

University of Groningen

Relative Best Response Dynamics in Finite and Convex Network Games

Govaert, Alain; Cenedese, Carlo; Grammatico, Sergio; Cao, Ming

Published in:
 Proceedings of the 58th IEEE Conference on Decision and Control

DOI:
[10.1109/CDC40024.2019.9029821](https://doi.org/10.1109/CDC40024.2019.9029821)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
 Publisher's PDF, also known as Version of record

Publication date:
 2019

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Govaert, A., Cenedese, C., Grammatico, S., & Cao, M. (2019). Relative Best Response Dynamics in Finite and Convex Network Games. In *Proceedings of the 58th IEEE Conference on Decision and Control* (pp. 3134-3139). [9029821] IEEE. <https://doi.org/10.1109/CDC40024.2019.9029821>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Relative Best Response Dynamics in finite and convex Network Games

Alain Govaert¹

Carlo Cenedese¹

Sergio Grammatico²

Ming Cao¹

Abstract—Motivated by theoretical and experimental economics, we propose novel evolutionary dynamics for games on networks, called the *h-Relative Best Response (h-RBR)* dynamics, that mixes the relative performance considerations of imitation dynamics with the rationality of best responses. Under such a class of dynamics, the players optimize their payoffs over the set of strategies employed by a time-varying subset of their neighbors. As such, the *h-RBR* dynamics share the defining non-innovative characteristic of imitation based dynamics and can lead to equilibria that differ from classic Nash equilibria. We study the asymptotic behavior of the *h-RBR* dynamics for both finite and convex games in which the strategy spaces are discrete and compact, respectively, and provide preliminary sufficient conditions for finite-time convergence to a generalized Nash equilibrium.

I. INTRODUCTION

Game theoretic scenarios in which players interact exclusively with a fixed group of neighbors traces back to the early 1990's when economists and biologists started to explore the effect of simple spatial structures in (probabilistic) decision making processes driven by rational best response processes and more biologically inspired imitation processes [1]–[3]. Later, the simple spatial structures were extended to arbitrary structures defined by graphs [4]–[6].

The long-run collective behavior of non-cooperative network games have been extensively studied for best response dynamics in which the players, given the history of plays of their neighbors, select a strategy that maximizes their own payoff. These extended research efforts have resulted in the identification of several classes of games that converge to a pure Nash equilibrium under a variety of such best response processes [7]–[10] and brought forth a number of algorithms that ensure convergence to an equilibrium [11]–[13]. Best response dynamics are “*innovative*” in the sense that, in order to optimize their payoffs, players are always able to select new strategies that are not played in the current strategy profile. They are in line with classic economic theories that support the idea that absolute optimization (or rational behavior) is a natural result of evolutionary forces [14]. Recently, the systems and control community has been interested in the analysis of dynamical systems driven by *imitation* [15]–[17]. Such dynamics are “*non-innovative*”: players can only select

strategies that already exist in the networked population. Therefore, non-innovative dynamics can lead to equilibrium concepts that differ from traditional Nash equilibria. In [18], [19], the authors studied an evolutionary process where the players, most of the time, choose a best response from the set of strategies that exist in the entire population strategy profile. In [19], this evolutionary process was simply referred to as imitation. Perhaps a more suitable name was proposed in [18], where such a revision was called a *Relative Best Response (RBR)*. In fact, RBR combines the non-innovative nature of pure imitation with the rationality of best response. Such dynamics match classic economic studies in which rather than absolute performance, the players opt for *relative performance*, that prove decisive in the long run [20]. Experimental evidences of such behaviour are documented in [21].

In this paper, we extend the RBR dynamics for finite games, with finite discrete strategy sets, to a spatial version where the players choose a best response from the current set of strategies of their neighbors. In this set-up, even though the feasible strategy sets are state-dependent, the players interact and relate their success exclusively with a *fixed* group of neighbours. So, the *replacement graph* [22], which determines the feasible strategy set of a player, and the *interaction graph* [22], which determines the payoff of a player, are fixed and equal. We generalize this spatial version of the RBR dynamics to the *h-RBR*, where players relate their success only to the subset of neighbors that obtain the *h*-highest payoffs. Such a generalized set-up is motivated by a pure imitation process in which, typically, only the strategies of the most successful neighbors are taken into account [3]. For *h-RBR*, next to a state-dependent feasible strategy set, the players relate their success to a state-dependent subset of their neighbors. This corresponds to a process in which the *replacement graph* is state-dependent and the *interaction graph* is fixed. Furthermore, a variation of the *h-RBR* dynamics is proposed for *convex* games, where the admissible strategy sets of the players are compact spaces that result from the convex hull of the strategies of the successful neighbors. For both finite and convex games, we analyze the asymptotic behavior of the *h-RBR* dynamics and provide sufficient conditions for reaching a generalized Nash equilibrium. To the best of the authors' knowledge, this is the first paper introducing the concept of (*h*-)RBR for both finite and convex games on arbitrarily connected networks. We believe that the RBR and *h-RBR* dynamics for games on networks proposed in this paper are complementary to the existing works on convergence of best response dynamics and provide an interesting mix between pure imitation and

This work was partially supported by the European Research Council (ERC-StG, 307207; ERC-StG COSMOS, 802348), and the Netherlands Organization for Scientific Research (NWO-VIDI, 14134; NWO-TOP OMEGA, 613.001.702; NWO-ESI-BIDA P2P-TALES, 647.003.003.)

