On Strategy Improvement Algorithms for Simple Stochastic Games[☆]

Rahul Tripathi[a,*], Elena Valkanova[a], V. S. Anil Kumar[b]

[a] *Department of Computer Science and Engineering, University of South Florida, Tampa, FL 33620, USA*
[b] *Department of Computer Science and Virginia Bioinformatics Institute, Virginia Tech., Blacksburg, VT 24061, USA*

Abstract

The study of *simple stochastic games* (SSGs) was initiated by Condon for analyzing the computational power of randomized space-bounded alternating Turing machines. The game is played by two players, MAX and MIN, on a directed multigraph, and when the play terminates at a sink vertex s, MAX wins from MIN a payoff $p(s) \in [0, 1]$. Condon proved that the problem SSG-VALUE—given a SSG, determine whether the expected payoff won by MAX is greater than $1/2$ when both players use their optimal strategies—is in NP $\cap$ coNP. However, the exact complexity of this problem remains open, as it is not known whether the problem is in P or is hard for some natural complexity class. In this paper, we study the computational complexity of a strategy improvement algorithm by Hoffman and Karp for this problem. The Hoffman-Karp algorithm converges to optimal strategies of a given SSG, but no nontrivial bounds were previously known on its running time. We prove a bound of $O(2^n/n)$ on the convergence time of the Hoffman-Karp algorithm, and a bound of $O(2^{0.78n})$ on a randomized variant. These are the first non-trivial upper bounds on the convergence time of these strategy improvement algorithms.

Keywords: algorithms, computational complexity, stochastic games, strategy improvement algorithms, optimal strategies

1. Introduction

Stochastic games, first studied by Shapley in 1953 [Sha53], are two-player graphical games that arise in a number of applications, including computational complexity theory, game theory, operations research, automated software verification, and reactive systems. Several variations of stochastic games have been studied, and an interesting restriction of this game model is the class of *simple stochastic games* (SSGs). Condon [Con92] initiated the study of SSGs for analyzing the computational power of randomized space-bounded alternating Turing machines. The game is played by two players, MAX and MIN, on a game board that is a directed multigraph $G = (V, E)$. The vertex set V of G is partitioned into disjoint subsets V_{MAX}, V_{MIN}, V_{AVE}, and V_{SINK}. Here,

[☆] A preliminary version of this paper appeared in *Proceedings of the 7th International Conference on Algorithms and Complexity* (2010) [TVK10].

[*] Corresponding author

Email addresses: tripathi@cse.usf.edu (Rahul Tripathi), evalkanova75@yahoo.com (Elena Valkanova), vsakumar@vt.edu (V. S. Anil Kumar)

V_{MAX} is the set of vertices controlled by MAX, V_{MIN} is the set of vertices controlled by MIN, and V_{AVE} is the set of vertices that allow stochastic transitions (edges). The vertices in V_{SINK}, called sink vertices, have zero outdegree, and each sink vertex x has a rational payoff $p(x) \in [0, 1]$.

To play a SSG, a token (or pebble) is initially placed on a designated *start* vertex. At each step of the play, the token is moved along a directed edge that is determined by the current position (i.e., vertex) of the token and the players' strategies (i.e., choices of an outgoing edge from vertices controlled by the players). Suppose that the token is on a vertex x. If $x \in V_{\text{MAX}}$ ($x \in V_{\text{MIN}}$), then the token is moved onto the end-vertex of the directed edge chosen by MAX (respectively, MIN); if $x \in V_{\text{AVE}}$, then the token is moved onto the end-vertex of a randomly chosen outgoing edge from x. The play stops when the token reaches a sink vertex s; at this point, MAX wins from MIN the payoff $p(s)$ associated with s (see Figure 1 for an example of a SSG). The goal of MAX is to maximize the expected payoff won from MIN, while the goal of MIN is to minimize it; see Section 2 for a formal description. The problem of solving a SSG, called SSG-VALUE, is to determine whether the expected payoff won by MAX is greater than $1/2$ when both players use their optimal strategies.

The problem SSG-VALUE is equivalent to the function problem STABLE-CIRCUIT [Jub05] in which the input consists of a circuit C rather than a two-player graphical game. The input circuit C is made of gates MAX, MIN, and AVG such that each gate has fan-in two and the output of a gate is allowed to be an input of a previous gate (i.e., internal feedbacks are allowed). The circuit C takes only two inputs, 0-input and 1-input, whose values are hard-wired to 0 and 1, respectively. Here, the gates of C are implicitly assumed to be ordered from 1 to $|C|$, where $|C|$ denotes the total number of gates in C. The output of the gate MAX is defined by $\text{MAX}(x, y) = \max\{x, y\}$, of the gate MIN is defined by $\text{MIN}(x, y) = \min\{x, y\}$, and of the gate AVG is defined by $\text{AVG}(x, y) = (x + y)/2$. Such a circuit is referred to as a MIN/MAX/AVG circuit.

