Stealthy deception in hypergames under informational asymmetry

Bahman Gharesifard Jorge Cortés

Abstract

This paper considers games of incomplete information, where one player (the deceiver) has an informational advantage over the other (the mark) and intends to employ it for belief manipulation. We use the formalism of hypergames to represent the asymmetric information available to players. This framework allows us to formalize various notions of belief manipulation that revolve around the idea of the deceiver being able to make the mark believe that a particular action has lost its advantageous character. In the case when the deceiver does not mind revealing information to the mark as the game evolves, we provide a necessary condition and a sufficient condition for deceivability. In the case when the deceiver acts in a stealthy way, i.e., restricts its actions to those that do not contradict the belief of the mark, we fully characterize when deception is possible and design the `worst-case max-strategy` to find a sequence of deceiving actions. Our correctness guarantees for this strategy are based on a precise characterization of the acyclic structure of subjective hypergames. An example illustrates our results.

I. Introduction

Informational asymmetry in strategic scenarios provide opportunities for manipulating beliefs or inducing certain desired perceptions. In this paper, we consider a class of games where one player (the deceiver) wishes to misrepresent certain information in order to gain a strategic advantage over the opponent (the mark). Formalizing such scenarios in a way which is mathematically precise presents many challenges. Here, we employ the language of hypergames. In our setup, the informational advantage of the deceiver allows it to anticipate the effect that its actions

Bahman Gharesifard and Jorge Cortés are with the Department of Mechanical and Aerospace Engineering, University of California, San Diego, {bgharesifard,cortes}@ucsd.edu. Preliminary versions of this work were presented at the 2011 American Control Conference, San Francisco, CA and the 2011 IEEE Conference on Decision and Control, Orlando, Florida.

will have on the mark's belief structure. In this sense, the deception goal can be understood as steering the evolution of a specific dynamical system into a desired set of outcomes. Scenarios of interest includes bargaining, cybersecurity, human decision making, and social interactions.

Literature review: The notion of hypergame goes back to [1] and is mostly used in the context of conflict analysis [2], [3]. In a hypergame, players can have different levels of perception about their opponents' game, in the sense that they might have perceptions about the opponents' preferences or about what the opponents think about their preferences, and so on. Hypergames are particularly useful in scenarios where players are absolutely certain about their opponents' perceptions, while these certainties may be mutually inconsistent. Some well-known examples of the application of hypergame analysis include the Normandy invasion and the Cuban missile crisis [3]. Hypergames are also relevant when players play security strategies or when the cost of risky actions is high, such as wartime negotiation [4] and cybersecurity [5].

The explicit modeling of misperception makes hypergames appropriate for the study of deception. Deception in the context of games with incomplete information has been scarcely studied.

The inconsistent structure of beliefs can lead to counterintuitive behaviors, as shown in [6]. The work [7] studies deception via strategic communication, in which a 'sophisticated' player sends either truthful or false messages to the opponents. The vulnerability of strategic decision makers to persuasion is investigated in [8]. The recent work [9] constructs a theory of deception for games with incomplete information where players form expectations about the average behavior of the other players based on past histories. Scenarios where one player has access to certain information and can distort it before it is passed on to others. Deception also arises in games of imperfect information are studied in [10] and [11]. Within this class, early references include [12], [13]. The works of [14], [15] and [16] provide examples of how informational asymmetries can be used to induce false perceptions in the opponent and lead to strategic deception. The works of [17] and [18] provide efficient deception-robust schemes for a class of discrete dynamic stochastic games under imperfect observations.

Statement of contributions: We consider games of incomplete information where players have different perceptions about the scenarios they are involved in. We study a class of 2-player hypergames where the deceiver has full information about the mark's game and intends to induce a certain belief in it. Each player's belief structure is encoded in a special class of digraph that we term H-digraph. The mark is rational, observes the actions taken by the deceiver, assumes it acts rationally (although it may not), and updates its H-digraph accordingly. From the deceiver's viewpoint, the mark's actions are rational. Since the deceiver does not have knowledge of the strategy that the mark follows to choose its actions, it adopts a probabilistic model. This framework sets the stage for the first contribution of the paper, which is the introduction of precise notions of deception to capture different forms of belief manipulation. These notions allow us to identify a necessary condition and a sufficient condition for manipulating a particular edge in the mark's H-digraph. Motivated by the observation that an action by the deceiver that contradicts the current mark's belief structure might trigger complex changes in its H-digraph, we next study scenarios where the deceiver purposefully restricts its actions to those that are aligned with the mark's beliefs. We term these actions stealthy. Our second contribution is the characterization of when deception via stealthy actions is possible. This result allows us to cast the search for stealthy sequences that achieve deception as the problem of finding a path in the mark's H-digraph between the current outcome and a desired goal set. Our third contribution is then the design of the `worst-case max-strategy` to find such a stealthy sequence of actions. The correctness guarantees for this strategy are based on the graph-theoretic properties of H-digraphs, that we also characterize. Specifically, our fourth contribution shows that the H-digraph associated to the perceived game of each player contains no weak improvement cycle which, when players preferences are strict, means that the H-digraph is acyclic. This result draws an interesting analogy with ordinal potential games. An example illustrates our results.

Organization: Section II presents preliminaries on graph theory and hypergames. Section III introduces various notions of deception and belief manipulation. Section IV presents a necessary condition for deceivability and a sufficient condition for surely deceivability in case the deceiver takes actions that may change the belief structure of the mark. We also present a full characterization of deceivability in the case when the deceiver acts in a stealthy way. Section V introduces the `worst-case max-strategy` to allow the deceiver to find a stealthy deceiving strategy. Section VI gather our conclusions and ideas for future work. Finally, Appendix A contains an instrumental result which unveils the acyclic structure of H-digraphs.

II. Preliminaries

In this section, we review basic notions from graph theory and hypergames. We begin with some basic notation. We let $\mathbb{R}$, $\mathbb{R}_{>0}$, and $\mathbb{R}_{\geq 0}$ denote the set of reals, positive reals, and non-negative reals, respectively. We denote by $\mathbb{Z}_{\geq k}$ the set of integers greater than or equal to $k \in \mathbb{R}$. We denote by $\mathbf{I}_{n\times n}$ the identity matrix in $\mathbb{R}^{n\times n}$, $n \in \mathbb{Z}_{\geq 1}$. A *preorder* $\succeq$ on a set X is a reflexive and transitive binary relation. We use $\sigma = (x_1, x_2, \ldots)$, where $x_1, x_2, \ldots \in X$, to denote a sequence of elements in X. The *length* of a finite sequence σ is the number of elements in σ.

A. Basic graph notions

We review some graph-theoretic notions following [19]. A directed graph, or simply 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, respectively, by $\mathcal{N}^{\text{in}}(v)$ and $\mathcal{N}^{\text{out}}(v)$. The *in-degree* and *out-degree* of v are the number of in-neighbors and out-neighbors of v, respectively. The (unweighted) *adjacency matrix* of G is the matrix $\text{Adj}(G) \in \mathbb{R}_{\geq 0}^{|V|\times|V|}$ defined as follows: for each $v_i, v_j \in V$, $\text{Adj}(G)_{ij} = 1$ if $(v_i, v_j) \in E$, and $\text{Adj}(G)_{ij} = 0$ otherwise. A *directed path* in a digraph, or in short path, is an ordered sequence of vertices so that any two consecutive vertices are an edge of the digraph. A vertex u is *reachable* from v if there exists a path starting at v and ending at u. A *cycle* in a digraph is a directed path that starts and ends at the same vertex and has no other repeated vertex. A digraph without any cycle is *acyclic*.

B. Hypergame theory

In this section, we review the basic notions of hypergame theory following [3], [20], [1]. A *(finite) game* [21], [22] is a triplet $\mathbf{G} = (V, \mathbf{S}_{\text{outcome}}, \mathbf{P})$ with the following elements: $V = \{A_1, \ldots, A_n\}$ is a set of $n \in \mathbb{Z}_{\geq 1}$ players, $\mathbf{S}_{\text{outcome}} = S_1 \times \ldots \times S_n$ is the outcome set with cardinality $N = |\mathbf{S}_{\text{outcome}}| \in \mathbb{Z}_{\geq 1}$, where S_i is a finite set of actions available to player $A_i \in V$, and $\mathbf{P} = (P_1, \ldots, P_n)$, with $P_i = (x_1, \ldots, x_N)^T \in \mathbf{S}_{\text{P}}$, the *preference vector* of player A_i, where $\mathbf{S}_{\text{P}} \subset \mathbf{S}_{\text{outcome}}^N$ denotes the set of all elements of $\mathbf{S}_{\text{outcome}}^N$ with pairwise different entries. We denote by π_i the natural projection of $\mathbf{S}_{\text{outcome}}$ onto the strategy set S_i of the ith player. We also use π_{-i} to denote the natural projection of $\mathbf{S}_{\text{outcome}}$ onto $S_{-i} = S_1 \times S_2 \times \cdots \times \hat{S}_i \times \cdots \times S_n$, where the hat denotes that S_i is excluded from the product. Each preference vector P_i is equipped with a preorder $\succeq_{P_i}$ such that, if x has a lower entry index that y in P_i, then $x \succeq_{P_i} y$. With this definition, the emphasis is put on the order of preferences among outcomes, rather than on the actual payoff that players obtain for each specific outcome.

A 0-level hypergame is simply a finite game. A 1-*level hypergame* with n players is a set $H^1 = \{\mathbf{G}_1, \ldots, \mathbf{G}_n\}$, where $\mathbf{G}_i = (V, (\mathbf{S}_{\text{outcome}})_i, \mathbf{P}_i)$, for $i \in \{1, \ldots, n\}$, is the subjective finite game of player $A_i \in V$, and V is a set of n players; $(\mathbf{S}_{\text{outcome}})_i = S_{1i} \times \ldots \times S_{ni}$, with S_{ji} the finite set of strategies available to A_j, as perceived by A_i; $\mathbf{P}_i = (P_{1i}, \ldots, P_{ni})$, with P_{ji} the preference vector of A_j, as perceived by A_i. Throughout this paper, we assume that all players have the same outcome set, and we denote it by $\mathbf{S}_{\text{outcome}}$. Being each subjective game a finite game, these preference vectors are equipped with a preorder. In a 1-level hypergame, each player $A_i \in V$ plays the game $\mathbf{G}_i$ with the perception that it is playing a game with complete information, which is not necessarily true. This is in contrast with Bayesian games [23], [24], where the players' perceptions about other player preferences are probabilistic. The definition of a 1-level hypergame can be extended to higher-level hypergames, where some of the players have access to additional information that allow them to form perceptions about other players' perceptions, other players' perceptions about them, and so on. One can, inductively, extend the definition of 1-level hypergame as follows: a k-*level hypergame* with n players, $k \geq 1$, 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 $i \in \{1, \ldots, n\}$.