¹ Engineering and Technology Institute Groningen (ENTEG), Faculty of Science and Engineering, University of Groningen, 9747 AG Groningen, The Netherlands. E-mail: a.govaert@rug.nl, c.cenedese@rug.nl and m.cao@rug.nl

² Delft Center for Systems and Control, TU Delft, 2628 CD Delft, The Netherlands. E-mail: s.grammatico@tudelft.nl

rational decision making.

II. NOTATION AND PRELIMINARIES

A. Notation

The set of real, positive, and non-negative numbers are denoted by \mathbb{R} , $\mathbb{R}_{>0}$, $\mathbb{R}_{\geq 0}$, respectively. The set of natural numbers is denoted by \mathbb{N} and the set of integers is denoted by \mathbb{Z} . For a square matrix $A \in \mathbb{R}^{n \times n}$, its transpose is denoted by A^\top , $[A]_i$ is the i -th row of the matrix and $[A]_{ij}$ the element in the i -th row and j -th column. Given two vectors $x, y \in \mathbb{R}^n$, $x > y$ ($x \geq y$) describes an element wise inequality. The identity matrix is denoted by $I_n \in \mathbb{R}^{n \times n}$. For a vector $s \in \mathbb{R}^n$ we denote the i^{th} element as s_i . For $s_1, \dots, s_N \in \mathbb{R}^n$ and $\mathcal{I} = \{1, \dots, N\}$, the collective vector is denoted as $\mathbf{s} := \text{col}(\{s_i\}_{i \in \mathcal{I}}) = [s_1^\top, \dots, s_N^\top]^\top$ and $\mathbf{s}_{-i} := \text{col}(\{s_j\}_{j \in \mathcal{I} \setminus \{i\}}) = [s_1^\top, \dots, s_{i-1}^\top, s_{i+1}^\top, \dots, s_N^\top]^\top$. Equivalently, we also use the notation $\mathbf{s} = (s_i, \mathbf{s}_{-i})$. A network with node set \mathcal{I} and edgese \mathcal{E} is indicated by $\mathcal{G} = (\mathcal{I}, \mathcal{E})$. For each $i \in \mathcal{I}$, the edgese $\mathcal{E} \subseteq \mathcal{I} \times \mathcal{I}$ defines a set of neighbors indicated by $\mathcal{N}_i = \{j \in \mathcal{I} : (i, j) \in \mathcal{E}\} \setminus \{i\}$. For a set $\mathcal{M} = \{m_1, \dots, m_N\}$ of N points in \mathbb{R}^n , organized in the collective vector $\mathbf{m} = \text{col}(\{m_i\}_{i \in \mathcal{M}}) \in \mathbb{R}^{Nn}$ we denote the convex hull of \mathcal{M} by $\text{conv}(\mathcal{M}) = \{(\alpha \otimes \mathbf{1}_n)^\top \mathbf{m} \mid \alpha \in \mathbb{R}_{\geq 0}^N, \alpha^\top \mathbf{1} = 1\}$. Given a set \mathcal{A} of M elements, its cardinality is indicated as $|\mathcal{A}| = M$. Given a set \mathcal{B} with N non-repeated elements, the single valued function $\max^k(\mathcal{B})$, where $k \leq N$, evaluates the k -th highest value in the set. For a given polyhedron \mathcal{C} , we denote the vertices of \mathcal{C} as $\text{vert}(\mathcal{C})$.

B. Network Games

Let us introduce the three main ingredients of non-cooperative network games: the network structure, the strategy space and the combined payoff function. The strategy space is defined for both finite games, and convex games, in which the strategy spaces are discrete and compact, respectively.

The network structure: Let $\mathcal{G} = (\mathcal{I}, \mathcal{E})$ be an undirected graph whose node set $\mathcal{I} = \{1, \dots, N\}$ represents players. The edge set $\mathcal{E} \subseteq \mathcal{I} \times \mathcal{I}$, represents the player interaction topology.

The strategy space: Let \mathcal{S}_i denote the set of strategies for player $i \in \mathcal{I}$. The strategy space of the game is defined as the Cartesian product of the strategy sets of the players, i.e., $\mathcal{S} = \prod_{i \in \mathcal{I}} \mathcal{S}_i$. A *strategy profile* of the game is an element of this set, hence a collective vector $\mathbf{s} := \text{col}(\{s_i\}_{i \in \mathcal{I}}) \in \mathcal{S}$.

The payoff functions: Let $\pi_i : \mathcal{S} \rightarrow \mathbb{R}$ indicate the local payoff function of player i . The combined payoff function $\pi : \mathcal{S} \rightarrow \mathbb{R}^N$ maps each strategy profile $\mathbf{s} \in \mathcal{S}$ to a payoff vector $\pi(\mathbf{s}) = \text{col}(\{\pi_i(\mathbf{s})\}_{i \in \mathcal{I}})$ whose elements correspond to the payoffs that the players receive for a single round interaction. In network games, the spatial structure is incorporated into the payoff function π . Thus, the network structure determined by the graph \mathcal{G} , the strategy space \mathcal{S} , and combined payoff function π define the network game as the triplet $\Gamma = (\mathcal{G}, \mathcal{S}, \pi)$. Throughout the paper, we consider the following assumption on the strategy sets.

