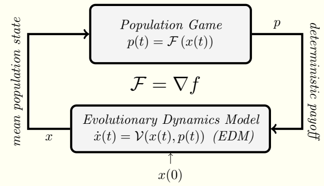
Exploring Counterclockwise Dissipativity

(Sandholm, 2001)

Nash stationarity
$\mathcal{V}(x,p) = 0 \iff x \in \arg\max_{\tilde{x}\in\mathbb{X}} \tilde{x}'p$

Positive correlation
$\mathcal{V}(x,p) \neq 0 \iff p'\mathcal{V}(x,p) > 0$

$+$



$$x(t) \underset{t\to\infty}{\longrightarrow} \mathbb{NE}(\mathcal{F}) = \mathbb{KKT}(f)$$

Contractive Games
(Hofbauer & Sandholm, 2009)

$$(x-\tilde{x})^T\big(\mathcal{F}(x) - \mathcal{F}(\tilde{x})\big) \leq 0$$

δ-passivity (Dynamic Payoffs)
(Fox & Shamma, 2013)
(Park, Martins, Shamma, tutorial 2019)
(Arcak & Martins, 2021)

Counterclockwise Dissipativity, Potential Games and Evolutionary Nash Equilibrium Learning

Nuno C. Martins, *Senior Member, IEEE*, Jair Certório, *Student Member, IEEE*, and Matthew S. Hankins, *Student Member, IEEE*

arXiv:2408.00647v1 [cs.GT] 1 Aug 2024

***Abstract*—We use system-theoretic passivity methods to study evolutionary Nash equilibria learning in large populations of agents engaged in strategic, non-cooperative interactions. The agents follow learning rules (rules for short) that capture their strategic preferences and a payoff mechanism ascribes payoffs to the available strategies. The population's aggregate strategic profile is the state of an associated evolutionary dynamical system. Evolutionary Nash equilibrium learning refers to the convergence of this state to the Nash equilibria set of the payoff mechanism. Most approaches consider memoryless payoff mechanisms, such as potential games. Recently, methods using δ-passivity and equilibrium independent passivity (EIP) have introduced dynamic payoff mechanisms. However, δ-passivity does not hold when agents follow rules exhibiting "imitation" behavior, such as in replicator dynamics. Conversely, EIP applies to the replicator dynamics but not to δ-passive rules. We address this gap using counterclockwise dissipativity (CCW). First, we prove that continuous memoryless payoff mechanisms are CCW if and only if they are potential games. Subsequently, under (possibly dynamic) CCW payoff mechanisms, we establish evolutionary Nash equilibrium learning for any rule within a convex cone spanned by imitation rules and continuous δ-passive rules.**

I. Introduction

Recent developments in population games and evolutionary dynamics [1], [2] have introduced new systematic methods to model and analyze the dynamics of strategic, non-cooperative interactions among large populations of learning agents. Learning rules (rules for short), also referred to as strategy revision protocols, describe how agents update their strategies based on the payoffs these strategies yield. The rules, which typically seek strategies with higher payoffs, can be programmed into artificial agents or represent the preferences or bounded rationality of humans and other natural agents.

The agents are nondescript and the population's aggregate strategic profile is represented as a population state vector whose entries are the proportions of the population selecting the available strategies. As agents revise their strategies, the population state changes over time according to an evolutionary dynamic model resulting from the rule in effect. A payoff mechanism generates the strategies' payoffs and might abstractly represent the specifics of the agents' interaction environments, such as in congestion games [3], or be designed by a coordinator to nudge the population state towards desired equilibria [4], [5]. This framework has been suitable for distributed optimization [6] and engineering systems [7], [8].

We seek to understand how populations that follow certain rules achieve and maintain Nash equilibria, a process we call evolutionary Nash equilibrium learning to distinguish it from Nash equilibrium seeking [9] that is used more generally when the learning process is not necessarily expressible using evolutionary dynamics, such as in [10]. Specifically, we say that a population evolutionarily learns Nash equilibria when its population state converges to the Nash equilibria set appropriately defined for the payoff mechanism. Establishing this property is crucial because, when it holds, the Nash equilibria set can be used to predict the long-term evolution of the population state. In certain applications where one has the authority to design the payoff mechanism, the Nash equilibria have inherent optimality properties for potential games [11, §3 and §5] and may even be selected to satisfy performance requirements [5], [12].

A. Background On Passivity Approaches

As evident from the synopsis [2], most research on evolutionary learning of Nash equilibria has focused on memoryless payoff mechanisms until recently. Of particular relevance is the focus on memoryless payoff mechanisms that are potential games [11]. Following this, contractive games were introduced [13], which have the desirable properties of concave potential games without necessarily being potential.

A novel approach utilizing system theoretic passivity methods extended the results in [13] to include dynamic payoff mechanisms [14] that are the dynamic analogs of contractive games. This extension was achieved through the introduction of the concept of δ-passivity, which was further explored in [15], [16] and applied in various design applications [5], [12]. Recent work established δ-passivity for convex cones of hybrid rules, formed as conic combinations of rules from canonical classes known to satisfy δ-passivity [17].

An important gap: Despite the significant success with δ-passivity approaches, [18, Proposition III.5] proved that the replicator rule associated with the well-known replicator dynamics is not δ-passive. More generally, no rule exhibiting "imitation" behavior has been shown to be δ-passive. Conversely, [19] proved that the replicator rule is equilibrium independent passive (EIP), a passivity notion not known to hold for the δ-passive rules in [17].

B. Main Objective

This paper aims to address the gap identified above by developing a system-theoretic passivity approach that is compatible with a broader set of rules, including both the continuous hybrid δ-passive rules described in [17] and rules exhibiting imitation behavior, such as the replicator rule. Specifically, we seek an approach that can accommodate conic combinations of imitation-based rules and the continuous δ-passive rules considered in [17]. This goal is particularly important when we are given or intend to design a payoff mechanism and seek evolutionary Nash equilibrium learning guarantees without precise knowledge of the population's rule. The broader the class

This work was supported by AFOSR Grant FA9550-23-1-0467 and NSF Grants 2139713 and 2135561.

Nuno C. Martins, Jair Certório and Matthew S. Hankins are with the Department of Electrical and Computer Engineering and the Institute For Systems Research at the University of Maryland, College Park (e-mail: nmartins@umd.edu, certorio@umd.edu, msh@umd.edu).

