Preprint submitted to Theory of Computing Systems
November 22, 2010

The Computational Complexity of Weak Saddles

**Felix Brandt · Markus Brill · Felix Fischer ·
Jan Hoffmann**

Abstract We study the computational aspects of weak saddles, an ordinal set-valued solution concept proposed by Shapley. Brandt et al. recently gave a polynomial-time algorithm for computing weak saddles in a subclass of matrix games, and showed that certain problems associated with weak saddles of bimatrix games are NP-hard. The important question of whether weak saddles can be *found* efficiently was left open. We answer this question in the negative by showing that finding weak saddles of bimatrix games is NP-hard, under polynomial-time Turing reductions. We moreover prove that recognizing weak saddles is coNP-complete, and that deciding whether a given action is contained in some weak saddle is hard for parallel access to NP and thus not even in NP unless the polynomial hierarchy collapses. Most of our hardness results are shown to carry over to a natural weakening of weak saddles.

Keywords Game Theory · Solution Concepts · Shapley's Saddles · Computational Complexity

1 Introduction

Saddle points, i.e., combinations of actions such that no player can gain by deviating, are one of the earliest solutions suggested in game theory (see, e.g., [25]). In two-player zero-sum games (henceforth *matrix games*), every saddle point happens to coincide with an optimal outcome both players can guarantee in the worst case and thus enjoys a very strong normative foundation. Unfortunately, however, not every matrix game possesses a saddle point. In order to remedy this situation, von Neumann [24] considered

F. Brandt · M. Brill
Institut für Informatik, Technische Universität München, 85748 Garching, Germany
E-mail: {brandtf,brill}@in.tum.de

F. Fischer
Harvard School of Engineering and Applied Sciences, Cambridge, MA 02138, USA
E-mail: fischerf@seas.harvard.edu

J. Hoffmann
Institut für Informatik, Ludwig-Maximilians-Universität München, 80538 München, Germany
E-mail: hoffmann@tcs.ifi.lmu.de

mixed—i.e., randomized—strategies and proved that every matrix game contains a mixed saddle point (or equilibrium) that moreover maintains the appealing normative properties of saddle points. The existence result was later generalized to arbitrary general-sum games by Nash [17], at the expense of its normative foundation. Since then, Nash equilibrium has commonly been criticized for its need for randomization, which may be deemed unsuitable, impractical, or even infeasible (see, e.g., [14, 15, 5]).

In two papers from 1953, Lloyd Shapley showed that existence of saddle points (and even uniqueness in the case of matrix games) can also be guaranteed by moving to *minimal sets* of actions rather than randomizations over them [21, 22].[1] Shapley defines a *generalized saddle point (GSP)* to be a tuple of subsets of actions of each player, such that every action not contained in the GSP is dominated by some action in the GSP, given that the remaining players choose actions from the GSP. A *saddle* is an inclusion-minimal GSP, i.e., a GSP that contains no other GSP. Depending on the underlying notion of dominance, one can define strict, weak, and very weak saddles. Shapley [23] showed that every matrix game admits a *unique* strict saddle. Duggan and Le Breton [10] proved that the same is true for the weak saddle in a certain subclass of symmetric matrix games that we refer to as *confrontation games*. While Shapley was the first to conceive weak GSPs, he was not the only one. Apparently unaware of Shapley's work, Samuelson [20] uses the very related concept of a *consistent pair* to point out epistemic inconsistencies in the concept of iterated weak dominance. Also, *weakly admissible sets* as defined by McKelvey and Ordeshook [15] in the context of spatial voting games are identical to weak GSPs. Other common *set-valued* concepts in game theory include *rationalizability* [3, 19] and *CURB sets* [1] (see also [16], pp. 88-91, for a general discussion of set-valued solution concepts).

In this paper we continue the study of the computational aspects of Shapley's saddles. Brandt et al. [5] recently gave polynomial-time algorithms for computing strict saddles in general games and weak saddles in confrontation games. Although it was shown that certain problems associated with weak saddles in bimatrix games are NP-complete, the question of whether weak saddles can be found efficiently was left open. We answer this question in the negative by showing that finding weak saddles is NP-hard. Moreover, we prove that recognizing weak saddles is coNP-complete, and that deciding whether an action is contained in a weak saddle of a bimatrix game is complete for parallel access to NP and thus not even in NP unless the polynomial hierarchy collapses. We finally demonstrate that our hardness results carry over to very weak saddles.

2 Related Work

In recent years, the computational complexity of game-theoretic solution concepts has come under increasing scrutiny. One of the most prominent results in this stream of research is that the problem of finding Nash equilibria in bimatrix games is PPAD-complete [7, 9], and thus unlikely to admit a polynomial-time algorithm. PPAD is a subclass of FNP, and it is obvious that Nash equilibria can be *recognized* in polynomial time. Interestingly, our results imply that this is not the case for weak saddles unless P=NP.

[1] The main results of the 1953 reports later reappeared in revised form [23].

Weak saddles rely on the elementary concept of weak dominance, whose computational aspects have been studied extensively in the form of *iterated* weak dominance [12, 8, 6]. In contrast to iterated dominance, saddles are based on a notion of stability reminiscent of Nash equilibrium and its various refinements. Weak saddles are also related to minimal covering sets, a concept that has been proposed independently in social choice theory [11, 10] and whose computational complexity has recently been analyzed [4, 2].

Brandt et al. [5] constructed a class of games that established a strong relationship between weak saddles and inclusion-maximal cliques in undirected graphs. Based on this construction and a reduction from the NP-complete problem CLIQUE, they showed that deciding whether there exists a weak saddle with a *certain number of actions* is NP-hard. This construction, however, did not permit any statements about the more important problems of *finding* a weak saddle, *recognizing* a weak saddle, or *deciding whether a certain action is contained* in some weak saddle.

3 Preliminaries

An accepted way to model situations of strategic interaction is by means of a *normal-form game* (see, e.g., [14]).

Definition 1 (Normal-Form Game) A (finite) *game in normal form* is a tuple $\Gamma = (N, (A_i)_{i\in N}, (p_i)_{i\in N})$ where $N = \{1, \ldots, n\}$ is a set of *players* and for each player $i \in N$, A_i is a nonempty finite set of *actions* available to player i, and $p_i : (\prod_{i\in N} A_i) \to \mathbb{R}$ is a function mapping each action profile (i.e., combination of actions) to a real-valued *payoff* for player i.

A *subgame* of a (normal-form) game $\Gamma = (N, (A_i)_{i\in N}, (p_i)_{i\in N})$ is a game $\Gamma' = (N, (A'_i)_{i\in N}, (p'_i)_{i\in N})$ where, for each $i \in N$, A'_i is a nonempty subset of A_i and $p'_i(a') = p_i(a')$ for all $a' \in \prod_{i=1}^{n} A'_i$. Γ is then called a *supergame* of Γ'.

In order to formally define Shapley's weak saddles, we need some additional notation. Let $A_N = (A_1, \ldots, A_n)$. For a tuple $S = (S_1, \ldots, S_n)$, write $S \subseteq A_N$ and say that S is a subset of A_N if $\emptyset \neq S_i \subseteq A_i$ for all $i \in N$. Further let $S_{-i} = \prod_{j\neq i} S_j$. For a player $i \in N$ and two actions $a_i, b_i \in A_i$ say that a_i *weakly dominates* b_i with respect to S_{-i}, denoted $a_i >_{S_{-i}} b_i$, if $p_i(a_i, s_{-i}) \geq p_i(b_i, s_{-i})$ for all $s_{-i} \in S_{-i}$, with at least one strict inequality.

Definition 2 (Weak Saddle) Let $\Gamma = (N, (A_i)_{i\in N}, (p_i)_{i\in N})$ be a game and $S = (S_1, \ldots, S_n) \subseteq A_N$. Then, S is a *weak generalized saddle point (WGSP)* of Γ if for each player $i \in N$ the following holds:

$$\text{For every } a_i \in A_i \setminus S_i \text{ there exists } s_i \in S_i \text{ such that } s_i >_{S_{-i}} a_i. \tag{1}$$

A *weak saddle* is a WGSP that contains no other WGSP.

In other words, every player i has a distinguished set S_i of actions such that for every action a_i that is not in the set S_i, there is some action in S_i that weakly dominates a_i, provided that the other players play only actions from their distinguished sets. Condition (1) will be called *external stability* in the following. A WGSP thus is a tuple that is externally stable for each player. Observe that the tuple A_N of all actions is always a WGSP, thereby guaranteeing existence of a weak saddle in every game. Weak

	b_1	b_2	b_3
a_1	0 2	1 1	0 1
a_2	1 1	1 1	1 2

Fig. 1 Example game with two weak saddles: $(\{a_1\}, \{b_1, b_2\})$ and $(\{a_1, a_2\}, \{b_2\})$. We follow the convention to write player 1's payoff in the lower left corner and player 2's payoff in the upper right corner of the corresponding matrix cell.

saddles do not have to be unique, as shown in the example in Figure 1. It is also not very hard to see that weak saddles are invariant under order-preserving transformations of the payoff functions and that every weak saddle contains a (mixed) Nash equilibrium.