Standing Assumption 1 (Identical strategy sets). *All players have the same strategy set \mathcal{S} , i.e., $\mathcal{S}_i = \mathcal{S}$ for all $i \in \mathcal{I}$.* \square

Standing Assumption 1 naturally allows players to imitate each other and is in fact common in imitation-like dynamics [3], [15], [16].

Convex and finite games: We say that Γ is a *finite* game if the strategy set of each player is a finite discrete set such that $\mathcal{S} \subset \mathbb{Z}$ and $\mathcal{S} \subset \mathbb{Z}^N$. We denote a finite game as Γ_f . On the other hand, we say Γ is a *convex* game if the strategy set of each player is a non-empty, convex subset of \mathbb{R}^n , i.e., $\mathcal{S} \subset \mathbb{R}^n$ and $\mathcal{S} \subset \mathbb{R}^{Nn}$. We denote a convex game by Γ_c . The convexity assumption over the strategy set for convex games is common in monotone games' literature [11], [23].

III. h -RELATIVE BEST RESPONSE DYNAMICS

Before defining the h -RBR dynamics, for the purpose of comparison, we give the definition of a best-response mapping.

Definition 1 (Best-Response mapping). *The best-response (BR) mapping of player $i \in \mathcal{I}$ is*

$$\mathcal{B}_i(\mathbf{s}_{-i}) := \underset{y \in \mathcal{S}}{\text{argmax}} \pi_i(y, \mathbf{s}_{-i}). \quad (1)$$

\square

The defining distinction of a *relative* best response is that, instead of optimizing over a fixed strategy set \mathcal{S} , player $i \in \mathcal{I}$ optimizes its payoffs over some *feasible* subset of \mathcal{S} that depends on the strategies of the neighbors of i and s_i itself. For a game Γ and a strategy profile $\mathbf{s} \in \mathcal{S}$, we denote the feasible strategy set for player $i \in \mathcal{I}$ by $\mathcal{F}_i(\mathbf{s}) \subseteq \mathcal{S}$. For a finite game Γ_f , the feasible strategy set of player $i \in \mathcal{I}$ is simply determined as the local set of strategies, i.e.,

$$\mathcal{F}_i^f(\mathbf{s}) := \{s_j \in \mathbf{s} \mid j \in \mathcal{N}_i\} \cup \{s_i\} \subseteq \mathcal{S}. \quad (2)$$

Instead, for a convex game Γ_c , the feasible strategy set is formed as the convex hull of a finite number of points. Thus, the strategy sets are convex and compact subsets of \mathbb{R}^n . Formally, the feasible strategy set for player $i \in \mathcal{I}$ is determined as

$$\mathcal{F}_i^c(\mathbf{s}) = \text{conv}(\mathcal{F}_i^f) \subseteq \mathcal{S}. \quad (3)$$

We are now ready to formalize the idea of RBR.

Definition 2 (Relative Best Response). *Given a game Γ , a relative best response of player $i \in \mathcal{I}$ is any strategy in the set*

$$\mathcal{B}_i^r(\mathbf{s}_{-i}) := \underset{y \in \mathcal{F}_i(\mathbf{s}) \subseteq \mathcal{S}}{\text{argmax}} \pi_i(y, \mathbf{s}_{-i}),$$

where $\mathcal{F}_i(\mathbf{s})$ is given by (2) for finite games and by (3) for convex games. \square

In imitation dynamics, in order to choose which neighbor's strategy to imitate, the players must have information about the strategies and the current payoffs of their neighbors. If this information is available to the players, an interesting generalization of a RBR is a decision process in which the feasible strategy set of player $i \in \mathcal{I}$ depends on a subset

of the neighbors that receive the h_i highest payoffs. In this case, the *relative success* of the neighbors of i will have an influence on the future strategy of player i , and $h_i \in \mathbb{N}$ is a measure for how restricting this relative success is for player i 's feasible strategy set. For all $i \in \mathcal{I}$, h_i cannot exceed the number of neighbours of player i , i.e., $0 < h_i \leq |\mathcal{N}_i|$. Before defining such a revision process, let us introduce some additional auxiliary sets. For some strategy profile $\mathbf{s} \in \mathcal{S}$, let us define the set of *distinct* payoffs obtained by the neighbors of i as $\mathcal{R}_i(\mathbf{s}) := \{\pi_j(\mathbf{s}) \mid j \in \mathcal{N}_i\}$, and define the set of neighbors that receive at least the h_i highest payoff as

$$\mathcal{H}_i(\mathbf{s}_{-i}, h_i) := \{j \in \mathcal{N}_i \mid \pi_j(\mathbf{s}) \geq \max^{h_i}(\mathcal{R}_i(\mathbf{s}))\},$$

where the notation $\max^{h_i}(\cdot)$ was introduced in Section II. Note that, it always holds that $|\mathcal{N}_i| \geq |\mathcal{H}_i(\mathbf{s}_{-i}, h_i)| \geq h_i$. Then, the set of strategies of these successful players is

$$\mathcal{M}_i(\mathbf{s}_{-i}, h_i) := \{s_j \in \mathcal{S} \mid j \in \mathcal{H}_i(\mathbf{s}_{-i}, h_i)\}. \quad (4)$$

In this case, for a finite game Γ_f , the feasible set of strategies is determined by

$$\forall i \in \mathcal{I} : \mathcal{F}_i^f(\mathbf{s}, h_i) := \{\mathcal{M}_i(\mathbf{s}_{-i}, h_i)\} \cup \{s_i\} \subseteq \mathcal{S}, \quad (5)$$

while for a convex game Γ_c , it is

$$\forall i \in \mathcal{I} : \mathcal{F}_i^c(\mathbf{s}, h_i) := \text{conv}\{\mathcal{F}_i^f(\mathbf{s}, h_i)\}. \quad (6)$$

Let $h = \text{col}(\{h_i\}_{i \in \mathcal{I}}) \in \mathbb{N}^N$. An h -RBR can now be formalized as follows.

