Evolution of players' misperceptions in hypergames under perfect observations

Bahman Gharesifard Jorge Cortés

Abstract—**This paper considers games of incomplete information and studies the evolution of the (not necessarily consistent) perceptions of the players using the framework of hypergames. The focus is on developing methods to modify the players' perception about other players' preferences by incorporating the lessons learned from observing their actions. If players are rational, our first update mechanism, called swap learning, is guaranteed to decrease the mismatch between a player's perception and the true payoff structure of other players. However, this method can lead to inconsistencies in the stability properties of the resulting perception. This motivates the introduction of a second update mechanism, called modified swap learning, that is guaranteed to produce a consistent perception. We also identify a class of hypergames for which modified swap is also guaranteed to decrease the mismatch in a player's perception. We introduce the novel notion of H-digraph as a useful tool to encode the information in a hypergame, and fully characterize how this digraph is affected by changes in the players' beliefs.**

I. Introduction

Belief manipulation plays a key role in many strategic situations. A proper understanding of the evolution of the perceptions of players about the game they are involved in is key to unravel how belief manipulation and deception may arise. In adversarial scenarios, it is common to encounter situations where the specific objective of any given individual are unknown or only partially known to the other players.

The goal of this paper is to develop methods that players can implement to modify their perception about other players' ultimate objectives and reason about the actions they take. In this context, the actions taken by a player and the implicit information they contain can be thought of as inputs to the dynamical system describing the evolution of the perceptions of other players. In that regard, controllability and reachability questions (i.e., is there a sequence of actions by one player that would make another player achieve certain perception) become relevant. We are also interested in characterizing how the stability properties of the game outcomes are affected by the evolution of the perceptions. Domains where these questions are relevant include social networks, modeling of human cultural behavior, cybersecurity, and financial markets.

Literature review: Deception and belief manipulation are rooted at an incorrect perception by a player about the true intentions or state of other players. Within the context of games, these situations can be modeled as games of either incomplete or imperfect information. In a game of incomplete information, players do not know the payoff structure of the other players and have an imprecise understanding about their objectives and true intentions. The usual approach, see e.g., [2], consists of transforming the game into one of imperfect information, where Nature decides the true type of the players according to some probability distribution that is known to all. This approach gives rise to Bayesian games [3], [4], where players try to learn from observations the true type of the opponents. Although games with incomplete information facilitate the modeling of uncertainty in players' beliefs, they do not account for a variety of asymmetric situations, such as some players being absolutely certain about other players' types and these certainties being mutually inconsistent, or scenarios where the full set of actions available to the opponents may not be known by some of the players. These restrictions have been pointed out in [5] for a nonzero sum game where players have subjective information structures, and the inconsistent structure of beliefs leads to counterintuitive behaviors. Furthermore, the differences in beliefs may in general not be smoothed out if the game is repeated infinitely many times. [6] demonstrates that there exist games with incomplete information in which players almost never learn to predict their opponents' behavior. Within the context of games of incomplete information, deception has not been studied in a systematic way with the exception of a few references. [7] studies deception via strategic communication, in which a 'sophisticated' player sends either truthful or false messages to the opponents. [8] investigates the vulnerability of strategic decision makers to persuasion. The work [9] constructs a theory of deception for games with incomplete information using the analogy-based sequential equilibrium approach [10], where players form expectations about the average behavior of other players based on histories. In such context, deception arises because boundedly rational players make incorrect inferences about the type of other players. Instead, in hypergames, deception arises because players choose to believe that their perceptions are correct.

In games of imperfect information, players observe only partially the actions taken by other players and therefore have uncertainty about the true state of the game, see e.g., [11], [12]. Early references on deception and deception-robustness in dynamic games with imperfect information include [13], [14]. The work [15] illustrates, in a particular example of a non-cooperative stochastic game, how a player has the potential of manipulating the information available to the opponent and can strategically deceive her. In [16], it is shown that asymmetric information has the potential to inject deception in a non-zero sum game. The work [17] presents an example of deception in a two-person zero-sum dynamic game with imperfect information. The works [18], [19] study deception

Bahman Gharesifard and Jorge Cortés are with the Department of Mechanical and Aerospace Engineering, University of California, San Diego, {bgharesifard,cortes}@ucsd.edu

and provide deception-robust schemes for a class of discrete dynamic stochastic games under imperfect observations.

Here, we consider games of incomplete information and, more specifically, the framework of hypergames [20], [21], [22], [23]. This approach allows us to consider situations where a player believes, whether it is true or not, that other players are of a certain type or have a specific set of actions available to them. This is in contrast with the explicit consideration of uncertainty about other players' types as typically done in games of incomplete information, see e.g. [2]. The introduction to the notion of hypergames goes back to [20] and was originally used to model conflicts [24]. An advantage of using hypergames instead of games with imperfect information is that they allow the possibility of explicitly modeling incorrect perceptions by some players about the intent of other players. Moreover, in hypergames, players can benefit from many levels of perception, in the sense that they can have perceptions about the other players' interpretations of the game, and also about the opponents' perception of their game and so on, see [21], [23]. Hypergames are also well suited to model scenarios where players play security strategies or when the cost of risky actions is high, such as wartime negotiation [25] and cybersecurity [26]. In the context of hypergames, few works [27], [28] have addressed the study of learning from observations. Throughout the paper, we make the simplifying assumption that the actions taken by the opponents are perfectly observed by the players.

Statement of contributions: The first contribution of the paper is the introduction of the basic notions of partial order, preference vector, rank, and H-digraph. These notions simplify the calculation of the equilibria of hypergames and their stability analysis. We also introduce the `H-digraph algorithm`, which provides a procedure for computing H-digraphs and characterize its complexity. The second contribution is the introduction of the swap learning method, which allows a player to update her own perception based on the information contained in the actions taken by other players. We use the misperception function as a measure of the mismatch between a player's perception and the true payoff structure of the other players. Assuming all players are rational, we show that the swap learning method ensures that the misperception function will decrease and that the players' perceptions will converge if they repeatedly use this strategy. On the other hand, we show that other plausible learning strategies, such as right-shift and left-shift learning, are not guaranteed to decrease the misperception function. The third contribution is the introduction of the notion of inconsistency in perceptions. Specifically, we show that the swap learning method can yield preference vectors that are inconsistent with the modified stability properties of the outcomes determined by the actions of other players. This leads us to propose a modified version of the swap learning method which is guaranteed to prevent any inconsistency in the perceptions. We establish a class of hypergames for which the modified strategy is also guaranteed to decrease the misperception function. Finally, the last contribution is the characterization of the evolution of the H-digraph under the swap learning method. We study the effect that the changes in the players' perceptions, determined by swap learning, have in the structure of their respective H-digraphs. These results provide a fast and inexpensive way for detecting outcomes which are not affected by a certain action and, more importantly, open the way to construct algorithmic procedures for belief manipulation. Throughout the paper, we illustrate our discussion with several examples.

Organization: Section II introduces a new framework for studying hypergames. The settlement game in Section III serves the dual purpose of illustrating the basic hypergame definitions and motivating the questions on learning. Section IV introduces the swap learning method to modify a player's perception by incorporating observations from other players' actions and studies its properties. Section V discusses the inconsistencies in perception that might arise under the swap update method and proposes a modified version. Section VI discusses the effect that the changes in the players' perceptions have in the structure of their respective H-digraphs. Section VII contains our conclusions and ideas for future work.

II. Hypergame theory

In this section, we review the basic notions of hypergame theory. Throughout the paper, unless otherwise noted, we assume that players are rational. Although most of the concepts can be found in [22], [21], [20], we have revised the discussion to provide a smooth presentation of the main ideas. We also introduce and analyze the novel concept of H-digraph.

A. Basic notions

A 0-level hypergame is a finite game, i.e., a triplet $\mathbf{G} = (V, \mathbf{S}_{\text{otcm}}, \mathbf{P})$, where V is a set of n players; $\mathbf{S}_{\text{otcm}} = S_1 \times \ldots \times S_n$ is the outcome set, where S_i is a finite set of strategies available to player $v_i \in V$, $i \in \{1, \ldots, n\}$; and $\mathbf{P} = (P_1, \ldots, P_n)$, with $P_i = (x_1, \ldots, x_N)^T \in \mathbf{S}_{\text{otcm}}^N$, $N = |\mathbf{S}_{\text{otcm}}|$ and $i \in \{1, \ldots, n\}$, is called the preference vector of player v_i. Each preference vector P_i is equipped with a preorder $\succeq_{P_i}$, i.e., a reflexive and transitive binary relation, such that, if x has a lower entry index that y in P_i, then $x \succeq_{P_i} y$. In this way, the emphasis is put on the order of preferences among outcomes, rather than on the actual payoff that players obtain for each specific outcome.

Definition 2.1: (1-level hypergame): A 1*-level* n*-person hypergame* is a set $H^1 = \{\mathbf{G}_1, \ldots, \mathbf{G}_n\}$, where $\mathbf{G}_i = (V, (\mathbf{S}_{\text{otcm}})_i, \mathbf{P}_i)$, $i \in \{1, \ldots, n\}$, is the subjective finite game of player $v_i \in V$, and

(i) V is a set of n players;
(ii) $(\mathbf{S}_{\text{otcm}})_i = S_{1i} \times \ldots \times S_{ni}$, where S_{ji} is the finite set of strategies available to v_j, as perceived by v_i;
(iii) $\mathbf{P}_i = (P_{1i}, \ldots, P_{ni})$, where P_{ji} is the preference vector of v_j, as perceived by v_i.

In a 1-level hypergame, each player $v_i \in V$ plays the 0-level hypergame $\mathbf{G}_i$ with the perception that she is playing a game with complete information, which is not necessarily true. The definition of a 1-level hypergame can be extended to high-level hypergames, where some of the players have access to some additional information that allow them to form perceptions about other players' beliefs, other players' perceptions about them, and so on. The following inductive definition allows modeling of multiple levels of perception.

Definition 2.2 (High-level hypergame): A *k-level n-person hypergame*, $k \geq 1$, is a set $H^k = \{H_1^{k_1}, \ldots, H_n^{k_n}\}$, where $k_i \leq k-1$ and at least one k_i is equal to $k-1$. The hypergame H^k is called *homogeneous* if $k_i = k-1$ for all $i \in \{1, \ldots, n\}$.

Assumption 2.3 (2-person 1-level hypergames): Here, we focus on 2-person 1-level hypergames. The results are extensible to 1-level hypergames with an arbitrary number of players, see Remark 2.13 later. 1-level hypergames are the simplest class where players have perceptions about their opponents' preferences. As the ensuing discussion shows, this scenario is already quite challenging, even though the perception about the opponent's preference is the only element susceptible of change. In high-level hypergames, however, players have to deal with multiple possibilities, including changing the perception that a player A has about the perception that another player B has about the original player A, and so on.•

B. Equilibria and stability

Next, we recall the notion of equilibria for hypergames [21]. Let us start by introducing some notation. For a 1-level hypergame H^1, we denote by $H_A^0 = (\mathrm{P}_{AA}, \mathrm{P}_{BA})$ the 0-level hypergame for A, where P_{AA} and P_{BA} are, respectively, the preferences of A and B perceived by A. Similarly, we define $H_B^0 = (\mathrm{P}_{AB}, \mathrm{P}_{BB})$. Here, we assume that players have no misperception in their own preferences and that all the 0-level hypergames have the same set of outcomes $\mathbf{S}_{\text{otcm}}$. Throughout the paper, we let $\mathbf{S}_{\mathrm{P}} \subset \mathbf{S}_{\text{otcm}}^N$, $N = |\mathbf{S}_{\text{otcm}}|$, denote the set of all elements of $\mathbf{S}_{\text{otcm}}^N$ with pairwise different entries. We denote by $\succeq_{\mathrm{P}_{IJ}}$ the binary relation on $\mathbf{S}_{\text{otcm}}$ corresponding to P_{IJ}, where $I, J \in \{A, B\}$ and by π_I the projection of $\mathbf{S}_{\text{otcm}}$ to the strategy set of player $I \in \{A, B\}$. For convenience, we define the restricted outcome set $\mathbf{S}_{\text{otcm}}|_{\pi_I(x)} = \{y \in \mathbf{S}_{\text{otcm}} \mid \pi_I(y) = \pi_I(x)\}$. We also find it useful to use I' to denote the opponent of I in $\{A, B\}$. The next definitions introduce the concepts of improvement and rational outcome.

Definition 2.4 (Improvement and rational outcome): Given two distinct outcomes $x, y \in \mathbf{S}_{\text{otcm}}$, y is an *improvement* from x for $I \in \{A, B\}$, perceived by $J \in \{A, B\}$ in H_J^0, if and only if $\pi_{I'}(y) = \pi_{I'}(x)$ and $y \succeq_{\mathrm{P}_{IJ}} x$. An outcome $x \in \mathbf{S}_{\text{otcm}}$ is *rational* for $I \in \{A, B\}$, perceived by $J \in \{A, B\}$ in H_J^0, if there exists no improvement from x for this player.

An outcome $x \in \mathbf{S}_{\text{otcm}}$ is a *pure Nash equilibrium of* H^1 if it is perceived as rational by A in H_A^0 and by B in H_B^0. This notion of equilibrium does not take into account the different perceptions of the players. This is best illustrated with an example. Suppose A has some perception about B's game and suppose A has an improvement y from x. According to the definition above, x is not a Nash equilibrium of the hypergame. However, if A believes that B has an improvement z from y such that $x \succ_{\mathrm{P}_{AA}} z$, then taking the action associated with y could lead A to an outcome less preferred than x. This mismatch can be addressed by extending the notion of Nash equilibrium using the concept of sequential rationality [22].