For two-player games, we can simplify notation and write $\Gamma = (A, B, p)$, where A is the set of actions of player 1, B is the set of actions of player 2, and $p : A \times B \to \mathbb{R} \times \mathbb{R}$ is the payoff function on the understanding that $p(a, b) = (p_1(a, b), p_2(a, b))$ for all $(a, b) \in A \times B$. A two-player game is often called a *bimatrix* game, as it can conveniently be represented as an $|A| \times |B|$ bimatrix M, i.e., a matrix with rows indexed by A, columns indexed by B, and $M(a, b) = p(a, b)$ for every action profile $(a, b) \in A \times B$. We will commonly refer to actions of players 1 and 2 by the rows and columns of this matrix, respectively.

For an action a and a weak saddle $S = (S_1, S_2)$, we will sometimes slightly abuse notation and write $a \in S$ if $a \in (S_1 \cup S_2)$. In such cases, whether a is a row or a column should be either clear from the context or irrelevant for the argument. This partial identification of S and $S_1 \cup S_2$ is also reflected in referring to S as a "set" rather than a "pair" or "tuple." When reasoning about the structure of the saddles of a game, the following notation will be useful. For two actions $x, y \in A \cup B$, we write $x \rightsquigarrow y$ if every weak saddle containing x also contains y. Observe that $\rightsquigarrow$ as a relation on $(A \cup B) \times (A \cup B)$ is transitive. We now identify two sufficient conditions for $x \rightsquigarrow y$ to hold.

Fact 1 *Let $\Gamma = (A, B, p)$ be a two-player game, $b \in B$ an action of player 2, and $a \in A$ an action of player 1. Then $b \rightsquigarrow a$ if one of the following two conditions holds:*[2]

(i) *a is the unique action maximizing $p_1(\cdot, b)$, i.e., $\{a\} = \arg\max_{a' \in A} p_1(a', b)$.*
(ii) *a maximizes $p_1(\cdot, b)$, and all actions maximizing $p_1(\cdot, b)$ yield identical payoffs for* all *opponent actions, i.e., $a \in \arg\max_{a' \in A} p_1(a', b)$ and $p_1(a_1, b') = p_2(a_2, b')$ for all $a_1, a_2 \in \arg\max_{a' \in A} p_1(a', b)$ and all $b' \in B$.*

Part *(i)* of the statement above can be generalized in the following way. An action a is in the weak saddle if it is a unique best response to a subset of saddle actions: if $\{b_1, \ldots, b_t\} \subseteq S$ and there is no $a' \in A \setminus \{a\}$ with $p_1(a', b_i) \geq p_1(a, b_i)$ for all $i \in [t]$, then $a \in S$.[3] In this case, we write $\{b_1, \ldots, b_t\} \rightsquigarrow a$. Moreover, for two sets of actions X and Y, we write $X \rightsquigarrow Y$ if $X \rightsquigarrow y$ for all $y \in Y$. For example, in the game of Figure 1, $b_1 \rightsquigarrow a_1 \rightsquigarrow b_2$, $\{b_2, b_3\} \rightsquigarrow a_2$, and $\{b_1, b_3\} \rightsquigarrow \{a_1, a_2\}$.

We assume throughout the paper that games are given explicitly, i.e., as tables containing the payoffs for every possible action profile. We will be interested in the following computational problems for a given game Γ:

[2] The statement remains true if the roles of the two players are reversed.
[3] For $n \in \mathbb{N}$, we denote $[n] = \{1, \ldots, n\}$.

- FindWeakSaddle: Find a weak saddle of Γ.
- IsWeakSaddle: Is a given tuple $(S_1, \dots, S_n)$ a weak saddle of Γ?
- UniqueWeakSaddle: Does Γ contain exactly one weak saddle?
- InWeakSaddle: Is a given action a contained in a weak saddle of Γ?
- InAllWeakSaddles: Is a given action a contained in *every* weak saddle of Γ?
- NontrivialWeakSaddle: Does Γ contain a weak saddle that does *not* consist of all actions?

We assume the reader to be familiar with the basic notions of complexity theory, such as polynomial-time many-one reductions, Turing reductions, and the related notions of hardness and completeness, and with standard complexity classes such as P, NP, and coNP (see, e.g., [18]). We will further use the complexity classes Σ_2^p and Θ_2^p. $\Sigma_2^p = \text{NP}^{\text{NP}}$ forms part of the second level of the polynomial hierarchy and consists of all problems that can be solved in polynomial time by a non-deterministic Turing machine with access to an NP oracle. $\Theta_2^p = \text{P}_{||}^{\text{NP}}$ consists of all problems that can be solved in polynomial time by a deterministic Turing machine with parallel (i.e., non-adaptive) access to an NP oracle.

4 Hardness Results for Weak Saddles

We will now derive various hardness results for weak saddles. We begin by presenting a general construction that transforms a Boolean formula φ into a bimatrix game Γ_φ, such that the existence of certain weak saddles in Γ_φ depends on the satisfiability of φ. This construction will be instrumental for each of the hardness proofs given in the sequel.

4.1 A General Construction

Let $\varphi = C_1 \wedge \dots \wedge C_m$ be a Boolean formula in conjunctive normal form (CNF) over a finite set $V = \{v_1, \dots, v_n\}$ of variables. Denote by $L = \{v_1, \overline{v}_1, \dots, v_n, \overline{v}_n\}$ the set of all *literals*, where a literal is either a variable or its negation. Each *clause* C_j is a set of literals. An *assignment* $\alpha \subseteq L$ is a subset of the literals with the interpretation that all literals in α are set to "true." Assignment α is *valid* if $\ell \in \alpha$ implies $\overline{\ell} \notin \alpha$ for all $\ell \in L$.[4] We say that α *satisfies* a clause C_j if α is valid and $C_j \cap \alpha \neq \emptyset$. An assignment that satisfies all clauses of φ will be called a *satisfying assignment* for φ. A satisfying assignment α will be called *minimal* if there does not exist a satisfying assignment α' with $\alpha' \subset \alpha$. A formula that has a satisfying assignment will be called *satisfiable*. Clearly, every satisfiable formula has at least one minimal satisfying assignment.

We assume without loss of generality that φ does not contain any *trivial* clauses, i.e., clauses that contain both a variable v and its negation $\overline{v}$, and that no literal is contained in every clause. The game $\Gamma_\varphi = (A, B, p)$ is defined in three steps.

Step 1. Player 1 has actions $\{a^*, d^*\} \cup C$, where $C = \{C_1, \dots, C_m\}$ is the set of clauses of φ. Player 2 has actions $B = \{b^*\} \cup L$, where L is the set of literals.[5] Payoffs are given by

---

[4] If $\ell = \overline{v}_i$, then $\overline{\ell} = v_i$.

[5] There shall be no confusion by identifying literals with corresponding actions of player 2, which will henceforth be called "literal actions" (or "literal columns").

	b^*	v_1	$\overline{v}_1$	v_2	$\overline{v}_2$	$\cdots$	v_n	$\overline{v}_n$
a^*	1 1	0 0	0 0	0 0	0 0	$\cdots$	0 0	0 0
d^*	0 0	1 1	1 1	1 1	1 1	$\cdots$	1 1	1 1
C_1	1 0	0 0	0 1	0 1	0 0	$\cdots$	0 1	0 1
C_2	1 0	0 1	0 0	0 1	0 1	$\cdots$	0 0	0 1
$\vdots$					$\vdots$			
C_m	1 0	0 1	0 1	0 1	0 1	$\cdots$	0 1	0 1

Fig. 2 Subgame of Γ_φ for a formula $\varphi = C_1 \wedge \cdots \wedge C_m$ with $v_1, \overline{v}_2 \in C_1$ and $\overline{v}_1, v_n \in C_2$.

- $p(a^*, b^*) = (1, 1)$,
- $p(d^*, \ell) = (1, 1)$ for all $\ell \in L$,
- $p(C_j, b^*) = (0, 1)$ for all $j \in [m]$,
- $p(C_j, \ell) = (1, 0)$ for all $j \in [m]$ and $\ell \in L \setminus C_j$, and
- $p(a, b) = (0, 0)$ otherwise.

An example of such a game is shown in Figure 2. Observe that $(\{a^*\}, \{b^*\})$ is a weak saddle, and thus no strict superset can be a weak saddle. Furthermore, row d^* dominates row C_j with respect to a set of columns $\{\ell_1, \ldots, \ell_t\} \subseteq L$ if and only if $\ell_i \in C_j$ for some $i \in [t]$. In particular, for a valid assignment α it holds that $d^* >_\alpha C_j$ if and only if α satisfies C_j. Another noteworthy property of this game is that no weak saddle contains any of the rows C_j, because $C_j \rightsquigarrow b^* \rightsquigarrow a^*$ for each $j \in [m]$.

The basic idea behind this construction is the following. The game Γ_φ will have a weak saddle containing row d^* if and only if φ is satisfiable. More precisely, we will show that whenever a weak saddle (S_1, S_2) contains d^*, the set S_2 of saddle columns is a minimal satisfying assignment. Such a saddle will be called an *assignment saddle*. In order to prove that assignment saddles only exist if φ is satisfiable, we need to ensure that a pair (S_1, S_2) with $d^* \in S_1$ and $S_2 = \alpha$ cannot be a weak saddle if α does not satisfy φ or if α is not a valid assignment. This is achieved by means of additional actions (see step 2 below), for which the payoffs are defined in such a way that every "wrong" (i.e., unsatisfying or invalid) assignment yields a set containing both a^* and b^*. Obviously, such a set can never be a weak saddle because it contains the weak saddle $(\{a^*\}, \{b^*\})$ as a proper subset. In fact, $(\{a^*\}, \{b^*\})$ will be the unique weak saddle in cases where there is no satisfying assignment.