A *circuit value* of a MIN/MAX/AVG circuit C is a vector $\vec{v} \in [0, 1]^{|C|}$ of $|C|$ values, where the i'th component $v_i \in [0, 1]$ is an assignment for the i'th gate of C. For convenience, assume that any circuit value $\vec{v}$ assigns the 0-input to 0 and the 1-input to 1. Given such a circuit C, a *circuit-wide* update function $F_C : [0, 1]^{|C|} \to [0, 1]^{|C|}$ is defined that takes a circuit value $\vec{v}$ and transforms it to another circuit value $F_C(\vec{v}) = \vec{w}$ as follows: For every $1 \le i \le |C|$, the component w_i of $\vec{w}$ is the maximum, minimum, or the average of the values assigned by $\vec{v}$ to the inputs of i'th gate of C when this gate is a MAX, MIN, or AVG, respectively. A circuit value $\vec{v}$ is said to be a *stable solution* if $F_C(\vec{v}) = \vec{v}$, i.e., if $\vec{v}$ is a fixed point of F_C. It is known that there always exists a stable solution of any MIN/MAX/AVG circuit C [Jub05] and that there is a unique *minimum stable solution* $\vec{m}$ of C that is minimum in the sense that, for any other stable solution $\vec{v}$ of C and for any index $1 \le i \le |C|$, it holds that $m_i \le v_i$. The function problem STABLE-CIRCUIT asks: given a MIN/MAX/AVG circuit C, find the minimum stable solution of C. The problems SSG-VALUE and STABLE-CIRCUIT are known to be polynomial-time Turing equivalent. Refer to [Poo03, Jub05] for more discussion on STABLE-CIRCUIT.

Condon [Con92] proved that SSG-VALUE is complete for the class of languages accepted by *logspace randomized alternating Turing machines* and that it belongs to the class NP ∩ coNP. Despite considerable interest, the complexity of SSG-VALUE is not fully resolved as it is unknown whether the problem is in P or is hard for some natural complexity class. The best known algorithms for SSG-VALUE are subexponential-time randomized algorithms of Ludwig [Lud95] and Halman [Hal07]. This puts SSG-VALUE among a small list of natural combinatorial problems in NP ∩ coNP that are not yet known to be in P; the sub-exponential upper bound makes this problem rarer still. Moreover, some other game problems, such as Parity Games (PGs), Mean Payoff Games (MPGs), and Discounted Payoff Games (DPGs), polynomial-time reduce to SSG-VALUE [Pur95, GW02, ZP96], and thus a polynomial-time algorithm for SSG-VALUE would imply a polynomial-time algorithm for all these games problems. PGs play an important role in verification and automata theory, while MPGs are useful in the design and analysis of algorithms for problems related to online job scheduling, finite-window online string matching, and selection with limited storage [ZP96].

Apart from the results in [Lud95, Hal07], very few additional upper bound results are known for SSG-VALUE. Condon [Con92] showed that restricted versions of SSG-VALUE consisting of only two classes of vertices (out of V_{MAX}, V_{MIN} and V_{AVE}) can be solved in polynomial time. This was extended by Gimbert et al. [GH09] who developed a fixed parameter tractable algorithm in terms of $|V_{\text{AVE}}|$, which runs in time $O(|V_{\text{AVE}}|! \cdot \text{poly}(n))$ where n is the size (in bits) of the input game. There have been numerous other algorithms for this problem based on the general approach of *strategy improvement*, which involves switching the choices of vertices that are not locally optimal. One well-studied strategy improvement algorithm is the Hoffman-Karp algorithm [HK66] (described in Section 3) in which choices of *all* locally non-optimal vertices are switched in each iteration until optimal strategies are found. Condon [Con92] showed that the Hoffman-Karp algorithm does converge to the optimum, though its convergence time is still not very well understood. The specific details of such strategy improvement algorithms are important, and as discussed by Condon [Con93], variants of Hoffman-Karp and other seemingly natural heuristics do not converge to the optimum.

The focus of this paper is on understanding the convergence time of the Hoffman-Karp algorithm. Let n denote $\min\{|V_{\text{MAX}}|, |V_{\text{MIN}}|\}$. We show that the convergence time of the Hoffman-Karp algorithm is $O(2^n/n)$. This is the first non-trivial upper bound on the convergence time of the Hoffman-Karp algorithm. We also consider a randomized variant of this algorithm, and show that the convergence time of the randomized algorithm is $O(2^{0.78n})$. While these bounds are still exponential, they represent an improved understanding of these strategy improvement algorithms. Our analyses extend those of Mansour and Singh [MS99] for policy iteration algorithms for Markov decision processes, which are special classes of stochastic games with a single player.