Definition 2.5 (Sequential rationality): Consider a 1-level hypergame H^1 between players A and B. An outcome $x \in \mathbf{S}_{\text{otcm}}$ is *sequentially rational* for $I \in \{A, B\}$ with respect to H_J^0, $J \in \{A, B\}$, if and only if for each improvement y for I, perceived by J in H_J^0, there exists an improvement z for I', perceived by J in H_J^0, such that $x \succ_{\mathrm{P}_{IJ}} z$. Whenever this holds, we say that the improvement z from y for I' *sanctions* the improvement y from x for I in H_J^0.

Note that the sanction z might itself not be sanction free for B. One could restrict sanctions to have this property at the cost of a more complex notion of sequential rationality. By definition, a rational outcome is also sequentially rational. We denote by $\mathbf{Seq}_I(H_J^0) \subset \mathbf{S}_{\text{otcm}}$ the set of all sequentially rational outcomes for player $I \in \{A, B\}$, as perceived by player $J \in \{A, B\}$ in H_J^0. An outcome $x \in \mathbf{S}_{\text{otcm}}$ is *unstable* for I with respect to H_J^0 if $x \in \mathbf{Seq}_I^c(H_J^0) = \mathbf{S}_{\text{otcm}} \backslash \mathbf{Seq}_I(H_J^0)$ and is an *equilibrium of* H_J^0 if $x \in \mathbf{Seq}_J(H_J^0) \cap \mathbf{Seq}_{J'}(H_J^0)$. For brevity, we sometimes omit the wording 'with respect to H_J^0' when it is clear from the context. An outcome x is an *equilibrium of* H^1 if $x \in \mathbf{Seq}_A(H_A^0) \cap \mathbf{Seq}_B(H_B^0)$. An outcome x can be an equilibrium for H^1 and not an equilibrium of H_A^0. Also, note that pure Nash equilibria of H^1 are equilibria of H^1.

The following results play an important role in the forthcoming discussion. For simplicity, we present them with respect to the player B in the game H_A^0. However, one can easily extend them for player I in the game H_J^0, $I, J \in \{A, B\}$.

Lemma 2.6: (Abundance of unstable outcomes): Assume $x \in \mathbf{S}_{\text{otcm}}$ is perceived as unstable for B by A in H_A^0. Then any other outcome $z \in \mathbf{S}_{\text{otcm}}$ such that $\pi_A(z) = \pi_A(x)$ and $x \succ_{\mathrm{P}_{BA}} z$ is also perceived as unstable for B by A in H_A^0.

Lemma 2.7: (Existence of rational outcomes): For $x \in \mathbf{S}_{\text{otcm}}$, either x is rational for B in H_A^0 or there exists an improvement y from x perceived by A for B in H_A^0 which is rational for B.

Since rational outcomes are also sequentially rational, Lemma 2.7 also shows the existence of sequentially rational outcomes. It can be shown [21] that every 0-level hypergame has an equilibrium outcome, which may not be unique. However, there exist high-level hypergames which do not contain any equilibrium outcome. Existence can be guaranteed, however, if one extends the notion of equilibria to include mixed strategies, see [29].

Remark 2.8: (Backward induction, subgame perfection, and sequential rationality): It is worth noting the difference between the notion of sequential rationality defined above and backward induction and subgame perfection [2]. To illustrate this point, given a player, say A, and an outcome x, consider the two-stage game where A acts first and B acts second. In general, the Nash subgame perfect equilibria of this game do not correspond to the sequentially rational outcomes given by Definition 2.5. Essentially, this is because sequential rationality cares about providing guarantees no matter the action of the opponent, whereas Nash equilibria cares about maximizing at each stage the expected payoff. Other notions of equilibria are also relevant for hypergames, see [22] for a discussion on the connections among them. The reason for the focus on sequential rationality here is the emphasis on secure actions and guaranteed payoffs based on the players' perceptions. •

C. H-digraphs

The stability analysis in hypergames is typically done by means of preference tables, see [21], [22]. Here, instead, we introduce an alternative method based on the novel notion of H-digraph. The H-digraph contains the information about the possible improvements from an outcome to another outcome, the equilibria, and the sanctions in a hypergame.

A digraph G is a pair (V, E), where V is a finite set, called the vertex set, and $E \subseteq V \times V$, called the edge set. Given $(u, v) \in E$, u is an *in-neighbor* of v and v is an *out-neighbor* of u. The set of in-neighbors and out-neighbors of v are denoted by $\mathcal{N}^{\text{in}}(v)$ and $\mathcal{N}^{\text{out}}(v)$, and their cardinalities are the *in-degree* and *out-degree* of v, respectively. $\mathcal{A}$ is an adjacency matrix for $G = (V, E)$ if the following holds: $a_{ij} > 0$ if and only if $(v_i, v_j) \in E$, for all $v_i, v_j \in V$. Before introducing the concept of H-digraph, we define the notion of rank.

Definition 2.9 (Rank): Let H^1 be a 1-level hypergame and consider the preference vector P_{IJ} in the hypergame H^0_J, $I, J \in \{A, B\}$. We assign to each outcome $x \in \mathbf{S}_{\text{otcm}}$ a positive number $\operatorname{rank}(x, \mathrm{P}_{IJ}) \in \mathbb{R}_{>0}$, called the *rank of outcome* x, such that, for each $\mathbf{S}_{\text{otcm}} \ni y \neq x$, we have that $\operatorname{rank}(y, \mathrm{P}_{IJ}) > \operatorname{rank}(x, \mathrm{P}_{IJ})$ if and only if $x \succ_{\mathrm{P}_{IJ}} y$.

According to this definition, players prefer the outcomes with lower ranks. Throughout the paper and without loss of generality, we use the set $\{1, \ldots, |\mathbf{S}_{\text{otcm}}|\}$ to rank the outcomes. Next, we introduce the notion of H-digraph.

Definition 2.10 (H-digraph): The *H-digraph* $\mathcal{G}_{H^0_A} = (\mathbf{S}_{\text{otcm}}, \mathcal{E}_{H^0_A})$ associated to H^0_A is defined by $(x, y) \in \mathcal{E}_{H^0_A}$ iff one of the following holds,

- there exists an improvement y from x for A for which there is no sanction of B in H^0_A;
- there exists an improvement y from x for B for which there is no sanction of A in H^0_A.

Moreover, each vertex $x \in \mathbf{S}_{\text{otcm}}$ is labeled with $(\operatorname{rank}(x, \mathrm{P}_{AA}), \operatorname{rank}(x, \mathrm{P}_{BA}))$.

Similarly, one can associate an H-digraph to H^0_B. The next result is an immediate consequence.

Lemma 2.11: (Stability notions via H-digraph): An outcome x is sequentially rational for A (respectively for B) if and only if $\mathcal{N}^{\text{out}}(x) \cap \mathbf{S}_{\text{otcm}}|_{\pi_B(x)} = \emptyset$ (respectively $\mathcal{N}^{\text{out}}(x) \cap \mathbf{S}_{\text{otcm}}|_{\pi_A(x)} = \emptyset$). Moreover, an outcome is an equilibrium for the hypergame H^0_A if and only if its out-degree in the associated H-digraph is zero.

Table 1 presents an algorithm to compute H-digraphs.

Lemma 2.12: (Computational complexity of the `H-digraph algorithm`*):* The computational complexity of the `H-digraph algorithm` is $\Theta(|\mathbf{S}_{\text{otcm}}|^2)$.

Proof: Note that $|\mathbf{S}_{\text{otcm}}| = nm$, where n, m are the number of actions of players I and J, respectively. Choose any action of I, and let $\{x_1, \ldots, x_m\}$ be all the outcomes who share this action. Without loss of generality, let $x_i \succ_{\mathrm{P}_{JI}} x_{i+1}$, for all $i \in \{1, \ldots, m\}$, $m \in \mathbb{Z}_{\geq 1}$. For each x_i, the algorithm compares the rank in P_{JI} of all the improvements for I from x_k, $k < i$, in H^0_I, to the rank of x_i in P_{JI}. Note that there are $(i-1)$ improvements from x_i perceived for J and for each of these improvements, there are at most $n = \frac{|\mathbf{S}_{\text{otcm}}|}{m}$ outcomes that need to be examined in P_{II} to draw a conclusion about the stability of x_i. As a result, the total number of computations required for the outcomes in $\mathbf{S}_{\text{otcm}}|_{\pi_I(x_1)}$ is in $\Theta(n \times \frac{m \times (m-1)}{2})$ and since π_I partitions $\mathbf{S}_{\text{otcm}}$ into n subsets, the total computation required is $\Theta(n^2 \times \frac{m \times (m-1)}{2}) = \Theta(|\mathbf{S}_{\text{otcm}}|^2)$. ■

Algorithm 1: The `H-digraph algorithm`

Goal: Compute the H-digraph $\mathcal{G}_{H^0_I}$
Input: $\mathbf{S}_{\text{otcm}}$, P_{II} and P_{JI}
Output: Adjacency matrix $\mathcal{A}^{\mathrm{H}}$ of $\mathcal{G}_{H^0_I}$
Initialization: associate matrices $\mathcal{A}^{\text{imp}}_I$ and $\mathcal{A}^{\text{imp}}_J$ to I and J, respectively, by assigning 1 to an entry (i, j) if there exists an improvement x_j from x_i for the corresponding player in H^0_I and zero otherwise; let $\mathcal{A}^{\mathrm{H}} = \mathbf{0}_{|\mathbf{S}_{\text{otcm}}| \times |\mathbf{S}_{\text{otcm}}|}$

foreach $x_i \in \mathbf{S}_{\text{otcm}}$ **do**
  foreach $K \in \{I, J\}$ **do**
    foreach $x_j \in \mathbf{S}_{\text{otcm}} \setminus \{x_i\}$ **do**
      if $(\mathcal{A}^{\text{imp}}_K)_{ij} \neq 0$ *and* $\nexists l \in \{1, \ldots, |\mathbf{S}_{\text{otcm}}|\}$ *such that* $(\mathcal{A}^{\text{imp}}_{K'})_{jl} \neq 0$*, where* $x_i \succeq_{\mathrm{P}_{KI}} x_l$ **then**
        $\mathcal{A}^{\mathrm{H}}_{ij} = 1$

Remark 2.13 (n-person hypergames): The notion of H-digraph can be extended to 0-level hypergames with a finite number n of players. Such H-digraphs are n-dimensional, with one dimension per individual player's action set. Sanctions are perceived with respect to all the opponents and edges correspond to sanction-free improvements. The time complexity grows with the number of outcomes, which in turn, grows with the number of players. •

Once an H-digraph is calculated with complexity as characterized by Lemma 2.12, if a change is done to the preference vectors of a player, the complexity of recomputing it decreases substantially. We will revisit this issue in Remark 6.3.

III. The settlement game

Here, we analyze in detail a hypergame to illustrate the notions introduced in Section II. The example also serves to motivate the questions addressed in the forthcoming discussion. Suppose two teams A and B are trying to deploy some resources in a field partitioned into four regions, North West (NW), North East (NE), South West (SW), and South East (SE). Each team has its own perception about the conditions in the field and, based on that, has some preferences for deploying the resources. Furthermore, each team has a perception about the opponent's intentions. We associate $\theta = [\theta_{A_1}, \theta_{A_2}, \theta_{B_1}, \theta_{B_2}]^T \in \{0, 1\}^4$ to each outcome, where

- θ_{A_1} is 0 if A chooses West and 1 otherwise; θ_{A_2} is 0 if A chooses North and 1 otherwise;
- θ_{B_1} is 0 if B chooses West and 1 otherwise; θ_{B_2} is 0 if B chooses North and 1 otherwise.

For example, $\theta = (0, 0, 1, 1)^T$ is associated to the outcome in which team A decides to settle in NW, while team B goes to SE. We associate a unique identifier $\operatorname{Ind}(\theta) \in \mathbb{Z}_{\geq 0}$ to θ as

$$\operatorname{Ind}(\theta) = \theta_{A_1} \times 2^0 + \theta_{A_2} \times 2^1 + \theta_{B_1} \times 2^2 + \theta_{B_2} \times 2^3.$$

Suppose the players' preferences and perceptions about each other's preferences are given by

$$\mathrm{P}_{AA} = (12, 9, 6, 3, 8, 4, 13, 1, 14, 2, 11, 7, 0, 5, 10, 15)^T,$$
$$\mathrm{P}_{BA} = (0, 5, 15, 10, 1, 2, 3, 7, 4, 6, 14, 13, 8, 11, 12, 9)^T,$$
$$\mathrm{P}_{BB} = (1, 2, 3, 7, 4, 6, 14, 13, 8, 11, 12, 9, 0, 5, 15, 10)^T,$$
$$\mathrm{P}_{AB} = (12, 9, 6, 3, 8, 4, 13, 1, 14, 2, 11, 7, 0, 5, 10, 15)^T.$$

We rank $\mathbf{S}_{\text{otcm}}$ with the integers $\{1, \ldots, |\mathbf{S}_{\text{otcm}}|\}$.