Step 2. We augment the action sets of both players. Player 1 has one additional row ℓ' for each literal $\ell \in L$.[6] Player 2 has one additional column y_i for each variable $v_i \in V$. Payoffs for profiles involving new actions are defined as follows (for an overview, refer to Figure 3):

- $p(a^*, y_i) = (1, 0)$ for all $i \in [n]$,

[6] Action ℓ' of player 1 and action ℓ of player 2 refer to the same literal, but we name them differently to avoid confusion.

- $p(\ell', \ell) = (2, 1)$ for all $\ell \in L$,
- $p(\ell', y_i) = (0, 1)$ for all $i \in [n]$ and $\ell' \in \{v_i', \overline{v}_i'\}$, and
- $p(a, b) = (0, 0)$ otherwise.

Observe that, by Fact 1 and the discussion following it, $\ell \rightsquigarrow \ell'$, $\{v_i', \overline{v}_i'\} \rightsquigarrow y_i$, and $y_i \rightsquigarrow a^* \rightsquigarrow b^*$ for each $\ell \in L$ and each $i \in [n]$. This means that no assignment saddle can contain both v_i as well as its negation $\overline{v}_i$.

There only remains one subtlety to be dealt with. In the game defined so far, there are weak saddles containing row d^*, whose existence is independent of the satisfiability of φ, namely $(\{d^*, \ell'\}, \{\ell\})$ for each $\ell \in L$. We destroy these saddles by using additional rows.

Step 3. We introduce new rows $r_1, \overline{r}_1, \ldots, r_n, \overline{r}_n$, one for each literal, with the property that $r_i \rightsquigarrow b^*$, and that r_i and $\overline{r}_i$ can only be weakly dominated (by v_i and $\overline{v}_i$, respectively) if at least one literal column other than v_i or $\overline{v}_i$ is in the saddle. For this, we define

- $p(r_i, b^*) = p(\overline{r}_i, b^*) = (0, 1)$ for all $i \in [n]$,
- $p(r_i, v_i) = p(\overline{r}_i, \overline{v}_i) = (2, 0)$ for all $i \in [n]$,
- $p(r_i, \ell) = p(\overline{r}_i, \ell) = (-1, 0)$ for all $i \in [n]$ and $\ell \in \{v_{i \bmod n+1}, \overline{v}_{i \bmod n+1}\}$, and
- $p(a, b) = (0, 0)$ otherwise.

The game Γ_φ now has action sets $A = \{a^*, d^*\} \cup C \cup L \cup \{r_1, \ldots, \overline{r}_n\}$ for player 1 and $B = \{b^*\} \cup L \cup \{y_1, \ldots, y_n\}$ for player 2. The size of Γ_φ thus is clearly polynomial in the size of φ. A complete example of such a game is given in Figure 3.

For a valid assignment α, define $S^\alpha = (\{d^*\} \cup \alpha, \alpha)$. It should be clear from the argumentation above that S^α is a weak generalized saddle point of Γ_φ if and only if α satisfies φ. In particular, S^α is a weak saddle if and only if α is a *minimal* satisfying assignment. To show that membership of a given action in a weak saddle is NP-hard, it suffices to show that there are no other weak saddles containing row d^*. We do so in the following section.

4.2 Membership is NP-hard

We now show that it is NP-hard to decide whether a given action is contained in some weak saddle.

Proposition 1 InWeakSaddle *is NP-hard, even for two-player games.*

Proof We give a reduction from SAT. For a CNF formula φ, we show that the game Γ_φ, defined in Section 4.1, has a weak saddle that contains action d^* if and only if φ is satisfiable. The direction from right to left is straightforward. If α is a minimal satisfying assignment for φ, then S^α is a weak saddle that contains d^*.

For the other direction, we will show that all weak saddles containing d^* are (essentially) assignment saddles. Let $S = (S_1, S_2)$ be a weak saddle of Γ_φ such that $d^* \in S_1$. We can assume that $S_2 \subseteq L$. If this was not the case, i.e., if there was a column $c \in \{b^*, y_1, \ldots, y_n\}$ with $c \in S_2$, then $c \rightsquigarrow a^* \rightsquigarrow b^*$, and $(\{a^*\}, \{b^*\})$ would be a smaller saddle contained in S, a contradiction. We will now show that

(i) $|S_2| \geq 2$,
(ii) $|\{v_i, \overline{v}_i\} \cap S_2| \leq 1$ for all $i \in [n]$, and
(iii) $C \cap S_1 = \emptyset$.

	b^*	v_1	$\overline{v}_1$	v_2	$\overline{v}_2$	$\cdots$	v_n	$\overline{v}_n$	y_1	y_2	$\cdots$	y_n
a^*	1 1					$\cdots$			0 1	0 1	$\cdots$	0 1
d^*		1 1	1 1	1 1	1 1	$\cdots$	1 1	1 1			$\cdots$	
C_1	1 0		0 1	0 1		$\cdots$	0 1	0 1			$\cdots$	
C_2	1 0	0 1		0 1	0 1	$\cdots$		0 1			$\cdots$	
$\vdots$						$\vdots$						
C_m	1 0	0 1	0 1	0 1	0 1	$\cdots$	0 1	0 1			$\cdots$	
v'_1		1 2				$\cdots$			1 0		$\cdots$	
$\overline{v}'_1$			1 2			$\cdots$			1 0		$\cdots$	
v'_2				1 2		$\cdots$				1 0	$\cdots$	
$\overline{v}'_2$					1 2	$\cdots$				1 0	$\cdots$	
$\vdots$						$\vdots$						
v'_n						$\cdots$	1 2				$\cdots$	1 0
$\overline{v}'_n$						$\cdots$		1 2			$\cdots$	1 0
r_1	1 0	0 2		0 -1	0 -1	$\cdots$					$\cdots$	
$\overline{r}_1$	1 0		0 2	0 -1	0 -1	$\cdots$					$\cdots$	
r_2	1 0			0 2		$\cdots$					$\cdots$	
$\overline{r}_2$	1 0				0 2	$\cdots$					$\cdots$	
$\vdots$						$\vdots$						
r_n	1 0	0 -1	0 -1			$\cdots$	0 2				$\cdots$	
$\overline{r}_n$	1 0	0 -1	0 -1			$\cdots$		0 2			$\cdots$	

Fig. 3 Game Γ_φ used in the proof of Proposition 1. Payoffs equal $(0, 0)$ unless specified otherwise. $S^\alpha = (\{d^*\} \cup \alpha, \alpha)$ is a weak generalized saddle point of Γ_φ if and only if α satisfies φ. For improved readability, thick lines are used to separate different types of actions.

For *(i)*, suppose that $|S_2| = 1$. Without loss of generality, $S_2 = \{v_i\}$. Then, both v'_i and r_i have to be in S_1, as they are maximal with respect to $\{v_i\}$. Together with $r_i \rightsquigarrow b^*$, this however contradicts the fact that $b^* \notin S_2$.

For *(ii)*, suppose that there exists $i \in [n]$ with $\{v_i, \overline{v}_i\} \subseteq S_2$. Then at least one of the rows v'_i or r_i *and* at least one of the rows $\overline{v}'_i$ or $\overline{r}_i$ is in the set S_1. Since $r_i \rightsquigarrow b^*$ as well as $\overline{r}_i \rightsquigarrow b^*$, and since $b^* \notin S_2$, we deduce that $\{v'_i, \overline{v}'_i\} \subseteq S_1$. On the other hand, $\{v'_i, \overline{v}'_i\} \rightsquigarrow y_i$, again contradicting $S_2 \subseteq L$.

For *(iii)*, merely observe that $C_j \rightsquigarrow b^*$ for all $j \in [m]$.

We now show that $d^* >_{S_2} C_j$ for all $j \in [m]$. Consider some $j \in [m]$. From *(iii)* we know that there exists a row $s \in S_1$ with $s >_{S_2} C_j$. We consider two cases. First, assume that $|\{\ell \in S_2 : p_1(C_j, \ell) = 1\}| \geq 2$. It follows from our assumption and from the definition of p_1 that d^* is the only row that can weakly dominate C_j with respect to S_2. If, on the other hand, $|\{\ell \in S_2 : p_1(C_j, \ell) = 1\}| \leq 1$, $d^* >_{S_2} C_j$ follows immediately from $S_2 \subseteq L$ and *(i)*.

Define the assignment $\alpha = S_2$ and note that by *(ii)*, α is valid. The fact that $d^* >_\alpha C_j$ implies that there exists $\ell \in \alpha$ with $p_1(C_j, \ell) = 0$, which means that $\ell \in C_j$. Thus α satisfies C_j for all $j \in [m]$. In other words, φ is satisfiable. □

4.3 Membership is coNP-hard

We have just seen that it is NP-hard to decide whether there exists a weak saddle containing a given action. In order to prove that this problem is also coNP-hard, we first show the following: given a game and an action c, it is possible to augment the game with additional actions such that every weak saddle of the augmented game that contains c contains all actions of this game.

Lemma 1 *Let $\Gamma = (A, B, p)$ be a two-player game and $c \in A \cup B$ an action of Γ. Then there exists a supergame $\Gamma^c = (A', B', p')$ of Γ with the following properties:*

(i) If S is a weak saddle of Γ^c containing c, then $S = (A', B')$.
(ii) If S is a weak saddle of Γ that does not *contain c, then S is a weak saddle of Γ^c.*
(iii) The size of Γ^c is polynomial in the size of Γ.