Example 2.1 (Sample hypergame): Consider a 1-level hypergame $H^1 = \{H_{A_1}^0, H_{A_2}^0\}$ with two players, A_1 and A_2, which choose their strategies from the sets $S_{A_1} = \{a_1, a_2, a_3\}$ and $S_{A_2} = \{b_1, b_2\}$, respectively. The outcome set is given by $\mathbf{S}_{\text{outcome}} = \{x_1, x_2, x_3, x_4, x_5, x_6\}$, where

$$x_1 = (a_1, b_1), \quad x_2 = (a_2, b_1), \quad x_3 = (a_3, b_1),$$
$$x_4 = (a_1, b_2), \quad x_5 = (a_2, b_2), \quad x_6 = (a_3, b_2).$$

Players A_1 and A_2's preference vectors are, respectively, given by

$$\mathrm{P}_{A_1A_1} = (x_1, x_5, x_6, x_2, x_3, x_4)^T \text{ and } \mathrm{P}_{A_2A_1} = (x_4, x_5, x_6, x_1, x_2, x_3)^T,$$
$$\mathrm{P}_{A_2A_2} = (x_4, x_5, x_2, x_1, x_3, x_6)^T \text{ and } \mathrm{P}_{A_1A_2} = \mathrm{P}_{A_1A_1}.$$

Note that player A_2's perception about the preferences of player A_1 is correct. The 0-level hypergames of each player are then given by $H^0_{A_1} = (\{A_1, A_2\}, \mathbf{S}_{\text{outcome}}, (\mathrm{P}_{A_2A_1}, \mathrm{P}_{A_1A_1}))$ and $H^0_{A_2} = (\{A_1, A_2\}, \mathbf{S}_{\text{outcome}}, (\mathrm{P}_{A_1A_2}, \mathrm{P}_{A_2A_2}))$. •

1) Sequential rationality: Consider a k-level hypergame H^k between players $\{A_1, \dots, A_n\}$ with outcome set $\mathbf{S}_{\text{outcome}}$. Without loss of generality and for simplicity, we assume that H^k is homogeneous. For $x \in \mathbf{S}_{\text{outcome}}$, we denote by $\mathbf{S}_{\text{outcome}}|_{\pi_i(x)}$ the set of outcomes $y \in \mathbf{S}_{\text{outcome}}$ such that $\pi_i(y) = \pi_i(x)$. For a sequence σ of length k on the set $\{A_1, \dots, A_n\}$, let $\mathrm{P}_{A_i\sigma}$, $i \in \{1, \dots, n\}$, denote the preferences of A_i as perceived by σ in H^k. For instance, $\mathrm{P}_{A_1A_2A_1}$ corresponds to what player A_1 perceives that player A_2 thinks about player A_1's preferences in a 2-level hypergame H^2. Given that the set of outcomes is common to all players, we use the shorthand notation $(\mathrm{P}_{A_1\sigma}, \dots, \mathrm{P}_{A_n\sigma})$ to denote the 0-level hypergame

$$H^0_\sigma = (\{A_1, \dots, A_n\}, \mathbf{S}_{\text{outcome}}, (\mathrm{P}_{A_1\sigma}, \dots, \mathrm{P}_{A_n\sigma})),$$

often referred to as the *subjective* hypergame perceived by σ. With a slight abuse of notation, $\succeq_{A_i\sigma}$ denotes the binary relation $\succeq_{\mathrm{P}_{A_i\sigma}}$ on $\mathbf{S}_{\text{outcome}}$.

Given two distinct outcomes $x, y \in \mathbf{S}_{\text{outcome}}$, $y \neq x$, y is an *improvement* from x for player A_i perceived by σ in H^0_σ if and only if $\pi_{-i}(y) = \pi_{-i}(x)$ and $y \succeq_{A_i\sigma} x$. An outcome $x \in \mathbf{S}_{\text{outcome}}$ is *rational* for player A_i in H^0_σ if there exists no improvement from x for this player. Finally, $x \in \mathbf{S}_{\text{outcome}}$ is *sequentially rational* for A_i in H^0_σ if and only if for each improvement y from x for A_i in H^0_σ there exists $z \in \mathbf{S}_{\text{outcome}}$ which *sanctions* y, i.e., $\pi_i(z) = \pi_i(y)$ and $x \succ_{A_i\sigma} z$ such that for all $j \in \{1, \dots, n\} \setminus \{i\}$, either $\pi_j(z) = \pi_j(y)$, or the outcome $z_{A_j} \in \mathbf{S}_{\text{outcome}}|_{\pi_i(y)}$, where $\pi_j(z_{A_j}) = \pi_j(z)$ and $\pi_l(z_{A_j}) = \pi_l(y)$, for all $l \in \{1, \dots, n\} \setminus \{j\}$, is an improvement from y in $\mathbf{S}_{\text{outcome}}|_{\pi_i(y)}$ for A_j. We denote the set of all sequentially rational outcomes for A_i in H^0_σ by $\mathbf{Seq}_{A_i}(H^0_\sigma)$. A rational outcome is also sequentially rational. A player is rational if it only takes actions associated to sanction-free improvements. It is known [3], [20] that all 0-level hypergames have at least one sequentially rational outcome.

2) H-digraphs: The notion of H-digraph introduced in [25], generalized here to n players, contains the information about the improvements from an outcome to other outcomes, the equilibria, and the sanctions. Consider a homogeneous k-level hypergame H^k between players $\{A_1, \dots, A_n\}$. Given σ, a sequence of length k on the set of player, and $i \in \{1, \dots, n\}$, we assign to each $x \in \mathbf{S}_{\text{outcome}}$ a positive number $\mathrm{rank}(x, \mathrm{P}_{A_i\sigma}) \in \mathbb{R}_{>0}$, called *rank*, such that, for

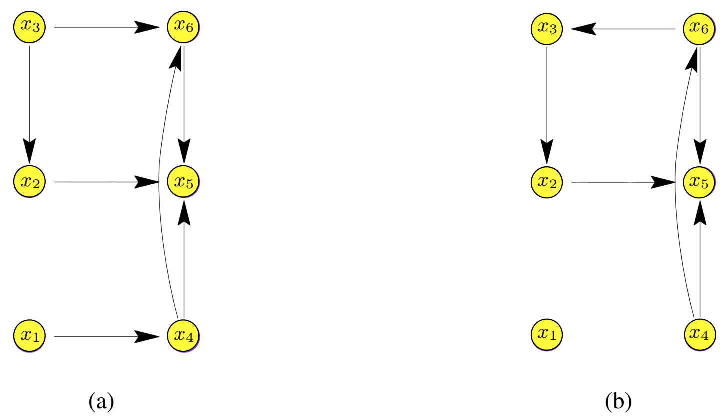

Fig. 1. H-digraphs of (a) player A_1 and (b) player A_2 in the hypergame described in Example 2.1. Vertical (respectively, horizontal) edges correspond to actions of player A_1 (respectively, player A_2).

each $\mathbf{S}_{\text{outcome}} \ni y \neq x$, we have $\text{rank}(y, \text{P}_{A_i\sigma}) > \text{rank}(x, \text{P}_{A_i\sigma})$ if $x \succ_{A_i\sigma} y$. The n-dimensional digraph $\mathcal{G}_{H^0_\sigma} = (\mathbf{S}_{\text{outcome}}, \mathcal{E}_{H^0_\sigma})$ is the *H-digraph* associated to the 0-level hypergame H^0_σ, where

- each vertex $x \in \mathbf{S}_{\text{outcome}}$ is labeled with $(\text{rank}(x, \text{P}_{A_1\sigma}), \ldots, \text{rank}(x, \text{P}_{A_n\sigma}))$, and
- (x, y) belongs to $\mathcal{E}_{H^0_\sigma}$ if and only if $\pi_i(x) \neq \pi_i(y)$, for some $i \in \{1, \ldots, n\}$, $\pi_{-i}(x) = \pi_{-i}(y)$, and there exists a perceived improvement y from x for player A_i in H^0_σ for which there exists no sanction of players A_{-i}, perceived by σ.

Figure 1 shows the H-digraphs of each player in Example 2.1.

Sequentially rational outcomes can be easily characterized by looking at the H-digraph. In fact, given $x \in \mathbf{S}_{\text{outcome}}$ and $i \in \{1, \ldots, n\}$, one has

$$x \in \mathbf{Seq}_{A_i}(H^0_\sigma) \quad \text{if and only if} \quad \mathcal{N}^{\text{out}}(x) \cap \mathbf{S}_{\text{outcome}}|_{\pi_{-i}(x)} = \emptyset.$$

It is worth mentioning that the complexity of computing H-digraphs is in $\Theta(|\mathbf{S}_{\text{outcome}}|^2)$, see [25].

3) Equilibria: An outcome $x \in \mathbf{S}_{\text{outcome}}$ is *unstable* for A_i, perceived by σ, in H^0_σ, $i \in \{1, \ldots, n\}$ if it is not sequentially rational and is an *equilibrium of* H^0_σ if it is sequentially rational for all players A_i, $i \in \{1, \ldots, n\}$, with respect to H^0_σ. Note that more than one equilibrium might exist. In terms of the H-digraph, an outcome is an equilibrium of H^0_σ if and only if its out-degree in $\mathcal{G}_{H^0_\sigma}$ is zero. An outcome is an equilibrium of H^k if it is sequentially rational in all H^0_σ, where $\sigma = A_iA_i \ldots A_i$, $i \in \{1, \ldots, n\}$ is a sequence of length k on $\{A_1, \ldots, A_n\}$. One can similarly define the notion of equilibrium for any intermediate level hypergame $H^{k_1}_\eta$, where $k_1 < k$ and η is sequence of length at most $k - 1$ on $\{A_1, \ldots, A_n\}$. Note that an outcome might be an equilibrium of H^k and not an equilibrium of $H_\eta^{k_1}$ and vice versa. For brevity, we sometimes omit the wording 'with respect to H_σ^0' and 'perceived by σ' when it is clear from the context. Other notions of equilibria have been considered for hypergames, see [3].