Definition 3 (*h -Relative Best Response*). Given a game Γ , a h -relative best response of player $i \in \mathcal{I}$ is any strategy in the set

$$\mathcal{B}_i^r(\mathbf{s}_{-i}, h_i) := \underset{y \in \mathcal{F}_i(\mathbf{s}, h_i)}{\text{argmax}} \pi_i(y, \mathbf{s}_{-i}), \quad (7)$$

where $\mathcal{F}_i(\mathbf{s})$ is given by (5) for finite games and by (6) for convex games. \square

Now that we have defined an h -RBR, let us introduce the *synchronous*, or parallel, evolutionary game dynamics that are associated with the h -RBR:

$$\forall i \in \mathcal{I} : s_i(t+1) \in \mathcal{B}_i^r(\mathbf{s}_{-i}(t), h_i). \quad (8)$$

We also define the *asynchronous* game dynamics via an activation sequence: at each time step $t \in \mathbb{N}$ for which $\mathbf{s}(t+1) \neq \mathbf{s}(t)$, there exists a unique player $i_t \in \mathcal{I}$ such that the collective dynamics satisfy

$$\text{if } i = i_t : \quad \mathbf{s}(t+1) = (s_i(t+1), \mathbf{s}_{-i}(t+1)) \in (\mathcal{B}_i^r(\mathbf{s}_{-i}(t), h_i), \mathbf{s}_{-i}(t)). \quad (9)$$

We will analyse the convergence of (8) in Section V for convex games, and that of (9) in Section IV for finite games.

Remark 1. When $h_i = |\mathcal{N}_i|$ for every $i \in \mathcal{I}$, the h -RBR coincides with RBR. For finite games, when $h_i = 1$, player i can only choose between its own strategy and the strategy of its most successful neighbors. Therefore, when for all $i \in \mathcal{I}$, $h_i = 1$ the admissible strategies of the h -RBR for

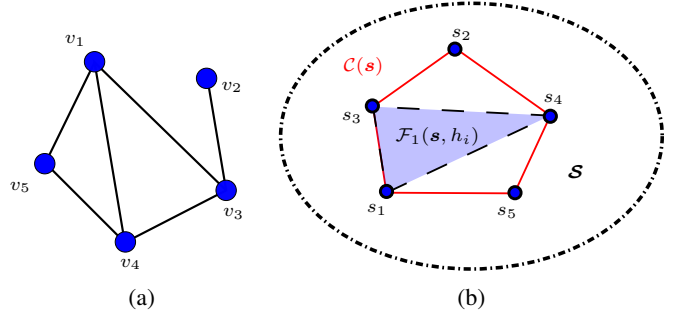


Fig. 1: Suppose the network is as in (a) such that $n = 5$. The set of strategies of the neighbors of 1 is $\mathcal{M}_1(\mathbf{s}_{-1}) = \{s_3, s_4, s_5\}$. Moreover, suppose that $\pi_4(\mathbf{s}) > \pi_3(\mathbf{s}) > \pi_2(\mathbf{s}) > \pi_5(\mathbf{s})$ and $h_1 = 2$. Then, $\mathcal{M}_1(\mathbf{s}_{-1}, 2) = \{s_4, s_5\}$, $\mathcal{F}_1^c(\mathbf{s}, 2) = \{s_4, s_5, s_1\}$ and the shaded area with the dashed border in (b) illustrates $\mathcal{F}_1^c(\mathbf{s}, 2)$. Moreover, $\mathcal{C}(\mathbf{s})$ from Equation (13) is indicated by the region with the red border.

finite games are exactly the admissible set of strategies in an unconditional imitation process. \square

Remark 2. From the definition of the feasible strategy sets in (5), it can be noticed that the evolutionary game dynamics described by (8) and (9) are non-innovative dynamics with time varying local strategy sets. Thus, the (h -)RBR dynamics differ significantly from traditional best response dynamics. For instance, when at some time t there exists a strategy $w \in \mathcal{S}$ such that $w \notin \mathcal{N}_i(\mathbf{s}(t))$, then $s_i(t+1) \neq w$ even if $w \in \mathcal{B}_i(\mathbf{s}_{-i})$. \square

A. Convergence Problem Statement

We devote the remainder of the paper to the study of the convergence properties of (8) and (9), where the convergence can be to an equilibrium point, or to a cycle. In the first case, all players in the network reach a decision with which they are satisfied. Thus, the decision process converges to a strategy profile which is invariant for the game dynamics, hence to an equilibrium strategy profile. First, let us postulate the following assumption, which ensures that players only switch to another strategy if they have an incentive to deviate. This ensures that, when they reach a *non-strict* equilibrium profile, the strategies do not change any further.