Proof Let $n = |A|$ and $m = |B|$. Without loss of generality, we may assume that all payoffs in Γ are positive and that c is a column, i.e., $p_\ell(a, b) > 0$ for all $(a, b) \in A \times B$, $\ell \in [2]$, and $c \in B$. Define $\lambda = \max_{a \in A} p_1(a, c) + 1$, such that λ is greater than the maximum payoff to player 1 in column c. Now, let Γ^c be a supergame of Γ with $n + m - 1$ additional rows and n additional columns, i.e., $\Gamma^c = (A', B', p')$, where $A' = A \cup \{a'_1, \ldots, a'_{n+m-1}\}$, $B' = B \cup \{b'_1, \ldots, b'_n\}$ and $p'|_{A \times B} = p$. Payoffs for action profiles not in $A \times B$ are shown in Figure 4.

For *(i)*, let $S = (S_1, S_2)$ be a weak saddle of Γ^c with $c \in S_2$. Using the second part of Fact 1, we get $c \rightsquigarrow A' \setminus A \rightsquigarrow B' \setminus \{c\} \rightsquigarrow A$. For *(ii)*, observe that our assumption about the payoffs in Γ implies that each additional action is dominated by each original action as long as c is not contained in the weak saddle. Finally, *(iii)* is immediate from the definition of Γ^c. □

We are now ready to show that InWeakSaddle is coNP-hard.

Proposition 2 InWeakSaddle *is coNP-hard, even for two-player games.*

Proof We give a reduction from UNSAT. For a given CNF formula φ, consider the game $\Gamma_\varphi^{b^*}$ obtained by augmenting the game Γ_φ defined in Section 4.1 in such a way that every weak saddle containing action b^* in fact contains all actions. We show that $\Gamma_\varphi^{b^*}$ has a weak saddle containing b^* if and only if φ is unsatisfiable.

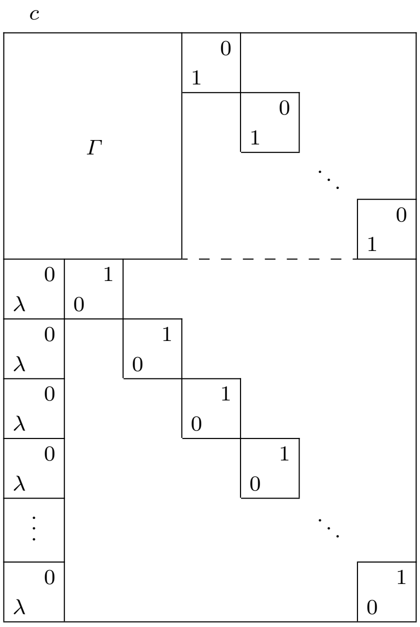

Fig. 4 Construction used in the proof of Lemma 1. Payoffs for new action profiles are $(0, 0)$ unless specified otherwise, and λ is chosen so as to maximize $p'_1(\cdot, c)$. Every weak saddle containing column c then equals the set of all actions.

For the direction from left to right, assume that there exists a weak saddle S with $b^* \in S$. By Lemma 1, S is trivial, i.e., equals the set of all actions. Furthermore, S must be the unique weak saddle of $\Gamma_\varphi^{b^*}$, because any other weak saddle would violate minimality of S. In particular, S^α cannot be a saddle for any assignment α, which by the discussion in Section 4.1 means that φ is unsatisfiable.

For the direction from right to left, assume that φ is unsatisfiable. Similar reasoning as in the proof of Proposition 1 shows that every weak saddle $S = (S_1, S_2)$ satisfies $S_2 \not\subseteq L$, i.e., S contains at least one column not corresponding to a literal. However, since $b \rightsquigarrow a^*$ for every column $b \in B \setminus L$ and $a^* \rightsquigarrow b^*$, we have that $b^* \in S_2$ for every weak saddle of $\Gamma_\varphi^{b^*}$. $\square$

The proof of Proposition 2 implies several other hardness results.

Corollary 1 *The following hold:*

- IsWeakSaddle *is coNP-complete.*
- InAllWeakSaddles *is coNP-complete.*
- UniqueWeakSaddle *is coNP-hard.*

All hardness results hold even for two-player games.

Proof Let φ be a Boolean formula, which without loss of generality we can assume to have either no satisfying assignment or more than one. (For any Boolean formula, this property can for example be achieved by adding a clause with two new variables, thereby multiplying the number of satisfying assignments by three.)

Recall the definition of the game $\Gamma_\varphi^{b^*}$ used in the proof of Proposition 2. It is easily verified that the following statements are equivalent: formula φ is unsatisfiable, $\Gamma_\varphi^{b^*}$ has a trivial weak saddle, $\Gamma_\varphi^{b^*}$ has a unique weak saddle, and b^* is contained in all weak saddles of $\Gamma_\varphi^{b^*}$. This provides a reduction from UNSAT to each of the problems above.

Membership of InAllWeakSaddles in coNP holds because any externally stable set that does not contain the action in question serves as a witness that this action is *not* contained in every weak saddle. For membership of IsWeakSaddle, consider a tuple S of actions that is *not* a weak saddle. Then either S is not externally stable, or there exists a proper subset of S that is externally stable. In both cases there is a witness of polynomial size. □

4.4 Finding a Saddle is NP-hard

A particularly interesting consequence of Proposition 2 concerns the existence of a nontrivial weak saddle. As we will see, the hardness of deciding the latter can be used to obtain a result about the complexity of the search problem.

Corollary 2 NontrivialWeakSaddle *is NP-complete. Hardness holds even for two-player games.*

Proof For membership in NP, observe that proving the existence of a nontrivial weak saddle is tantamount to finding a proper subset of the set of all actions that is externally stable. By definition, every such subset is guaranteed to contain a weak saddle. Obviously, external stability can be checked in polynomial time.

Hardness is again straightforward from the proof of Proposition 2, since the game $\Gamma_\varphi^{b^*}$ has a nontrivial weak saddle if and only if formula φ is satisfiable. □

Corollary 3 FindWeakSaddle *is NP-hard under polynomial-time Turing reductions, even for two-player games.*

Proof Suppose there exists an algorithm that computes some weak saddle of a game in time polynomial in the size of the game. Such an algorithm could obviously be used to solve the NP-hard problem NontrivialWeakSaddle in polynomial time. Just run the algorithm once. If it returns a nontrivial saddle, the answer is "yes." Otherwise the set of all actions must be the unique weak saddle of the game, and the answer is "no." □

4.5 Membership is Θ_2^p-hard

Now that we have established that InWeakSaddle is both NP-hard and coNP-hard, we will raise the lower bound to Θ_2^p. Wagner provided a sufficient condition for Θ_2^p-hardness that turned out to be very useful (see, e.g., [13]).

Lemma 2 (Wagner [26]) *Let S be an NP-complete set, and let T be an arbitrary set. If there exists a polynomial-time computable function f such that*

$$\|\{i : x_i \in S\}\| \text{ is odd} \iff f(x_1, \ldots, x_{2k}) \in T \tag{2}$$

for all $k \geq 1$ and all strings $x_1, \ldots, x_{2k}$ satisfying $x_{j-1} \in S$ whenever $x_j \in S$ for every j with $1 < j \leq 2k$, then T is Θ_2^p-hard.

We now apply Wagner's Lemma to show Θ_2^p-hardness of InWeakSaddle.

Theorem 1 InWeakSaddle *is Θ_2^p-hard, even for two-player games.*

Proof We apply Lemma 2 with $S = \text{SAT}$ and $T = \text{InWeakSaddle}$. Fix an arbitrary $k \geq 1$ and let $\varphi_1, \ldots, \varphi_{2k}$ be $2k$ Boolean formulas such that satisfiability of φ_j implies satisfiability of φ_{j-1}, for each j, $1 < j \leq 2k$.

We will now define a polynomial-time computable function f which maps the given $2k$ Boolean formulas to an instance of InWeakSaddle such that (2) is satisfied. For *odd* $i \in [2k]$, let $\Gamma_i = (A_i, B_i, p_i)$ be the game Γ_{φ_i} as defined in the proof of Proposition 1, with decision row d^* renamed as d_i. Recall that this game has a weak saddle containing d_i if and only if φ_i is satisfiable. Analogously, for *even* $i \in [2k]$, let $\Gamma_i = (A_i, B_i, p_i)$ be the game $\Gamma_{\varphi_i}^{d_i}$ as defined in the proof of Proposition 2, with decision column b^* renamed as d_i. Thus, Γ_i has a weak saddle containing d_i if and only if φ_i is *unsatisfiable*. For all $i \in [2k]$, we may without loss of generality assume that all payoffs in Γ_i are positive and strictly smaller than some $K \in \mathbb{N}$, and that the decision action d_i of game Γ_i is a row, i.e., $0 < p_\ell(a, b) < K$ for all $(a, b) \in A_i \times B_i$ and $\ell \in [2]$, and $d_i \in A_i$.[7]

Now define the game Γ by combining the games Γ_i, $i \in [2k]$, with one additional row z_i and two additional columns c_i^1 and c_i^2 for each $i \in [2k]$, as well as a decision row d^*, i.e., $\Gamma = (A, B, p)$ where $A = \bigcup_{i=1}^{2k} A_i \cup \{z_1, \ldots, z_{2k}\} \cup \{d^*\}$ and $B = \bigcup_{i=1}^{2k} B_i \cup \bigcup_{i=1}^{2k} \{c_i^1, c_i^2\}$. For $a \in A_i$ and $b \in B_j$, payoffs are defined as $p(a, b) = p_i(a, b)$ if $i = j$ and $p(a, b) = (0, 0)$ otherwise. Furthermore, for $b \in \bigcup B_j$, let $p(z_i, b) = (0, 1)$ for all $i \in [2k]$ and $p(d^*, b) = (0, 1)$. The definition of p on profiles containing a new column c_i^ℓ, $i \in [2k]$, $\ell \in [2]$ is quite complicated, and we recommend consulting Figure 5 for an overview. Player 2 has only two distinct payoffs for these columns: For $a \in A$ and $\ell \in [2]$,

$$p_2(a, c_i^\ell) = \begin{cases} K & \text{if } a = d_i \\ 0 & \text{otherwise.} \end{cases}$$

Recall that all payoffs in the games Γ_i are smaller than K, such that the payoff for player 2 in the profiles (d_i, c_i^1) and (d_i, c_i^2) is maximal in Γ.