2. Preliminaries

We now discuss the basic concepts and notations needed for the rest of the paper. We follow the definitions presented in [Som05] for the most part.

Definition 2.1 (simple stochastic games). *A SSG is a two-player game between players* MAX *and* MIN*. The game is played on a* game board *that is a directed multigraph $G = (V, E)$. The vertex set V is partitioned into disjoint subsets V_{MAX}, V_{MIN}, V_{AVE}, and* V_{SINK}*. All vertices of G, except those of* V_{SINK}*, have exactly two outgoing edges. The vertices belonging to* V_{SINK} *have no outgoing edge. The vertex set $V_{\text{POS}} =_{df} V_{\text{MAX}} \cup V_{\text{MIN}} \cup V_{\text{AVE}}$ represents the set of all game positions and the edge set E denotes a possible move in the game. One vertex from V_{POS} is called* start vertex*. We call $x \in V$ a* MAX-position *if $x \in V_{\text{MAX}}$, a* MIN-position *if $x \in V_{\text{MIN}}$, an* AVE-position *if $x \in V_{\text{AVE}}$, and a* sink *if $x \in \text{V}_{\text{SINK}}$. Each sink x has a rational* payoff $p(x) \in [0, 1]$*. For every $x \in V_{\text{AVE}}$, every edge leaving x is labeled with a rational probability such that the sum of probabilities over all edges leaving x is one; the probability associated with an edge $(x, y) \in E$, for $x \in V_{\text{AVE}}$, is denoted by $q(x, y)$. The outgoing edges from any vertex $x \in V_{\text{MAX}} \cup V_{\text{MIN}}$ are unlabeled.*

In what follows, we refer to a SSG by the name of its game board, i.e., the multigraph G. Thus, saying that G is a SSG means that the SSG is defined on a game board G that is a directed multigraph. We denote the game G starting at vertex x by $G(x)$.

Definition 2.1 generalizes the usual definition of SSGs in the papers [Con92, Con93, Lud95, Hal07] in that the former definition allows sink vertices to have rational payoffs whereas the latter considers only two types of sink vertices, 0-sink and 1-sink, but no payoffs. The 0-sink and the 1-sink vertices in the usual definition correspond to the sink vertex with payoff 0 and the sink vertex with payoff 1, respectively, in our definition. However, as discussed by Condon [Con92], the problem SSG-VALUE corresponding to these two definitions of SSGs are polynomial-time equivalent.

A strategy of a player specifies the choices made by the player during a play of the game. A strategy is *pure* if the choices are made in a deterministic manner, and is *mixed* if the choices are made according to some probability distribution. A pure strategy may depend on time, and, more generally, it may depend on the history of a play. On the other hand, a *stationary* (or *memoryless*) strategy depends only on the current state of a play, and is independent of both time and history. Condon [Con92] showed that both MAX and MIN have *optimal strategies* (Definition 2.7) that are pure and stationary. Therefore, in this paper, we consider only pure, stationary strategies of the two players.

Definition 2.2 (strategies). *A strategy σ of* MAX *is a function $\sigma : V_{\mathrm{MAX}} \to V$ such that, for every $x \in V_{\mathrm{MAX}}$, we have $(x, \sigma(x)) \in E$. Similarly, a strategy τ of* MIN *is a function $\tau : V_{\mathrm{MIN}} \to V$ such that, for every $x \in V_{\mathrm{MIN}}$, we have $(x, \tau(x)) \in E$.*

We sometimes represent any MAX-strategy σ by the set $\{(x, y) \mid x \in V_{\mathrm{MAX}} \text{ and } y = \sigma(x)\}$ and represent any MIN-strategy τ by the set $\{(x, y) \mid x \in V_{\mathrm{MIN}} \text{ and } y = \tau(x)\}$. The rules of playing a SSG according to a MAX-strategy σ and a MIN-strategy τ are formally described in Definition 2.3.