Standing Assumption 2 (*Incentive to deviate*). For Γ , $s_i(t) \neq s_i(t+1)$ only if there exists $y \neq s_i(t)$ such that

$$y \in \mathcal{F}_i(\mathbf{s}, h_i) \text{ and } \pi_i(y, \mathbf{s}_{-i}(t)) > \pi_i(s_i(t), \mathbf{s}_{-i}(t)).$$

When updating is synchronous this is an if and only if statement. \square

For the h -RBR dynamics, the local feasible strategy set for each player is constrained by the strategies of the other players and hence the equilibrium strategy profiles of these dynamics correspond to a *Generalized Nash Equilibria* (GNE) [24].

Definition 4 (Generalized Nash Equilibrium). *The strategy profile $s^* \in \mathcal{S}$ is a GNE for Γ , if for all $i \in \mathcal{I}$*

$$s_i^* \in \mathcal{B}_i^r(s_{-i}^*, h_i), \quad (10)$$

where the feasible strategy set $\mathcal{F}_i(s^*, h_i)$ of a finite game and convex game are given by (5) and (6), respectively. \square

Note that, in the convex game case, our GNE problem is not jointly convex [25].

Alternatively, the multi-agent decision process may converge to a cycle, in which the players periodically adjust their strategies.

Definition 5 (Convergence to a cycle). *We say that the sequence of strategy profiles $(s(t))_{t \in \mathbb{N}} \in \mathcal{S}$ converges to a cycle if there exist $\bar{t}, T \in \mathbb{N}$ such that*

$$s(t+T) = s(t), \forall t \geq \bar{t}, \quad (11)$$

where $T \in \mathbb{N}$ is called period of the cycle.

For the asynchronous evolutionary dynamics in (9) we assume that the activation sequence ensures that at any time step, each player is guaranteed to be active at some finite future time.

Standing Assumption 3. *Every sequence of activated players $(i_t)_{t \in \mathbb{N}}$ driving the asynchronous dynamics (9) is persistent, i.e., for every player $j \in \mathcal{I}$ and every time $t \in \mathbb{N}$, there exists some finite time $\bar{t} > t$ at which player j is active again, i.e., $i_{\bar{t}} = j$.* \square

IV. CONVERGENCE IN FINITE NETWORK GAMES

In this section, we study the convergence of the asynchronous h -RBR dynamics (9) to a GNE for finite network games. First, we define two sets that will prove useful in the analysis of the h -BRB dynamics in finite and convex games. For an initial strategy profile $s(0)$, let us denote the set that contains all strategies that are employed by at least one player in the initial strategy profile by $\mathcal{S}_0 := \cup_{i \in \mathcal{I}} \{s_i(0)\}$, and let $\mathcal{S}_0 := \mathcal{S}_0^N$. The set \mathcal{S}_0 is called the *support* of $s(0)$ in [18]. The key property of \mathcal{S}_0 is that it is positively invariant with respect to the h -RBR dynamics (9), due to their non-innovative nature. To study the convergence properties of finite games under the asynchronous h -RBR dynamics we use the theory of potential games [7]. Consider the following definition of a potential like function.

Definition 6 (\mathcal{S}_0 -potential function). *A function $P : \mathcal{S} \rightarrow \mathbb{R}$ is a \mathcal{S}_0 -potential function for Γ_f and some $s(0) \in \mathcal{S}$, if for every $i \in \mathcal{I}$, $s_i, s'_i \in \mathcal{S}_0$ and $s_{-i} \in \mathcal{S}_0^{N-1}$, the following implication holds:*

$$\pi_i(s'_i, s_{-i}) - \pi_i(s_i, s_{-i}) > 0 \Rightarrow P(s'_i, s_{-i}) - P(s_i, s_{-i}) > 0. \quad (12)$$

If such a function exists, then we call Γ_f a relative potential game with respect to \mathcal{S}_0 . \square

Remark 3. *When $s(0) \in \mathcal{S}$ is such that $\mathcal{S}_0 = \mathcal{S}$, then Definition 6 is recovers the definition of a generalized ordinal potential function and a generalized ordinal potential game*

[7, Sec. 2]. In this case, the implication in (12) needs to be satisfied on the entire strategy space \mathcal{S} to ensure convergence of the innovative best response dynamics to a pure Nash equilibrium. \square

We are now ready to present the main result for finite games that relies on the existence of a \mathcal{S}_0 -potential function.

Proposition 1. *Suppose that Γ_f is a relative potential game with respect to \mathcal{S}_0 . Then, for all $s(0) \in \mathcal{S}_0$ the asynchronous h -RBR dynamics defined in (9) converge to a GNE in finite time.* \square

Proof. The proof is based on the non-innovative feature of h -RBR dynamics and the fact that (9) together with Assumption 2 generates improvements paths over a finite action space. Details are omitted due to space limitations. \blacksquare

Corollary 1. *For any finite generalized ordinal potential game, the asynchronous h -RBR dynamics converge globally to a GNE.* \square

Remark 4. *Let E, W, G, B , represent the class of exact, weighted, generalized ordinal and best response potential games, respectively. Because $E \subseteq W \subset G$, Proposition 1 and Corollary 1 imply that for the classes of exact, weighted and ordinal potential games, the asynchronous h -RBR dynamics converge to a GNE. For finite games it is known that for another general class of games, called best response potential games [9, Def. 2.1], the asynchronous best response dynamics converge to a pure Nash equilibrium. For this class of games there exists a function, common for all the players, whose set of maximizers coincides with the set of maximizers of each player's payoff function. In general, $\mathcal{B}_i(s_{-i}) \neq \mathcal{B}_i^r(s_{-i}, h_i)$, and thus the existence of a best response potential function for Γ does not imply convergence of the h -RBR dynamics.*

V. CONVERGENCE IN CONVEX NETWORK GAMES

In this section, we analyze the convergence of the dynamics in (8) for a convex game Γ_c . First, let us define an auxiliary set, namely the convex hull of the support of some given s :

$$\mathcal{C}(s) := \text{conv}(\cup_{i \in \mathcal{I}} \{s_i\}), \quad (13)$$

see Figure 1b for an example. In the following, we prove that the set $\mathcal{C}(s)$ converges to a fixed polyhedron $\bar{\mathcal{C}}$. Furthermore, we analyse the convergence for convex games with linear payoff functions.