Example 2.2: (Example 2.1 revisited): Consider the 1-level hypergame $H^1 = \{H_{A_1}^0, H_{A_2}^0\}$ of Example 2.1. Let us illustrate the notions introduced so far. Observing the H-digraph of player A_2, see Figure 1(b), the outcome x_3 is perceived as unstable for A_1 by player A_2, since player A_1 is perceived to have a sanction-free improvement to x_2 from x_3. Instead, the outcome x_5 is perceived as sequentially rational for A_1 by A_2, since there is no sanction-free improvement to any other outcome for A_1. The set of equilibria of H^2 is given by $\{x_1, x_5\} = \mathbf{Seq}_{A_1}(H_{A_1}^0) \cap \mathbf{Seq}_{A_2}(H_{A_2}^0)$. •

III. Modeling deception and belief manipulation

This section introduces the notions of deception and belief manipulation upon which we base our analysis. The scenario we consider is as follows: a deceiver D and a mark M are engaged in a hypergame. The deceiver has full knowledge of the preferences of the mark. Both players can perfectly observe the actions taken by the other player. According to Section II-B, the mark's belief structure is encoded in its H-digraph. It is therefore natural to formally define belief manipulation in terms of specific changes that the deceiver wants to induce in the topology of this digraph. Before doing so, however, two mechanisms need to be specified: how the mark incorporates the lessons learned from the actions taken by the deceiver into its H-digraph and how the deceiver models the strategy followed by the mark to choose its actions.

After formally introducing in Section III-A the elements of the problem scenario, we describe the two mechanisms mentioned above in Sections III-B and III-C, respectively. With all the ingredients in place, Section III-D introduces the notions of deception.

A. Problem setup

Consider a 2-player, 2-level hypergame $H^2 = \{H_M^0, H_D^1\}$, where $H_D^1 = \{H_{MD}^0, H_{DD}^0\}$ is such that $H_{MD}^0 = H_M^0$ (i.e., the deceiver has full information about the mark's preferences). Note that, because of the special form of H^2, its equilibria are exactly the same as the equilibria of $H_D^1 = \{H_M^0, H_D^0\}$. In line with the notion of sequential rationality, we assume throughout this paper that, in all sequences of outcomes, players take actions sequentially, one after the other. Note that scenarios where one player takes multiple actions before the other player can be accommodated within this assumption. A finite sequence of outcomes $\sigma = (x_0, x_1, \ldots, x_k)$ is called *stealthy* if $(x_\ell, x_{\ell+1}) \in \mathcal{E}_{H_M^0}$, for all $\ell < k-1$, and $(x_{k-1}, x_k) \notin \mathcal{E}_{H_M^0}$. Note that, in a stealthy sequence, the mark's belief structure is not contradicted by the outcomes up to the last one.

When convenient, we use the notation σ_D and σ^D to denote, respectively, sequences where the deceiver is the first and last to take an action. The notation σ_D^D then means that the deceiver is the first and last to take an action. Similar notations can be defined for the mark. Given a finite sequence $\sigma = (x_0, x_1, \ldots, x_k)$, we say that $z \in \mathrm{S}_{\text{outcome}}$ is aligned with σ at time $\ell \in \{1, \ldots, k\}$ if $z = x_\ell$. Without loss of generality, we assume that the deceiver is the first to take an action.

B. Mark's learning via swap update

Here, we describe the mechanism that the mark implements to update its perception given the actions taken by the deceiver. Following [25], we employ the swap learning method because of its simplicity and because any other learning mechanism on hypergames can be described as a composition of swap updates. From the point of view of the mark, using swap learning is reasonable given that this method is guaranteed to decrease its misperception provided that the other player behaves rationally. Swap learning can also be defined for the deceiver, but in this case, given that it has full information about the other player's game, it does not make sense to implement it.

Consider first the case where the mark has made a single observation, corresponding to an action by the deceiver changing the outcome from x to y. The mark thus concludes that the deceiver prefers $(\pi_M(x), \pi_D(y))$ (which in this case is equal to y) over x. Therefore, the mark can incorporate this information into its hypergame and update its perception about the preferences of the deceiver. We start by an algebraic construction. Let V be a set of cardinality N and let

$W \subset V^N$ with pairwise different elements. For $x_1, x_2 \in V$, let $\mathrm{swap}_{x_1 \mapsto x_2} : W \to W$ be

$$(\mathrm{swap}_{x_1 \mapsto x_2}(v))_k = v_k \quad \text{if } v_k \neq x_1, x_2,$$

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

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

We refer to $\mathrm{swap}_{x_1 \mapsto x_2}$ as the x_1 *to* x_2 *swap map*. The *swap learning map* $\mathbf{Sw}_{x,y}^{M} : \mathbf{S}_{\mathrm{P}} \to \mathbf{S}_{\mathrm{P}}$ for the mark is given by

$$\mathbf{Sw}_{x,y}^{M}(\mathrm{P}) = \mathrm{swap}_{x \mapsto (\pi_M(x), \pi_D(y))}(\mathrm{P}). \tag{1}$$

Next, consider the case where the mark has made multiple observations of actions by the deceiver. Let $\mathcal{O}_{DM}$ denote the mark's observation set, composed of a collection of binary relations in P_{DD}. The preference vector P_{DM} is compatible with the observation set $\mathcal{O}_{DM}$ if the binary relations in $\mathcal{O}_{DM}$ hold in P_{DM}. As established in [25], in this case the mark can use any sorting algorithm which employs swap updates of the form (1) as basic operations to obtain a preference vector which is compatible with its observation set. We denote the corresponding swap update map for M by $\mathbf{Sw}_{\mathcal{O}_{DM}}^{M}$ and the resulting H-digraph after the mark updates its perception by $\mathbf{Sw}_{\mathcal{O}_{DM}}^{M}(\mathcal{G}_{H_M^0})$. Also, given a sequence of outcomes σ starting from $x \in \mathbf{S}_{\text{outcome}}$, we sometimes use, with a slight abuse of notation, $\mathbf{Sw}_{x,\sigma}^{M}$ to denote the mapping $\mathbf{Sw}_{\mathcal{O}_{DM}}^{M}$, where $\mathcal{O}_{DM}$ is the observation set of the mark corresponding to σ.

According to the swap learning method, an action taken by the deceiver that contradicts the mark's perception will trigger an update of its preference vectors. Given that the deceiver has the informational advantage of knowing a priori the full mark's game, it can indeed pre-compute the effect that any particular action will have on the mark's beliefs before executing it. However, as mentioned in Section II-B2, this procedure becomes computationally expensive. Instead, if the outcomes of the game correspond to a stealthy sequence $\sigma_D = (x_0, x_1, \dots, x_k)$, $k \in \mathbb{Z}_{\geq 1}$, then $\mathbf{Sw}_{x_{\ell-1},x_\ell}^{M}(\mathcal{E}_{H_M^0}) = \mathcal{E}_{H_M^0}$, for all $\ell \in \{1, \dots, k-1\}$, i.e., the mark does not see its perception contradicted up to the $(k-1)$th outcome. At the last outcome, instead,

$$\mathbf{Sw}_{x_{k-1},x_k}^{M}(\mathcal{E}_{H_M^0}) = \mathbf{Sw}_{x_0,\sigma_D}^{M}(\mathcal{E}_{H_M^0}) \neq \mathcal{E}_{H_M^0}. \tag{2}$$

In addition to the associated computational savings for the deceiver (the H-digraph remains the same), stealthy sequences have the advantage of not warning the mark about the possibility of deception since its perceptions are not contradicted up to the last moment.

C. Modeling the mark's actions via probability distributions

Although the deceiver has complete information about the mark's game, it does not know the specific strategy that the mark follows to choose its actions. For instance, if multiple sanction-free improvements are available to the mark, it might not necessarily pick its most preferred sequentially rational outcome (a less favorite improvement now may allow its to achieve a larger payoff in the future). Unaware of the mark's strategy, the deceiver assigns a probability distribution to the edges of the H-digraph of the mark as follows. Let $\mathbb{P}_{MD}(X_{n+1} = y \mid X_n = x)$, for $y \in \mathbf{S}_{\text{outcome}}|_{\pi_D(x)}$, denote the probability that the outcome of the hypergame changes from x to y by the action $\pi_M(y)$ of the mark, as perceived by the deceiver. Given what the deceiver knows about the mark's game, we have that for all $(x, y) \notin \mathcal{E}_{H_M^0}$, $\mathbb{P}_{MD}(X_{n+1} = y \mid X_n = x) = 0$. Note that, for all $x \in \mathbf{S}_{\text{outcome}}$, $\sum_{y \in \mathbf{S}_{\text{outcome}}|_{\pi_D(x)}} \mathbb{P}_{MD}(X_{n+1} = y \mid X_n = x) = 1$. The probability distribution $\mathbb{P}_{MD}$ is selected by the deceiver by applying some rule (e.g., 'assign more probability to the most preferred outcome') to the H-digraph of the opponent. The results of the paper are independent of the specific rule used and so we leave it unspecified.