Definition 2.3 (a play of a SSG). *Let G be a SSG, σ be a* MAX-*strategy, τ be a* MIN-*strategy, and $x \in V_{\mathrm{POS}}$ be a game position. Before we begin playing $G(x)$, a token is placed on x. At each step of the play, the token is moved from current vertex $y \in V$ to a neighboring vertex according to the following rules:*

- *If $y \in V_{\mathrm{MAX}}$, then* MAX *takes a turn and moves the token from y to $\sigma(y)$.*
- *If $y \in V_{\mathrm{MIN}}$, then* MIN *takes a turn and moves the token from y to $\tau(y)$.*
- *If $y \in V_{\mathrm{AVE}}$, then none of the players take any turn. Instead, the token is moved from y to a neighboring vertex chosen randomly from a distribution $q(y, \cdot)$ over the neighbors of y.*

If the current vertex $y \in \mathrm{V}_{\mathrm{SINK}}$, then the play stops. At this point, MAX *wins a* payoff $p(y) \in [0, 1]$ *from* MIN (*equivalently,* MIN *looses $p(y)$ to* MAX). *A* play *of $G(x)$ is, therefore, defined to be a maximal path starting from x that the token takes.*

For any strategy α of a player $P \in \{\mathrm{MAX}, \mathrm{MIN}\}$, we say that a play x_0, x_1, x_2, $\ldots$ of $G(x_0)$ confirms *to α if, for every $x_i \in V_P$, we have $x_{i+1} = \alpha(x_i)$.*

Because the move from any AVE-position is probabilistic, the payoff won by MAX is a random variable in the range $[0, 1]$. The objective of MAX is to choose a strategy σ that maximizes the expected payoff, while the objective of MIN is to choose a strategy τ that minimizes the expected payoff.

For every choice of start vertex x and strategies σ and τ of the players MAX and MIN, respectively, the expected payoff $v_{\sigma,\tau}(x)$ that MAX wins is defined as the weighted sum of payoffs $p(s)$ over all sink vertices s, where the weight contributed by a sink vertex s is the probability that a play of $G(x)$ confirming to σ and τ stops at s. For any strategies σ, τ of the players, the *expected payoff vector* $v_{\sigma,\tau}$ contains $v_{\sigma,\tau}(x)$ for every $x \in V_{\text{POS}}$.

Definition 2.4 (expected payoffs). *Let σ and τ be strategies of the players* MAX *and* MIN*, respectively. Let $q_{\sigma,\tau}(x, s)$ denote the probability that a play of $G(x)$ confirming to σ and τ stops at a sink vertex s. The* expected payoff vector $v_{\sigma,\tau} : V_{\text{POS}} \to [0, 1]$ *of G corresponding to σ and τ is defined as follows: For every $x \in V_{\text{POS}}$,*

$$v_{\sigma,\tau}(x) = \sum_{s \in \mathrm{V}_{\text{SINK}}} q_{\sigma,\tau}(x, s) \cdot p(s).$$

It is implicit in the above definition that $v_{\sigma,\tau}(x) = 0$ if no play of $G(x)$ confirming to σ and τ is finite. A *stopping* SSG G is a SSG in which starting at any initial position x, any play of $G(x)$ stops at a sink vertex with probability one regardless of the strategies used by the players. Stopping SSGs are known to have certain desirable properties (e.g., the existence of a unique optimal value vector). Condon [Con92] showed that there is a polynomial-time transformation that, given any SSG G, constructs a new stopping SSG G' such that G' is as good as G for studying the problem SSG-VALUE. (See Lemma 2.24 for a precise statement relating the two SSGs.) Since the focus of this paper is on studying algorithms for the problem SSG-VALUE, we will henceforth restrict our attention to only stopping SSGs.

Definition 2.5 (stopping SSGs). *A SSG G is said to be a* stopping *SSG if, for every* MAX-*strategy σ and* MIN-*strategy τ and for any position x, there is a play of $G(x)$ confirming to σ and τ that stops at a sink vertex* (*in other words, with probability one, any play of $G(x)$ confirming to σ and τ is finite*).

Corresponding to any MAX-strategy σ and MIN-strategy τ in a SSG G, there is a multigraph $G_{\sigma,\tau}$ defined as follows: $G_{\sigma,\tau}$ is obtained from G by removing from each MAX-position x the outgoing edge (x, z) for which $z \neq \sigma(x)$ and by removing from each MIN-position y the outgoing edge (y, w) for which $w \neq \tau(y)$. Thus, $G_{\sigma,\tau}$ has exactly one outgoing edge from each MAX-position and each MIN-position. For every position x in G and $G_{\sigma,\tau}$, any play of $G(x)$ confirming to σ and τ can be considered as an equivalent play in $G_{\sigma,\tau}$ starting at x. The only distinct point is that whenever the token is at $y \in V_{\text{MAX}} \cup V_{\text{MIN}}$, the next move in G is determined by a player strategy whereas the next move in $G_{\sigma,\tau}$ is onto the end-vertex of the unique outgoing edge from y.