A. *Convergence of the feasible strategy set*

From the equations (6) and (13), it can be noticed that

$$\cup_{i \in \mathcal{I}} \mathcal{F}_i^c(s, h_i) \subseteq \mathcal{C}(s), \quad \forall i \in \mathcal{I}. \quad (14)$$

Since the players can only choose a strategy in the convex hull of the local strategies, the set $\mathcal{C}(s(t))$ cannot grow over time, hence it must converge to some static set. Let us formalize this statement in the following Lemma.

Lemma 1 (Convergence of feasible strategy set). *Consider the game Γ_c under the dynamics (8), then the following statements hold:*

- $\mathcal{C}(s(t+1)) \subseteq \mathcal{C}(s(t))$ for every $t \in \mathbb{N}$,
- $\lim_{t \rightarrow \infty} \mathcal{C}(s(t)) = \bar{\mathcal{C}} \neq \emptyset$.

Proof. The proof is omitted due to space limitations ■

The above lemma highlights that also in convex games the h -RBR is non-innovative and the dynamics are positively invariant with respect to the set $\mathcal{C}(s(0))$. This feature is used in Section V-B to prove the convergence of the h -RBR.

Remark 5. *The fact that the set $\mathcal{C}(s(t))$ converges to $\bar{\mathcal{C}}$, does not imply that every set $\mathcal{F}_i^c(s(t), h_i)$ is converging. This implication holds if (14) is an equality for all $t \in \mathbb{N}$, e.g., if each player communicates to all other players (complete communication graph).*

B. Linear payoff function

By focusing on linear payoff functions, we are able to prove finite time convergence of h -RBR dynamics for convex games. For all $i \in \mathcal{I}$, the linear payoff functions are

$$\pi_i(y, s_{-i}(t)) := s(t)^\top C_i^\top y, \quad \forall i \in \mathcal{I}, \forall t \in \mathbb{N}, \quad (15)$$

where $C_i \in \mathbb{R}^{n \times Nn}$. If $j \notin \mathcal{N}_i$, then $[C_i]_{lj} = 0$ for all $l \in \mathcal{I}$, thus the cost function considers only the local strategies of the neighbours. The following lemma guarantees that for each player $i \in \mathcal{I}$ and at each time t , there exists an h -RBR that is a corner point of the player's feasible strategy set.

Lemma 2. *Consider Γ_c with a payoff (15), for every time $t \in \mathbb{N}$ and every $i \in \mathcal{I}$ it holds that*

$$\mathcal{B}_i^r(s_{-i}(t), h_i) \cap \text{vert}(\mathcal{F}_i(s(t), h_i)) \neq \emptyset. \quad (16)$$

Proof. The proof is omitted due to space limitations. ■

From the previous lemma, we have that the choice of the future strategy of a player can always be found between those of the best performing neighbors and the player's own strategy. This motivates us to postulate the following assumption, which is met by adopting the simplex algorithm [26, Ch. 3] for solving the set of h -relative best responses.

Assumption 4. *In (8) each player $i \in \mathcal{I}$ chooses its future strategy as a corner point of its feasible strategy set, i.e., $s_i(t+1) \in \mathcal{B}_i^r(s_{-i}(t), h_i) \cap \text{vert}(\mathcal{F}_i(s(t), h_i))$. If there exist multiple corner points that are optimal, the players consistently choose one such corner point.* □

Assumption 4 connects the h -RBR in convex games with linear payoff functions to h -RBR dynamics in finite games. In fact, the set of all the strategies $\{s_i(t)\}_{i \in \mathcal{I}}$ played at time t is a subset of the strategies played at the previous time instant, i.e., $\{s_i(t)\}_{i \in \mathcal{I}} \subseteq \{s_i(t-1)\}_{i \in \mathcal{I}}$ for all $t \in \mathbb{N}$. Thus, for all $t \geq 0$ and all $i \in \mathcal{I}$, $s_i(t) \in \mathcal{S}_0$. In other words, it can be seen as a finite game in which $\mathcal{S} = \mathcal{S}_0$.

Assumption 4 enables us to prove that for the convex game Γ_c the dynamics converge to a cycle or to an equilibrium

Proposition 2. *If Assumption 4 holds, then the dynamics in (8) converge, for any initial strategy profile $s(0) \in \mathcal{S}$, to a cycle or to a point in \mathcal{S} .* □

Proof. From Lemma 2 and Assumption 4, it follows that $\{s_i(t+1)\}_{i \in \mathcal{I}} \subseteq \{s_i(t)\}_{i \in \mathcal{I}}$, for all t . The proof can be completed using the fact that the set of possible feasible strategies is finite and players consistently choose one optimal corner point. ■

It is worth noting that if Assumption 4 is relaxed such that players choose randomly between optimal corner points, then the strategy profile can exhibit oscillations with variable periods. In the case in which the strategies are scalars, i.e., $\mathcal{S} \subseteq \mathbb{R}$, it is possible to refine the result in Proposition 2. Specifically, under some condition on the linear payoff functions we can show convergence of the dynamics to a strategy profile in \mathcal{S} . Similar to what was done in Section IV, we will exploit the proprieties of \mathcal{S}_0 and \mathcal{S}_0 to prove convergence to an equilibrium.