The deceiver can choose its own actions based on its preferences in any way it sees fit. For later use, we formally describe this via a probability distribution $\mathbb{P}_D$ on any action $\pi_D(y)$ which changes the outcome from x to y. Note that this can, in particular, be a vector with one entry of 1 and the rest 0, and that it can be re-selected at each round of the game. Since players only use the current state of the game to decide about their next action, the sequence of repeated outcomes of the game is a Markov chain, possibly time-varying as the H-digraph of the mark can evolve with time. We gather here a few notational conventions that will be useful later. The *probability transition kernel* $T_{\mathbb{P}}$ of a probability distribution $\mathbb{P}$ is

$$T_{\mathbb{P}}(x_\ell, x_{\ell'}) = \mathbb{P}(X_{n+1} = x_\ell \,|\, X_n = x_{\ell'}),$$

where $x_\ell, x_{\ell'} \in \mathbf{S}_{\text{outcome}}$. One can inductively define

$$T_{\mathbb{P}}^k(x_\ell, x_{\ell'}) := \mathbb{P}(X_{n+k} = x_\ell \,|\, X_n = x_{\ell'}).$$

If there exists $k \in \mathbb{Z}_{\geq 1}$ with $T_{\mathbb{P}}^{k}(x_\ell, x_{\ell'}) > 0$, the outcome x_ℓ is *reachable* from $x_{\ell'}$. We denote the set of all outcomes reachable from $x_{\ell'}$ with respect to the transition probability $T_{\mathbb{P}}$ by

$$\mathcal{R}_{T_{\mathbb{P}}}(x_{\ell'}) = \{x_\ell \in \Omega \mid \exists k_\ell \in \mathbb{Z}_{\geq 1},\ T_{\mathbb{P}}^{k_\ell}(x_\ell, x_{\ell'}) > 0\}.$$

Finally, note that the fact that the H-digraph does not change when the hypergame is played according to a stealthy sequence implies that, in such cases, the distribution $\mathbb{P}_{MD}$ does not change either. This motivates the definition of a subset $\mathscr{S}^{\mathbb{P}_{MD}}(\mathbf{S}_{\text{outcome}})$ of sequences with

$$\mathscr{S}^{\mathbb{P}_{MD}}(\mathbf{S}_{\text{outcome}}) = \{(x_0, x_1, x_2, \ldots) \mid T_{\mathbb{P}_{MD}}(x_{\ell+1}, x_\ell) > 0, \forall \ell \in \mathbb{Z}_{\geq 0} \text{ such that } \pi_D(x_\ell) = \pi_D(x_{\ell+1})\}.$$

If $\sigma \in \mathscr{S}^{\mathbb{P}_{MD}}(\mathbf{S}_{\text{outcome}})$, we call σ a $\mathbb{P}_{MD}$*-sequence*. With this definition, the deceiver perceives a positive probability to the actions of the mark contained in σ. From now on, when we use the term ‘stealthy sequence’ we mean ‘stealthy $\mathbb{P}_{MD}$-sequence’.

D. Surely and strong edge-deceivability

We are finally ready to introduce precise mathematical notions of belief manipulation in the scenario described in Section III-A. Our notions revolve around the deceivability of edges in the mark’s H-digraph that correspond to sanction-free improvements that it could potentially take.

Definition 3.1 (Edge-deceivability): An edge $(x, y) \in \mathcal{E}_{H_M^0}$, $\pi_D(x) = \pi_D(y)$, is *deceivable* by the deceiver in H_M^0 from $x_0 \in \mathbf{S}_{\text{outcome}}$ if there exists a sequence of outcomes $\sigma_D = (x_0, x_1, x_2, \ldots, x_{2k+1})$, $k \in \mathbb{Z}_{\geq 0}$, where

(i) $(x_{2\ell-1}, x_{2\ell}) \in \mathbf{Sw}^M_{x_0, \sigma_D(\ell-1)}(\mathcal{E}_{H_M^0})$, where $\sigma_D(\ell - 1) = (x_0, x_1, x_2, \ldots, x_{2\ell-1})$, and

(ii) $T_{\mathbb{P}_{MD}}(x_{2\ell}, x_{2\ell-1}) > 0$,

for all $\ell \in \{1, \ldots, k\}$, such that $(x, y) \notin \mathbf{Sw}^M_{x_0,\sigma_D}(\mathcal{E}_{H_M^0})$. We refer to σ_D as a *deceiving sequence*.

Let us elaborate on the properties of the deceiving sequence σ_D in the above definition: (i) states that the mark uses its updated H-digraph and takes an action to shift the outcome to a sanction-free improvement; (ii) states that the deceiver perceives a positive probability to the actions of the mark contained in σ_D. There is an abuse of notation due to the fact that $\mathbb{P}_{MD}$ can change with the evolution of the H-digraph. Also, here we have assumed that the deceiver takes the last action. This is without loss of generality; if the edge (x, y) has been deceived, a further action by the mark will have no effect on this fact.

We let $E^{D,x_0}_{\text{dec}}(H^0_M) \subseteq \mathcal{E}_{H^0_M}$ denote the set of all deceivable edges in H^0_M from x_0. We say that (x, y) is *surely deceivable* by the deceiver in H^0_M from x_0 if it is deceivable with probability one, i.e., with (ii) in Definition 3.1 substituted by (ii') $T_{\mathbb{P}_{MD}}(x_{2\ell}, x_{2\ell-1}) = 1$. We denote the set of all such edges by $E^{D,x_0}_{\text{sdec}}(H^0_M) \subseteq \mathcal{E}_{H^0_M}$. The following definition captures the situation where an edge is deceivable regardless of the initial condition.

Definition 3.2: (Strong edge-deceivability): The edge (x, y) is *strong deceivable* by the deceiver in H^0_M if it is deceivable from any outcome $x_0 \in \mathbf{S}_{\text{outcome}}$ and is *surely strong deceivable* if it is strong deceivable with probability one. The set of strong deceivable and surely strong deceivable edges are denoted, respectively, by $E^D_{\text{stdec}}(H^0_M)$ and $E^D_{\text{sstdec}}(H^0_M)$.

Definitions 3.1 and 3.2 are the basic building blocks towards the deceiver being able to make an unstable outcome sequentially rational for the mark. One can indeed define in a similar way the notion of outcome-deceivability: an outcome is deceivable if all the out-edges corresponding to the opponent's sanction-free improvements can be deceived. Throughout the paper, we focus on edge-deceivability.

In the above definitions, the sequence of outcomes is arbitrary. As have already discussed, this might require a substantial computational effort on the part of the deceiver. This problem can be addressed by restricting the set of candidate sequences to be stealthy. The problems we are then interested in solving are the following:

(i) given an edge in the H-digraph that corresponds to an action of the mark, what are the set of initial outcomes from which the edge is surely deceivable by the deceiver? Similarly, when is an outcome strong deceivable?
(ii) when is it possible to perform deception via a stealthy sequence?
(iii) given answers to the previous questions, how can the deceiver find a (stealthy) strategy to deceive the mark?

Before addressing these questions, we illustrate some of the notions introduced so far with an example.

Example 3.3: (Notions of deceivability in hypergames): Consider a 2-level hypergame $H^2 = \{H^0_M, H^1_D\}$ between the mark and the deceiver, with $H^0_{MD} = H^0_M$ and outcome set $\mathbf{S}_{\text{outcome}} = S_M \times S_D = \{1, \dots, 50\}$, where S_M and S_D are the action sets of the mark and the deceiver, respectively, and $|S_M| = 5$ and $|S_D| = 10$. The preference vectors P_{MM} and P_{DM} are shown in Figure 2. The H-digraph $\mathcal{G}_{H_M^0}$ is shown in Figure 3. Regarding the actions of the mark, the

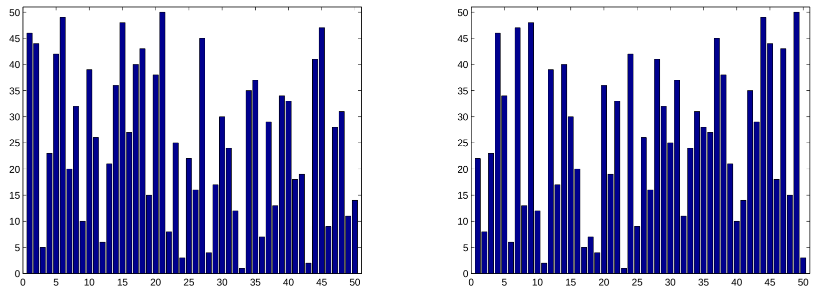

Fig. 2. Preference vectors P_{MM} (left) and P_{DM} (right). The horizontal axis shows the outcomes and the vertical axis shows the rank of outcomes.

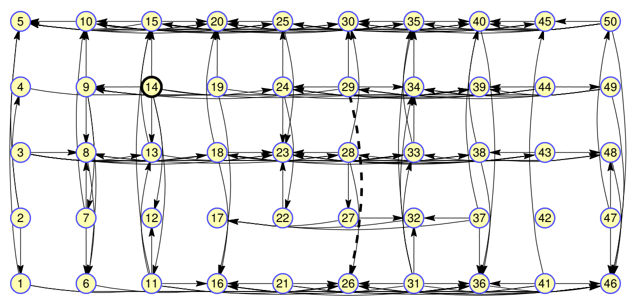

Fig. 3. H-digraphs $\mathcal{G}_{H_M^0}$. The mark plays rows, the deceiver plays columns. The deceiver intends to remove the edge $(29, 26) \in \mathcal{E}_{H_M^0}$ (left plot, dashed) via a stealthy sequence, starting from outcome 14.

deceiver perceives that outcomes with lower rank in P_{MM} have higher probability of occurring. Formally, the deceiver assigns the probability

$$T_{\mathbb{P}_{MD}}(\ell', \ell) = \frac{50 - \mathrm{rank}(\ell', \mathrm{P}_{MM})}{\sum_{l \in \mathcal{N}^{\mathrm{out}}(\ell) \cap \mathbf{S}_{\mathrm{outcome}}|_{\pi_D(\ell)}} (50 - \mathrm{rank}(l, \mathrm{P}_{MM}))},$$

to the event that the outcome changes from ℓ to ℓ' by the action $\pi_M(\ell')$ of the mark, where $\ell' \in \mathbf{S}_{\mathrm{outcome}}|_{\pi_D(\ell)}$. The deceiver now wishes to find out if the edge $(29, 26) \in \mathcal{E}_{H_M^0}$ is deceivable in H_M^0 from $14 \in \mathbf{S}_{\mathrm{outcome}}$, i.e, if $(29, 26) \in E_{\mathrm{dec}}^{D, x_0}(H_M^0)$. The deceiver also aims to remove this edge (Figure 3, dashed) via a stealthy sequence, starting from outcome 14. •

IV. When is it possible to perform deception?

In this section, we identify a necessary condition and a sufficient condition for edge-deceivability by means of an arbitrary sequence of outcomes. When, instead, the sequence is restricted to be stealthy, we are able to provide a full characterization of when deceivability is possible.

A. Deceivability conditions

We start this section by stating several relationships between the edge-deceivability notions introduced in Section III-D.

Lemma 4.1 (Deceivability inclusions): For all $x_0 \in \mathbf{S}_{\text{outcome}}$, the following inclusions hold

$$E^{D}_{\text{sstdec}}(H^0_M) \subseteq E^{D,x_0}_{\text{sdec}}(H^0_M),\ E^{D}_{\text{stdec}}(H^0_M) \subseteq E^{D,x_0}_{\text{dec}}(H^0_M).$$

The proof follows directly from Definitions 3.1 and 3.2. We next identify a necessary condition for edge-deceivability.