Exploring Counterclockwise Dissipativity

$$\mathbf{u} : [0,\infty) \to \mathbb{U} \longrightarrow \boxed{\mathfrak{G} : \mathbf{u} \mapsto \mathbf{y}} \longrightarrow \mathbf{y} : [0,\infty) \to \mathbb{Y}$$

$\mathbf{u} \in \mathscr{U}$ and $\mathbf{y} \in \mathscr{Y}$, where $\mathscr{U}$ and $\mathscr{Y}$ are the input and output sets.

We assume all elements of $\mathscr{U}$ and $\mathscr{Y}$ are differentiable a.e.

We assume u(t) and y(t) have the same dimension, so we can write $u'(t)y(t)$

D. Angeli, "Systems with counterclockwise input-output dynamics," in IEEE Transactions on Automatic Control, vol. 51, no. 7, pp. 1130-1143, July 2006.

The system $\mathfrak{G}$ is counterclockwise dissipative (CCW) when:

$$\inf_{T>0,\mathbf{u}\in\mathscr{U}} \int_0^T u'(t)\dot{y}(t)dt \; > \; -\infty$$

Exploring Counterclockwise Dissipativity

$\mathbf{u} : [0, \infty) \to \mathbb{U}$ → $\mathfrak{G} : \mathbf{u} \mapsto \mathbf{y}$ → $\mathbf{y} : [0, \infty) \to \mathbb{Y}$

$\mathbf{u} \in \mathscr{U}$ and $\mathbf{y} \in \mathscr{Y}$, where $\mathscr{U}$ and $\mathscr{Y}$ are the input and output sets.

We assume all elements of $\mathscr{U}$ and $\mathscr{Y}$ are differentiable a.e.

We assume u(t) and y(t) have the same dimension, so we can write $u'(t)y(t)$


A. Lanzon and I. R. Petersen, "Stability Robustness of a Feedback Interconnection of Systems With Negative Imaginary Frequency Response," in IEEE Transactions on Automatic Control, vol. 53, no. 4, pp. 1042-1046, May 2008.

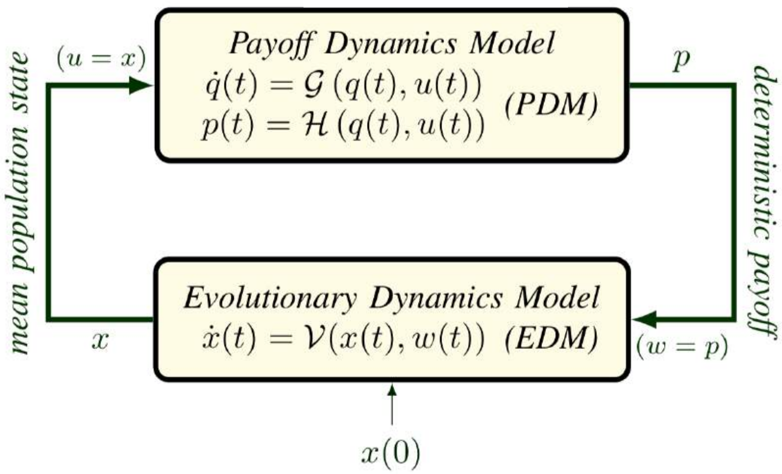
A linear time invariant system $\mathfrak{G}$ with rational, proper and stable transfer function matrix $G(s)$ is negative imaginary (NI) when:

$$j\Big(G(jw) - G'(-jw)\Big) \succeq 0.$$

An NI system is CCW.

Exploring Counterclockwise Dissipativity

$$\mathfrak{F} : \mathbf{x} \mapsto \mathbf{p}$$



Input and output sets

It suffices to consider Lipschitz continuous (w.r.t. time) population state trajectories $\mathbf{x}$, leading to the set

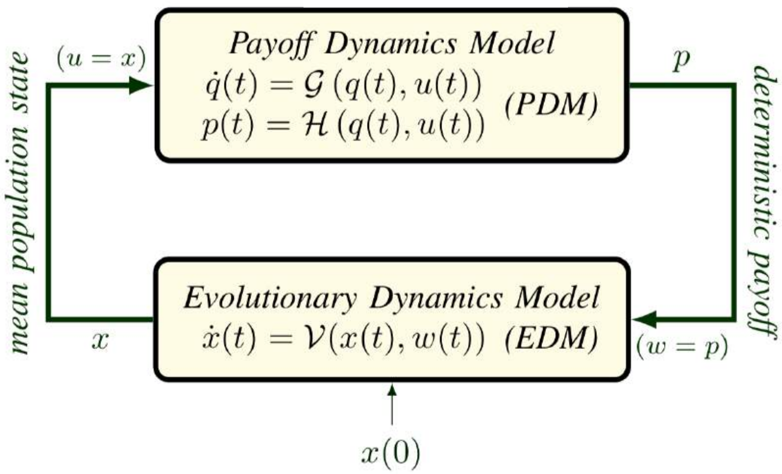
$$\mathscr{X} := \left\{ \mathbf{x} : [0, \infty] \to \mathbb{X} \;\middle|\; \mathbf{x} \text{ is Lipschitz continuous} \right\}.$$

It also suffices to consider Lipschitz continuous (w.r.t time) payoff trajectories $\mathbf{p}$, leading to the set

$$\mathscr{P} := \left\{ \mathbf{p} : [0, \infty] \to \mathbb{R}^n \;\middle|\; \mathbf{p} \text{ is Lipschitz continuous} \right\}.$$

Exploring Counterclockwise Dissipativity

$$\mathfrak{F} : \mathbf{x} \mapsto \mathbf{p}$$



Input and output sets

It suffices to consider Lipschitz continuous (w.r.t. time) population state trajectories $\mathbf{x}$, leading to the set

$$\mathscr{X} := \left\{ \mathbf{x} : [0, \infty] \to \mathbb{X} \mid \mathbf{x} \text{ is Lipschitz continuous } \right\}.$$

It also suffices to consider Lipschitz continuous (w.r.t time) payoff trajectories $\mathbf{p}$, leading to the set

$$\mathscr{P} := \left\{ \mathbf{p} : [0, \infty] \to \mathbb{R}^n \mid \mathbf{p} \text{ is Lipschitz continuous} \right\}.$$