Proposition 3. *Suppose Assumption 4 holds and Γ_c with $\mathcal{S} \subseteq \mathbb{R}$ has payoff functions as in (15). Assume that for all $i \in \mathcal{I}$, the vector $C_i^\top \in \mathbb{R}^N$ is such that $C_i \bar{s} \in \mathbb{R}$ has the same sign for every $\bar{s} \in \mathcal{S}_0$. Then, for any initial condition $s(0) \in \mathcal{S}_0$ the sequence of profile strategies $(s(t))_{t \in \mathbb{N}}$, generated by the dynamics (8), converge to a GNE.* □

Proof. The proof relies on the fact that the finite sets $\mathcal{I}^+ := \{i \in \mathcal{I} \mid C_i s > 0, \forall s \in \mathcal{S}_0\}$ and $\mathcal{I}^- := \{j \in \mathcal{I} \mid C_j s < 0, \forall s \in \mathcal{S}_0\}$ are monotonically increasing, respectively decreasing over time. Details are omitted due to space limitations. ■

Corollary 2. *Under the same conditions as in Proposition 3, assume additionally that for all $i \in \mathcal{I}$, the vector $C_i^\top \in \mathbb{R}^N$ is such that $C_i \bar{s} \in \mathbb{R}$ has the same sign for every $\bar{s} \in \mathcal{S}$. Then the sequence of profile strategies $(s(t))_{t \in \mathbb{N}}$, generated by the dynamics (8), converge globally to a GNE.* □

VI. NUMERICAL SIMULATION

In this section we shortly explore the behavior of the synchronous h -RBR dynamics (8) in a classic *Rock-Scissors-Paper* (RSP) game on a square lattice network with $N = 25$ players. In the RSP game strategy set is $\mathcal{S} = \{e_1, e_2, e_3\}$, where e_i is the i 'th column of the 3×3 Identity matrix, consequently $\mathcal{S} = \{e_1, e_2, e_3\}^N$. At each time step, each individual plays a RSP game with its neighbors. The total payoff of player i is the sum of each payoff obtained from each pairwise interaction is $\pi_i(s) = \sum_{j \in \mathcal{N}_i} s_i^\top M_i s_j$, where $M_i \in \mathbb{R}^{3 \times 3}$ is a *circulant* payoff matrix of player $i \in \mathcal{I}$, given by

$$\forall i \in \mathcal{I}: \quad M_i = \begin{bmatrix} a_i & b_i & c_i \\ c_i & a_i & b_i \\ b_i & c_i & a_i \end{bmatrix}, \quad b_i > a_i \geq c_i.$$

At each time step, all the agents update their strategies in parallel. Figure 2 shows a typical behavior of such a spatial RSP game under three different synchronous dynamics: myopic best response associated to the Best-Response

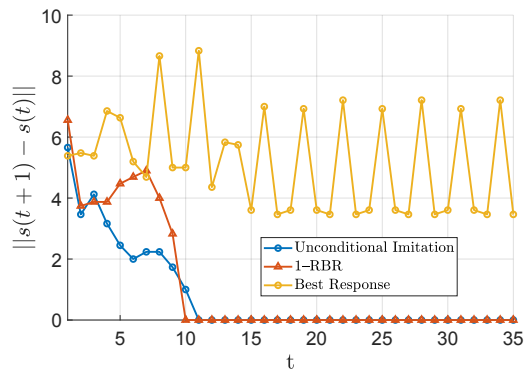


Fig. 2: The difference in consecutive strategy profiles of synchronous myopic best response, h -RBR and ‘imitate-the-best’ dynamics for a Rock-Scissors-Paper game on a square lattice. In the simulation, $n = 5$ and $h = 1$.

mapping in Definition (1), h -RBR (8) and ‘imitate-the-best’ dynamics (defined in [3]). In the latter, each player updates its strategy by imitating the strategy of a neighbor with the highest payoff. In Figure (2), it can be seen that from time $t = 20$, the classic myopic best response dynamics is cycling. This is in line with the behavior of two player RSP games, that have a unique mixed Nash equilibrium. For the h -RBR dynamics, after 7 rounds of parallel plays, one of the strategies ceases to exist in the strategy profile. The non-innovative nature of the h -RBR dynamics subsequently allow the dynamics to converge to a GNE at time $t = 10$. A similar behavior is observed in the trajectory generated by synchronous imitation dynamics that converge to an equilibrium strategy profile at time $t = 11$.

These numerical simulations highlight the key differences between best response and h -RBR dynamics and show that the non-innovative nature of the h -RBR can result in a behavior that resembles a pure imitation process.