Figure 1(a) and (b) show the H-digraphs associated to each team's hypergame. For instance, in Figure 1(a), there is no outgoing edge from 0 to 4, 8, and 12, which, according to Lemma 2.11, means that 0 is perceived as sequentially rational for B in H_A^0. Let us analyze what happens if players play this hypergame. Team A hopes for the equilibrium 3 and moves to SE. Team B also perceives 3 as the best equilibrium and so moves to NW. The result of the game does not reveal any new information about the misperceptions, in the sense that none of the teams would do anything differently if they got the chance to play it again.

Next, consider the same setup as above with a new set of preferences for B,

$$\mathrm{P}'_{BB} = (13, 14, 12, 8, 9, 11, 2, 1, 3, 4, 7, 6, 15, 10, 0, 5)^T.$$

Figure 1(c) shows the new H-digraph associated to B's hypergame. Team A hopes for 3 and so plays the action $\pi_A(3)$. Similarly, B hopes for the equilibrium 12 and thus plays the action $\pi_B(12)$. The result of a one-stage play is 15, which is unstable for A in H_A^0 and B in H_B^0. If any of them got the chance to move again, they could find an improvement to a sequentially rational outcome and select the action associated to it. For example, B could take the action $\pi_B(11)$.

We are interested in understanding what the players could have observed, at each round of play, about their misperception of the opponent's game. For example, consider A's perception. Initially, A thinks that 15 is (sequentially) rational for B. This can be observed in Figure 1(a), where 15 has no outgoing edge to 3, 7, or 11. Based on the action $\pi_B(11)$, A could *learn*: (i) outcome 15 is not sequentially rational for B; (ii) B prefers outcome 11 to outcome 15, i.e., $15 \prec_{\mathrm{P}'_{BB}} 11$. Player A could use these observations to improve her perception about B's game. These are the kind of questions that motivate our developments below.

IV. Decreasing misperception by observations

In this section, we investigate methods that allow players to update their own perceptions based on the information contained in the actions taken by other players. To make this precise, consider a 1-level hypergame H^1 with two players A and B. Suppose players take actions sequentially. Denote by $\mathcal{O}_{BA}$ the *observation set* of player A, that is, the set of binary relations in P_{BB} observed by A. We say that the preference vector P_{BA} is *compatible* with the observation set $\mathcal{O}_{BA}$ if the binary relations in $\mathcal{O}_{BA}$ hold in P_{BA}. Similar definitions can be made for player B's preferences.

In general, players seek for strategies to update their perceptions so that their preference vectors are compatible with their observation sets. This is the problem treated here. Section IV-A considers learning when only a single observation has been made and Section IV-B analyzes its effect on misperceptions. This discussion sets the basis for analyzing the case of multiple observations in Section IV-C.

A. Learning from a single observation

In most of the following, we analyze the hypergame from the viewpoint of A. An analogous discussion can be carried out for B. Suppose B takes an action that changes the outcome from $x \in \mathbf{S}_{\text{otcm}}$ to $y \in \mathbf{S}_{\text{otcm}}$, with $x \neq y$. Then, A deduces that B prefers y over x. Therefore, A can incorporate this information into her hypergame and update her perception about the preferences of B. This section explores the suitability of several methods to incorporate this information.

1) Swap update: In the second part of the settlement example of Section III, the players' change of actions leads to a shift in the outcomes from 15 to 11; thus A concludes that B prefers the outcome 11 to 15. Player A originally has the perception $15 \succ_{\mathrm{P}_{BA}} 3 \succ_{\mathrm{P}_{BA}} 7 \succ_{\mathrm{P}_{BA}} 11$ about $\mathbf{S}_{\text{otcm}}|_{\pi_A(15)}$. After moving from outcome 15 to 11, it would appear reasonable for A to interchange the positions of 15 and 11 in her belief about B's preferences: $11 \succ_{\mathrm{P}_{BA}} 3 \succ_{\mathrm{P}_{BA}} 7 \succ_{\mathrm{P}_{BA}} 15$. We call this *swap learning*. We formally define this map next.

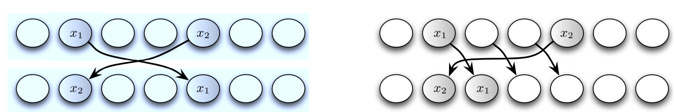
Definition 4.1 (Swap map): Let V be a set of cardinality N and let W be the subset of V^N with pairwise different elements. For $x_1, x_2 \in V$, define $\mathrm{swap}_{x_1 \mapsto x_2} : W \to W$ by

$$(\mathrm{swap}_{x_1 \mapsto x_2}(v))_i = \begin{cases} v_j & \text{if } v_i = x_1, v_j = x_2 \text{ and } i < j, \\ v_i & \text{if } v_i = x_1, v_j = x_2 \text{ and } i > j, \end{cases}$$

$$(\mathrm{swap}_{x_1 \mapsto x_2}(v))_j = \begin{cases} v_i & \text{if } v_i = x_1, v_j = x_2 \text{ and } i < j, \\ v_j & \text{if } v_i = x_1, v_j = x_2 \text{ and } i > j, \end{cases}$$

and $(\mathrm{swap}_{x_1 \mapsto x_2}(v))_k = v_k$ if $v_k \neq x_1, x_2$. We refer to $\mathrm{swap}_{x_1 \mapsto x_2}$ as the x_1 *to* x_2 *swap map*.

Figure 2(a) shows the effect of the swap map for a vector v with $v_i = x_1$, $v_j = x_2$, and $i < j$. We are now ready to define the swap learning map acting on the preference vectors.

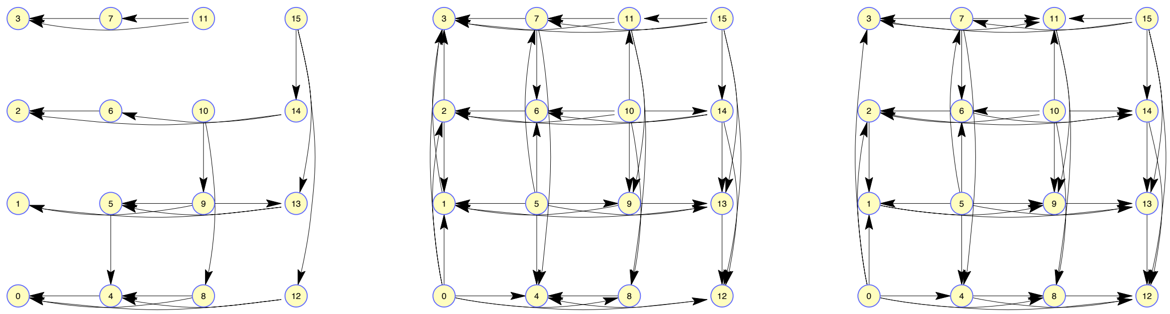


(a) (b)

Fig. 2. Effect of (a) the swap map and (b) the right-shift map on a vector.

Definition 4.2 (Swap learning): Let H^1 be a 1-level hypergame with two players A and B and suppose B takes an action that changes the outcome from x to y. Then the *swap learning map* $\mathbf{Sw}_{x,y}^A : \mathbf{S}_{\mathrm{P}} \to \mathbf{S}_{\mathrm{P}}$ for A is $\mathbf{Sw}_{x,y}^A(\mathrm{P}) = \mathrm{swap}_{x \mapsto y}(\mathrm{P})$.

2) Right-shift learning: In the second part of the settlement example of Section III, when the outcomes change from 15 to 11, A could instead update her belief about B's preferences as follows: $11 \succ_{\mathrm{P}_{BA}} 15 \succ_{\mathrm{P}_{BA}} 3 \succ_{\mathrm{P}_{BA}} 7$. Note that with this update, unlike the swap learning, A employs the information $11 \succ_{\mathrm{P}_{BB}} 15$, while still believing that B prefers 15 to outcomes 3 and 7. We call this *right-shift learning*. We formally define this map next.



(a) (b) (c)

Fig. 1. H-digraphs for the hypergames (a) $(\mathrm{P}_{AA}, \mathrm{P}_{BA})$, (b) $(\mathrm{P}_{AB}, \mathrm{P}_{BB})$, and (c) $(\mathrm{P}_{AB}, \mathrm{P}'_{BB})$.

Definition 4.3 (Right-shift map): Let V be a set of cardinality N and let W be the subset of V^N with pairwise different elements. For $x_1, x_2 \in V$, define $\text{r-shift}_{x_1 \mapsto x_2} : W \to W$ by

$$(\text{r-shift}_{x_1 \mapsto x_2}(v))_i = \begin{cases} v_j & \text{if } v_i = x_1, v_j = x_2 \text{ and } i < j, \\ v_i & \text{if } v_i = x_1, v_j = x_2 \text{ and } i > j, \end{cases}$$

$$(\text{r-shift}_{x_1 \mapsto x_2}(v))_l = \begin{cases} v_{l-1} & \text{if } v_i = x_1, v_j = x_2 \\ & \text{and } i < l \leq j, \\ v_l & \text{if } v_i = x_1, v_j = x_2 \\ & \text{and } j \leq l < i, \end{cases}$$

and $(\text{r-shift}_{x_1 \mapsto x_2}(v))_k = v_k$ if $v_i = x_1, v_j = x_2$ and $k < i$ or $k > j$. We refer to $\text{r-shift}_{x_1 \mapsto x_2}$ as the *x_1 to x_2 right-shift map*.

Figure 2(b) shows the effect of the right-shift map for a vector v with $v_i = x_1$, $v_j = x_2$, and $i < j$. Next, we show that the right-shift map corresponds to a composition of swap maps.

Lemma 4.4: (Right-shift map as a composition of swap maps): Let V be a set of cardinality N and let W be the subset of V^N with pairwise different elements. For $x_1, x_2 \in V$,

$$\text{r-shift}_{x_1 \mapsto x_2}(v) = \\ \text{swap}_{v_{j-(j-i-1)} \mapsto x_1} \circ \cdots \circ \text{swap}_{v_{j-1} \mapsto x_1} \circ \text{swap}_{x_1 \mapsto x_2}(v),$$

where $v_i = x_1$ and $v_j = x_2$.

A right-shift map $\text{r-shift}_{x_1 \mapsto x_2}$ acting on $W \subset U$ can be extended to a map $\overline{\text{r-shift}}_{x_1 \mapsto x_2}$ acting on U by prescribing that $\overline{\text{r-shift}}_{x_1 \mapsto x_2}$ fixes all elements of $U \setminus W$.

Definition 4.5 (Right-shift learning): Let H^1 be a 1-level hypergame with two players A and B and suppose B takes an action that changes the outcome from x to y. The *right-shift learning map* $\textbf{R-Sh}^A_{x,y} : \mathbf{S}_{\mathrm{P}} \to \mathbf{S}_{\mathrm{P}}$ for A is given by

$$\textbf{R-Sh}^A_{x,y}(\mathrm{P}) = \overline{\text{r-shift}}_{x \mapsto y}(\mathrm{P})$$

where $\overline{\text{r-shift}}_{x \mapsto y}$ is the x to y right-shift map on $\mathbf{S}_{\text{otcm}}|_{\pi_A(x)}$ extended to $\mathbf{S}_{\text{otcm}}$.

It is also possible to define the notion of left-shift learning map, in which the player trusts that her initial belief about the relative ranks with respect to the second outcome is correct.

B. Effect of learning on misperception

Here, our objective is to understand the effect of the learning maps introduced above on the player's perception. To that goal, we introduce a function that compares the rank of each outcome in the preference vector for B in H^0_A to its rank in B's true preference vector in H^0_B.

Definition 4.6 (Misperception function): Let H^1 be a hypergame with outcome set $\mathbf{S}_{\text{otcm}}$. The *misperception function* $\mathcal{L}_{BA} : \mathbf{S}_{\mathrm{P}} \to \mathbb{R}_{\geq 0}$ of A about B's game is

$$\mathcal{L}_{BA}(P) = \sum_{i=1}^{N} |\text{rank}(x_i, P_{BB}) - \text{rank}(x_i, P)|$$

An analogous definition can be given for the misperception function $\mathcal{L}_{AB}$ of B about A's game. The next result shows that swap learning can only decrease the misperception.

Theorem 4.7: (The misperception does not increase under swap learning): Consider a 1-level hypergame H^1 between players A and B. Suppose B takes an action such that the outcome of the hypergame changes from x_i to x_j. Then $\mathcal{L}_{BA}(\mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA})) \leq \mathcal{L}_{BA}(\mathrm{P}_{BA})$.