Well-behaved learning rules satisfy:

Nash stationarity

$$\mathcal{V}(x, p) = 0 \iff x \in \arg\max_{\tilde{x} \in \mathbb{X}} \tilde{x}' p$$

Positive correlation

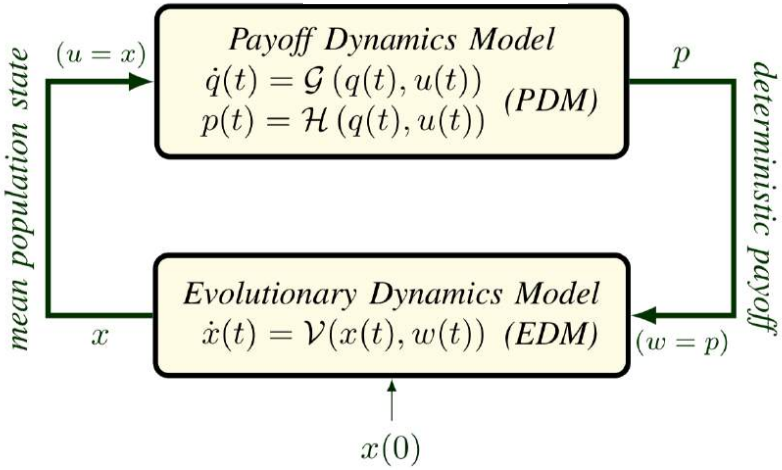
$$\mathcal{V}(x, p) \neq 0 \iff p' \mathcal{V}(x, p) > 0$$

Any well-behaved rule leads to a CCW (EDM).

$$\inf_{T > 0, \mathbf{p} \in \mathscr{P}} \int_0^T p'(t) \dot{x}(t) dt \geq 0$$

Exploring Counterclockwise Dissipativity

$$\mathfrak{F} : \mathbf{x} \mapsto \mathbf{p}$$



Input and output sets

It suffices to consider Lipschitz continuous (w.r.t. time) population state trajectories $\mathbf{x}$, leading to the set

$$\mathscr{X} := \left\{ \mathbf{x} : [0, \infty] \to \mathbb{X} \;\middle|\; \mathbf{x} \text{ is Lipschitz continuous } \right\}.$$

It also suffices to consider Lipschitz continuous (w.r.t time) payoff trajectories $\mathbf{p}$, leading to the set

$$\mathscr{P} := \left\{ \mathbf{p} : [0, \infty] \to \mathbb{R}^n \;\middle|\; \mathbf{p} \text{ is Lipschitz continuous} \right\}.$$

A payoff mechanism $\mathfrak{F}$ is CCW when the following holds:

$$\inf_{T>0, \mathbf{x} \in \mathscr{X}} \int_0^T x'(t)\dot{p}(d)dt > -\infty$$

CCW $\mathfrak{F}$ and Well-Behaved Rules: Main Result

Well-behaved rules:

Nash stationarity

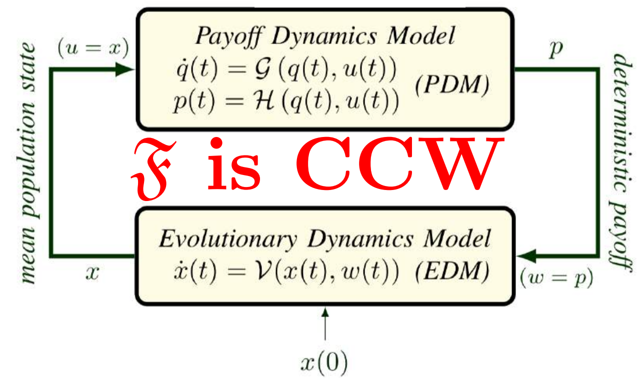
$$\mathcal{V}(x,p) = 0 \iff x \in \arg\max_{\tilde{x} \in \mathbb{X}} \tilde{x}' p$$