Figure 1(a) shows an example of a SSG G. Figure 1(b) shows the multigraph $G_{\sigma,\tau}$ corresponding to the MAX-strategy $\sigma = \{(1, 2)\}$ and MIN-strategy $\tau = \{(2, 4), (3, 2)\}$. For this choice of player strategies, any play of $G(1)$ stops at the sink vertex 6 with probability one and the expected payoff vector $v_{\sigma,\tau}$ is $(2/3, 2/3, 2/3, 2/3, 2/3)$. Figure 1(c) shows the multigraph $G_{\sigma,\tau}$ corresponding to the MAX-strategy $\sigma = \{(1, 2)\}$ and MIN-strategy $\tau = \{(2, 3), (3, 2)\}$. For this choice of player strategies, no play of $G(1)$ is finite and the expected payoff vector $v_{\sigma,\tau}$ is $(0, 0, 0, 2/3, 1/3)$. Therefore, the SSG in Figure 1(a) is a *non-stopping* SSG. An example of a stopping SSG is shown in Figure 2.

Given any strategy τ of MIN, a best response (i.e., optimal) strategy $\sigma = \sigma(\tau)$ of MAX w.r.t. τ, if it exists, is one that, for every game position $x \in V_{\text{POS}}$, assures the *maximum* payoff

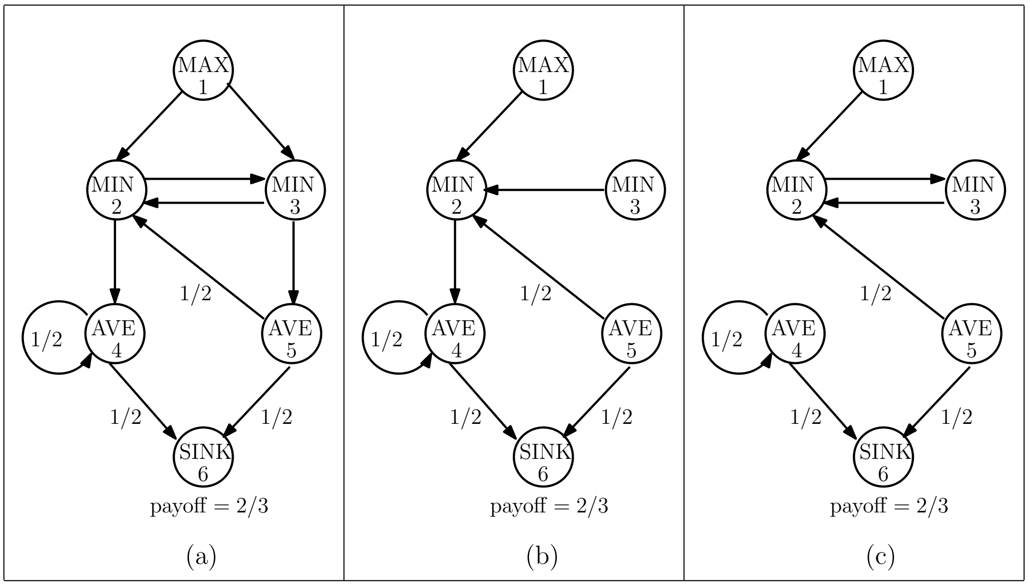

Figure 1: (a) A *non-stopping* SSG G with five game positions and one sink vertex. The game positions are vertex 1 (MAX-position), vertices 2 and 3 (MIN-positions), and vertices 4 and 5 (AVE-positions). The only sink vertex, 6, has payoff 2/3. The probability associated with each edge leaving some AVE-position is 1/2. (b) the multigraph $G_{\sigma,\tau}$ corresponding to $\sigma = \{(1,2)\}$ and MIN-strategy $\tau = \{(2,4),(3,2)\}$. (c) the multigraph $G_{\sigma,\tau}$ corresponding to $\sigma = \{(1,2)\}$ and MIN-strategy $\tau = \{(2,3),(3,2)\}$.

for MAX over all choices of MAX-strategies, i.e., $v_{\sigma,\tau}(x) = \max_{\sigma'} v_{\sigma',\tau}(x)$. Similarly, given any strategy σ of MAX, a best response strategy $\tau = \tau(\sigma)$ of MIN w.r.t. σ, if it exists, is one that, for every game position $x \in V_{\text{POS}}$, assures the *minimum* payoff for MAX over all choices of MIN-strategies, i.e., $v_{\sigma,\tau}(x) = \min_{\tau'} v_{\sigma,\tau'}(x)$. Howard [How60] showed that such strategies always exist for a stopping SSG, and therefore, $\sigma(\tau)$ and $\tau(\sigma)$ are well defined in this case. Derman's [Der70] LP formulation can be used for constructing these best response strategies for any stopping SSG in polynomial time.