Proof: Let $x_i \succeq_{\mathrm{P}_{BA}} x_j$ (otherwise, the swap learning map is trivial and the result follows). For $x_k \in \mathbf{S}_{\text{otcm}}|_{\pi_A(x_i)}$ let $r_k = \text{rank}(x_k, \mathrm{P}_{BB})$ and $a_k = \text{rank}(x_k, \mathrm{P}_{BA})$, and, up to relabeling the outcomes, suppose that $a_l \leq a_k$ if and only if $l < k$. Under the swap learning map,

$$\begin{aligned} \Delta\mathcal{L}_{BA} &= \mathcal{L}_{BA}(\mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA})) - \mathcal{L}_{BA}(\mathrm{P}_{BA}) \\ &= (|r_i - a_j| + |r_j - a_i|) - (|r_j - a_j| + |r_i - a_i|). \end{aligned}$$

Since B is rational and has changed her action such that the outcome shifted from x_i to x_j, we have $r_j \leq r_i$. If $a_i = a_j$ or $r_i = r_j$, then $\Delta\mathcal{L}_{BA} = 0$. Next, suppose $a_i < a_j$ and $r_j < r_i$. Then one of the following cases will happen

- if $r_i - a_j \geq 0$, $r_i - a_i > 0$, $r_j - a_i < 0$, and $r_j - a_j < 0$, then $\Delta\mathcal{L}_{BA} = 2(a_i - a_j) < 0$;
- if $r_i - a_j \geq 0$, $r_i - a_i > 0$, $r_j - a_i > 0$, and $r_j - a_j \geq 0$, then $\Delta\mathcal{L}_{BA} = 0$;
- if $r_i - a_j \geq 0$, $r_i - a_i > 0$, $r_j - a_i > 0$, and $r_j - a_j < 0$, then $\Delta\mathcal{L}_{BA} = 2(r_j - a_j) < 0$;
- if $r_i - a_j < 0$, $r_i - a_i \geq 0$, $r_j - a_i \geq 0$, and $r_j - a_j < 0$, then $\Delta\mathcal{L}_{BA} = 2(r_j - r_i) < 0$;
- if $r_i - a_j < 0$, $r_i - a_i < 0$, $r_j - a_i < 0$, and $r_j - a_j < 0$, then $\Delta\mathcal{L}_{BA} = 0$;

and the result follows. ■

However, the misperception can potentially increase under right-shift learning, as shown next.

Proposition 4.8: (The misperception can increase under right-shift learning): Consider a 1-level hypergame H^1 between players A and B. Suppose B takes an action that changes the outcome from x_i to x_j. If $\operatorname{rank}(x_j, \mathrm{P}_{BA}) < \operatorname{rank}(x_j, \mathrm{P}_{BB})$, then

$$\mathcal{L}_{BA}(\mathbf{R\text{-}Sh}^A_{x_i,x_j}(\mathrm{P}_{BA})) \geq \mathcal{L}_{BA}(\mathrm{P}_{BA}).$$

Proof: Note that the only part of $\mathbf{S}_{\text{otcm}}$ affected by the right-shift learning are the outcomes in $\mathbf{S}_{\text{otcm}}|_{\pi_A(x_i)}$ which do not have ranks lower than x_i or higher than x_{i+l}. Therefore, without loss of generality, we can assume $\mathbf{S}_{\text{otcm}}|_{\pi_A(x_i)} = \{x_i, x_{i+1}, x_{i+2}, \ldots, x_{i+l}\}$, where $x_i \succ_{\mathrm{P}_{BA}} x_{i+1} \succ_{\mathrm{P}_{BA}} x_{i+2} \succ_{\mathrm{P}_{BA}} \ldots \succ_{\mathrm{P}_{BA}} x_{i+l}$, and B takes an action that changes the outcome from x_i to x_{i+l}. For $x_k \in \mathbf{S}_{\text{otcm}}|_{\pi_A(x_i)}$, where $k \in \mathbb{Z}_{\geq 1}$, let $r_k = \operatorname{rank}(x_k, \mathrm{P}_{BB})$ and $a_k = \operatorname{rank}(x_k, \mathrm{P}_{BA})$. We compute the change in A's misperception about B's game as follows,

$$\begin{aligned}\Delta\mathcal{L}_{BA} =& \mathcal{L}_{BA}(\mathbf{R\text{-}Sh}^A_{x_i,x_j}(\mathrm{P}_{BA})) - \mathcal{L}_{BA}(\mathrm{P}_{BA}) \\ =& \sum_{k=i}^{i+l-1} (|r_k - a_{k+1}| - |r_k - a_k|) \\ &+ (|r_{i+l} - a_i| - |r_{i+l} - a_{i+l}|).\end{aligned}$$

By assumption, we have $r_{i+l} - a_{i+l} > 0$. Since $a_i < a_{i+l}$, we have $r_{i+l} - a_i > 0$. Thus

$$\Delta\mathcal{L}_{BA} = a_{i+l} - a_i + \sum_{k=i}^{i+l-1} (|r_k - a_{k+1}| - |r_k - a_k|).$$

Moreover, $\sum_{k=i}^{i+l-1}(|r_k - a_{k+1}| - |r_k - a_k|) \geq a_i - a_{i+l}$, since for each $i \leq k \leq i+l-1$ we have $|r_k - a_{k+1}| - |r_k - a_k| \geq -|a_k - a_{k+1}| = a_k - a_{k+1}$. As a result, $\Delta\mathcal{L}_{BA} \geq 0$. ■

Note that $\Delta\mathcal{L}_{BA} = 0$ in the proof of Proposition 4.8 if and only if $r_k \geq a_{k+1}$, for all $i \leq k \leq i+l-1$. Since the true preference of B is independent of A's perception about it, it is not difficult to find examples for which the misperception function strictly increases. Even though a right-shift map can be described as a composition of swap maps (cf. Lemma 4.4), Proposition 4.8 does not contradict Theorem 4.7. This is because only the first swap map in the description corresponds to a change in outcomes caused by the action of the other player, while the rest of swap maps do not. One can prove a similar version of Proposition 4.8 for left-shift learning. Given these results, we focus on swap learning.

C. Learning from multiple observations and convergence of perceptions

Here we investigate the behavior of the hypergame when players repeatedly use the swap update map to update their perceptions. We assume that players take actions sequentially and, at each round, each player chooses an action that she believes will shift the outcome to a sequentially rational one for her (note that this outcome does not necessarily need to be the best sequentially rational outcome).

In the case of multiple observations, one should note that the composition of swap update maps associated to the individual observations does not result in general in a preference vector that is compatible with the observation set. As a simple illustration, suppose the perception of A specifies that $x \succ_{\mathrm{P}_{BA}} z \succ_{\mathrm{P}_{BA}} y$. If A observes first $z \succ_{\mathrm{P}_{BB}} y$ and then $y \succ_{\mathrm{P}_{BB}} x$, the composition of the two swap updates will result in $y \succ_{\mathrm{P}_{BA}} z \succ_{\mathrm{P}_{BA}} x$, which is not compatible with $\mathcal{O}_{BA}$. However, an additional swap update between y and z would result in the correct perception $z \succ_{\mathrm{P}_{BA}} y \succ_{\mathrm{P}_{BA}} x$. Therefore, players must use a sorting algorithm to guarantee that the binary relations in their observation set are respected. Such sorting algorithm should only employ swap updates that are compatible with their observation set (an example of such algorithm is, for instance, the bubble sort algorithm [30]). The resulting swap update maps for A and B are denoted by $\mathbf{Sw}^A_{\mathcal{O}_{BA}}$ and $\mathbf{Sw}^B_{\mathcal{O}_{AB}}$, respectively. By definition, and using Theorem 4.7, these swap update maps do not increase the corresponding misperception function.

The dynamical system that results from A using swap learning to update her perception about B's game is

$$\mathrm{P}_{BA}(l+1) = \mathbf{Sw}^A_{\mathcal{O}_{BA}(l)}(\mathrm{P}_{BA}(l)).$$

Here, $\mathcal{O}_{BA}(l)$ denotes the observation set of players A at round $l \in \mathbb{Z}_{\geq 0}$ and $\mathrm{P}_{BA}(0) = \mathrm{P}_{BA}$ is the initial perception of A about B's game. We refer to this dynamical system as $(\mathrm{P}_{BA}, \mathbf{Sw}^A)$. A similar equation characterizes the evolution $(\mathrm{P}_{AB}, \mathbf{Sw}^B)$ for B. The following convergence analysis is valid for any initial outcome, and therefore, is independent of the method used by the players to choose their initial actions.

Theorem 4.9: (Convergence of evolutions under swap learning): Suppose A and B are playing a 1-level hypergame with strict preferences, are rational, and are using the swap learning method to update their perceptions. Then, the evolutions defined by $(\mathrm{P}_{BA}, \mathbf{Sw}^A)$ and $(\mathrm{P}_{AB}, \mathbf{Sw}^B)$ for the hypergames H^0_A and H^0_B converge to some preference vectors P^*_{BA} and P^*_{AB}, respectively. Furthermore, the induced sequences $\{\mathcal{L}_{BA}(l) = \mathcal{L}_{BA}(\mathrm{P}_{BA}(l))\}_{l\geq 0}$ and $\{\mathcal{L}_{AB}(l) = \mathcal{L}_{AB}(\mathrm{P}_{AB}(l))\}_{l\geq 0}$ are monotonically convergent.

Proof: Here, we give the proof for the evolution $(\mathrm{P}_{BA}, \mathbf{Sw}^A)$; a similar argument works for $(\mathrm{P}_{AB}, \mathbf{Sw}^B)$. Given the definition of misperception function, the sequence $\{\mathcal{L}_{BA}(l)\}_{l\geq 0}$ is positive and bounded from below. Since the sequence is non-decreasing, cf. Theorem 4.7, convergence follows. However, this does not necessarily mean that the evolution $(\mathrm{P}_{BA}, \mathbf{Sw}^A)$ is convergent. To establish this, we need to show that, after a certain number of rounds, the misperception being constant implies that $\mathbf{Sw}^A$ becomes the identity. Suppose B takes an action such that the outcome changes from $x(l)$ to $x(l+1)$. Then,

$$\operatorname{rank}(x(l), \mathrm{P}_{BB}) > \operatorname{rank}(x(l+1), \mathrm{P}_{BB}).$$

By rationality and since the preferences are strict, B will never take an action which changes the outcome from $x(l+1)$ to $x(l)$ in future rounds. Hence, the set of possible swap learning maps available to each player is finite, and $\mathbf{Sw}^A$ becomes the identity after finitely many rounds. ■

Remark 4.10 (Non-strict preferences): Theorem 4.9 can be generalized with minimal changes to hypergames with non-strict preferences. This is because if B takes an action that changes the outcome from $x(l)$ to $x(l+1)$, she will only take an action from $x(l+1)$ back to $x(l)$ if these outcomes are equally preferred. A can easily detect this and not perform further swaps involving these outcomes. In the rest of the paper, for simplicity, we assume all preferences are strict. •

In general, the final value of the misperception in Theorem 4.9 is not necessarily zero. This is typical of hypergames whose outcome set has a large cardinality, because the evolution of the hypergame may finish in an equilibrium where none of the players is willing to change her action any more, whereas large parts of the outcome set remain unexplored.

Example 4.11 (The settlement game revisited): Recall the settlement game introduced in Section III. One can compute the initial misperception of A about B's game to be $\mathcal{L}_{BA}(\mathrm{P}_{BA}) = 120$. After B takes the action $\pi_B(11)$, A, using the swap learning map, updates her perception about B to be

$$\mathbf{Sw}^A_{15,11}(\mathrm{P}_{BA}) = (0,5,11,10,1,2,3,7,4,6,14,13,8,15,12,9)^T,$$

with $\mathcal{L}_{BA}(\mathbf{Sw}^A_{15,11}(\mathrm{P}_{BA})) = 106$. This decrease in the value of the misperception function is consistent with Theorem 4.7. Since outcome 11 is an equilibrium of H^1, A's perception converges to $\mathbf{Sw}^A_{15,11}(\mathrm{P}_{BA})$, as predicated by Theorem 4.9. Observe that, after swap update, 11 and 15 are perceived by A as sequentially rational and unstable for B, respectively. The resulting perception of A not only correctly reflects the fact that B prefers 11 over 15, but also correctly encodes the stability properties of both outcomes. The latter, however, may not hold in general. Under swap update, the stability of outcomes may not be consistent with the action taken by the opponent. This is what motivates Section V. •

Remark 4.12 (Extensions to n-person hypergames revisited): Following up on Remark 2.13, the basis for the extension of the methods and results presented above to an n-person scenario is the following: when a player A_i observes an action taken by other player A_j, she updates its perception reasoning on the 2-dimensional plane that corresponds to A_i and A_j, leaving the edges corresponding to the remaining $(n-2)$ dimensions unchanged. •

V. Detecting the Inconsistencies in Perception

Even though the swap update method introduced in Section IV is guaranteed to decrease the misperception of a player, it could lead to inconsistencies in perceptions about the other players' preferences. To make this point clear, consider the hypergame H^0_A and suppose B takes an action which changes the outcome from x_i to x_j. Because of the rationality of B, moving from x_i to x_j implies that x_i is unstable and x_j is sequentially rational in H^0_B. These two pieces of information are not captured in general by the swap update method, which instead simply takes care of updating the perception of A to assert that B prefers x_j to x_i. In other words, it is possible that the stability properties of x_i and x_j as computed by A with her updated perceptions and as observed from the action taken by B do not match. This discussion is also valid for the case when B does not change its action (because x_i is sequentially rational for her) while at the same time x_i is perceived as unstable for B by A. Here, we develop a learning procedure that addresses this problem.

Throughout the section, we present the results from the viewpoint of A. An analogous discussion can be carried out for B. We focus primarily on the case when B changes its action (Remark 5.10 discusses the case when B does not change its action). Recall that if $x_i \prec_{\mathrm{P}_{BA}} x_j$, the swap map is the identity map and hence no change in perception occurs. Thus, we deal with the case $x_i \succ_{\mathrm{P}_{BA}} x_j$.

A. Inconsistency in perception

Here we study all the cases that can occur under swap learning regarding the consistency between a player's perception and the stability properties of the outcomes as implied by the actions taken by the other player. We summarize the possible scenarios in Table I. For each case, we refer to the corresponding result.