The payoffs for player 1 are defined in order to connect the games Γ_{2i-1} and Γ_{2i}, for each $i \in [k]$. We need some notation. For $i \in [2k]$, let i° be $i+1$ if i is odd and $i-1$ if i is even. Thus, each pair $\{i, i^\circ\}$ is of the form $\{2j-1, 2j\}$ for some j. For $a \in \bigcup A_j$, define

$$p_1(a, c_i^\ell) = \begin{cases} 1 & \text{if } \ell = 1 \text{ and } a \in A_i \\ 2 & \text{if } \ell = 1 \text{ and } a \in A_{i^\circ} \\ 0 & \text{otherwise.} \end{cases}$$

Furthermore, for $i, j \in [2k]$, let

$$p_1(z_j, c_i^1) = \begin{cases} 1 & \text{if } j = i \\ 0 & \text{otherwise,} \end{cases} \qquad \text{and} \qquad p_1(z_j, c_i^2) = \begin{cases} 0 & \text{if } j = i^\circ \\ 1 & \text{otherwise.} \end{cases}$$

Finally, let $p_1(d^*, c_i^1) = 0$ and $p_1(d^*, c_i^2) = 1$ for all $i \in [2k]$.

An example of the game Γ for the case $k = 2$ is depicted in Figure 5, where we assume without loss of generality that each d_i is the first row of Γ_i.

[7] Adding a positive number to every payoff does not change the dominance relation between the actions. As the minimum payoff in Γ_i is -1, adding a number greater than 1 suffices. If d_i is a column, as in the proof of Proposition 2, we can simply transpose the game by exchanging the two players.

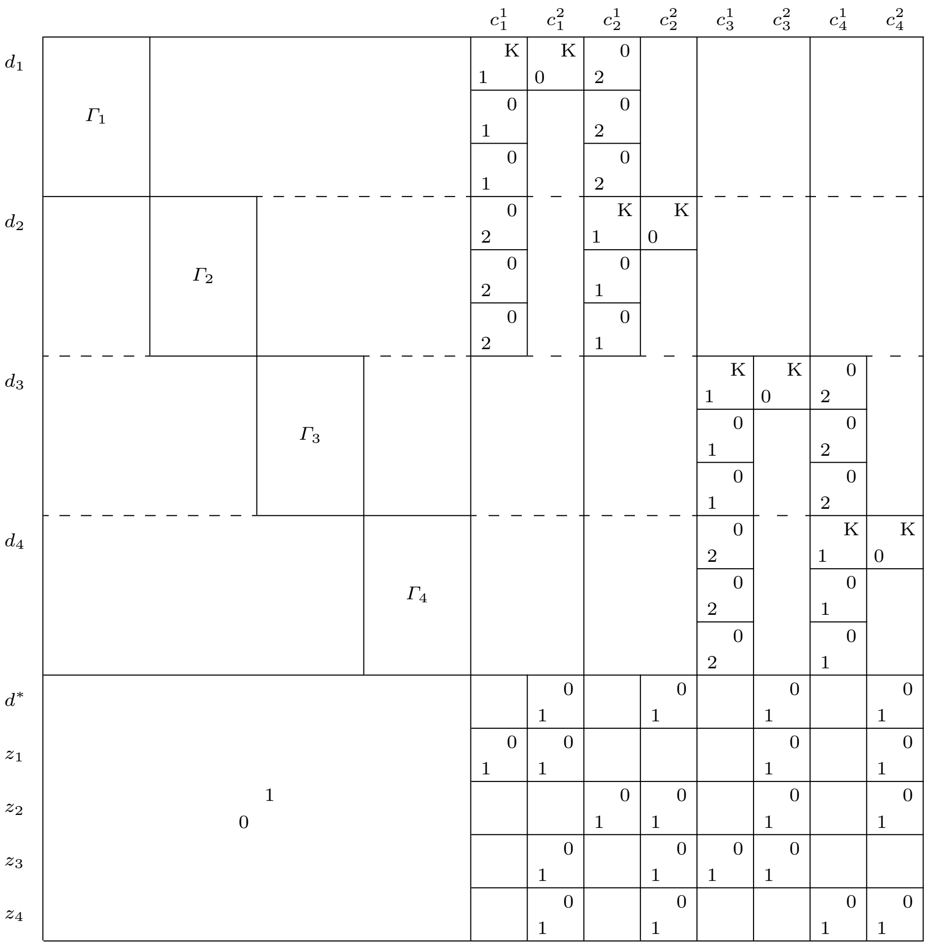

Fig. 5 Game Γ used in the proof of Theorem 1. Payoffs are $(0, 0)$ unless specified otherwise. Γ has a weak saddle containing row d^* if and only if both Γ_1 and Γ_2 *or* both Γ_3 and Γ_4 have a weak saddle containing their respective decision row d_i.

The following facts are readily appreciated.

Fact 2 *If S is a weak saddle of Γ_i and $d_i \notin S$, then S is also a weak saddle of Γ.*

For a weak saddle $S = (S_1, S_2)$ of Γ and $i \in [2k]$, define $S^i = (S_1 \cap A_i, S_2 \cap B_i)$ as the intersection of S with Γ_i.

Fact 3 *If S is a weak saddle of Γ, then S^i is either a weak saddle of Γ_i or empty.*

For Fact 2 it suffices to check external stability. For Fact 3, observe that our assumption that $p_\ell(a, b) > 0$ implies that weak domination with respect to a subset of $A_i \cup B_i$ can only occur among actions belonging to $A_i \cup B_i$. Therefore, if some action profile in $A_i \times B_i$ is contained in a weak saddle, all actions of Γ_i *not* contained in the saddle must be dominated by some saddle action of the same subgame Γ_i.

In order to be able to apply Lemma 2, we now prove (2), which here amounts to showing the following equivalence:

$$\|\{i : \varphi_i \in \mathrm{SAT}\}\| \text{ is odd} \iff \mathit{\Gamma} \text{ has a weak saddle } S \text{ with } d^* \in S. \qquad (3)$$

For the direction from left to right, assume that there is an odd number i such that φ_i is satisfiable and $\varphi_{i^\circ} = \varphi_{i+1}$ is not. Then, there exist weak saddles S^i and S^{i° of the games $\mathit{\Gamma}_i$ and $\mathit{\Gamma}_{i^\circ}$, respectively, such that $d_i \in S^i$ and $d_{i^\circ} \in S^{i^\circ}$. Define $S = S^i \cup S^{i^\circ} \cup \{d^*, z_1, \ldots, z_{2k}\} \cup \{c_i^1, c_i^2, c_{i^\circ}^1, c_{i^\circ}^2\}$. We claim that S is a weak saddle of $\mathit{\Gamma}$. The proof consists of two parts.

First, we have to show that S is externally stable, i.e., that all actions not in the saddle are weakly dominated by saddle actions. To see this, let $a \in A_j$ be a row that is not in S. If $j \notin \{i, i^\circ\}$, then a is weakly dominated by *every* saddle row because it yields payoff 0 to player 1 against any saddle column. If, on the other hand, $j \in \{i, i^\circ\}$, then a is weakly dominated by the same row that weakly dominates it in the subgame $\mathit{\Gamma}_j$. The argument for non-saddle columns $b \in \bigcup_j B_j$ is analogous. Moreover, every column c_j^ℓ with $j \notin \{i, i^\circ\}$ is weakly dominated by the saddle columns $c_i^1, c_i^2, c_{i^\circ}^1$, and $c_{i^\circ}^2$.

Second, we have to show that S is inclusion-minimal, i.e., that no proper subset of S is a weak saddle of $\mathit{\Gamma}$. Let $\tilde{S} \subseteq S$ be a weak saddle. By Fact 3 and the observation that $\tilde{S}^i$ cannot be empty, we know that $\tilde{S}^i = S^i$, as otherwise inclusion-minimality of S^i in $\mathit{\Gamma}_i$ would be violated. In particular, $d_i \in \tilde{S}^i$, which implies that $\{c_i^1, c_i^2\} \subseteq \tilde{S}$. The same reasoning for i° shows that $\tilde{S}^{i^\circ} = S^{i^\circ}$ and $\{c_{i^\circ}^1, c_{i^\circ}^2\} \subseteq \tilde{S}$. Now, $\{c_i^1, c_i^2\} \rightsquigarrow z_i$ and $\{c_{i^\circ}^1, c_{i^\circ}^2\} \rightsquigarrow z_{i^\circ}$. Furthermore, all rows z_j with $j \notin \{i, i^\circ\}$, as well as d^*, are in $\tilde{S}$, because they are all maximal and identical with respect to S. Here, maximality is due to the fact that they are the only rows that yield a positive payoff to player 1 against both saddle columns c_i^2 and $c_{i^\circ}^2$. Thus $\tilde{S} = S$, meaning that S is indeed inclusion-minimal.