VII. CONCLUSION AND FINAL REMARKS

We have proposed the h -RBR dynamics for finite and convex games as a mixture of rational best responses and imitation dynamics. Moreover, we have shown conditions under which these dynamics converge to a generalized Nash equilibrium. In its current form, the h -RBR dynamics are deterministic, i.e., all players always select an h -RBR. A natural extension would be to consider a process in which the probability of a non- h -RBR declines exponentially in the loss of payoff. Such a setup would result in a learning rule where the constraints on the strategy sets of players are a function depending on the strategies of the opponents. This approach is a variation of the log-linear learning process for incomplete games studied in [27]. It can be shown that for finite games, such a perturbed version of the h -RBR is a regularly perturbed Markov chain [4]. We leave characterization of the stochastically stable equilibria of this process as future research.

A natural extension of the results presented for the h -RBR in convex games is to study the convergence of more

general payoff function. Finally, the h -RBR can be linked to opinion dynamics model in which the averaging of opinions is performed over a state-dependent subset of neighbors.

REFERENCES

- [1] L. E. Blume, “The statistical mechanics of strategic interaction,” *Games and Economic Behavior*, vol. 5, no. 3, pp. 387–424, 1993.
- [2] G. Ellison, “Learning, local interaction, and coordination,” *Econometrica: Journal of the Econometric Society*, pp. 1047–1071, 1993.
- [3] M. A. Nowak and R. M. May, “The spatial dilemmas of evolution,” *International Journal of Bifurcation and Chaos*, vol. 3, no. 01, pp. 35–78, 1993.
- [4] H. P. Young, *Individual strategy and social structure: An evolutionary theory of institutions*. Princeton University Press, 2001.
- [5] E. Lieberman, C. Hauert, and M. A. Nowak, “Evolutionary dynamics on graphs,” *Nature*, vol. 433, no. 7023, p. 312, 2005.
- [6] G. Szabó and G. Fath, “Evolutionary games on graphs,” *Physics Reports*, vol. 446, no. 4-6, pp. 97–216, 2007.
- [7] D. Monderer and L. S. Shapley, “Potential games,” *Games and Economic Behavior*, vol. 14, no. 1, pp. 124–143, 1996.
- [8] P. Dubey, O. Haimanko, and A. Zapechelnuk, “Strategic complements and substitutes, and potential games,” *Games and Economic Behavior*, vol. 54, no. 1, pp. 77–94, 2006.
- [9] M. Voorneveld, “Best-response potential games,” *Economics Letters*, vol. 66, no. 3, pp. 289–295, 2000.
- [10] H. P. Young, “The evolution of conventions,” *Econometrica: Journal of the Econometric Society*, pp. 57–84, 1993.
- [11] S. Grammatico, “Proximal dynamics in multi-agent network games,” *IEEE Transactions on Control of Network Systems*, 2018.
- [12] L. Pavel, *Game theory for control of optical networks*. Springer Science & Business Media, 2012.
- [13] P. Yi and L. Pavel, “An operator splitting approach for distributed generalized nash equilibria computation,” *Automatica*, vol. 102, pp. 111–121, 2019.
- [14] M. Friedman and M. Friedman, *Essays in positive economics*. University of Chicago Press, 1953.
- [15] L. Zino, G. Como, and F. Fagnani, “On imitation dynamics in potential population games,” in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 757–762.
- [16] J. Barreiro-Gomez and H. Tembine, “Constrained evolutionary games by using a mixture of imitation dynamics,” *Automatica*, vol. 97, pp. 254–262, 2018.
- [17] A. Govaert, P. Ramazi, and M. Cao, “Convergence of imitation dynamics for public goods games on networks,” in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 4982–4987.
- [18] D. K. Levine and W. Pesendorfer, “The evolution of cooperation through imitation,” *Games and Economic Behavior*, vol. 58, no. 2, pp. 293–315, 2007.
- [19] F. Vega-Redondo, “The evolution of Walrasian behavior,” *Econometrica*, vol. 65, no. 2, pp. 375–384, 1997.
- [20] A. A. Alchian, “Uncertainty, evolution, and economic theory,” *Journal of Political Economy*, vol. 58, no. 3, pp. 211–221, 1950.
- [21] D. D. Davis, “Advance production and Cournot outcomes: an experimental investigation,” *Journal of Economic Behavior & Organization*, vol. 40, no. 1, pp. 59–79, 1999.
- [22] H. Ohtsuki, M. A. Nowak, and J. M. Pacheco, “Breaking the symmetry between interaction and replacement in evolutionary dynamics on graphs,” *Physical Review Letters*, vol. 98, no. 10, p. 108106, 2007.
- [23] P. Yi and L. Pavel, “A distributed primal-dual algorithm for computation of generalized nash equilibria via operator splitting methods,” in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 3841–3846.
- [24] G. Debreu, “A social equilibrium existence theorem,” *Proceedings of the National Academy of Sciences*, vol. 38, no. 10, pp. 886–893, 1952.
- [25] F. Facchinei and C. Kanzow, “Generalized Nash equilibrium problems,” *4or*, vol. 5, no. 3, pp. 173–210, 2007.
- [26] D. Bertsimas and J. Tsitsiklis, *Introduction to Linear Optimization*, 1st ed. Athena Scientific, 1997.
- [27] J. R. Marden and J. S. Shamma, “Revisiting log-linear learning: Asynchrony, completeness and payoff-based implementation,” *Games and Economic Behavior*, vol. 75, no. 2, pp. 788–808, 2012.
- [28] R. T. Rockafellar and R. J.-B. Wets, *Variational analysis*. Springer Science & Business Media, 2009, vol. 317.