Lemma 5.1: (In a restricted outcome set, an unstable outcome cannot have a rank lower than a sequentially rational one): Suppose player B takes an action which changes the outcome from x_i to x_j, where $x_i \succ_{\mathrm{P}_{BA}} x_j$. Then x_i and x_j cannot be perceived simultaneously as sequentially rational and unstable in $(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$, respectively.

Proof: By virtue of Lemma 2.6, a sequentially rational outcome x_i cannot have a higher rank than an unstable outcome x_j whenever $\pi_I(x_i) = \pi_I(x_j)$, $I \in \{A, B\}$. ■

Next, we characterize two cases for which the swap learning does not create inconsistencies.

Lemma 5.2: (Preservation of correct perception under swap learning): Suppose player B takes an action which changes the outcome from x_i to x_j, where $x_i \succ_{\mathrm{P}_{BA}} x_j$.

(i) If x_i is perceived by A as an unstable outcome for B in H^0_A, then it is also perceived as unstable in $(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$.

(ii) If x_j is perceived by A as a sequentially rational outcome for B in H^0_A, then it is also perceived as sequentially rational for B in $(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$.

Proof: We show (i) first. Suppose x_i is perceived as unstable for B in H^0_A. By definition, there exists a perceived improvement y from x_i for B without any sanction of A. Since $\mathrm{rank}(x_i, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA})) > \mathrm{rank}(x_i, \mathrm{P}_{BA})$, y is also a perceived improvement from x_i for B without any sanction of A; thus x_i remains unstable for B in $(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$. Next, we show (ii). Suppose x_j is perceived as sequentially rational for B in H^0_A. By definition, there exists no perceived improvement for B from the outcome x_j without sanction of A, i.e., there exists no outcome y, $\pi_A(y) = \pi_A(x_j)$, that B can move to from x_j such that $\mathrm{rank}(x_j, \mathrm{P}_{BA}) > \mathrm{rank}(y, \mathrm{P}_{BA})$ without a sanction of A. Since $\mathrm{rank}(x_j, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA})) < \mathrm{rank}(x_j, \mathrm{P}_{BA})$, there is no improvement for B from the outcome x_j without sanction of A. ■

The next result identifies a case where swap learning modifies correctly A's perception about x_i. The proof follows from the definition of sequential rationality and Lemma 2.6.

Lemma 5.3: (Correction of perceptions under swap learning): Suppose player B takes an action which changes the outcome from x_i to x_j, where $x_i \succ_{\mathrm{P}_{BA}} x_j$. Suppose that x_i is perceived as sequentially rational for B in H^0_A and there exists an outcome y, where $\pi_A(y) = \pi_A(x_j)$, perceived as

	$x_i \in \mathbf{Seq}_B(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$	$x_i \in \mathbf{Seq}^c_B(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$
$x_j \in \mathbf{Seq}_B(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$	Inconsistent (Lemma 5.4)	Consistent (Lemmas 5.2 and 5.3)
$x_j \in \mathbf{Seq}^c_B(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$	Never happens (Lemma 5.1)	Inconsistent (Lemma 5.4)

TABLE I
POSSIBLE PERCEPTIONS OF A ABOUT THE STABILITY OF OUTCOMES x_i AND x_j AFTER APPLYING SWAP LEARNING.

unstable for B in H^0_A with $\mathrm{rank}(y, \mathrm{P}_{BA}) < \mathrm{rank}(x_j, \mathrm{P}_{BA})$. Then x_i is unstable in the game $(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$.

The next result captures two interesting situations: one in which x_j is perceived as unstable (respectively, one in which x_i is perceived as sequentially rational) in H^0_A and remains unstable (respectively sequentially rational) after applying the swap learning map, thus giving rise to a contradiction in the perceptions of A about the game of B.

Lemma 5.4: (Inconsistency in perceptions under swap learning): Suppose player B takes an action which changes the outcome from x_i to x_j, where $x_i \succ_{\mathrm{P}_{BA}} x_j$.

(i) The outcome x_j is perceived as unstable in $(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$ if and only if x_i is unstable for B in H^0_A.

(ii) If x_i is perceived as sequentially rational for B in H^0_A and there exists a sequentially rational outcome y, where $\pi_A(y) = \pi_A(x_j)$ and $\mathrm{rank}(y, \mathrm{P}_{BA}) > \mathrm{rank}(x_j, \mathrm{P}_{BA})$, then x_i remains sequentially rational for B in $(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$.

Proof: Both statements follow from Lemma 2.6. We only describe the proof of (i). Suppose x_i is unstable for B in H^0_A. By Lemma 2.6, x_j is also unstable for B in H^0_A. By assumption, there exists a perceived improvement from x_i to an outcome y for player B without sanction of A in H^0_A such that $\mathrm{rank}(y, \mathrm{P}_{BA}) < \mathrm{rank}(x_i, \mathrm{P}_{BA})$. Since $\mathrm{rank}(x_j, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA})) = \mathrm{rank}(x_i, \mathrm{P}_{BA})$, x_j remains unstable for B. The converse follows similarly. ∎

B. Modified swap learning method

Here, we investigate how a player can include the information gathered from the contradictions in her perception under swap learning (cf. Lemma 5.4) to learn more about the other player's game. We introduce a modified version of the swap leaning method that prevents any inconsistency in perceptions. Under this method, A assumes that B has perfect information about her game and is convinced that any inconsistency is due to her lack of knowledge about B's game.

As we did for swap update in Section IV, we first describe the strategy for the case with a single observation and later discuss in Remark 5.11 the extension to multiple observations. To formally define the method, we first need to present some auxiliary results.

Lemma 5.5: (Existence of sequentially rational sanction-free improvements): Consider a 1-level hypergame between A and B. Let B take an action that changes the outcome from x_i to x_j, where $x_i \succ_{\mathrm{P}_{BA}} x_j$, and suppose both x_i and x_j are perceived as unstable for B in $(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$. Define

$$\begin{aligned}I^{BA}_{x_j} = \{&y \in \mathbf{S}_{\mathrm{otcm}}|_{\pi_A(x_j)} \mid y \text{ sanc-free improv from } x_j \text{ for } B,\\ &\text{seq. rational in } (\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))\} \subset \mathbf{S}_{\mathrm{otcm}}|_{\pi_A(x_j)}.\end{aligned}$$

Then $I^{BA}_{x_j} \neq \emptyset$.

The proof of this result follows from Lemma 2.7. For each $y \in I^{BA}_{x_j}$, let I^{AA}_y denote the set of improvements from y in P_{AA}. For convenience, denote by

$$\begin{aligned}\mathrm{lwt}(S, \mathrm{P}) &= \mathrm{argmin}_{o \in S} \mathrm{rank}(o, \mathrm{P}),\\ \mathrm{hght}(S, \mathrm{P}) &= \mathrm{argmax}_{o \in S} \mathrm{rank}(o, \mathrm{P}),\end{aligned}$$

the outcomes in $S \subset \mathbf{S}_{\mathrm{otcm}}$ with lowest and highest rank in P, respectively. Note that there exists no outcome in $I^{BA}_{x_j} \backslash \{\mathrm{hght}(I^{BA}_{x_j}, \mathrm{P}_{BA})\}$ which is rational for A in $(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$, since otherwise $\mathrm{hght}(I^{BA}_{x_j}, \mathrm{P}_{BA})$ would be unstable for B.

Definition 5.6 (Outcome w): With the assumptions of Lemma 5.5, let $o = \mathrm{hght}(I^{BA}_{x_j}, \mathrm{P}_{BA})$. For each $y \in I^{BA}_{x_j}$, if $y \neq o$, let

$$w_y = \mathrm{lwt}(\{w \in I^{AA}_y \mid w \prec_{\mathrm{P}_{BA}} o\}, \mathrm{P}_{BA});$$

if $y = o$ is rational in P_{AA}, let $w_y = o$; if $y = o$ is not rational in P_{AA} and $\{w \in I^{AA}_y \mid w \prec_{\mathrm{P}_{BA}} o\} \neq \emptyset$, let

$$w_y = \mathrm{lwt}(\{w' \in I^{AA}_y \mid w' \prec_{\mathrm{P}_{BA}} o\}, \mathrm{P}_{BA});$$

and otherwise, let $w_y = \mathrm{hght}(I^{AA}_y, \mathrm{P}_{BA})$. Then, let

$$w = \mathrm{lwt}(\{w_y \mid y \in I^{BA}_{x_j}\}, \mathrm{P}_{BA}).$$

Definition 5.6 plays an important role in modified swap learning to correct the inconsistency of x_j being perceived as unstable. The next result, instead, is relevant to correct the inconsistency of x_i being perceived as sequentially rational.

Lemma 5.7: (Outcome z): Consider a 1-level hypergame between players A and B. Suppose B takes an action that changes the outcome from x_i to x_j, where $x_i \succ_{\mathrm{P}_{BA}} x_j$, and suppose both x_i and x_j are perceived as sequentially rational for B in $(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$. Then there exists an improvement $z \in \mathbf{S}_{\mathrm{otcm}}|_{\pi_B(x_j)}$ from x_j for A in $(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$.

Proof: Suppose otherwise; then x_j is an improvement from x_i for B in $(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$ such that there is no sanction of A against it, i.e., x_i is unstable for B in $(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$, which is a contradiction. ∎

We are now ready to introduce the modified swap learning method.

Definition 5.8 (Modified swap learning): Consider a 1-level hypergame between players A and B. Suppose B takes an action that changes the outcome from $x_i \in \mathbf{S}_{\mathrm{otcm}}$ to $x_j \in \mathbf{S}_{\mathrm{otcm}}$, where $x_i \succ_{\mathrm{P}_{BA}} x_j$. The *modified swap learning map* $\mathbf{MSw}^A_{x_i,x_j} : \mathbf{S}_{\mathrm{P}} \to \mathbf{S}_{\mathrm{P}}$ is

- if $x_i \in \mathbf{Seq}^c_B(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}))$ and $x_j \in \mathbf{Seq}_B(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}))$, then

$$\mathbf{MSw}^A_{x_i,x_j}(\mathrm{P}) = \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}),$$

- if $x_i, x_j \in \mathbf{Seq}^c_B(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}))$, then
$$\mathbf{MSw}^A_{x_i,x_j}(\mathrm{P}) = \mathbf{Sw}^A_{w,x_j} \circ \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}),$$
where w is defined in Definition 5.6.
- if $x_i, x_j \in \mathbf{Seq}_B(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}))$, then
$$\mathbf{MSw}^A_{x_i,x_j}(\mathrm{P}) = \mathbf{Sw}^A_{x_i,z} \circ \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}),$$
where $z \in \mathbf{S}_{\text{otcm}}|_{\pi_B(x_j)}$ is the outcome with the highest rank, with respect to $\mathbf{Sw}^A_{x_i,x_j}(\mathrm{P})$, which satisfies the conditions of Lemma 5.7.

According to Lemma 5.1, the case $x_i \in \mathbf{Seq}_B(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}))$ and $x_j \in \mathbf{Seq}^c_B(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}))$ will never occur.

In Definition 5.8, the choice of y with highest rank makes the perception of player A consistent with the least amount of change in its preference vector. However, the choice of z with the highest rank is necessary for the following result to hold.

Proposition 5.9: (Modified swap learning results in no inconsistency): Consider a 1-level hypergame between players A and B. Suppose B takes an action which shifts the outcome from x_i to x_j, where $x_i \succ_{\mathrm{P}_{BA}} x_j$. Then, under the modified swap learning, outcomes x_i and x_j are perceived by A, respectively, as unstable and sequentially rational for player B in $(\mathrm{P}_{AA}, \mathbf{MSw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$.