For the direction from right to left, let S be a weak saddle of $\mathit{\Gamma}$ with $d^* \in S$. From the definition of $p_2(d^*, \cdot)$, we infer that $S \cap \bigcup_j B_j \neq \emptyset$, which in turn implies that $S \cap \bigcup_j A_j \neq \emptyset$. We can now deduce that there is at least one column $c_i^\ell \in S$, as otherwise row d^* would always yield 0 against all saddle actions and $S \setminus \{d^*\}$ would be externally stable. Now observe that for any $i \in [2k]$, the definition of $p_2(\cdot, c_i^\ell)$ implies that every weak saddle of $\mathit{\Gamma}$ contains either none or both of the columns c_i^1 and c_i^2. We thus have $\{c_i^1, c_i^2\} \subseteq S$. Furthermore, $z_i \in S$ because $\{c_i^1, c_i^2\} \rightsquigarrow z_i$. However, z_i must not weakly dominate d^* with respect to S, because otherwise $S \setminus \{d^*\}$ would be externally stable. This means there has to be a saddle column $c \in S$ with $p_1(z_i, c) < p_1(d^*, c)$. The only column satisfying this property is $c_{i^\circ}^2$, which means that both $c_{i^\circ}^2$ and, by the same argument as above, $c_{i^\circ}^1$ are contained in S. Now that both c_i^1 and $c_{i^\circ}^1$ are in S, at least one row from each of the games $\mathit{\Gamma}_i$ and $\mathit{\Gamma}_{i^\circ}$ has to be a saddle action, i.e., $S^i \neq \emptyset$ and $S^{i^\circ} \neq \emptyset$. By Fact 3, we conclude that S^i and S^{i° are weak saddles of $\mathit{\Gamma}_i$ and $\mathit{\Gamma}_{i^\circ}$, respectively.

It remains to be shown that $d_i \in S^i$ and $d_{i^\circ} \in S^{i^\circ}$. If $d_i \notin S^i$, then by Fact 2, $S^i \subset S$ would be a weak saddle of $\mathit{\Gamma}$, contradicting inclusion-minimality of S. The argument for S^{i° is analogous. It finally follows from the construction that φ_i is satisfiable and φ_{i° is unsatisfiable,[8] which completes the proof of (3). By Lemma 2, InWeakSaddle is $\mathit{\Theta}_2^p$-hard. □

[8] Here we have assumed without loss of generality that $i < i^\circ$, i.e., i is odd and $i^\circ = i + 1$ is even.

We conclude this section by showing that Σ_2^p is an upper bound for the membership problem.

Proposition 3 InWeakSaddle *is in* Σ_2^p.

Proof Let $\Gamma = (N, (A_i)_{i\in N}, (p_i)_{i\in N})$ be a game and $d^* \in \bigcup_i A_i$ a designated action. First observe that we can verify in polynomial time whether a subset of A_N is externally stable. We can guess a weak saddle S containing d^* in nondeterministic polynomial time and verify its minimality by checking that none of its subsets are externally stable. This places InWeakSaddle in $\mathrm{NP}^{\mathrm{coNP}} = \mathrm{NP}^{\mathrm{NP}}$ and thus in Σ_2^p. □

5 Very Weak Saddles

A natural weakening of weak dominance is *very weak dominance*, which does not require a strict inequality in addition to the weak inequalities (see, e.g., [14]). Thus, in particular, two actions that always yield the same payoff very weakly dominate each other. Formally, for a player $i \in N$ and two actions $a_i, b_i \in A_i$ we say that a_i *very weakly dominates* b_i with respect to S_{-i}, denoted $a_i \geq_{S_{-i}} b_i$, if $p_i(a_i, s_{-i}) \geq p_i(b_i, s_{-i})$ for all $s_{-i} \in S_{-i}$. Based on this modified notion of dominance, one can define the very weak analog of the weak saddle (cf. Definition 2).

Definition 3 (Very Weak Saddle) Let $\Gamma = (N, (A_i)_{i\in N}, (p_i)_{i\in N})$ be a game and $S = (S_1, \ldots, S_n) \subseteq A_N$. Then, S is a *very weak generalized saddle point (VWGSP)* of Γ if for each player $i \in N$ the following condition holds:

For every $a_i \in A_i \setminus S_i$ there exists $s_i \in S_i$ such that $s_i \geq_{S_{-i}} a_i$.

A *very weak saddle* is a VWGSP that contains no other VWGSP.

Computational problems for very weak saddles are defined analogously to their counterparts for weak saddles. It turns out that most of our results for the latter also hold for the former.

Theorem 2 *The following hold:*

- InVeryWeakSaddle *is NP-hard.*
- InVeryWeakSaddle *is coNP-hard.*
- IsVeryWeakSaddle *is coNP-complete.*
- InAllVeryWeakSaddles *is coNP-complete.*
- UniqueVeryWeakSaddle *is coNP-hard.*
- NontrivialVeryWeakSaddle *is NP-complete.*
- FindVeryWeakSaddle *is NP-hard under Turing reductions.*

All hardness results hold even for two-player games.

The proof of this theorem is deferred to the appendix. It is worth noting here, however, that the results for very weak saddles do not follow in an obvious way from those for weak saddles, or vice versa. While the proofs are based on the same general idea, and again on one core construction, there are some significant technical differences.

An argument analogous to that for InWeakSaddle shows that InVeryWeakSaddle is in Σ_2^p. On the other hand, Θ_2^p-hardness of InVeryWeakSaddle appears much harder to obtain. In particular, the construction in the proof of Theorem 1 uses pairs of actions c_i^1 and c_i^2 that are identical from the point of view of player 2, and argues that every weak saddle must contain either none or both of them. This argument no longer goes through for very weak saddles, because c_i^1 and c_i^2 very weakly dominate each other, and indeed there are very weak saddles that contain only one of the two actions. Additional insights will therefore be required to raise the lower bound for InVeryWeakSaddle.

6 Conclusion

In the early 1950s, Shapley proposed an ordinal set-valued solution concept known as the weak saddle. We have shown that weak saddles are computationally intractable even in bimatrix games. As it turned out, not only *finding* but also *recognizing* weak saddles is computationally hard. This distinguishes weak saddles from Nash equilibrium, iterated dominance, and most other game-theoretic solution concept we are aware of. Three of the most challenging open problems concern the complexity of weak saddles in matrix games, the gap between Θ_2^p and Σ_2^p for InWeakSaddle, and a complete characterization of the complexity of FindWeakSaddle.

Acknowledgements This material is based on work supported by the Deutsche Forschungsgemeinschaft under grants BR 2312/3-2, BR 2312/3-3, BR 2312/6-1, and FI 1664/1-1. We gratefully acknowledge the support of the TUM Graduate School's Faculty Graduate Center CeDoSIA at Technische Universität München, Germany. We further thank the anonymous reviewers for helpful comments.

References

1. K. Basu and J. Weibull. Strategy subsets closed under rational behavior. *Economics Letters*, 36:141–146, 1991.
2. D. Baumeister, F. Brandt, F. Fischer, J. Hoffmann, and J. Rothe. The complexity of computing minimal unidirectional covering sets. In *Proceedings of the 7th International Conference on Algorithms and Complexity (CIAC)*, number 6078 in Lecture Notes in Computer Science (LNCS), pages 299–310. Springer-Verlag, 2010.
3. B. Bernheim. Rationalizable strategic behavior. *Econometrica*, 52(4):1007–1028, 1984.
4. F. Brandt and F. Fischer. Computing the minimal covering set. *Mathematical Social Sciences*, 56(2):254–268, 2008.
5. F. Brandt, M. Brill, F. Fischer, and P. Harrenstein. Computational aspects of Shapley's saddles. In *Proceedings of the 8th International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*, pages 209–216, 2009.
6. F. Brandt, M. Brill, F. Fischer, and P. Harrenstein. On the complexity of iterated weak dominance in constant-sum games. *Theory of Computing Systems*. Forthcoming.
7. X. Chen, X. Deng, and S.-H. Teng. Settling the complexity of computing two-player Nash equilibria. *Journal of the ACM*, 56(3), 2009.
8. V. Conitzer and T. Sandholm. Complexity of (iterated) dominance. In *Proceedings of the 6th ACM Conference on Electronic Commerce (ACM-EC)*, pages 88–97. ACM Press, 2005.