Positive correlation

$$\mathcal{V}(x,p) \neq 0 \iff p' \mathcal{V}(x,p) > 0$$

&

$\mathfrak{F} : \mathbf{x} \mapsto \mathbf{p}$

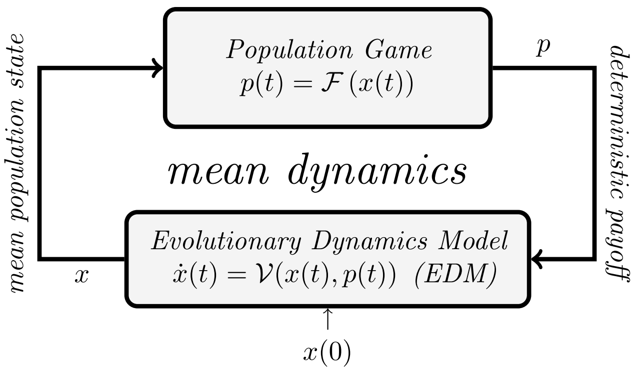


THEOREM

$$x(t) \underset{t \to \infty}{\longrightarrow} \mathbb{NE}(\mathcal{F})$$

evolutionary Nash equilibrium learning

CCW Payoff Mechanisms: Memoryless Case



potential games

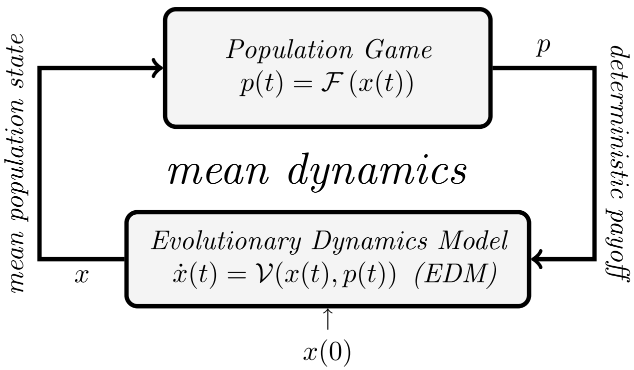
robust

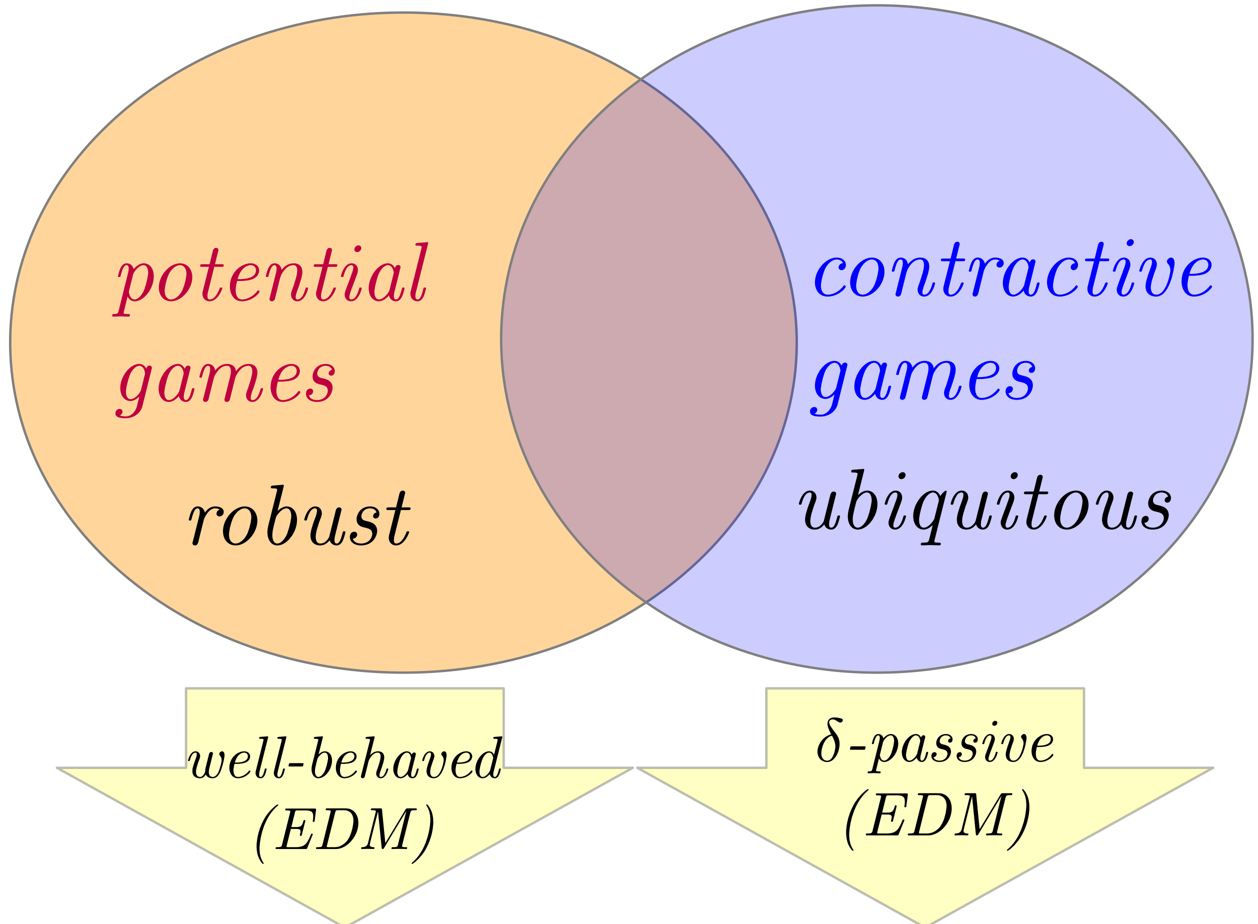
well-behaved (EDM)