Proof: By Definition 5.8, we need to consider three cases. If $x_i \in \mathbf{Seq}^c_B(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$ and $x_j \in \mathbf{Seq}_B(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$, the result holds trivially. If $x_i, x_j \in \mathbf{Seq}^c_B(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$, then the action of $\mathbf{Sw}^A_{w,x_j}$ does not have any impact on the stability of x_i. Note that the rank of x_j in $\mathbf{Sw}^A_{w,x_j} \circ \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA})$ is lower than in $\mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA})$. Moreover, by Definition 5.6, for each improvement perceived for B in $(\mathrm{P}_{AA}, \mathbf{MSw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$, there exists a sanction perceived by A, since w has the lowest rank among all outcomes $\{w_y \mid y \in I^{BA}_{x_j}\}$. Thus we conclude that x_j is perceived as sequentially rational for B in $(\mathrm{P}_{AA}, \mathbf{MSw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$. Finally, suppose $x_i, x_j \in \mathbf{Seq}_B(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$. The action of $\mathbf{Sw}^A_{x_i,z}$ does not have any impact on the stability of x_j (note that x_i, x_j are preferred by A to z in $\mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA})$). Moreover, since z is the outcome with highest rank in $\mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA})$ which is an improvement from x_j for A, the improvement x_j from x_i is perceived as sanction free in $(\mathrm{P}_{AA}, \mathbf{MSw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$ for B. Therefore, x_i is unstable in $(\mathrm{P}_{AA}, \mathbf{MSw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$. ■

Remark 5.10: (No change of action by the opponent): Consider the case when B does not change its action and hence $x_j = x_i$. If x_i was perceived by A as sequentially rational, then no inconsistency arises. On the contrary, if A perceived x_i as unstable for B, then an inconsistency arises with the observation that x_i is sequentially rational for B. Player A can still use the modified swap map to make her perception consistent. According to Definition 5.8, this case corresponds to the second bullet. After the modified swap update, x_i is perceived by A as sequentially rational for B, resolving the inconsistency. •

Remark 5.11: (The case of multiple observations): The modified swap update method can also be extended to the case of multiple observations. Since the perception of players about the stability of the outcomes can change along the evolution, in our definition of the modified swap map $\mathbf{MSw}^A_{\mathcal{O}_{BA}}$, players only remove inconsistencies related to the last action taken by their opponents. After implementing the modified swap update on the latest preference vector using the last observed action, the player uses $\mathbf{Sw}^A_{\mathcal{O}_{BA}}$ to make her preference vector compatible with $\mathcal{O}_{BA}$. •

Example 5.12: (Consistent perception under modified swap update): Consider a 1-level hypergame $H^1 = \{H^0_A, H^0_B\}$ between A and B with outcome set $\mathbf{S}_{\text{otcm}} = \{x_1, x_2, x_3, x_4\}$. Let
$$\mathrm{P}_{AA} = (x_2, x_3, x_1, x_4)^T, \quad \mathrm{P}_{BA} = (x_2, x_3, x_1, x_4)^T,$$
$$\mathrm{P}_{BB} = (x_1, x_3, x_2, x_4)^T, \quad \mathrm{P}_{AB} = \mathrm{P}_{AA}$$
Figures 3(a) and (b) show the H-digraphs associated to these hypergames. Initially, suppose A takes the action $\pi_A(x_2)$ and

$$\begin{tikzcd} x_1 \arrow[r] \arrow[d] & x_2 \\ x_3 & x_4 \arrow[l] \arrow[u] \end{tikzcd} \qquad \begin{tikzcd} x_1 \arrow[d] & x_2 \arrow[l] \\ x_3 & x_4 \arrow[l] \arrow[u] \end{tikzcd} \qquad \begin{tikzcd} x_1 \arrow[d] & x_2 \arrow[l] \\ x_3 & x_4 \arrow[l] \arrow[u] \end{tikzcd}$$

(a) (b) (c)

Fig. 3. H-digraphs associated to (a) H^0_A, (b) H^0_B, and (c) H^0_A after applying $\mathbf{MSw}^A_{x_1,x_1}$, respectively.

B takes the action $\pi_B(x_1)$ and thus the first outcome is x_1. Suppose players play this game sequentially and B is the first one to move. Based on her preferences, B does not take any action from x_1. Hence, A observes that x_1 is sequentially rational for B, unlike its initial perception. If A uses swap learning (the identity map in this case), this will result in an inconsistent perception. However, if A uses modified swap learning, then $\mathbf{MSw}^A_{x_1,x_1}(\mathrm{P}_{BA}) = (x_1, x_3, x_2, x_4)^T$, which is consistent with the action taken by B. Figure 3(c) shows the new H-digraph for A, which coincidentally matches the one associated to H^0_B. •

C. Decreasing misperception via modified swap learning

In general, the modified swap learning method is not guaranteed to decrease the misperception function. This is a consequence of the fact that player A is convinced that any inconsistency is due to her lack of knowledge about B's game, whereas indeed such inconsistencies may be due to B's misperception about A's game. The following result shows that, under the assumption that B has perfect information about A's game and always chooses the sequentially rational outcome, then A, using the modified swap learning method, decreases her misperception in the sense of Definition 4.6, while preventing inconsistency in her perceptions.

Theorem 5.13: (Misperception function and modified swap learning): Consider a 1-level hypergame between players A and B, where $\mathrm{P}_{AB} = \mathrm{P}_{AA}$. Suppose B takes an action which changes the outcome from x_i to x_j, where $x_i \succ_{\mathrm{P}_{BA}} x_j$. If x_j is perceived as sequentially rational in $(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$, then the misperception function $\mathcal{L}_{BA}$ does not increase under modified swap learning.

Proof: If $x_i \in \mathbf{Seq}_B^c(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$, since $x_j \in \mathbf{Seq}_B(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$, the result follows from Theorem 4.7 since, in this case, the actions of the modified swap map and the swap map coincide. Otherwise, suppose $x_i \in \mathbf{Seq}_B(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$ and let z be given as in Definition 5.8. By Lemma 5.7, $z \in \mathbf{S}_{\mathrm{otcm}}|_{\pi_B(x_j)}$ is an improvement from x_j for player A in $(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$. Note that necessarily

$$\mathrm{rank}(z, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA})) > \mathrm{rank}(x_i, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA})),$$

since otherwise, x_j would be perceived as a sanction-free improvement for B from x_i in $(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$ and thus x_i would be unstable for B in $(\mathrm{P}_{AA}, \mathbf{Sw}^A_{x_i,x_j}(\mathrm{P}_{BA}))$, a contradiction. The outcome x_i is unstable for B in H_B^0 and there exists an improvement x_j from x_i for B in H_B^0 without any sanction by player A in H_B^0. Since, by assumption, $\mathrm{P}_{AB} = \mathrm{P}_{AA}$, we deduce that $\mathrm{rank}(z, \mathrm{P}_{BB}) < \mathrm{rank}(x_i, \mathrm{P}_{BB})$, otherwise x_i would be sequentially rational for B in H_B^0. Thus, by Theorem 4.7, the function $\mathcal{L}_{BA}$ does not increase. ■

If the assumptions of Theorem 5.13 hold at every step, then one can establish a similar result to Theorem 4.9 regarding the convergence of the perceptions under modified swap learning.

VI. How do changes in perception affect the H-digraph?

Here, we study the effect that the changes in the players' perceptions have in the structure of their respective H-digraphs. In contrast to the previous discussion, we study the impact in the preferences on the whole set of outcomes, instead of only on the outcomes that are swapped. One byproduct of this study is computational efficiency for regenerating an H-digraph after changes have occurred. We only consider changes in the preference vectors due to a swap update since the effect of any learning mechanism can be described as a composition of swaps.

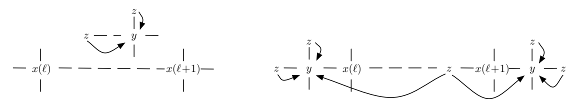
Let us introduce some notation. We denote by $\mathcal{G}_{H_A^0}(0)$ the initial H-digraph associated to player A's hypergame. Suppose at round $l \in \mathbb{Z}_{\geq 1}$ the outcome changes from $x(l)$ to $x(l+1)$ by an action of B. If A does not change the order of these two outcomes, then the H-digraph remains the same. If, instead, A swaps the order of the two outcomes to update her perception about B, then a new H-digraph $\mathcal{G}_{H_A^0}(l+1)$ is formed. For convenience, we denote by $\mathcal{N}_l^{\mathrm{in}}(x)$ and $\mathcal{N}_l^{\mathrm{out}}(x)$, respectively, the set of in- and out-neighbors of $x \in \mathbf{S}_{\mathrm{otcm}}$ in $\mathcal{G}_{H_A^0}(l)$. In this discussion, the term 'new hypergame' refers to the hypergame associated to A's new perception once a change has been done. To study the changes of the H-digraph, it is enough to describe how the in- and out-neighbors of each outcome change. The following result captures the outcomes whose in-neighbors are not affected by the changes in A's perception.

Proposition 6.1: (Sufficient conditions for invariance of in-neighboring structure of an outcome): Suppose player B takes an action that changes the outcome from $x(l)$ to $x(l+1)$. Let

$$\mathcal{M}_{BA}(x(l), x(l+1)) = \{y \in \mathbf{S}_{\mathrm{otcm}} \mid x(l) \succeq_{\mathrm{P}_{BA}(l)} y \succeq_{\mathrm{P}_{BA}(l)} x(l+1)\}.$$

If $y \notin (\mathcal{M}_{BA}(x(l), x(l+1)) \cap \mathbf{S}_{\mathrm{otcm}}|_{\pi_A(x(l))}) \cup \mathbf{S}_{\mathrm{otcm}}|_{\pi_B(x(l))} \cup \mathbf{S}_{\mathrm{otcm}}|_{\pi_B(x(l+1))}$, then $\mathcal{N}_l^{\mathrm{in}}(y) = \mathcal{N}_{l+1}^{\mathrm{in}}(y)$.

Proof: We start by showing that the statement holds for $y \notin \mathbf{S}_{\mathrm{otcm}}|_{\pi_A(x(l))} \cup \mathbf{S}_{\mathrm{otcm}}|_{\pi_B(x(l))} \cup \mathbf{S}_{\mathrm{otcm}}|_{\pi_B(x(l+1))}$. Figure 4(a) shows such an outcome y in a generic H-digraph. Let $z \in \mathcal{N}_l^{\mathrm{in}}(y)$. If $z \in \mathbf{S}_{\mathrm{otcm}}|_{\pi_A(y)}$, then y is an improvement



(a) (b)

Fig. 4. Part of an H-digraph $\mathcal{G}_{H_A^0}$, where A and B play rows and columns, respectively. (a) shows a case where $y \notin \mathbf{S}_{\mathrm{otcm}}|_{\pi_A(x(l))} \cup \mathbf{S}_{\mathrm{otcm}}|_{\pi_B(x(l))} \cup \mathbf{S}_{\mathrm{otcm}}|_{\pi_B(x(l+1))}$ and (b) shows cases where $y \in \mathbf{S}_{\mathrm{otcm}}|_{\pi_A(x(l))}$ with $y \succ_{\mathrm{P}_{BA}(l)} x(l)$ or $y \prec_{\mathrm{P}_{BA}(l)} x(l+1)$.

from z for player B in the hypergame $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l))$ without any sanction of player A. Since, by assumption, $z \neq x(l)$, $x(l+1)$, player B is also perceived to have an improvement y from z, with respect to the preference vector $\mathrm{P}_{BA}(l+1)$, without any sanction from player A; thus $z \in \mathcal{N}_{l+1}^{\mathrm{in}}(y)$. Now suppose $z \in \mathbf{S}_{\mathrm{otcm}}|_{\pi_B(y)}$. Since, by assumption, the ranking of the outcomes in $\mathbf{S}_{\mathrm{otcm}}|_{\pi_A(y)}$ is the same with respect to $\mathrm{P}_{BA}(l)$ and $\mathrm{P}_{BA}(l+1)$, player A still has an improvement y from z, with respect to the preference vector P_{AA}, without any perceived sanction from player B; thus $z \in \mathcal{N}_{l+1}^{\mathrm{in}}(y)$. This proves that $\mathcal{N}_l^{\mathrm{in}}(y) \subseteq \mathcal{N}_{l+1}^{\mathrm{in}}(y)$. A similar argument shows the converse inclusion; thus $\mathcal{N}_l^{\mathrm{in}}(y) = \mathcal{N}_{l+1}^{\mathrm{in}}(y)$.

To complete the proof, we show that if $y \in \mathbf{S}_{\mathrm{otcm}}|_{\pi_A(x(l))}$ such that $y \succ_{\mathrm{P}_{BA}(l)} x(l)$ or $y \prec_{\mathrm{P}_{BA}(l)} x(l+1)$, then $\mathcal{N}_l^{\mathrm{in}}(y) \subseteq \mathcal{N}_{l+1}^{\mathrm{in}}(y)$, see Figure 4(b). Let $z \in \mathcal{N}_l^{\mathrm{in}}(y)$. If $z \in \mathbf{S}_{\mathrm{otcm}}|_{\pi_A(y)}$, since by assumption $y \succ_{\mathrm{P}_{BA}(l)} x(l)$ or $y \prec_{\mathrm{P}_{BA}(l)} x(l+1)$, the possible sanctions of player A against the perceived improvement y of player B from z stay the same after swapping $x(l)$ and $x(l+1)$, and therefore $z \in \mathcal{N}_{l+1}^{\mathrm{in}}(y)$. If $z \in \mathbf{S}_{\mathrm{otcm}}|_{\pi_B(y)}$, since $y \succ_{\mathrm{P}_{BA}(l)} x(l)$ or $y \prec_{\mathrm{P}_{BA}(l)} x(l+1)$, the perceived sanctions of player B are the same in hypergames $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l))$ and $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l+1))$; thus we conclude that $z \in \mathcal{N}_{l+1}^{\mathrm{in}}(y)$. A similar argument shows that the converse holds, yielding $\mathcal{N}_l^{\mathrm{in}}(y) = \mathcal{N}_{l+1}^{\mathrm{in}}(y)$. ■

Next, we identify the outcomes whose out-neighbors in the H-digraph do not change.