Lemma 4.2: (Necessary condition on the edge for edge-deceivability): Let $x_0 \in \mathbf{S}_{\text{outcome}}$ and assume $(x, y) \in E^{D,x_0}_{\text{dec}}(H^0_M)$. Then, $\mathsf{H}^M_{\text{dec}}(x, y) := \{u \in \mathbf{S}_{\text{outcome}}|_{\pi_M(y)} \mid u \prec_{\mathrm{P}_{MM}} x\} \neq \emptyset$.

Proof: We reason by contradiction. Suppose $u \succeq_{\mathrm{P}_{MM}} x$ for all $u \in \mathbf{S}_{\text{outcome}}|_{\pi_M(y)}$. Therefore, the deceiver has no sanction against the improvement from x to y for the mark, and thus the edge $(x, y) \in \mathbf{Sw}^M_{x_0,\sigma}(\mathcal{E}_{H^0_M})$, for any sequence of outcomes σ and any initial outcome $x_0 \in \mathbf{S}_{\text{outcome}}$. ■

Given the result in Lemma 4.1, the condition identified in Lemma 4.2 is also necessary for strong deceivability. Next, we identify a condition on the initial outcome which guarantees surely deceivability of an edge.

Lemma 4.3: (Sufficient condition on the initial outcome for surely deceivability): Let $(x, y) \in \mathcal{E}_{H^0_M}$, with $\pi_D(x) = \pi_D(y)$, such that $\mathsf{H}^M_{\text{dec}}(x, y) \neq \emptyset$. Then $(x, y) \in E^{D,\tilde{y}}_{\text{sdec}}(H^0_M)$, for all $\tilde{y} \in \mathsf{T}^M_{\text{dec}}(y) := \{w \in \mathbf{S}_{\text{outcome}}|_{\pi_M(y)} \mid w \succeq_{\mathrm{P}_{DM}} y\}$.

Proof: Note that $z \prec_{\mathrm{P}_{DM}} y$ for $z \in \mathsf{H}^M_{\text{dec}}(x, y)$, since otherwise, the improvement y from x of the mark would be sanctioned by the perceived improvement z from y of the deceiver and this would imply $(x, y) \notin \mathcal{E}_{H^0_M}$. Suppose the deceiver takes an action from $\tilde{y} \in \mathsf{T}^M_{\text{dec}}(y)$ that changes the outcome to $z \in \mathsf{H}^M_{\text{dec}}(x, y)$. Since $(\tilde{y}, z) \notin \mathcal{E}_{H^0_M}$, the mark uses the swap learning map to update its perceptions about the deceiver. But then $(x, y) \notin \mathbf{Sw}^M_{\tilde{y},z}(\mathcal{E}_{H^0_M})$, since the outcome z

with $z \succ_{\mathbf{Sw}^M_{\tilde{y},z}(\mathrm{P}_{DM})} y$ is now perceived by the mark as a sanction of the deceiver against the improvement y from x by the mark. As a result, $\sigma_D = (\tilde{y}, z)$ is a deceiving sequence for the deceiver and thus the result follows. ■

Next, we turn our attention to stealthy sequences of outcomes. The following result shows that the two conditions identified above for arbitrary sequences essentially provide a full characterization of when deception is possible using stealthy sequences.

Theorem 4.4: (Characterization of deceivability via stealthy sequences): Let $x_0 \in \mathbf{S}_{\text{outcome}}$ and $(x, y) \in \mathcal{E}_{H_M^0}$, $\pi_D(x) = \pi_D(y)$. The following are equivalent:

(i) (x, y) is deceivable from x_0 via a stealthy sequence;

(ii) $\mathsf{H}^M_{\text{dec}}(x, y) \neq \emptyset$ and

$$\mathcal{T}^M_{\text{dec}}(y, x_0) := \mathsf{T}^M_{\text{dec}}(y) \cap \big(\{x_0\} \cup \mathcal{R}_{T_{\mathbb{P}_{MD}} T_{\mathbb{P}_D}}(x_0)\big) \neq \emptyset,$$

for a probability distribution $\mathbb{P}_D$ such that $\mathbb{P}_D(X_{n+1} = z \mid X_n = r) > 0$ for any $(r, z) \in \mathcal{E}_{H_M^0}$.

Proof: We first show that (i) implies (ii). Suppose $(x, y) \in \mathcal{E}_{H_M^0}$. First of all, note that since, by assumption, (x, y) is deceivable from x_0, the necessary conditions of Lemma 4.2 for (x, y) hold, i.e., $\mathsf{H}^M_{\text{dec}}(x, y) \neq \emptyset$. If $x_0 \in \mathsf{T}^M_{\text{dec}}(y)$, then $\mathcal{T}^M_{\text{dec}}(y, x_0) \neq \emptyset$ and the result follows. Suppose $x_0 \notin \mathsf{T}^M_{\text{dec}}(y)$. By Definition 3.1, there exists a stealthy sequence $\sigma_D = (x_0, \ldots, x_{k-1}, x_k)$ such that $(x, y) \notin \mathbf{Sw}^M_{x_0, \sigma_D}(\mathcal{E}_{H_M^0})$ and $\mathfrak{S} = (x_0, \ldots, x_{k-1})$ is a path in $\mathcal{G}_{H_M^0}$. Note that $x_{k-1} \in \mathsf{T}^M_{\text{dec}}(y)$ and $x_k \in \mathsf{H}^M_{\text{dec}}(x, y)$, since otherwise, the action of the deceiver from x_{k-1} to x_k is sanction free and thus, by definition of a swap learning map and equation (2), $(x, y) \in \mathbf{Sw}^M_{x_0, \sigma_D}(\mathcal{E}_{H_M^0})$, a contradiction with the assumption. Next, let $\tilde{y} = x_{k-1}$ (observe that $k \geq 3$). We show that $\tilde{y} \in \mathcal{R}_{T_{\mathbb{P}_{MD}} T_{\mathbb{P}_D}}(x_0)$. Suppose $\mathbb{P}_D$ is a probability distribution such that $\mathbb{P}_D(X_{n+1} = z \mid X_n = r) > 0$, for all r and all $z \in \mathbf{S}_{\text{outcome}}|_{\pi_M(r)}$ with $(r, z) \in \mathcal{E}_{H_M^0}$. By definition of a $\mathbb{P}_{MD}$-sequence, $T_{\mathbb{P}_{MD}}(x_\ell, x_{\ell+1}) > 0$ and $\pi_D(x_\ell) = \pi_D(x_{\ell+1})$, for all $\ell \in \{1, \ldots, k\}$. Thus there is a strictly positive probability that $\tilde{y}$ is reachable from x_0 via the path $\mathfrak{S} = (x_0, \ldots, x_{k-2}, \tilde{y})$, $\pi_D(x_{k-2}) = \pi_D(\tilde{y})$. Thus there exists some $K \in \mathbb{Z}_{\geq 1}$ such that

$$(T_{\mathbb{P}_{MD}} T_{\mathbb{P}_D})^K (\tilde{y}, x_0) > 0,$$

i.e., $\tilde{y} \in \mathcal{R}_{T_{\mathbb{P}_{MD}} T_{\mathbb{P}_D}}(x_0)$. As a result, $\mathcal{T}^M_{\text{dec}}(y, x_0) \neq \emptyset$.

Conversely, let us show that (ii) implies (i). The results hold by Lemma 4.3 if $x_0 \in \mathsf{T}^M_{\text{dec}}(y)$. Suppose $x_0 \notin \mathsf{T}^M_{\text{dec}}(y)$. We need to show that there exists a stealthy sequence of outcomes σ_D such that $(x, y) \notin \mathbf{Sw}^M_{x_0,\sigma_D}(\mathcal{E}_{H^0_M})$. First, note that, by Lemma 4.2, there exists an outcome $z \in \mathbf{S}_{\text{outcome}}|_{\pi_M(y)}$ such that $z \prec_{\mathrm{P}_{MM}} x$. By assumption, there exists an outcome $\tilde{y} \in \mathsf{T}^M_{\text{dec}}(y)$ that can be reached from x_0, for a probability distribution $\mathbb{P}_D$ described above, i.e., there exists a path $\mathfrak{S} = (x_0, \ldots, x_{k-1}, \tilde{y})$ in $\mathcal{G}_{H^0_M}$ such that $T_{\mathbb{P}_{MD}}(x_{\ell+1}, x_\ell) > 0$, for all $\ell \in \{0, \ldots, k-2\}$ with $\pi_D(x_\ell) = \pi_D(x_{\ell+1})$. If the deceiver takes an action that changes the outcome from $\tilde{y}$ to z, then, by definition, z is a sanction against the perceived improvement y from x for the mark; thus $(x, y) \notin \mathbf{Sw}^M_{\tilde{y},z}(\mathcal{E}_{H^0_M})$. Next, we define $\sigma_D = (x_0, \ldots, x_{k-1}, \tilde{y}, z)$, which is a stealthy sequence starting from x_0, since $\mathfrak{S} = (x_0, \ldots, x_{k-1}, \tilde{y})$ is a path in $\mathcal{G}_{H^0_M}$ and by construction $(\tilde{y}, z)$ is not an edge in $\mathcal{G}_{H^0_M}$. Finally, since $(x, y) \notin \mathbf{Sw}^M_{\tilde{y},z}(\mathcal{E}_{H^0_M})$, σ_D is also a deceiving sequence, as claimed. ■

The choice of $\mathbb{P}_D$ in Theorem 4.4(ii) ensures that all actions of the deceiver are considered when determining if a stealthy sequence exists to deceive the mark. Once such sequence is found, the deceiver can assign probability one to each of its actions prescribed in the sequence.

Theorem 4.4 shows that, given $x_0 \in \mathbf{S}_{\text{outcome}}$, any action of the deceiver from $\mathcal{T}^M_{\text{dec}}(y, x_0)$ to $\mathsf{H}^M_{\text{dec}}(x, y)$ removes the edge (x, y) from the H-digraph $\mathcal{G}_{H^0_M}$. One can then characterize the set of all initial outcomes from which the edge (x, y) is deceivable as

$$\mathcal{I}^M_{\text{dec}}(x, y) = \{x_0 \in \mathbf{S}_{\text{outcome}} \mid \mathsf{H}^M_{\text{dec}}(x, y) \neq \emptyset \neq \mathcal{T}^M_{\text{dec}}(y, x_0)\}.$$

Consequently, finding a stealthy sequence from $x_0 \in \mathcal{I}^M_{\text{dec}}(x, y)$ is equivalent to finding a path in the H-digraph $\mathcal{G}_{H^0_M}$ that reaches $\mathcal{T}^M_{\text{dec}}(y, x_0)$ from x_0. Since the outcome of the hypergame is influenced by the actions of the mark and these, from the point of view of the deceiver, are probabilistic, a reasonable strategy is to find a stealthy sequence that maximizes the minimum probability of achieving the deception goal. It must be observed that the success of such strategy relies on the particular structure of H-digraphs, which is what we tackle next.