9. C. Daskalakis, P. Goldberg, and C. Papadimitriou. The complexity of computing a Nash equilibrium. *SIAM Journal on Computing*, 39(1):195–259, 2009.
10. J. Duggan and M. Le Breton. Dutta's minimal covering set and Shapley's saddles. *Journal of Economic Theory*, 70:257–265, 1996.
11. B. Dutta. Covering sets and a new Condorcet choice correspondence. *Journal of Economic Theory*, 44:63–80, 1988.
12. I. Gilboa, E. Kalai, and E. Zemel. The complexity of eliminating dominated strategies. *Mathematics of Operations Research*, 18(3):553–565, 1993.
13. E. Hemaspaandra, L. Hemaspaandra, and J. Rothe. Exact analysis of Dodgson elections: Lewis Carroll's 1876 voting system is complete for parallel access to NP. *Journal of the ACM*, 44(6):806–825, 1997.
14. R. D. Luce and H. Raiffa. *Games and Decisions: Introduction and Critical Survey*. John Wiley & Sons Inc., 1957.
15. R. D. McKelvey and P. C. Ordeshook. Symmetric spatial games without majority rule equilibria. *The American Political Science Review*, 70(4):1172–1184, 1976.
16. R. B. Myerson. *Game Theory: Analysis of Conflict*. Harvard University Press, 1991.
17. J. F. Nash. Non-cooperative games. *Annals of Mathematics*, 54(2):286–295, 1951.
18. C. H. Papadimitriou. *Computational Complexity*. Addison-Wesley, 1994.
19. D. Pearce. Rationalizable strategic behavior and the problem of perfection. *Econometrica*, 52(4):1029–1050, 1984.
20. L. Samuelson. Dominated strategies and common knowledge. *Games and Economic Behavior*, 4:284–313, 1992.
21. L. Shapley. Order matrices. I. Technical Report RM-1142, The RAND Corporation, 1953.
22. L. Shapley. Order matrices. II. Technical Report RM-1145, The RAND Corporation, 1953.
23. L. Shapley. Some topics in two-person games. In M. Dresher, L. S. Shapley, and A. W. Tucker, editors, *Advances in Game Theory*, volume 52 of *Annals of Mathematics Studies*, pages 1–29. Princeton University Press, 1964.
24. J. von Neumann. Zur Theorie der Gesellschaftspiele. *Mathematische Annalen*, 100: 295–320, 1928.
25. J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.
26. K. Wagner. More complicated questions about maxima and minima, and some closures of NP. *Theoretical Computer Science*, 51:53–80, 1987.

A Proofs for Very Weak Saddles

As in the case of weak saddles we define, for each Boolean formula φ, a two-player game Γ_φ that admits certain types of very weak saddles if and only if φ is satisfiable. Let $\varphi = C_1 \wedge \cdots \wedge C_m$ be a 3-CNF formula[9] over variables $v_1, \ldots, v_n$, where $C_i = \{\ell_{i,1}, \ell_{i,2}, \ell_{i,3}\}$. Call a pair $\{\ell_{i,j}, \ell_{i',j'}\}$ of literal occurrences *conflicting*, and write $[\ell_{i,j}, \ell_{i',j'}]$, if $i \neq i'$ and $\ell_{i,j} = \overline{\ell_{i',j'}}$.

[9] A formula in 3-CNF is a CNF formula where every clause consists of exactly three literals. The problems SAT and UNSAT remain NP-hard and coNP-hard, respectively, even for this restricted class of formulas. While the construction works for arbitrary CNF formulas, we employ 3-CNFs for ease of notation.

	$\ell_{1,1}$	$\ell_{1,2}$	$\ell_{1,3}$	$\ell_{2,1}$	$\ell_{2,2}$	$\ell_{2,3}$	$\cdots$	$\ell_{m,1}$	$\ell_{m,2}$	$\ell_{m,3}$
C_1	1 / 0	1 / 0	1 / 0				$\cdots$	0 / 1	0 / 1	0 / 1
C_2	0 / 1	0 / 1	0 / 1	1 / 0	1 / 0	1 / 0	$\cdots$			
C_3				0 / 1	0 / 1	0 / 1	$\cdots$			
$\vdots$					$\vdots$					
C_m							$\cdots$	1 / 0	1 / 0	1 / 0
$[\ell_{1,1}, \ell_{2,3}]$	0 / 1					0 / 1	$\cdots$			
$[\ell_{1,2}, \ell_{2,1}]$		0 / 1		0 / 1			$\cdots$			
$[\ell_{1,1}, \ell_{m,1}]$	0 / 1						$\cdots$	0 / 1		
$\vdots$					$\vdots$					

Fig. 6 Game Γ_φ for a formula φ with $C_1 = v_1 \vee \overline{v}_2 \vee v_3$, $C_2 = \overline{v}_1 \vee v_2 \vee v_4$, and $C_m = \overline{v}_1 \vee \overline{v}_2 \vee v_4$. Payoffs are $(0, 0)$ unless specified otherwise.

Define the bimatrix game $\Gamma_\varphi = (A, B, p)$ as follows. The set A of actions of player 1 comprises the set $C = \{C_1, \dots, C_m\}$ of clauses of φ as well as one additional action for each conflicting pair $[\ell_{i,j}, \ell_{i',j'}]$ of literals.[10] The set B of actions of player 2 is the set of all literal occurrences, i.e., $B = \bigcup_{j=1}^{m} \{\ell_{j,1}, \ell_{j,2}, \ell_{j,3}\}$. Payoffs are given by

$$p(C_i, \ell_{p,q}) = \begin{cases} (0,1) & \text{if } p = i, \\ (1,0) & \text{if } p = i \bmod m + 1, \\ (0,0) & \text{otherwise,} \end{cases}$$

and

$$p([\ell_{i,j}, \ell_{i',j'}], \ell_{p,q}) = \begin{cases} (1,0) & \text{if } i = p \text{ and } j = q, \\ (1,0) & \text{if } i' = p \text{ and } j' = q, \\ (0,0) & \text{otherwise.} \end{cases}$$

An example of a game Γ_φ is shown in Figure 6.

Consider a very weak saddle (S_1, S_2) of Γ_φ. We will exploit the following three properties, which are easy consequences of Fact 1:

(I) If $C_i \in S_1$ for some $i \in [m]$, then $\ell_{i,j} \in S_2$ for some $j \in [3]$.
(II) If $\ell_{i,j} \in S_2$ for some $i \in [m]$ and $j \in [3]$, then $C_{i \bmod m+1} \in S_1$ or $[\ell_{i,j}, \ell_{i',j'}] \in S_1$ for some $i' \in [m]$ and $j' \in [3]$.
(III) For two conflicting literals $\ell_{i,j}$ and $\ell_{i',j'}$, we have $\{\ell_{i,j}, \ell_{i',j'}\} \rightsquigarrow [\ell_{i,j}, \ell_{i',j'}]$.

The idea underlying the definition of Γ_φ is formalized in the following lemma.

Lemma 3 *Γ_φ has a very weak saddle (S_1, S_2) with $S_1 = C$ if and only if φ is satisfiable.*

Proof For the direction from left to right, consider a saddle (S_1, S_2) with $S_1 = C$. By *(III)*, S_2 does not include any conflicting literals and thus defines a valid assignment for φ. Moreover,

[10] We identify $[\ell_{i,j}, \ell_{i',j'}]$ and $[\ell_{i',j'}, \ell_{i,j}]$ and thus have only one action per conflicting pair.

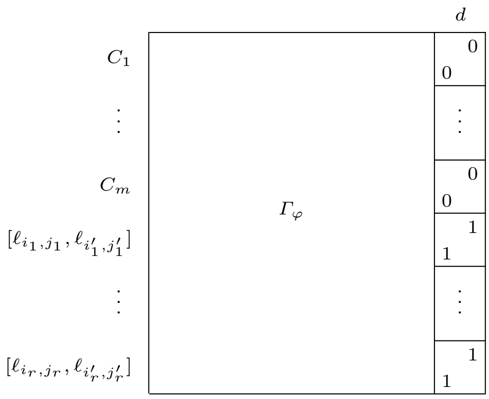

Fig. 7 Game Γ'_φ used in the proof of Lemma 4.

(I) ensures that $|\{\ell_{i,1}, \ell_{i,2}, \ell_{i,3}\} \cap S_2| \geq 1$ for each $i \in [m]$, which means that this assignment satisfies φ.

For the direction from right to left, let α be a satisfying assignment of φ and $f : [m] \to [3]$ a function such that $\ell_{i,f(i)} \in \alpha$ for all $i \in [m]$. It is then easily verified that $(C, \bigcup_{i=1}^{m}\{\ell_{i,f(i)}\})$ is a very weak saddle of Γ_φ. ☐

In the following we define two bimatrix games Γ'_φ and Γ''_φ that extend Γ_φ with new actions in such a way that properties *(I)*, *(II)*, and *(III)* still hold. In particular, statements similar to Lemma 3 will hold for Γ'_φ and Γ''_φ. The game Γ'_φ is then used to prove the NP-hardness of InVeryWeakSaddle, while Γ''_φ is used in the proofs of all other hardness results.

The game Γ'_φ, shown in Figure 7, is defined by adding a column d to Γ_φ. Payoffs for the new action profiles are defined as $p(C_i, d) = (0, 0)$ for all $i \in [m]$, and $p([\ell_{i,j}, \ell_{i',j'}], d) = (1, 1)$ for each conflicting pair.

Lemma 4 *Γ'_φ has a very weak saddle (S_1, S_2) with $C_1 \in S_1$ if and only if φ is satisfiable.*

Proof By Lemma 3, Γ_φ has a very weak saddle (S_1, S_2) with $S_1 = C$ if and only if φ is satisfiable. Since $p_2(C_i, d) = 0$ for all $i \in [m]$, this property still holds for Γ'_φ.