Proposition 6.2: (Sufficient conditions for invariance of out-neighboring structure of an outcome): Suppose player B takes an action that changes the outcome from $x(l)$ to $x(l+1)$. Let $x_{AA}^{\min} = \mathrm{argmin}_{z \in \{x(l), x(l+1)\}}\{\mathrm{rank}(z, \mathrm{P}_{AA})\}$, and $x_{AA}^{\max} = \mathrm{argmax}_{z \in \{x(l), x(l+1)\}}\{\mathrm{rank}(z, \mathrm{P}_{AA})\}$. If $y \notin \mathcal{M}_{BA}(x(l), x(l+1))$ and any of the following holds,

(i) $y \succ_{\mathrm{P}_{AA}} x_{AA}^{\min}$;
(ii) $y \prec_{\mathrm{P}_{AA}} x_{AA}^{\max}$ and $y \notin \mathbf{S}_{\mathrm{otcm}}|_{\pi_B(x_{AA}^{\min})} \cup \mathbf{S}_{\mathrm{otcm}}|_{\pi_B(x_{AA}^{\max})}$;
(iii) $y \in \mathbf{S}_{\mathrm{otcm}}|_{\pi_B(x_{AA}^{\max})}$ and $x_{AA}^{\max} \in \mathcal{N}_l^{\mathrm{out}}(y)$;
(iv) $y \in \mathbf{S}_{\mathrm{otcm}}|_{\pi_B(x_{AA}^{\min})}$ and $x_{AA}^{\min} \notin \mathcal{N}_l^{\mathrm{out}}(y)$;

then $\mathcal{N}_l^{\mathrm{out}}(y) = \mathcal{N}_{l+1}^{\mathrm{out}}(y)$.

Proof: We present the proof for the case $x(l) \succ_{\mathrm{P}_{AA}} x(l+1)$ (the proof for the case $x(l+1) \succ_{\mathrm{P}_{AA}} x(l)$ follows similarly). Thus $x_{AA}^{\min} = x(l)$ and $x_{AA}^{\max} = x(l+1)$. We begin by noting that if $y \notin \mathcal{M}_{BA}(x(l), x(l+1))$, any outcome which is perceived as a sanction-free improvement from y for B in $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l))$ is also perceived as a sanction-free improvement from y for this player in $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l+1))$. Thus, to complete the proof, we need to show that an outcome z is a sanction-free improvement from y for A in $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l))$ if and only if z is a sanction-free improvement from y for A in $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l+1))$. We prove this result for each of the cases identified in the statement. Let z be an improvement from y for A in $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l))$.

Consider case (i). If $z \notin \mathbf{S}_{\text{otcm}}|_{\pi_A(x(l))}$, since the new perceived improvements for B can only change in $\mathbf{S}_{\text{otcm}}|_{\pi_A(x(l))}$, B has a perceived sanction against the improvement z from y for A in $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l))$ if and only if such sanction exists in $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l+1))$. If $z \in \mathbf{S}_{\text{otcm}}|_{\pi_A(x(l))}$, since $y \succ_{\mathrm{P}_{AA}} x(l)$, the perceived sanctions of B against the improvement z from y are the same in $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l))$ and $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l+1))$.

Consider case (ii). Since $y \notin \mathbf{S}_{\text{otcm}}|_{\pi_B(x(l))} \cup \mathbf{S}_{\text{otcm}}|_{\pi_B(x(l+1))}$, we have that $z \neq x(l), x(l+1)$. If $z \succ_{\mathrm{P}_{BA}(l)} x(l)$ or $z \prec_{\mathrm{P}_{BA}(l)} x(l+1)$, then it is clear that there exists a perceived sanction against the improvement z of A in $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l))$ if and only if such a sanction exists against this improvement in $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l+1))$. Next, suppose $x(l) \succ_{\mathrm{P}_{BA}(l)} z \succ_{\mathrm{P}_{BA}(l)} x(l+1)$. Note that the only new perceived improvement from z for B is $x(l+1)$ and since $y \prec_{\mathrm{P}_{AA}} x(l+1)$, this does not affect the set of sanction-free improvements from y.

Consider case (iii). If $z \neq x(l+1)$, then, since the new perceived improvements of B can only change in $\mathbf{S}_{\text{otcm}}|_{\pi_A(x(l))}$, B has a perceived sanction against the improvement z from y for A in $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l))$ if and only if such sanction exists in $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l+1))$. By assumption, $x(l+1) \in \mathcal{N}_l^{\text{out}}(y)$, i.e., there exists no sanction of B against the improvement $x(l+1)$ from y for A. Since $\operatorname{rank}(x(l+1), \mathrm{P}_{BA}(l+1)) < \operatorname{rank}(x(l+1), \mathrm{P}_{BA}(l))$, we conclude that $x(l+1) \in \mathcal{N}_{l+1}^{\text{out}}(y)$.

Finally, consider case (iv). If $z \neq x(l)$, then, since the new perceived improvements of B can only change in $\mathbf{S}_{\text{otcm}}|_{\pi_A(x(l))}$, B has a perceived sanction against the improvement z from y for A in $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l))$ if and only if such sanction exists in $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l+1))$. It remains to show that $x(l) \notin \mathcal{N}_{l+1}^{\text{out}}(y)$. This holds since $\operatorname{rank}(x(l), \mathrm{P}_{BA}(l+1)) > \operatorname{rank}(x(l), \mathrm{P}_{BA}(l))$ and $x(l) \notin \mathcal{N}_l^{\text{out}}(y)$. ■

Propositions 6.1 and 6.2 give necessary conditions for an outcome to have different in- or out-neighbors under a change in A's perception about B. These results are important in the sense that B, without having access to the belief structure of A, can a priori establish which outcomes are guaranteed not to be affected in A's perception by an action of B. Conversely, if an outcome belongs to either one of the sets identified in the results, it is possible for A to update her belief structure in such a way that the neighboring structure of the outcome changes. Therefore, the results capture the best conclusion that B can draw without having access to the belief structure of A.

Remark 6.3: (Reducing the complexity of recomputing H-digraphs): A consequence of Proposition 6.1 is the simplification in the complexity of computing the new H-digraph that results from the changes in A's perception. Assuming the original H-digraph is available, one only needs to compute the changes in the in-neighboring structure of the outcomes characterized in Proposition 6.1. The number of these outcomes is $O(2n+m)$, where n and m are the number of actions available to A and B, respectively. Therefore, the complexity of modifying the H-digraph is $O(nm(2n+m))$, which is smaller that the complexity of computing it from scratch, cf. Lemma 2.12. •

Next, we turn our attention to the outcomes whose in- and out-neighbors are susceptible of change. Since the new out-neighbors can be identified via the new in-neighbors, we only study how the in-neighboring structure changes.

Theorem 6.4 (Changes of the in-neighboring structure): Suppose player B takes an action that changes the outcome from $x(l)$ to $x(l+1)$. The following holds,

(i) if $y \in \mathbf{S}_{\text{otcm}}|_{\pi_B(x(l))}$, then $\mathcal{N}_{l+1}^{\text{in}}(y) \subseteq \mathcal{N}_l^{\text{in}}(y)$;

(ii) if $y \in \mathbf{S}_{\text{otcm}}|_{\pi_B(x(l+1))}$, then $\mathcal{N}_l^{\text{in}}(y) \subseteq \mathcal{N}_{l+1}^{\text{in}}(y)$;

(iii) if $y \in \mathcal{M}_{BA}(x(l), x(l+1)) \cap \mathbf{S}_{\text{otcm}}|_{\pi_A(x(l))}$, then

- (a) $x(l) \in \mathcal{N}_{l+1}^{\text{in}}(y)$ if and only if $x(l+1) \in \mathcal{N}_l^{\text{in}}(y)$;
- (b) for $z \in \mathbf{S}_{\text{otcm}}|_{\pi_A(y)}$, $z \in \mathcal{N}_l^{\text{in}}(y) \setminus \{x(l+1)\}$ if and only if $z \in \mathcal{N}_{l+1}^{\text{in}}(y) \setminus \{x(l)\}$;
- (c) for $z \in \mathbf{S}_{\text{otcm}}|_{\pi_B(y)}$,
  - if $z \succ_{\mathrm{P}_{AA}} x_{AA}^{\min}$, then $z \notin \mathcal{N}_l^{\text{in}}(y) \cup \mathcal{N}_{l+1}^{\text{in}}(y)$;
  - if $z \prec_{\mathrm{P}_{AA}} x_{AA}^{\max}$, then $z \in \mathcal{N}_l^{\text{in}}(y)$ if and only if $z \in \mathcal{N}_{l+1}^{\text{in}}(y)$;
  - if $x_{AA}^{\min} \succ_{\mathrm{P}_{AA}} z \succ_{\mathrm{P}_{AA}} x_{AA}^{\max}$ and $x_{AA}^{\min} = x(l)$, then $z \notin \mathcal{N}_{l+1}^{\text{in}}(y)$;
  - if $x_{AA}^{\min} \succ_{\mathrm{P}_{AA}} z \succ_{\mathrm{P}_{AA}} x_{AA}^{\max}$ and $x_{AA}^{\min} = x(l+1)$, then $z \notin \mathcal{N}_l^{\text{in}}(y)$.

Proof: We first show (i). Let $y \in \mathbf{S}_{\text{otcm}}|_{\pi_B(x(l))}$ and $z \notin \mathcal{N}_l^{\text{in}}(y)$. Two things can happen:

- when $z \in \mathbf{S}_{\text{otcm}}|_{\pi_B(y)}$, two further possibilities might arise. If $y \neq x(l)$, then either y is not an improvement from z for A or there is a perceived sanction of B against the improvement y from z. Either of the cases will still hold after swapping $x(l)$ and $x(l+1)$ by A, and therefore $z \notin \mathcal{N}_{l+1}^{\text{in}}(y)$. If $y = x(l)$, the same reasoning plus the fact that $\operatorname{rank}(x(l), \mathrm{P}_{BA}(l)) < \operatorname{rank}(x(l), \mathrm{P}_{BA}(l+1))$ implies that $z \notin \mathcal{N}_{l+1}^{\text{in}}(y)$.
- when $z \in \mathbf{S}_{\text{otcm}}|_{\pi_A(y)}$, then either y is not an improvement from z for B or there is a perceived sanction by A against the improvement y from z. Either of the cases will still hold after the swap and thus $z \notin \mathcal{N}_{l+1}^{\text{in}}(y)$.

Next we show (ii). Let $y \in \mathbf{S}_{\text{otcm}}|_{\pi_B(x(l+1))}$ and suppose $z \in \mathcal{N}_l^{\text{in}}(y)$. Two things can happen:

- when $z \in \mathbf{S}_{\text{otcm}}|_{\pi_B(y)}$, two further possibilities might arise. If $y \neq x(l+1)$, then it is clear that $z \in \mathcal{N}_{l+1}^{\text{in}}(y)$, since there is no new sanction for B for the improvement y from z of A. If $y = x(l+1)$, the same reasoning plus the fact that $\operatorname{rank}(x(l+1), \mathrm{P}_{BA}(l)) > \operatorname{rank}(x(l+1), \mathrm{P}_{BA}(l+1))$ implies that $z \in \mathcal{N}_{l+1}^{\text{in}}(y)$.
- when $z \in \mathbf{S}_{\text{otcm}}|_{\pi_A(y)}$, then, since the improvement y from z for B remains free of sanctions, we conclude that $z \in \mathcal{N}_{l+1}^{\text{in}}(y)$.

Finally, we show part (iii). We start by (a). Suppose a perceived improvement y from $x(l+1)$ exists for B in the game $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l))$ without sanction of A. Then, since

$\operatorname{rank}(x(l), \mathrm{P}_{BA}(l+1)) = \operatorname{rank}(x(l+1), \mathrm{P}_{BA}(l))$, the improvement y from $x(l)$ with respect to the preference vector $\mathrm{P}_{BA}(l+1)$ is also sanction free. The converse follows similarly. Thus $x(l) \in \mathcal{N}_{l+1}^{\text{in}}(y)$ if and only if $x(l+1) \in \mathcal{N}_{l}^{\text{in}}(y)$. A similar argument shows that (b) holds. To end the proof, we show that (c) holds. Let $z \in \mathbf{S}_{\text{otcm}}|_{\pi_B(y)}$. Note that if $y \prec_{\mathrm{P}_{AA}} z$, all the statements hold trivially, since y is not an improvement from z for A. Thus we need to prove the results for $y \succ_{\mathrm{P}_{AA}} z$. If $z \succ_{\mathrm{P}_{AA}} x_{AA}^{\min}$, then any improvement y from z for A is sanctioned by the perceived improvement $x(l)$ from y for B in the game $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l))$. Similarly, any improvement y from z for A is sanctioned by the perceived improvement $x(l+1)$ from y for B in the game $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l+1))$; thus $z \notin \mathcal{N}_{l}^{\text{in}}(y) \cup \mathcal{N}_{l+1}^{\text{in}}(y)$. Now, suppose $z \prec_{\mathrm{P}_{AA}} x_{AA}^{\max}$. The only new perceived improvement from y for B is $x(l+1)$. Since $x(l+1) \succ_{\mathrm{P}_{AA}} z$, this improvement does not create any new sanction against the improvement y from z for A. Similarly, the only removed perceived improvement from y for B is $x(l)$. Since $x(l) \succ_{\mathrm{P}_{AA}} z$, $x(l)$ is not a sanction of B in $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l+1))$; thus $z \in \mathcal{N}_{l}^{\text{in}}(y)$ if and only if $z \in \mathcal{N}_{l+1}^{\text{in}}(y)$. Next, suppose $x_{AA}^{\min} \succ_{\mathrm{P}_{AA}} z \succ_{\mathrm{P}_{AA}} x_{AA}^{\max}$. If $x_{AA}^{\min} = x(l)$, then $x(l+1)$ is a perceived sanction of B in $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l+1))$ and thus $z \notin \mathcal{N}_{l+1}^{\text{in}}(y)$. If $x_{AA}^{\min} = x(l+1)$, then $x(l)$ is a perceived sanction of B in $(\mathrm{P}_{AA}, \mathrm{P}_{BA}(l))$ and thus $z \notin \mathcal{N}_{l}^{\text{in}}(y)$. This completes the proof. ■

If the action taken by B is aligned with B's game as perceived by A, i.e., if $x(l+1) \succ_{\mathrm{P}_{BA}(l)} x(l)$, then in Propositions 6.1 and 6.2, and in Theorem 6.4, the sets prescribed by $\succeq_{\mathrm{P}_{BA}}$ are empty. This is consistent with the fact that no change in A's perception occurs in this case.