V. The worst-case max-strategy

Here, we provide an algorithmic approach that can be used by the deceiver to determine a stealthy sequence to deceive the mark.

Consider the scenario described in Section III. Suppose at time $\ell \geq 0$ the outcome of the 2-player 2-level hypergame is $x(\ell)$. Without loss of generality, assume that the deceiver takes actions when $\ell \in 2\mathbb{Z}_{\geq 0}$ and the mark takes actions when $\ell \in 2\mathbb{Z}_{\geq 0} + 1$. In this situation, Theorem 4.4 characterizes the edges of the H-digraph of the mark that are deceivable by the deceiver via a stealthy sequence. In this scenario, a reasonable strategy for the deceiver at each round is to take an action that maximizes the minimum probability of achieving the deception goal. Informally, this strategy can be described as follows:

[Informal description]: Initially, the deceiver has a stealthy sequence (possibly empty) stored in its memory (along which it would like the hypergame to evolve). At each round,

(i) if there is a deceiving action that takes the current outcome to $\mathsf{H}_{\text{dec}}^{M}(x, y)$, the deceiver takes it to deceive the mark, c.f. Lemma 4.3;

(ii) otherwise, the deceiver checks if the mark's last action is aligned with the stored sequence. If it is, the deceiver takes the next action prescribed by the sequence. If it is not (and this includes the case when the stored sequence is empty), the deceiver considers the outcomes $w \in \mathbf{S}_{\text{outcome}}$ where it can take the game to by an action aligned with the mark's H-digraph, and computes the stealthy sequence with minimum probability of reaching an outcome in $\mathcal{T}_{\text{dec}}^{M}(y, x_0)$ from each w (this is instantiated as the empty sequence if no such sequence exists). The deceiver stores the sequence with maximum probability. If this probability is positive, the deceiver takes the action prescribed by it, otherwise the algorithm terminates without success.

We call this strategy the `worst-case max-strategy` and, after introducing some notions, formally describe it in Algorithm 1.

To model the fact that the outcome of the hypergame is influenced by the actions of the mark, let us introduce the map $\Phi_{\mathbb{P}_{MD}} : \mathscr{S}^{\mathbb{P}_{MD}}(\mathbf{S}_{\text{outcome}}) \to \mathbb{R}$,

$$\Phi_{\mathbb{P}_{MD}}(x_0, \dots, x_k) = \sum_{\substack{\ell=0 \\ \pi_D(x_\ell)=\pi_D(x_{\ell+1})}}^{k-1} \ln\left(T_{\mathbb{P}_{MD}}(x_{\ell+1}, x_\ell)\right). \tag{3}$$

This map captures the probability of reaching an outcome via a $\mathbb{P}_{MD}$-sequence.

The rationale behind the name '`worst-case max-strategy`' is made explicit next.

Algorithm 1: `worst-case max-strategy`

Input: $\mathcal{G}_{H_M^0}$, $\mathbb{P}_{MD}$, $(x,y) \in \mathcal{E}_{H_M^0}$, $x_0 \in \mathsf{S}_{\text{outcome}}$, $\mathcal{N}^{\text{out}}(x_0) \cap \mathsf{S}_{\text{outcome}}|_{\pi_M(x_0)} \neq \emptyset$

Initialization: $\alpha^{\text{maxmin}} = -\infty$, $\sigma_D = \emptyset$, $x(0) = x_0$

check $\mathsf{H}_{\text{dec}}^M(x,y) \neq \emptyset$; else, announce (x,y) not deceivable

At time: $\ell \in 2\mathbb{Z}_{\geq 0}$

if $x(\ell) \in \mathsf{T}_{\text{dec}}^M(y)$ **then**
    take action that makes $x(\ell+1) \in \mathsf{H}_{\text{dec}}^M(x,y)$
else
    if $\sigma_D \neq \emptyset$ and $x(\ell)$ is aligned with σ_D **then**
        take action prescribed by σ_D
    else
        foreach $w \in \mathsf{S}_{\text{outcome}}|_{\pi_M(x(\ell))}$, $(x(\ell), w) \in \mathcal{E}_{H_M^0}$ **do**
            $\alpha^{\text{min}} = +\infty$
            foreach $\tilde{y} \in \mathsf{T}_{\text{dec}}^M(y)$ **do**
                if there is path in $\mathcal{G}_{H_M^0}$ from w to $\tilde{y}$ **then**
                    find σ_M^M from w to $\tilde{y}$ minimizing $\Phi_{\mathbb{P}_{MD}}$
                    if $\Phi_{\mathbb{P}_{MD}}(\sigma_M^M) \leq \alpha^{\text{min}}$ **then**
                        $\alpha^{\text{min}} = \Phi_{\mathbb{P}_{MD}}(\sigma_M^M)$
                    end
                end
            end
            if $\alpha^{\text{min}} \neq +\infty$ and $\alpha^{\text{min}} \geq \alpha^{\text{maxmin}}$ **then**
                $\alpha^{\text{maxmin}} = \alpha^{\text{min}}, \quad \eta = \sigma$
            end
        end
        if $\alpha^{\text{maxmin}} \neq -\infty$ **then**
            $\sigma_D = (x(\ell), \eta)$ take action prescribed by σ_D
        else
            announce (x,y) is not deceivable from $x(\ell)$
        end
    end
end

Lemma 5.1: (Algorithm 1 maximizes the minimum probability of deception): The following are equivalent:

(i) $\sigma_D = (x_0, x_1, x_2, \ldots, x_{2k}) \in \mathscr{S}^{\mathbb{P}_{MD}}(\mathbf{S}_{\text{outcome}})$, where $k \in \mathbb{Z}_{\geq 1}$, $x_{2k} \in \mathsf{T}^M_{\text{dec}}(y)$, and $(x_\ell, x_{\ell+1}) \in \mathcal{E}_{H^0_M}$ for $\ell \in \{0, \ldots, 2k-1\}$, is a minimizer of $\Phi_{\mathbb{P}_{MD}}$;

(ii) σ_D corresponds to the longest path from x_0 to $x_{2k} \in \mathsf{T}^M_{\text{dec}}(y)$, in $(\mathbf{S}_{\text{outcome}}, \mathcal{E}_{H^0_M}, \mathcal{M}_{H^0_M})$, where, for $\ell, \ell' \in \{1, \ldots |\mathbf{S}_{\text{outcome}}|\}$, $(\mathcal{M}_{H^0_M})_{\ell\ell'} = |\ln(T_{\mathbb{P}_{MD}}(z_{\ell'}, z_\ell))|$, if $\pi_D(z_\ell) = \pi_D(z_{\ell'})$, and is zero otherwise.

Note that, in Lemma 5.1, (i) is equivalent to stating that σ_D is a minimizer of $\Pi_{i=1}^k T_{\mathbb{P}_{MD}}(x_{2i}, x_{2i-1})$, and (ii) implies that finding solutions to the `worst-case max-strategy` is equivalent to finding a longest path on a digraph. By virtue of Theorem A.2, this problem is well-posed and can be solved efficiently.

Next, we show that the `worst-case max-strategy` is complete, in the sense that it always finds a stealthy sequence that deceives a surely deceivable edge.

Theorem 5.2: (Surely deceivable edges via `worst-case max-strategy`): The edge $(x, y) \in \mathcal{E}_{H^0_M}$, $\pi_D(x) = \pi_D(y)$, is surely deceivable from $x_0 \in \mathbf{S}_{\text{outcome}}$ via a stealthy sequence of D iff $\mathsf{H}^M_{\text{dec}}(x, y) \neq \emptyset$ and either $x_0 \in \mathsf{T}^M_{\text{dec}}(y)$ or

$$\max_{x_1 \in \mathbf{S}_{\text{outcome}}|_{\pi_M(x_0)}} \min_{\sigma_D} \Phi_{\mathbb{P}_{MD}}(\sigma_D) = 0,$$

where $\sigma_D = (x_0, x_1, x_2, \ldots, x_{2k}) \in \mathscr{S}^{\mathbb{P}_{MD}}(\mathbf{S}_{\text{outcome}})$, $k \in \mathbb{Z}_{\geq 1}$, $x_{2k} \in \mathsf{T}^M_{\text{dec}}(y)$, $(x_\ell, x_{\ell+1}) \in \mathcal{E}_{H^0_M}$, $\ell \in \{0, \ldots, 2k-1\}$.

Proof: Suppose $(x, y) \in \mathcal{E}_{H^0_M}$, $\pi_D(x) = \pi_D(y)$, is surely deceivable from x_0 via a stealthy sequence. By Lemma 4.2, $\mathsf{H}^M_{\text{dec}}(x, y) \neq \emptyset$. Suppose $x_0 \notin \mathsf{T}^M_{\text{dec}}(y)$. By Theorem 4.4 and by the definition of surely deceivability, there exists a sequence of outcomes $\sigma_D = (x_0, x_1, x_2, \ldots, x_{2k}) \in \mathscr{S}^{\mathbb{P}_{MD}}(\mathbf{S}_{\text{outcome}})$, $k \in \mathbb{Z}_{\geq 1}$, $x_{2k} \in \mathsf{T}^M_{\text{dec}}(y)$, where

(i) $(x_\ell, x_{\ell+1}) \in \mathcal{E}_{H^0_M}$, for all $\ell \in \{0, \ldots, 2k-1\}$;

(ii) $T_{\mathbb{P}_{MD}}(x_{2\ell}, x_{2\ell-1}) = 1$, for all $\ell \in \{1, \ldots, k\}$.

By (3), $\Phi_{\mathbb{P}_{MD}}(\sigma_D) = 0$. Since $T_{\mathbb{P}_{MD}}(x_{2\ell}, x_{2\ell-1}) = 1$, for all $\ell \in \{1, \ldots, k\}$, if the deceiver chooses its sequential actions aligned with σ_D at each time, then the sequence will reach x_{2k} with probability one, and thus it is the unique sequence starting at x_0 reaching x_{2k} which includes x_1. This, along with the fact that $\Phi_{\mathbb{P}_{MD}}(\sigma_D) \leq 0$ for any σ_D, proves the result.