Definition 2.6 (best response strategies). *A strategy σ of* MAX *is said to be* optimal with respect to *a strategy τ of* MIN *if, for every $x \in V_{\text{POS}}$,*

$$v_{\sigma,\tau}(x) = \max_{\sigma'} v_{\sigma',\tau}(x).$$

Similarly, a strategy τ of MIN *is said to be* optimal with respect to *a strategy σ of* MAX *if, for every $x \in V_{\text{POS}}$,*

$$v_{\sigma,\tau}(x) = \min_{\tau'} v_{\sigma,\tau'}(x).$$

Strategies σ and τ are *optimal strategies* if each is a best response strategy w.r.t. the other. Condon [Con92] showed that every stopping SSG has optimal strategies.

Definition 2.7 (optimal strategies). *Let σ and τ be strategies of* MAX *and* MIN*, respectively. Strategies σ and τ are* optimal at $x \in V_{\text{POS}}$ *if, for any strategy σ' of* MAX *and for any strategy*

τ' *of* MIN*, it holds that* $v_{\sigma',\tau}(x) \leq v_{\sigma,\tau}(x) \leq v_{\sigma,\tau'}(x)$*. Strategies* σ *and* τ *are* optimal *if they are optimal at every* $x \in V_{\mathrm{POS}}$*.*

The expected payoff vector corresponding to a pair of optimal strategies is called an *optimal value vector*. Condon [Con92] showed that every stopping SSG has an optimal value vector, which is *unique*.

Definition 2.8 (optimal value vector). *For any pair of optimal strategies* σ, τ *at* $x \in V_{\mathrm{POS}}$*, the expected payoff* $v_{\mathrm{opt}}(x) =_{df} v_{\sigma,\tau}(x)$ *is called an* optimal value *of* $G(x)$*. If, for every* $x \in V_{\mathrm{POS}}$*, there exist optimal strategies at x, then a vector* $v_{\mathrm{opt}} : V_{\mathrm{POS}} \to [0,1]$ *of optimal values is called an* optimal value vector *of G.*

We call any mapping $v : V_{\mathrm{POS}} \to [0,1]$ a *value vector*. Sometimes, we extend the domain of such a value vector v to $\mathrm{V_{SINK}}$ and define the mapping of any sink vertex s as the payoff $p(s) \in [0,1]$. We use $\overline{v}$ to denote this extension of v. For any MAX-strategy σ and MIN-strategy τ, this extension of a value vector $v_{\sigma,\tau}$ is denoted by $\overline{v}_{\sigma,\tau}$.

Given a value vector v, we define *v-stable* and *v-switchable* positions in Definitions 2.9 and 2.10, respectively. Notice that only a position $x \in V_{\mathrm{MAX}} \cup V_{\mathrm{MIN}}$ can possibly be a v-switchable position and that no position can both be v-switchable and v-stable.

Definition 2.9 (stable positions). *Let* $v : V_{\mathrm{POS}} \to [0,1]$ *be a value vector. For any* $x \in V_{\mathrm{POS}}$*, we say that x is v-*stable *if either* $x \in V_{\mathrm{MAX}}$ *and* $v(x) = \max\{\overline{v}(y) \mid (x,y) \in E\}$*, or* $x \in V_{\mathrm{MIN}}$ *and* $v(x) = \min\{\overline{v}(y) \mid (x,y) \in E\}$*, or* $x \in V_{\mathrm{AVE}}$ *and* $v(x) = \sum_{(x,y)\in E} q(x,y) \cdot \overline{v}(y)$*.*

The *v-switchable* positions, defined in Definition 2.10, will be useful in constructing an improved value vector from v (see Lemma 3.5).

Definition 2.10 (switchable positions). *Let* $v : V_{\mathrm{POS}} \to [0,1]$ *be a value vector. For any* $x \in V_{\mathrm{MAX}} \cup V_{\mathrm{MIN}}$*, we say that x is v-*switchable *if either* $x \in V_{\mathrm{MAX}}$ *and* $v(x) < \max\{\overline{v}(y) \mid (x,y) \in E\}$*, or* $x \in V_{\mathrm{MIN}}$ *and* $v(x) > \min\{\overline{v}(y) \mid (x,y) \in E\}$*.*

Given a value vector v, we say that a player strategy is *v-greedy* if the strategy makes locally optimal choice w.r.t. v at every position of the player.