$$x(t) \underset{t\to\infty}{\longrightarrow} \mathbb{NE}(\mathcal{F})$$

evolutionary Nash equilibrium learning

CCW Payoff Mechanisms: Memoryless Case

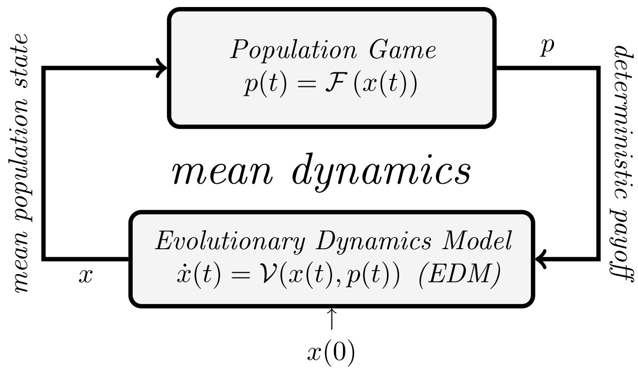


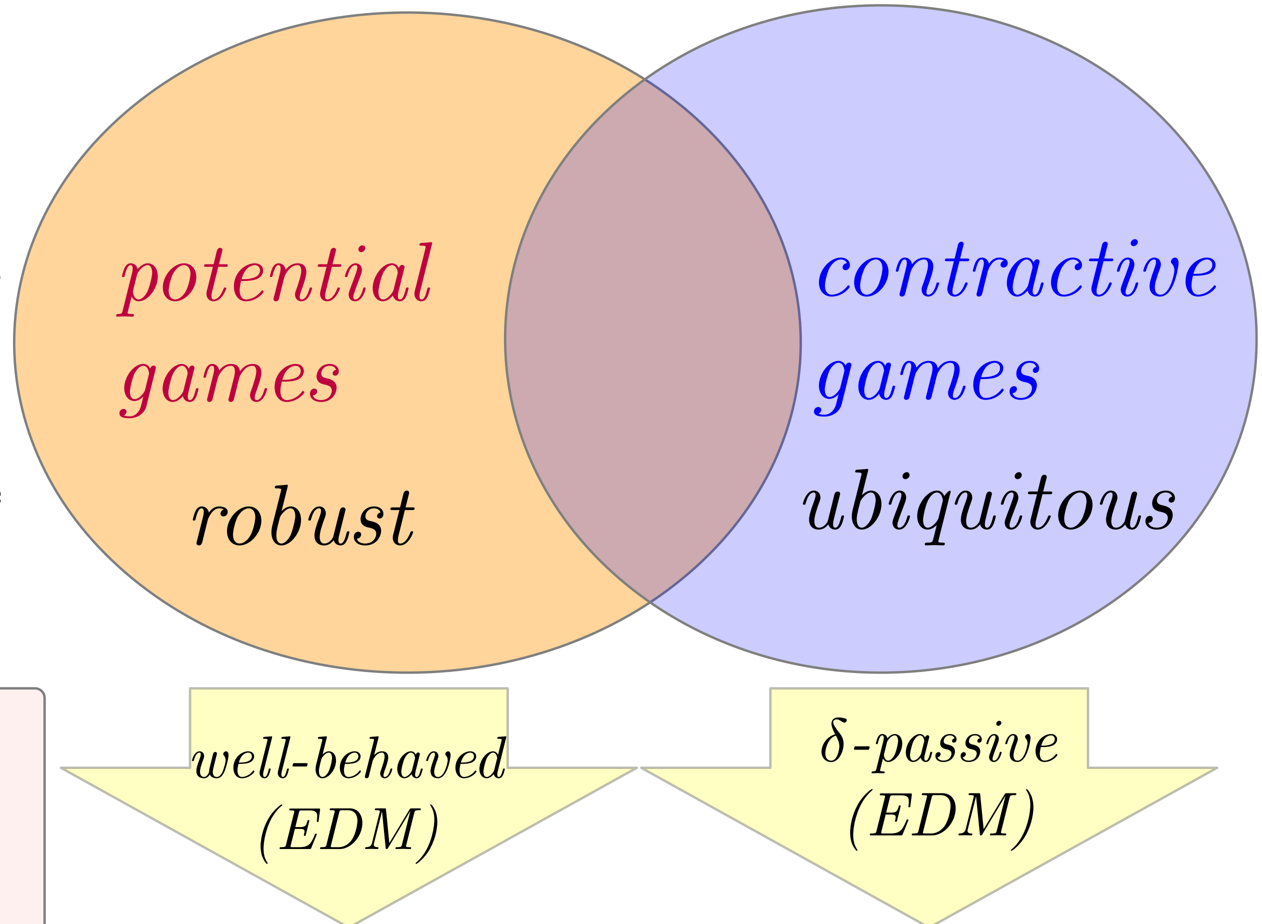


$$x(t) \xrightarrow[t\to\infty]{} \mathbb{NE}(\mathcal{F})$$

evolutionary Nash equilibrium learning

CCW Payoff Mechanisms: Memoryless Case





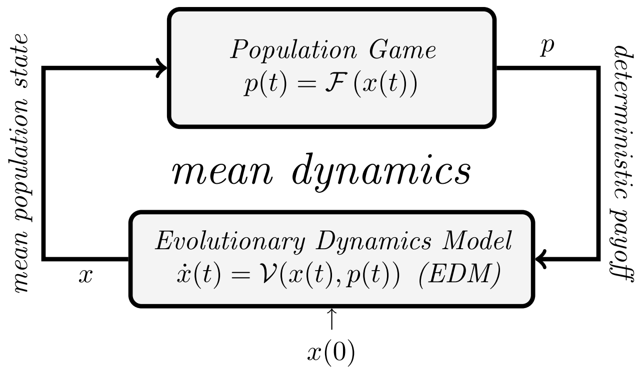
A payoff mechanism $\mathfrak{F}$ is CCW when the following holds:

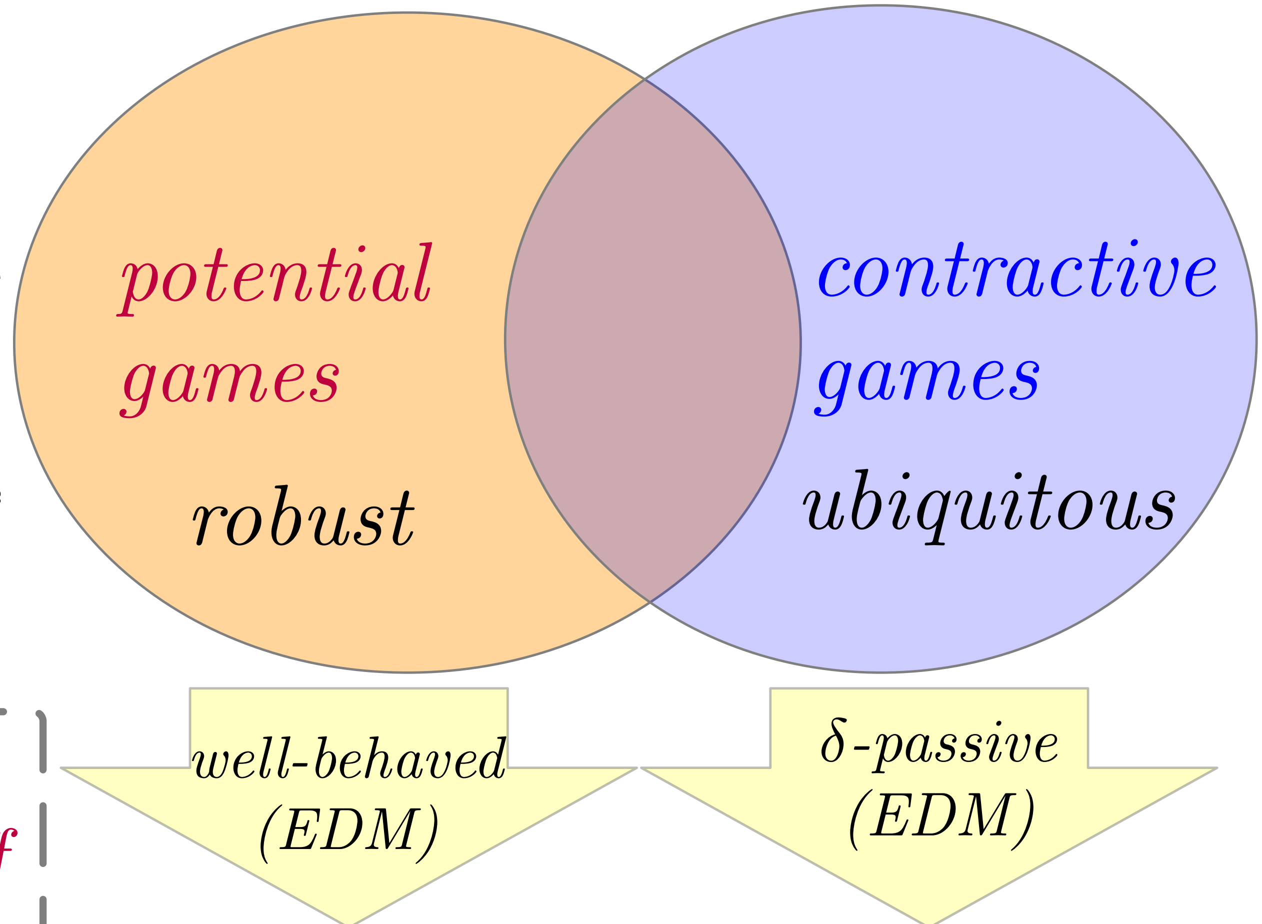
$$\inf_{T>0,\mathbf{x}\in\mathscr{X}} \int_0^T x'(t)\dot{p}(d)dt > -\infty$$