It remains to be shown that if (S_1, S_2) is a very weak saddle of Γ'_φ with $C_1 \in S_1$, then $S_1 = C$. If $C_1 \in S_1$, properties *(I)* and *(II)* imply that $C_i \in S_1$ for all $i \in [m]$. On the other hand, observe that $[\ell_{i,j}, \ell_{i',j'}] \rightsquigarrow d$ for every conflicting pair $[\ell_{i,j}, \ell_{i',j'}]$ and that $(\{[\ell_{i,j}, \ell_{i',j'}]\}, \{d\})$ is a very weak saddle. Obviously, this is the only very weak saddle containing $[\ell_{i,j}, \ell_{i',j'}]$. Therefore, a saddle containing C_1 does not contain any row that corresponds to a conflicting pair. ☐

To show the remaining hardness results, we define the bimatrix game Γ''_φ as another supergame of Γ_φ. In addition to the properties *(I)*, *(II)*, and *(III)*, Γ''_φ will have the following new property:

(IV) For every row $[\ell_{i,j}, \ell_{i,j'}]$ that corresponds to a conflicting pair, it is true that $[\ell_{i,j}, \ell_{i,j'}] \rightsquigarrow a$ for every action a of Γ''_φ.

Let r denote the number of conflicting pairs of φ and rename the actions of $\Gamma_\varphi = (A, B, p)$ in such a way that $A = \{a_1, \ldots, a_{m+r}\}$ with $C_i = a_i$ for all $i \in [m]$ and $B = \{b_1, \ldots, b_{3m}\}$. To obtain the game Γ''_φ shown in Figure 8, we augment Γ_φ by s additional columns $d_1, \ldots, d_s$ and s additional rows $e_1, \ldots, e_s$, where $s = \max(|A|, |B|) + 1$. Payoffs for new action profiles are defined as follows:

- $p(e_i, d_j) = (2, 0)$ if $j = i$,
- $p(e_i, d_j) = (0, 2)$ if $j = i \bmod s + 1$,
- $p(e_i, b_j) = (0, 1)$ if $i \in \{j, j + 1\}$,
- $p(a_i, d_j) = (1, 0)$ if $j \in \{i, i + 1\}$,

	b_1	b_2	$\cdots$	b_{3m}	d_1	d_2	d_3	d_4	$\cdots$	d_{s-1}	d_s
$a_1 = C_1$	1 / 0	1 / 0	$\cdots$	0 / 1	0 / 1	0 / 1			$\cdots$		
$a_2 = C_2$	0 / 1	0 / 1	$\cdots$			0 / 1	0 / 1		$\cdots$		
$a_3 = C_3$			$\cdots$				0 / 1	0 / 1	$\cdots$		
$\vdots$					$\vdots$						
$a_m = C_m$			$\cdots$	1 / 0					$\cdots$		
$a_{m+1} = [\ell_{i_1,j_1}, \ell_{i'_1,j'_1}]$		0 / 1	$\cdots$		1 / 0				$\cdots$		
$a_{m+2} = [\ell_{i_2,j_2}, \ell_{i'_2,j'_2}]$	0 / 1		$\cdots$		1 / 0				$\cdots$		
$\vdots$					$\vdots$						
$a_{m+r} = [\ell_{i_r,j_r}, \ell_{i'_r,j'_r}]$			$\cdots$		1 / 0				$\cdots$	0 / 1	0 / 1
e_1	1 / 0		$\cdots$		0 / 2	2 / 0			$\cdots$		
e_2	1 / 0	1 / 0	$\cdots$			0 / 2	2 / 0		$\cdots$		
e_3		1 / 0	$\cdots$				0 / 2	2 / 0	$\cdots$		
e_4			$\cdots$					0 / 2	$\cdots$		
$\vdots$					$\vdots$						
e_s		$\cdots$			2 / 0				$\cdots$		0 / 2

Fig. 8 Game Γ''_φ used in the proof of Theorem 2. Payoffs are $(0, 0)$ unless specified otherwise.

- $p([\ell_{i,j}, \ell_{i',j'}], d_1) = (0, 1)$ for all conflicting pairs, and
- $p(a, b) = (0, 0)$ otherwise.

Note that Γ''_φ satisfies properties *(I)*, *(II)*, and *(III)*, since $p_2(C_i, d_j) = 0$ for all $i \in [m]$ and $p_1(e_i, b) = 0$ for all $b \in B$. We can thus prove the following lemma analogously to Lemma 3.

Lemma 5 *Γ''_φ has a very weak saddle (S_1, S_2) with $S_1 = C$ if and only if φ is satisfiable.*

In order to prove that Γ''_φ satisfies property *(IV)*, note that $[\ell_{i,j}, \ell_{i',j'}] \rightsquigarrow d_1$ for every conflicting pair $[\ell_{i,j}, \ell_{i',j'}]$. Furthermore, we have $d_i \rightsquigarrow e_i$ for every $i \in [s]$, and $e_i \rightsquigarrow d_{i+1}$ for every $i \in [s-1]$. It therefore follows from the transitivity of $\rightsquigarrow$ that $[\ell_{i,j}, \ell_{i',j'}] \rightsquigarrow d_k$ and $[\ell_{i,j}, \ell_{i',j'}] \rightsquigarrow e_k$ for every $[\ell_{i,j}, \ell_{i',j'}] \in A$ and every $k \in [s]$. Finally, by construction, $\{d_i, d_{i+1}\} \rightsquigarrow a_i$ for all $i \in \{1, \ldots, |A|\}$, and $\{e_i, e_{i+1}\} \rightsquigarrow b_i$ for all $i \in \{1, \ldots, |B|\}$. Since $s > \max(|A|, |B|)$, this implies *(IV)*.

Lemma 6 *Γ''_φ has a nontrivial very weak saddle if and only if φ is satisfiable.*

Proof If φ is satisfiable, then by Lemma 5 there exists a nontrivial very weak saddle.

Conversely assume that φ is not satisfiable. Then by *(IV)* there is no nontrivial saddle (S_1, S_2) with $[\ell_{i,j}, \ell_{i',j'}] \in S_1$ for a conflicting pair $[\ell_{i,j}, \ell_{i',j'}]$. By Lemma 5, there is no saddle (S_1, S_2) with $S_1 = C$. Furthermore, it follows from *(I)*, *(II)*, and *(IV)* that there cannot be a saddle (S_1, S_2) with $S_1 \subset C$. It remains to show that a nontrivial very weak saddle cannot contain any of the new actions e_i or d_j. As mentioned above, $d_i \rightsquigarrow e_i$ and $e_i \rightsquigarrow d_{i \bmod s+1}$ for all $i \in [s]$. Hence we can conclude—analogously to the proof of *(IV)*—that $d_i \rightsquigarrow a$ and $e_i \rightsquigarrow a$ for every action a. Thus, d_i and e_i cannot be part of a nontrivial saddle for any $i \in [s]$. □

We are now ready to prove Theorem 2.

Theorem 2 *The following hold:*

(i) InVeryWeakSaddle *is NP-hard.*
(ii) InVeryWeakSaddle *is coNP-hard.*
(iii) IsVeryWeakSaddle *is coNP-complete.*
(iv) InAllVeryWeakSaddles *is coNP-complete.*
(v) UniqueVeryWeakSaddle *is coNP-hard.*
(vi) NontrivialVeryWeakSaddle *is NP-complete.*
(vii) FindVeryWeakSaddle *is NP-hard under Turing reductions.*

All hardness results hold even for two-player games.

Proof Let φ be a Boolean formula and let Γ'_φ and Γ''_φ be the games defined above.

(i) NP-hardness of InVeryWeakSaddle can be shown by a reduction from 3-SAT. Lemma 4 shows that Γ'_φ has a very weak saddle containing C_1 if and only if φ is satisfiable.

(ii) coNP-hardness of InVeryWeakSaddle can be shown by a reduction from 3-UNSAT. Consider the game Γ''_φ and assume without loss of generality that φ has at least one pair of conflicting literals. It follows from property *(IV)* and Lemma 6 that each row that corresponds to a conflicting pair is contained in a very weak saddle of Γ''_φ, namely the trivial one, if and only if φ is unsatisfiable.

(iii) A minor modification of the coNP algorithm for IsWeakSaddle shows that IsVeryWeakSaddle is in coNP. We show coNP-hardness by a reduction from 3-UNSAT. It follows from Lemma 6 that the set of all actions of Γ''_φ is a very weak saddle if and only if φ is unsatisfiable.

(iv) The proof of coNP-membership of InAllVeryWeakSaddles is similar to the proof of coNP-membership of InAllWeakSaddles. Hardness follows from the same argument as in *(ii)*.

(v) coNP-hardness of UniqueVeryWeakSaddle can be shown by a reduction from 3-UNSAT. Consider the game Γ''_φ and assume without loss of generality that φ has either none or more than one satisfying assignment. Then, if φ is satisfiable, Γ''_φ has multiple very weak saddles, each of them corresponding to a particular satisfying assignment. If on the other hand φ is unsatisfiable, Γ''_φ has only the trivial very weak saddle.

(vi) The proof of NP-membership of NontrivialVeryWeakSaddle is similar to the proof of NP-membership of NontrivialWeakSaddle. NP-hardness of the problem follows from a reduction from 3-SAT. Lemma 6 shows that Γ''_φ has a nontrivial very weak saddle if and only if φ is satisfiable.

(vii) NP-hardness of FindVeryWeakSaddle can be shown in the same way as that of FindWeakSaddle.

□