Remark 6.5 (Belief manipulation and deception): If B has complete information about A's game H_A^0, then she can use the `H-digraph algorithm` to study the changes in the belief structure of A and possibly manipulate it. The results presented above are helpful because they narrow down the outcomes on which an action of B would have an effect on. This opens the way for algorithmic approaches to belief manipulation in hypergames. Also importantly, the results capture the outcomes that B does not have direct control over and for which she may need a sequence of actions, instead of a single one, to manipulate A's belief. •

Example 6.6 (An example of deception): Here, we present an example in which one of the players has perfect information about the other player's game and is aware of this fact, while the second player is trying to update his misperceptions by observing the actions of her opponent. We show how the player with perfect information may be able to deceive the opponent. Our discussion follows the scenario presented in Example 5.12. Note that in the 1-level hypergame introduced in the example, B has perfect information about A but is not aware of it. To model this fact, we consider instead a 2-level hypergame $H^2 = \{H_A^0, H_B^1\}$, with $H_B^1 = \{H_A^0, H_B^0\}$. In particular, $\mathrm{P}_{ABB} = \mathrm{P}_{AAB} = \mathrm{P}_{AB} = \mathrm{P}_{AA}$, $\mathrm{P}_{BAB} = \mathrm{P}_{BA}$. We assume that A is using a modified swap learning scheme to update her perceptions about B. We show that B can deceive A so that eventually A believes that the outcome x_1, the best outcome for B, is an equilibrium. As in Example 5.12, the initial outcome is x_1. B gets the first chance to move and does not take any action. A observes this and uses modified swap learning to update her perception as $\mathbf{MSw}_{x_1,x_1}^A(\mathrm{P}_{BA}) = (x_1, x_3, x_2, x_4)^T$. Note that x_1 is unstable for A and hence, in her turn, takes an action that changes the outcome from x_1 to x_3. Outcome x_3 is sequentially rational for B in H_B^0, but

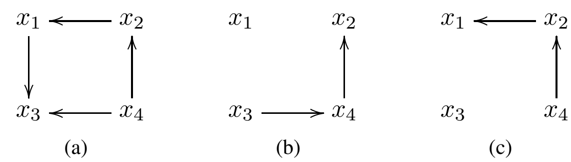


Fig. 5. H-digraph of H_A^0 after applying (a) $\mathbf{MSw}_{x_1,x_1}^A$, (b) $\mathbf{MSw}_{\mathcal{O}_{BA}(1)}^A$, and (c) $\mathbf{MSw}_{\mathcal{O}_{BA}(2)}^A$, respectively.

B prefers the outcome x_1 to x_3. Therefore, with the intention of deceiving A, B takes an irrational action that changes the outcome to x_4. Since there is no prior observation about B, by adding this information to her observation set $\mathcal{O}_{BA}(1)$, A updates her perception about B as follows,

$$\begin{aligned}\mathbf{MSw}_{\mathcal{O}_{BA}(1)}^A(\mathrm{P}_{BA}) &= \mathbf{MSw}_{x_3,x_4}^A\left(\mathbf{MSw}_{x_1,x_1}^A(\mathrm{P}_{BA})\right)\\ &= (x_1, x_4, x_2, x_3)^T.\end{aligned}$$

As a result, x_1 becomes sequentially rational for A. Next, A takes an action that changes the outcome to x_2. Finally, B takes an action that changes the outcome to x_1. Therefore, A changes her perception about B to

$$\begin{aligned}\mathbf{MSw}_{\mathcal{O}_{BA}(2)}^A(\mathrm{P}_{BA}) &= \mathbf{MSw}_{x_2,x_1}^A\left(\mathbf{MSw}_{\mathcal{O}_{BA}(1)}^A(\mathrm{P}_{BA})\right)\\ &= (x_1, x_4, x_3, x_2)^T,\end{aligned}$$

where the map $\mathbf{MSw}_{x_2,x_1}^A$ is compatible with the observation set $\mathcal{O}_{BA}(1)$. Thus the hypergame converges to x_1. This evolution is shown in Figure 5(a)-(c). This example raises some interesting questions, including the potential use by B of general algorithmic techniques to perform deception and by A of an analysis similar to the one in Section VI to detect the possibility of deception. •

VII. Conclusions

We have studied adversarial situations where players' perceptions about the game they are involved in might be inconsistent and evolving. We have introduced the swap learning method to allow players to incorporate into their beliefs the information gained from observing the opponents' actions. A player that uses this method decreases her misperception at the cost of potentially incurring in inconsistencies in her perception. This has motivated the introduction of the modified swap learning method, which yields consistent beliefs and, under the assumption that the opponent has perfect information and plays her best strategy, also decreases the misperception. Using the newly introduced notion of H-digraph, we have fully characterized how a player's perception is affected by the actions taken by other players.

The methods discussed here attribute the origin of the misperception on the player doing the update. Numerous avenues for future research appear open, including the exploration of other learning schemes and extensions to high-level hypergames. Learning methods at the other extreme of the spectrum, where inconsistencies are blamed on the opponents' misperceptions, and in the middle of the spectrum, via the construction of hypergames of higher level, are also worth exploring. Another direction of research is the study of learning under imperfect observation and the use of probabilistic methods to update the preference vectors for the opponents. It is also worth investigating how misperception can be decreased by departing from sequentially rational outcomes when the cost of such irrational actions is not prohibitive. We also plan to use our results on the evolution of H-digraphs in the design of deception and deception-robust strategies.

ACKNOWLEDGMENTS

This work was partially supported in part by Award FA9550-10-1-0499. We wish to thank the anonymous referees and the associate editor for their valuable comments.

REFERENCES

[1] B. Gharesifard and J. Cortés, "Evolution of the perception about the opponent in hypergames," in *IEEE Conf. on Decision and Control*, (Atlanta, Georgia), pp. 1076–1081, 2010.

[2] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA: MIT Press, 1991.

[3] J. C. Harsanyi, "Bargaining in ignorance of the opponents' utility function," *Journal of Conflict Resolution*, vol. 6, no. 1, pp. 29–38, 1962.

[4] J. C. Harsanyi, "Games with incomplete information played by 'Bayesian' players, i-iii," *Management Science*, vol. 50, no. 12 Supplement, pp. 1804–1817, 2004.

[5] D. Teneketzis and D. A. Castañón, "Informational aspects of class of subjective games of incomplete information: static case," *Journal of Optimization Theory & Applications*, vol. 54, no. 2, pp. 413–422, 1987.

[6] D. P. Foster and H. P. Young, "On the impossibility of predicting the behavior of rational agents," *Proceedings of the National Academy of Sciences*, vol. 98, no. 222, pp. 12848–12853, 2001.

[7] V. P. Crawford, "Lying for strategic advantage: Rational and boundedly rational misrepresentation of intentions," *American Economic Review*, vol. 93, pp. 133–149, 2003.

[8] S. Mullainathan, J. Schwartzstein, and A. Shleifer, "Coarse thinking and persuasion," *Quarterly Journal of Economics*, vol. 123, no. 5, pp. 577–619, 2008.

[9] D. Ettinger and P. Jehiel, "A theory of deception," *American Economic Journal: Microeconomics*, vol. 2, pp. 1–20, 2010.

[10] P. Jehiel and F. Koessler, "Revising games of incomplete information with analogy-based expectations," *Journal of Economic Theory*, vol. 62, pp. 533–557, 2008.

[11] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*. SIAM, 2 ed., 1999.

[12] A. Kott and W. McEneaney, eds., *Adversarial Reasoning: Computational Approaches to Reading the Opponent's Mind*. Boca Raton, FL: Chapman and Hall/CRC, 2007.

[13] P. Bernhard, A. L. Colomb, and G. Papavassilopoulos, "Rabbit and hunter game: Two discrete stochastic formulations," *International Journal of Applied Mathematics and Computer Science*, vol. 13, pp. 205–225, 1987.

[14] G. J. Olsder and G. P. Papavassilopoulos, "About when to use a searchlight," *Journal of Mathematical Analysis and Applications*, vol. 136, pp. 466–487, 1988.

[15] J. P. Hespanha, "Application and value of deception," in *Adversarial Reasoning: Computational Approaches to Reading the Opponent's Mind* (A. Kott and W. M. McEneaney, eds.), (Boca Raton, FL), pp. 145–165, Chapman and Hall/CRC, 2007.

[16] D. Li and J. B. C. Jr., "Information, decision-making and deception in games," *Decision Support Systems*, vol. 47, pp. 518–527, 2009.

[17] D. A. Castañón, P. R. Chandler, and M. Pachter, "A game of deception," in *IEEE Conf. on Decision and Control*, (Paradise Island, Bahamas), pp. 3364–3369, 2004.

[18] W. M. McEneaney, "Some classes of imperfect information finite state space stochastic games with finite-dimensional solutions," *Applied Mathematics and Optimization*, vol. 50, no. 2, pp. 87–118, 2004.

[19] W. M. McEneaney and R. Singh, "Robustness against deception," in *Adversarial Reasoning: Computational Approaches to Reading the Opponent's Mind* (A. Kott and W. McEneaney, eds.), (Boca Raton, FL), pp. 167–207, Chapman and Hall/CRC, 2007.

[20] P. G. Bennett, "Towards a theory of hypergames," *OMEGA*, vol. 5, no. 6, pp. 749–751, 1977.

[21] N. M. Fraser and K. W. Hipel, *Conflict Analysis: Models and Resolutions*. Systems Science and Engineering, Oxford, UK: Elsevier, 1984.

[22] M. Wang, K. W. Hipel, and N. M. Fraser, "Solution concepts in hypergames," *Applied Mathematics and Computation*, vol. 34, pp. 147–171, 1989.

[23] R. Vane and P. Lehner, "Using hypergames to increase planned payoff and reduce risk," *Autonomous Agents and Multi-Agent Systems*, vol. 5, pp. 365–380, 2002.

[24] P. G. Bennett, "Hypergames, the development of an approach to modeling conflicts," *Futures*, vol. 12, pp. 489–507, 1980.

[25] B. L. Slantchev, "The principle of convergence in wartime negotiations," *American Political Science Review*, vol. 97, no. 4, pp. 621–632, 2003.

[26] J. T. House and G. Cybenko, "Hypergame theory applied to cyber attack and defense," in *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense IX*, vol. 7666, (Orlando, Florida), pp. 766604–766611, April 2010.

[27] B. Chaib-dara, "Hypergame analysis in multiagent environments," in *AAAI Spring Symposia*, (Stanford, CA), pp. 147–149, Mar. 2001.

[28] R. Vane, "Advances in hypergame theory," in *Proceedings of the Eight Workshop on Game Theoretic and Decision Theoretic Agents*, (Hakodate, Japan), 2006.

[29] Y. Sasaki, N. Kobayashi, and K. Kijima, "Mixed extension of hypergames and its applications to inspection games," in *Proceedings of the 51nd Annual Meeting of the ISSS*, (Tokyo, Japan), Aug. 2007.

[30] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*. MIT Press, 2 ed., 2001.

Bahman Gharesifard received his undergraduate degree in Mechanical Engineering, in 2002, and his M.Sc. degree in Control and Dynamics, in 2005, from Shiraz University, Iran. He received the Ph.D. degree in Mathematics from Queen's University, Canada, in 2009. Since then he has been a postdoctoral research associate with the Department of Mechanical and Aerospace Engineering at University of California, San Diego. His research interests include autonomy, distributed optimization, sensor networks, social and economic networks, game theory, geometric control and mechanics, and Riemannian geometry. He is the recipient of R. S. McLaughlin Fellowship award in 2008.

Jorge Cortés received the Licenciatura degree in mathematics from the Universidad de Zaragoza, Spain, in 1997, and the Ph.D. degree in engineering mathematics from the Universidad Carlos III de Madrid, Spain, in 2001. He held postdoctoral positions at the University of Twente, The Netherlands, and the University of Illinois at Urbana-Champaign, USA. He was an Assistant Professor with the Department of Applied Mathematics and Statistics at the University of California, Santa Cruz from 2004 to 2007. He is currently an Associate Professor in the Department of Mechanical and Aerospace Engineering, University of California, San Diego. His research interests include cooperative control, game theory, spatial estimation, distributed optimization, and geometric mechanics. He is the author of *Geometric, Control and Numerical Aspects of Nonholonomic Systems* (Springer-Verlag, 2002) and co-author (together with F. Bullo and S. Martínez) of *Distributed Control of Robotic Networks* (Princeton University Press, 2009). He received a NSF CAREER award in 2006 and was the recipient of the 2006 Spanish Society of Applied Mathematics Young Researcher Prize. He has co-authored papers that won the 2008 IEEE Control Systems Outstanding Paper Award and the 2009 SIAM Review SIGEST selection from SIAM Journal on Control and Optimization. He is a IEEE Control Systems Society Distinguished Lecturer (2010-2012).