Conversely, if $x_0 \in \mathsf{T}^M_{\text{dec}}(y)$, since $\mathsf{H}^M_{\text{dec}}(x,y) \neq \emptyset$, the result follows from Lemma 4.3. Suppose $x_0 \notin \mathsf{T}^M_{\text{dec}}(y)$. Then, by assumptions, σ_D is a stealthy sequence from x_0 which reaches $x_{2k} \in \mathcal{T}^M_{\text{dec}}(y, x_0)$, with probability one. Since $\mathsf{H}^M_{\text{dec}}(x,y) \neq \emptyset$, the result follows by Theorem 4.4 and the definition of surely deceivability. ■

The following results demonstrates that the `worst-case max-strategy` can also characterize the surely deceivable edges when the opponent is using a best-response strategy. The proof is similar to Theorem 5.2 and is omitted here.

Proposition 5.3: (Best-response strategies and the `worst-case max-strategy`): If the mark takes the sanction-free action associated to its most preferred outcome at all times and the deceiver knows about this, then $(x,y) \in E^{D,x_0}_{\text{sdec}}(H^0_M)$, $x_0 \in \mathbf{S}_{\text{outcome}}$, via a stealthy sequence of the deceiver if and only if $\mathsf{H}^M_{\text{dec}}(x,y) \neq \emptyset$ and either $x_0 \in \mathsf{T}^M_{\text{dec}}(y)$ or

$$\max_{x_1 \in \mathbf{S}_{\text{outcome}}|_{\pi_M(x_0)}} \min_{\sigma_D} \Phi_{\mathbb{P}^*_{MD}}(\sigma_D) = 0,$$

where $\mathbb{P}^*_{MD}$ assigns one to the edges of $\mathcal{G}_{H^0_M}$ associated to the most preferred sanction-free actions of the mark and $\sigma_D = (x_0, x_1, x_2, \ldots, x_{2k}) \in \mathscr{S}^{\mathbb{P}_{MD}}(\mathbf{S}_{\text{outcome}})$, $k \in \mathbb{Z}_{\geq 1}$, $x_{2k} \in \mathsf{T}^M_{\text{dec}}(y)$, $(x_\ell, x_{\ell+1}) \in \mathcal{E}_{H^0_M}$ for all $\ell \in \{0, \ldots, 2k-1\}$.

Remark 5.4 (Strong deceivability): The execution of the `worst-case max-strategy` from all the outcomes in $\mathbf{S}_{\text{outcome}}$ fully characterizes the set $\mathcal{I}^M_{\text{dec}}(x,y)$. Note that, by definition, $\mathcal{I}^M_{\text{dec}}(x,y) = \mathbf{S}_{\text{outcome}}$ iff (x,y) is strongly deceivable via a stealthy sequence. •

Example 5.5: (Example 3.3 revisited): Consider the scenario discussed in Example 3.3. Suppose the game initially starts at outcome $x_0 = 14$ and the deceiver wishes to deceive the mark by removing the edge $(29, 26) \in \mathcal{E}_{H^0_M}$ via a stealthy sequence. Since

$$\mathcal{H}^M_{\text{dec}}(29, 26) = \{z \in \mathbf{S}_{\text{outcome}}|_{\pi_M(26)} \mid z \prec_{\mathrm{P}_{MM}} 29\} = \{11, 31, 41\} \neq \emptyset,$$

the necessary condition of Lemma 4.2 for $(29, 26)$ is satisfied. According to Theorem 4.4, we compute

$$\mathsf{T}^M_{\text{dec}}(26) = \{w \in \mathbf{S}_{\text{outcome}}|_{\pi_M(26)} \mid w \succeq_{\mathrm{P}_{DM}} 26\} = \{1, 6, 26, 36\}.$$

The actions of the deceiver from 14 aligned with the mark's H-digraph are

$$\mathcal{N}^{\text{out}}(14) \cap \mathbf{S}_{\text{outcome}}|_{\pi_M(14)} = \{9, 24, 39\}.$$

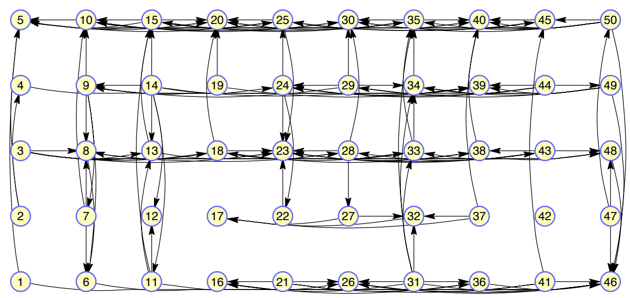

Fig. 4. $\mathbf{Sw}^{M}_{36,11}(\mathcal{G}_{H^0_M})$ is shown. The mark plays rows, the deceiver plays columns, and P_{MM} and P_{DM} are given in Figure 2. The deceiver intends to remove the edge $(29, 26) \in \mathcal{E}_{H^0_M}$ (left plot, dashed) via a stealthy sequence, starting from outcome 14. After reaching the outcome 36, the edge $(29, 26) \in \mathcal{E}_{H^0_M}$ is removed (right plot) by the action $\pi_D(11)$ of the deceiver.

By executing the `worst-case max-strategy`, the deceiver finds that the action that maximizes the minimum probability of reaching any of the outcomes in $\mathsf{T}^M_{\text{dec}}(26)$ is $\pi_D(24)$, where it perceives that the repeated play of the game will reach outcome 36 via the path

$$\mathfrak{S} = (14, 24, 25, 40, 36),$$

with probability 0.52. Note that, since the outcome 36 is reachable via $\mathfrak{S}$ and belongs to $\mathsf{T}^M_{\text{dec}}(26)$, by definition $36 \in \mathcal{T}^M_{\text{dec}}(26, 14)$. In particular, the characterization of deceivability in Theorem 4.4 is satisfied. If the repeated play goes according to the deceiver's perception, after reaching 36, the deceiver takes an action that changes the outcome to any of the outcomes in $\mathcal{H}^M_{\text{dec}}(29, 26)$, e.g., if the deceiver chooses to take the action $\pi_D(11)$ (note that $(36, 11) \notin \mathcal{E}_{H^0_M}$), then the mark's H-digraph after updating its perception via swap learning is shown in Figure 4. If the mark takes an action not aligned with the sequence $\mathfrak{S}$ at any round, according to the `worst-case max-strategy`, the deceiver will recompute the stealthy sequence and take the ensuing action accordingly. •

VI. Conclusions

We have studied scenarios of active deception in 2-person 2-level hypergames with asymmetric information between a deceiver and a mark. Building on the framework of hypergames and its explicit modeling of player misperception, we have introduced novel formal notions that capture various forms of deception and belief manipulation. We have provided a necessary condition and a sufficient condition for deceivability for the case when the deceiver might take actions that contradict the perception of the mark. In addition, we have fully characterized when deception is possible for the case when, instead, the deceiver acts in a stealthy way and only takes actions aligned with its opponent's perception. Finally, we have designed the `worst-case max-strategy` that the deceiver can use to find a stealthy sequence that maximizes the minimum probability of achieving its deception goal. The correctness of this strategy follows from the fact that the H-digraph associated to a finite subjective hypergame does not contain any weak improvement cycle. Future work will study the impact of signaling costs on the deceiver, the design of strategies for performing outcome deceivability deception via non-stealthy strategies, and the applications of our results to human behavior modeling.

REFERENCES

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

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

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

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

[5] 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.

[6] 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.

[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] V. J. Baston and F. A. Bostock, "Deception games," *International Journal of Game Theory*, vol. 17, pp. 129–134, 1988.

[11] K. T. Lee and K. L. Teo, "A game with distorted information," *Naval Research Logistics*, vol. 40, pp. 993–1001, 1993.

[12] 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.

[13] 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.

[14] 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.

[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. Cruz Jr., "Information, decision-making and deception in games," *Decision Support Systems*, vol. 47, pp. 518–527, 2009.

[17] 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.

[18] 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.

[19] F. Bullo, J. Cortés, and S. Martínez, *Distributed Control of Robotic Networks*. Applied Mathematics Series, Princeton University Press, 2009. Electronically available at http://coordinationbook.info.

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

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

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

[23] H. P. Young, *Strategic Learning and its Limits*. Oxford, UK: Oxford University Press, 2004.

[24] L. Smith and P. Sorensen, "Pathological outcomes of observational learning," *Econometrica*, vol. 68, no. 2, pp. 371–398, 2000.

[25] B. Gharesifard and J. Cortés, "Evolution of players' misperceptions in hypergames under perfect observations," *IEEE Transactions on Automatic Control*, vol. 57, no. 7, pp. 1641–1656, 2012.

[26] M. Voorneveld and H. Norde, "A characterization of ordinal potential games," *Games and Economic Behavior*, vol. 19, pp. 235–242, 1997.

[27] Y. M. Ermoliev and S. D. Flaam, "Repeated play of potential games," *Cybernetics and Systems Analysis*, vol. 38, pp. 355–367, 2002.

APPENDIX

In this appendix we study the topological structure of H-digraphs and show that they are acyclic. Because the results of this section are of independent interest, and have relevance beyond the deception scenario considered in this paper, we carry over our discussion for k-level n-player hypergames. We start by defining two special sequences of outcomes.

Definition A.1: (Nondeteriorating paths and weak improvement cycles in subjective hypergames): A strategic path $\mathfrak{S} = (x_1, x_2, \dots)$ in $\mathrm{S}_{\text{outcome}}$ is *nondeteriorating* for H^0_σ if $(x_\ell, x_{\ell+1}) \in \mathcal{E}_{H^0_\sigma}$, for all $\ell \in \mathbb{Z}_{\geq 1}$. A finite sequence of outcomes $\mathfrak{S} = (x_1, x_2, \dots, x_m, x_1)$, $m \in \mathbb{Z}_{\geq 1}$, is a *weak improvement cycle* for H^0_σ if it is nondeteriorating and $x_{\ell+1} \succ_{A_i\sigma} x_\ell$ for some $\ell \in \{1, \dots, m-1\}$ and $i \in \{1, \dots, n\}$.