Definition 2.11 (greedy strategies). *Let* $v : V_{\mathrm{POS}} \to [0,1]$ *be a value vector. A* MAX*-strategy* σ *is said to be v-*greedy at $x \in V_{\mathrm{MAX}}$ *if* $\overline{v}(\sigma(x)) = \max\{\overline{v}(y) \mid (x,y) \in E\}$*. Similarly, a* MIN*-strategy* τ *is said to be v-*greedy at $x \in V_{\mathrm{MIN}}$ *if* $\overline{v}(\tau(x)) = \min\{\overline{v}(y) \mid (x,y) \in E\}$*. For* $P \in \{\mathrm{MAX}, \mathrm{MIN}\}$*, a strategy of P is said to be v-*greedy *if it is v-greedy at each* $x \in V_P$*.*

Condon [Con92] introduced an operator F_G corresponding to any SSG G. This operator allows to give an alternate characterization of an optimal value vector of G (see Lemmas 2.18 and 2.19).

Definition 2.12 ([Con92]). *Let* σ *and* τ *be strategies of* MAX *and* MIN*, respectively. An operator* $F_G : (V_{\mathrm{POS}} \to [0,1]) \to (V_{\mathrm{POS}} \to [0,1])$ *is defined as follows: For every* $v : V_{\mathrm{POS}} \to [0,1]$*, we have* $F_G(v) = w$ *such that, for every* $x \in V_{\mathrm{POS}}$*,*

$$w(x) = \begin{cases} \max\{\overline{v}(y) \mid (x,y) \in E\} & \text{if } x \in V_{\mathrm{MAX}}, \\ \min\{\overline{v}(y) \mid (x,y) \in E\} & \text{if } x \in V_{\mathrm{MIN}}, \\ \sum_{(x,y)\in E} q(x,y) \cdot \overline{v}(y) & \text{if } x \in V_{\mathrm{AVE}}. \end{cases}$$

Given any MAX-strategy σ and MIN-strategy τ, we can associate an operator $F_{\sigma,\tau}$ defined below. This operator allows to give an alternate characterization of the expected payoff vector $v_{\sigma,\tau}$ (see Corollary 2.15).

Definition 2.13 ([Som05]). *Let σ and τ be strategies of* MAX *and* MIN*, respectively. Corresponding to σ and τ, we define an operator $F_{\sigma,\tau} : (V_{\mathrm{POS}} \to [0,1]) \to (V_{\mathrm{POS}} \to [0,1])$ as follows: For every $v : V_{\mathrm{POS}} \to [0,1]$, we have $F_{\sigma,\tau}(v) = w$ such that, for every $x \in V_{\mathrm{POS}}$,*

$$w(x) = \begin{cases} \overline{v}(\sigma(x)) & \text{if } x \in V_{\mathrm{MAX}}, \\ \overline{v}(\tau(x)) & \text{if } x \in V_{\mathrm{MIN}}, \\ \sum_{(x,y)\in E} q(x,y) \cdot \overline{v}(y) & \text{if } x \in V_{\mathrm{AVE}}. \end{cases}$$

Lemma 2.14 shows that, given strategies σ and τ of the two players in a stopping SSG, the expected payoff vector $v_{\sigma,\tau}$ is the unique solution to a system of linear equations, and so it can be computed in polynomial time.

Lemma 2.14 ([Con92]). *Let G be a stopping SSG, and let σ and τ be strategies of* MAX *and* MIN*, respectively. Then there are a $|V_{\mathrm{POS}}| \times |V_{\mathrm{POS}}|$ matrix $Q_{\sigma,\tau}$ with entries in $A =_{df} \{0,1\} \cup \{q(x,y) \mid x \in V_{\mathrm{AVE}}$ and $(x,y) \in E\}$ and a $|V_{\mathrm{POS}}|$-vector $b_{\sigma,\tau}$ with entries in $\{a \cdot p(x) \mid a \in A$ and $x \in \mathrm{V}_{\mathrm{SINK}}\}$ such that $v_{\sigma,\tau}$ is the* unique solution *to the matrix equation $v_{\sigma,\tau} = Q_{\sigma,\tau} v_{\sigma,\tau} + b_{\sigma,\tau}$. Also, $I - Q_{\sigma,\tau}$ is invertible, all entries of $(I - Q_{\sigma,\tau})^{-1}$ are non-negative, and the entries along the diagonal are strictly positive.*

As a direct consequence of Lemma 2.14, we get Corollary 2.15. Note that a *fixed point* of an operator F is an element of the domain of F such that $F(x) = x$. The fact that $v_{\sigma,\tau}$ is a fixed point of $F_{\sigma,\tau}$ also appears in [Som05, Proposition 2.4].