$$x(t) \xrightarrow[t\to\infty]{} \mathbb{NE}(\mathcal{F})$$

evolutionary Nash equilibrium learning

CCW Payoff Mechanisms: Memoryless Case





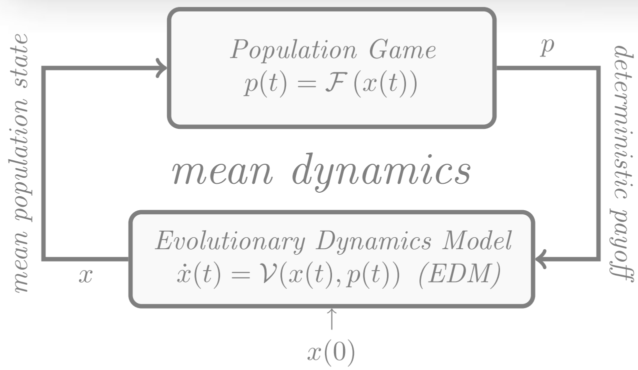
$$x(t) \xrightarrow[t \to \infty]{} \mathbb{NE}(\mathcal{F})$$

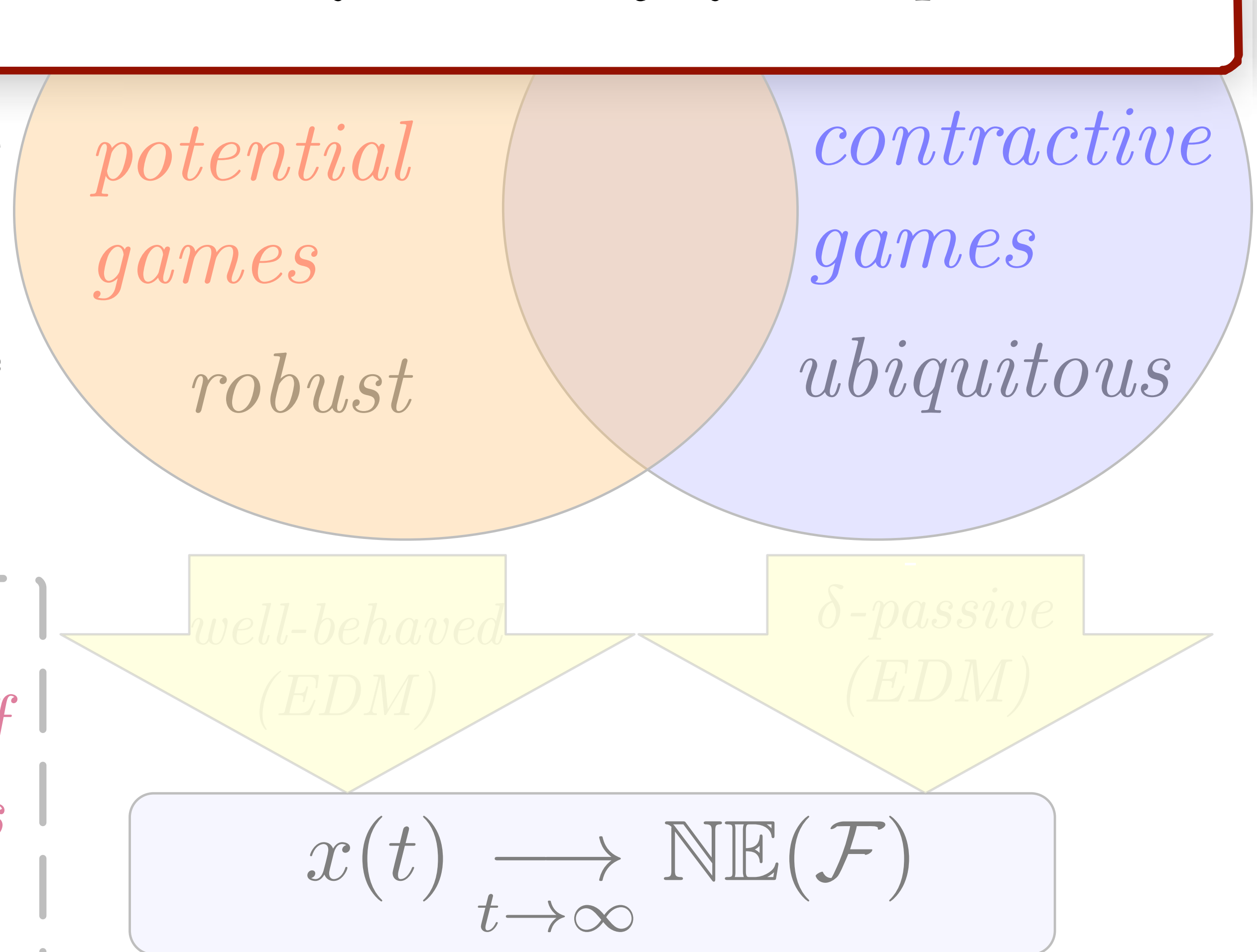
evolutionary Nash equilibrium learning

Where are CCW memoryless payoff mechanism in this picture?