Note that, if the preorders $\succ_{A_i\sigma}$, $i \in \{1, \ldots, n\}$, are all strict, every nondeteriorating path of the form $(x_1, x_2, \ldots, x_m, x_1)$ is a weak improvement cycle. We next characterize the acyclic structure of 0-level hypergames with two players.

Theorem A.2: (Subjective hypergames with two players contain no weak improvement cycle): Consider a k-level hypergame H^k between players A_1 and A_2. Let H^0_σ be a 0-level subjective hypergame perceived by σ, a sequence of length at most k on $\{A_1, A_2\}$. Then H^0_σ contains no weak improvement cycle.

Proof: We reason by contradiction. Suppose $\mathfrak{S} = (x_1, x_2, x_3, \ldots, x_p, x_1)$ is a weak improvement cycle for H^0_σ. Without loss of generality, we assume that players take alternate turns to take actions along the path. In other words, for $1 \leq \ell \leq p-2$, if $\pi_1(x_\ell) \neq \pi_1(x_{\ell+1})$ (resp. $\pi_2(x_\ell) \neq \pi_2(x_{\ell+1})$), then $\pi_2(x_{\ell+1}) \neq \pi_2(x_{\ell+2})$ (resp. $\pi_1(x_{\ell+1}) \neq \pi_1(x_{\ell+2})$). Our assumption is justified by the fact that, if $\pi_1(x_\ell) \neq \pi_1(x_{\ell+1}) \neq \pi_1(x_{\ell+2})$, then $x_{\ell+2}$ is a perceived improvement from x_ℓ for player A_1 and thus the outcome $x_{\ell+1}$ can be removed from the path $\mathfrak{S}$, with the result corresponding to a weak improvement cycle for H^0_σ. Note that, in particular, our assumption implies $p \geq 4$ is even.

Suppose A_2 is the first player to take an action, i.e., $\pi_2(x_1) \neq \pi_2(x_2)$ (the reasoning for the case when the first player is A_1 is analogous). Since $\mathfrak{S}$ is a weak improvement cycle, $x_2 \succeq_{A_2\sigma} x_1$. Moreover, since $\pi_1(x_2) \neq \pi_1(x_3)$, we have that $x_3 \succeq_{A_1\sigma} x_2$. As a result, we deduce that $x_3 \succeq_{A_2\sigma} x_1$; otherwise, A_2's perceived improvement x_2 from x_1 is not sanction-free. With a similar argument, one can deduce that, for $\ell \in \{1, \ldots, \frac{p-2}{2}\}$,

(i) $x_{2\ell}, x_{2\ell+1} \succeq_{A_2\sigma} x_{2\ell-1}$;

(ii) $x_{2\ell+1}, x_{2\ell+2} \succeq_{A_1\sigma} x_{2\ell}$;

(iii) $x_p, x_1 \succeq_{A_2\sigma} x_{p-1}$ and $x_1, x_2 \succeq_{A_1\sigma} x_p$.

Since $\mathfrak{S}$ is an improvement cycle, there must exist at least one $l \in \{1, \ldots, p-1\}$ such that either $x_{l+1} \succ_{A_1\sigma} x_l$ with $\pi_1(x_l) \neq \pi_1(x_{l+1})$ or $x_{l+1} \succ_{A_2\sigma} x_l$ with $\pi_2(x_l) \neq \pi_2(x_{l+1})$. Assume we are in the first case, i.e., l is odd, (the argument for the case when l is even is the same). Then, using (ii), one concludes that $x_p \succ_{A_1\sigma} x_2$, which contradicts (iii). ■

When the preorders $\succ_{A_i\sigma}$, $i \in \{1, \ldots, n\}$, are all strict, the existence of no weak improvement cycle implies that the associated H-digraph is, in fact, acyclic. We generalize the result above to the case of an arbitrary number of players using an inductive procedure.

Theorem A.3: (Subjective hypergames contain no weak improvement cycle): Consider a k-level hypergame H^k with n players $\{A_1, \ldots, A_n\}$. Then none of the subjective 0-level hypergame H^0_σ, where σ is a sequence of length at most k on $\{A_1, \ldots, A_n\}$, contains a weak improvement cycle.

Proof: Let $\{A_1, A_2, \ldots, A_n\}$ be a set of $n \in \mathbb{Z}_{\geq 3}$ players and $H^0_\sigma = (\mathrm{P}_{A_1\sigma}, \ldots, \mathrm{P}_{A_n\sigma})$. We denote by $\mathbf{S}^{\text{reachable}}_{\text{outcome}}|_{\pi_i(x)} \subseteq \mathbf{S}_{\text{outcome}}|_{\pi_i(x)}$ the set of all outcomes in $\mathbf{S}_{\text{outcome}}|_{\pi_i(x)}$ which can be reached from $x \in \mathbf{S}_{\text{outcome}}$ in the digraph $\mathcal{G}_{H^0_\sigma}$ by a directed path whose vertices belong to $\mathbf{S}_{\text{outcome}}|_{\pi_i(x)}$. Consider a sequence of outcomes $\mathfrak{S} = (x_1, x_2, \ldots, x_m)$ for H^0_σ, with $m \in \mathbb{Z}_{\geq 1}$. Similarly to the two players' case, without loss of generality, we assume that if player A_i takes an action that changes the outcome from x_ℓ to $x_{\ell+1}$, $\ell \in \{1, \ldots, m-1\}$, then a different player A_j, $j \in \{1, \ldots, n\} \setminus \{i\}$ takes the next action.

We proceed with the proof by induction on n. By Theorem A.2, the claim holds for $n = 2$. Suppose that the claim holds for any subjective 0-level hypergame with $n = N - 1$ players, and let us show that it also holds when $n = N$. If we fix the action of one player, say A_i, then players A_{-i} are playing a 0-level hypergame with $N - 1$ players, which contains no weak improvement cycle by the assumption of induction.

Without loss of generality, assume that player A_2 is the player that takes the action changing the outcome from x_1 to x_2. Let $\mathfrak{S}' = (x_1, x_2, \ldots, x_{m'})$ be a sequence of outcomes of largest cardinality m' with the property that, for all $\ell' \in \{1, \ldots, m'-1\}$, $x_{\ell'}$ and $x_{\ell'+1}$ are two consecutive outcomes in $\mathfrak{S}$ such that $\pi_2(x_{\ell'}) \neq \pi_2(x_{\ell'+1})$. Note that, by the induction assumption, for any $x_{\ell'}, x_{\ell'+1} \in \mathfrak{S}'$, the corresponding subsequence $(x_{\ell'}, \ldots, x_{\ell'+1})$ of $\mathfrak{S}$ contains no weak improvement cycle. Thus it is enough to show that $\mathfrak{S}'$ cannot be a weak improvement cycle. We reason by contradiction. Assume then that $\mathfrak{S}'$ is a weak improvement cycle. First, we claim that $x_{\ell'+1} \succeq_{A_2\sigma} x_{\ell'}$, for all $x_{\ell'}, x_{\ell'+1} \in \mathfrak{S}'$. For any outcome $x_l \in \mathfrak{S} \cap \mathbf{S}^{\text{reachable}}_{\text{outcome}}|_{\pi_2(x_{\ell'})}$, $l \in \{1, \ldots, m'\}$, we have that $x_l \succeq_{A_2\sigma} x_\ell$. In particular, there exists an outcome $x^*_l \in \mathfrak{S} \cap \mathbf{S}^{\text{reachable}}_{\text{outcome}}|_{\pi_2(x_{\ell'})}$ such that $\pi_{-2}(x^*_l) = \pi_{-2}(x_{\ell'+1})$, $x_{\ell'+1} \succeq_{A_2\sigma} x^*_l$, and $x^*_l \succeq_{A_2\sigma} x_\ell$. Thus we conclude that $x_{\ell'+1} \succeq_{A_2\sigma} x_{\ell'}$, as claimed. By a similar argument, one can conclude that $x_1 \succeq_{A_2\sigma} x_{m'}$. Since $\mathfrak{S}'$ is a weakly improvement cycle by hypothesis of contradiction, there exist at least two consecutive outcomes $x, y \in \mathfrak{S}'$, such that player A_2 is perceived to strictly prefer y to x. Using now an argument similar to the one in Theorem A.2, this leads to a contradiction. ■

Note that a corollary of Theorem A.3 is the known fact [20] that 0-level hypergames have at least one equilibrium.

Remark A.4: (Connection to ordinal potential games): A game $\mathbf{G} = (V, \mathbf{S}_{\text{outcome}}, \mathbf{P})$ is *ordinal potential*, cf. [26], if there exists a real-valued function $\mathcal{P} : \mathbf{S}_{\text{outcome}} \to \mathbb{R}$ such that for all $i \in \{1, \ldots, n\}$ and $a_i, b_i \in S_i$, we have $(a_i, a_{-i}) \succ_{P_i} (b_i, a_{-i})$ if and only if $\mathcal{P}(a_i, a_{-i}) > \mathcal{P}(b_i, a_{-i})$. The function $\mathcal{P}$ is called the *ordinal potential function* for $\mathbf{G}$. It is known [26] that $\mathbf{G}$ is ordinal potential if and only if it does not have any weak improvement cycle. Suppose $\mathcal{G}_{H^0_\sigma}$ is the H-digraph associated to a subjective hypergame H^0_σ with n players $V = \{A_1, \ldots, A_n\}$ and let $\mathbf{G} = (V, \mathbf{S}_{\text{outcome}}, \mathbf{P})$ be the game defined by $x_2 \succeq_{P_i} x_1$ with $\pi_{-i}(x_1) = \pi_{-i}(x_2)$, $\pi_i(x_1) \neq \pi_i(x_2)$ for $v_i \in V$ if and only if $(x_1, x_2) \in \mathcal{G}_{H^0_\sigma}$. Then, $\mathbf{G}$ is an ordinal potential game since, by Theorem A.3, the digraph $\mathcal{G}_{H^0_\sigma}$ contains no weak improvement cycle. •

An *improving adjustment scheme* in H^0_σ is any method that, given an initial outcome $x_1 \in \mathbf{S}_{\text{outcome}}$, generates a nondeteriorating sequence of outcomes $\mathfrak{S} = (x_1, x_2, \ldots)$. A best-response scheme is a special case of this notion, see, for example, [27]. We state next an immediate consequence of Theorem A.3 which captures how each individual player may learn the equilibria of its subjective hypergame.

Corollary A.5: (Learning in subjective hypergames): Any improving adjustment scheme in H^0_σ converges to an equilibrium.