CCW Payoff Mechanisms: Memoryless Case

Theorem: *A game is CCW if and only if it is potential*



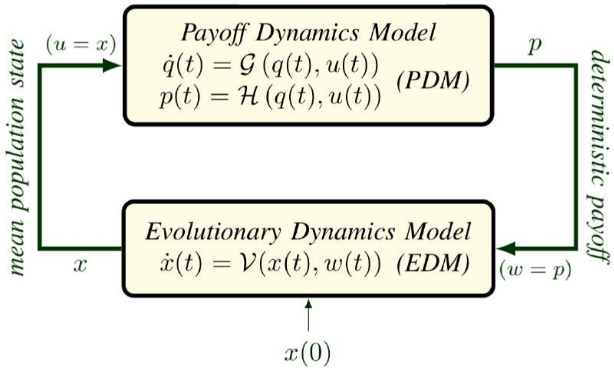


Where are CCW memoryless payoff mechanism in this picture?

evolutionary Nash equilibrium learning

CCW Payoff Mechanisms: Dynamic Case

$\mathfrak{F} : \mathbf{x} \mapsto \mathbf{p}$



$-\mathfrak{F}$ is δ-passive

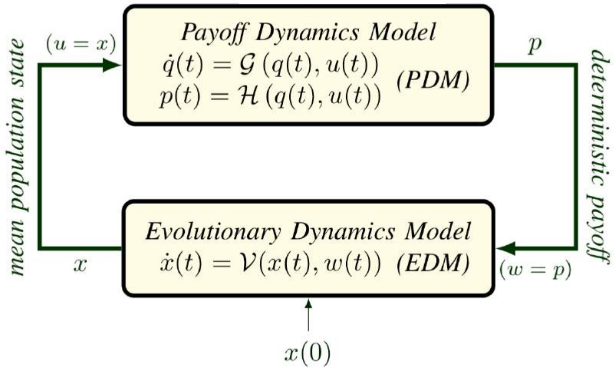
δ-passive (EDM)

$$x(t) \underset{t \to \infty}{\longrightarrow} \mathbb{NE}(\mathcal{F})$$

evolutionary Nash equilibrium learning

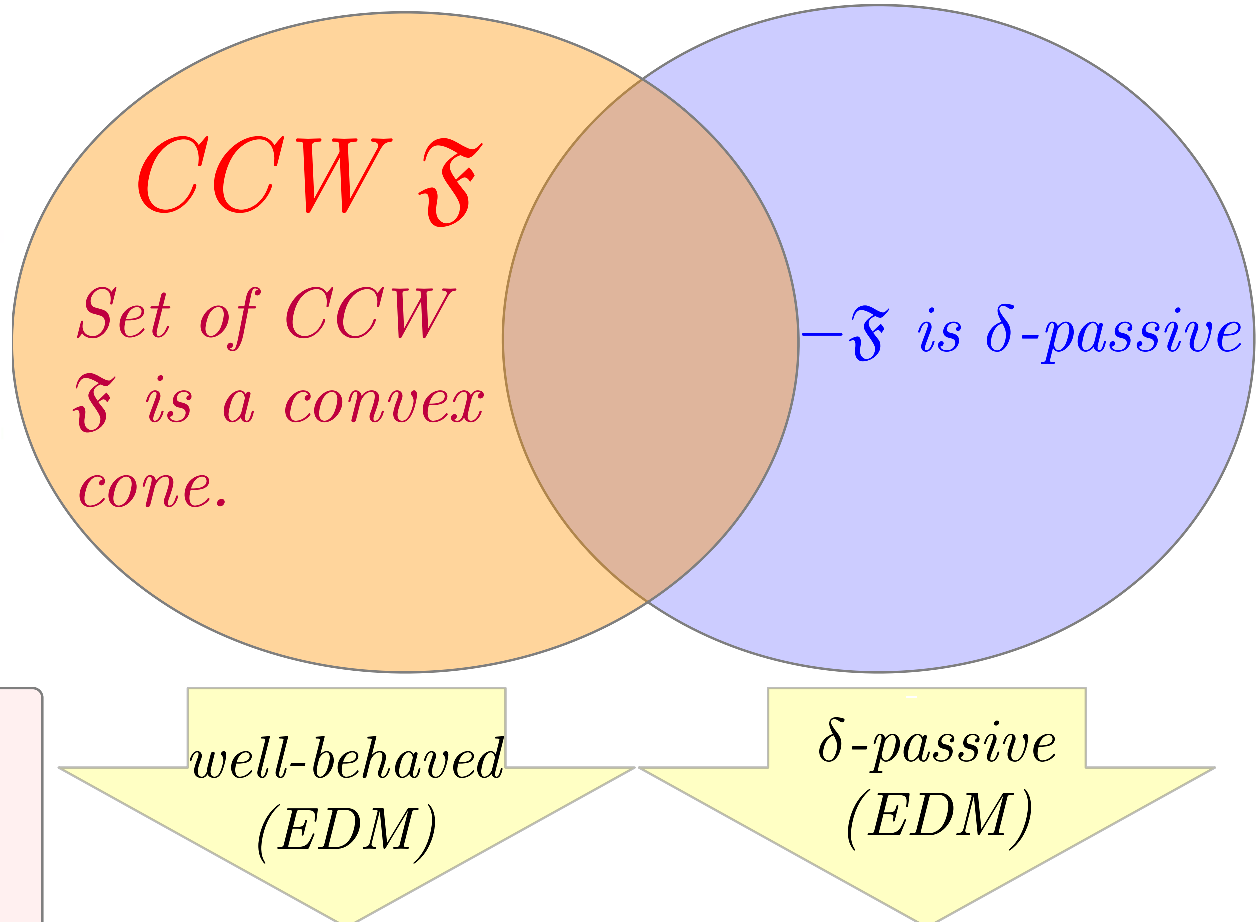
CCW Payoff Mechanisms: Dynamic Case

$\mathfrak{F} : \mathbf{x} \mapsto \mathbf{p}$



A payoff mechanism $\mathfrak{F}$ is CCW when the following holds:

$$\inf_{T>0, \mathbf{x} \in \mathscr{X}} \int_0^T x'(t)\dot{p}(d)dt > -\infty$$



$$x(t) \xrightarrow[t \to \infty]{} \mathbb{NE}(\mathcal{F})$$

evolutionary Nash equilibrium learning

CCW Payoff Mechanisms: Dynamic Case

Set of CCW $\mathfrak{F}$ is a convex cone.

EXAMPLE: For a potential game $\mathcal{F}$, a given time-constant $\lambda^{-1} > 0$, b in $\mathbb{R}^n$, and $A = A'$ in $\mathbb{R}^{n \times n}$, consider

$$\dot{q}(t) = \lambda\big(Ax(t) + b - q(t)\big), \qquad t \geq 0, \quad q(0) = 0,$$
$$p(t) = \mathcal{F}\big(x(t)\big) + k\lambda\big(Ax(t) + b - q(t)\big)$$

where k is a real constant satisfying $kA \preceq 0$.

A payoff mechanism $\mathfrak{F}$ is CCW when the following holds:

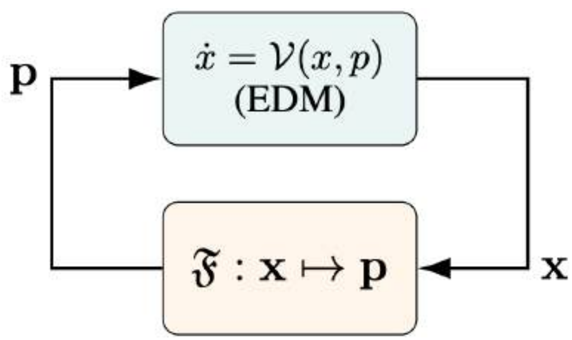
$$\inf_{T>0, \mathbf{x} \in \mathscr{X}} \int_0^T x'(t)\dot{p}(d)dt > -\infty$$

The example could be used to model:

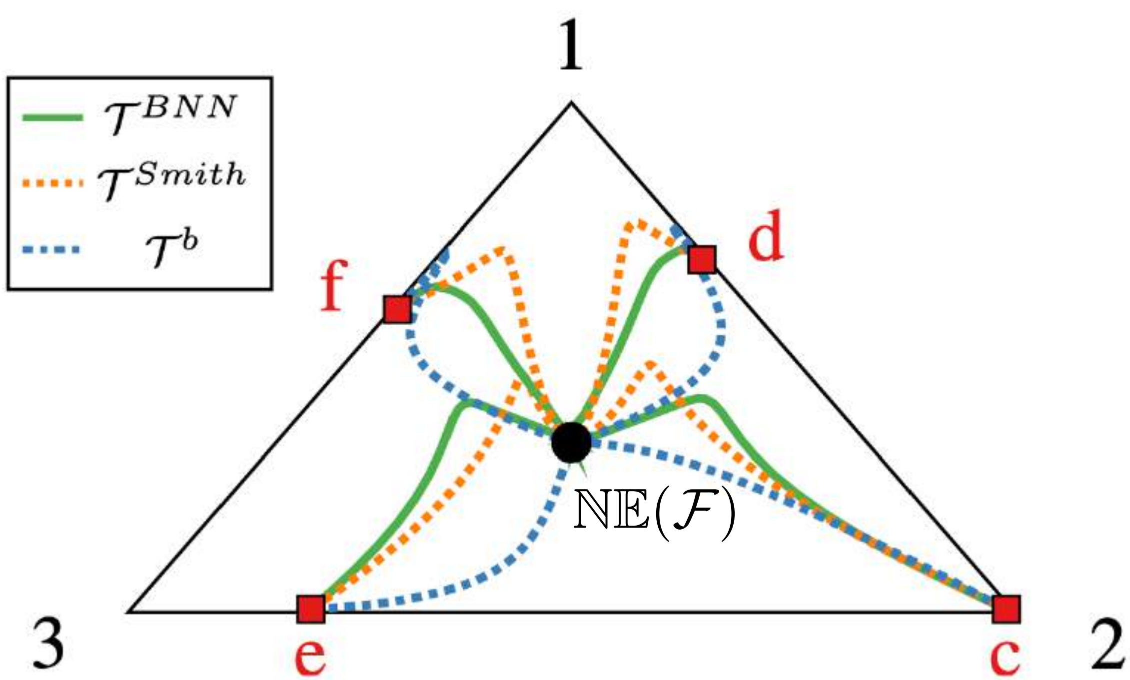
- *potential games disturbed by NI dynamics;*
- *anticipatory learning;*
- *payoff information diffusion.*

Simulation Example (3 strategies)

well-behaved learning rule $\mathcal{T}^{\mathrm{b}}_{ij}(x,p) = \underbrace{x_j[p_j - p_i]_+}_{\mathcal{T}^{\mathrm{R}}} + 0.01\,\underbrace{[p_j - p_i]_+}_{\mathcal{T}^{\mathrm{Smith}}},$



CCW anticipatory learning $\mathfrak{F}$



Other Contributions In The Reported Period

(Sandholm, 2001)

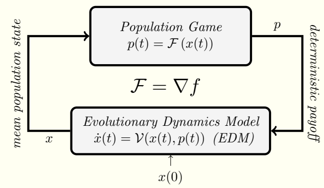
Nash stationarity

$\mathcal{V}(x,p) = 0 \iff x \in \arg\max_{\tilde{x} \in \mathbb{X}} \tilde{x}'p$

Positive correlation

$\mathcal{V}(x,p) \neq 0 \iff p'\mathcal{V}(x,p) > 0$

$+$



$$x(t) \underset{t \to \infty}{\longrightarrow} \mathbb{NE}(\mathcal{F}) = \mathbb{KKT}(f)$$

Strategy-dependent Rates and Erlang Revision Dynamics

★**(Kara, Martins, 2023)**

(Kara, Martins, Arcak, 2022)

Nash learning with constrained switching

★**(Kara, Martins, 2024)**

Contractive Games

(Hofbauer & Sandholm, 2009)

$$(x - \tilde{x})^T \big(\mathcal{F}(x) - \mathcal{F}(\tilde{x})\big) \leq 0$$

★ *Accepted or published*

δ-passivity (Dynamic Payoffs)

(Arcak & Martins, 2021)

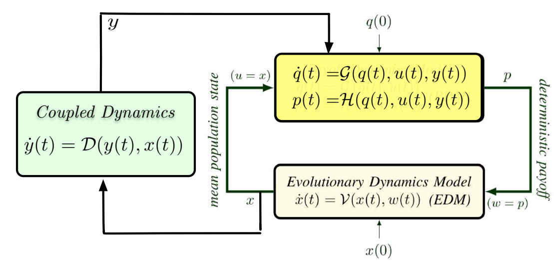
(Park, Martins, Shamma, tutorial 2019)

(Fox & Shamma, 2013)

Coupled Dynamics

★**(Martins, Certorio, La, Arcak, 2024)**

(Martins, Certorio, La, 2023)



Thanks for support by AFOSR
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Questions ?