Corollary 2.15. *Let G be a* stopping *SSG and let σ and τ be strategies of* MAX *and* MIN*, respectively. The expected payoff vector $v_{\sigma,\tau}$ is the* unique fixed point *of the operator $F_{\sigma,\tau}$. Moreover, $v_{\sigma,\tau}$ can be computed in polynomial time by solving a system of $|V_{\mathrm{POS}}|$ linear equations.*

Proof Lemma 2.14 implies that $v_{\sigma,\tau}$ is the unique solution to the equation $v_{\sigma,\tau} = Q_{\sigma,\tau} v_{\sigma,\tau} + b_{\sigma,\tau}$. Notice that, by Definition 2.13, for any value vector v, $F_{\sigma,\tau}$ satisfies $F_{\sigma,\tau}(v) = Q_{\sigma,\tau} v_{\sigma,\tau} + b_{\sigma,\tau}$. Thus, it follows that $v_{\sigma,\tau}$ is the unique solution to the equation $F_{\sigma,\tau}(v) = v$, and so is the unique fixed point of $F_{\sigma,\tau}$. Next, computing $v_{\sigma,\tau}$ requires solving the linear system $v_{\sigma,\tau} = Q_{\sigma,\tau} v_{\sigma,\tau} + b_{\sigma,\tau}$ that takes polynomial time. ☐ (Corollary 2.15)

2.1. Existence of Best Response Strategies

Lemma 2.16 gives an alternate characterization of best response strategies. Its proof is similar to the proof of correctness of Howard's algorithm [How60] presented in [Con92, Lemma 4].

Lemma 2.16. *Let G be a stopping SSG, and let σ, τ be strategies of* MAX *and* MIN*, respectively, in G. Then the following statements hold:*

1. *σ is optimal w.r.t. τ if and only if, for every $x \in V_{\mathrm{MAX}}$, we have $v_{\sigma,\tau}(x) = \max\{\overline{v}_{\sigma,\tau}(y) \mid (x,y) \in E\}$.*
2. *τ is optimal w.r.t. σ if and only if, for every $x \in V_{\mathrm{MIN}}$, we have $v_{\sigma,\tau}(x) = \min\{\overline{v}_{\sigma,\tau}(y) \mid (x,y) \in E\}$.*

Lemma 2.17 shows that, for stopping SSGs with only MIN- and AVE-positions, an optimal strategy of MIN can be found in polynomial time.

Lemma 2.17 ([Der70]). *Let G be a* stopping *SSG with no* MAX*-positions. Then an optimal strategy of* MIN (*w.r.t. this trivial* MAX*-strategy σ*) *can be found in polynomial time by solving the following linear program:*

$$\begin{aligned}
&\text{Maximize} \sum_{x \in V} \overline{v}(x) && \\
&\text{Subject to} \quad \overline{v}(x) \leq \overline{v}(y) && \text{if } x \in V_{\mathrm{MIN}} \text{ and } (x,y) \in E, \\
&\qquad\qquad\quad \overline{v}(x) = \sum_{(x,y)\in E} q(x,y) \cdot \overline{v}(y) && \text{if } x \in V_{\mathrm{AVE}}, \\
&\qquad\qquad\quad \overline{v}(x) = p(x) && \text{if } x \in \mathrm{V}_{\mathrm{SINK}}.
\end{aligned}$$

Let $\overline{v} : V \to [0,1]$ be an optimal solution of this linear program and let $v : V_{\mathrm{POS}} \to [0,1]$ be such that $v(x) = \overline{v}(x)$, for every $x \in V_{\mathrm{POS}}$. Then, any v-greedy strategy of MIN *is also his optimal strategy.*

Since the size of the linear program (LP) in Lemma 2.17 is polynomial in $|V|$, this LP can be solved in time polynomial in $|V|$ using any polynomial-time LP solver. In a similar way, for any stopping SSG that has no MIN-positions, it is possible to design a minimizing linear program for computing an optimal strategy of MAX in polynomial time. Extending the proof technique used in Lemma 2.17, it can be shown that, for any stopping SSG and for any MAX-strategy σ, an optimal MIN-strategy $\tau(\sigma)$ w.r.t. σ can be found in polynomial time. Similarly, for any stopping SSG and for any MIN-strategy τ, we can find an optimal MAX-strategy $\sigma(\tau)$ w.r.t. τ in polynomial time.

Henceforward, $\tau(\sigma)$ refers to such an optimal MIN-strategy w.r.t. any MAX-strategy σ, and $\sigma(\tau)$ refers to such an optimal MAX-strategy w.r.t. any MIN-strategy τ. For every MAX-strategy σ, we write S_σ to denote the set of all $v_{\sigma,\tau(\sigma)}$-switchable positions, and, for every MIN-strategy τ, we write T_τ to denote the set of all $v_{\sigma(\tau),\tau}$-switchable positions of G. Note that $S_\sigma \subseteq V_{\mathrm{MAX}}$ and $T_\tau \subseteq V_{\mathrm{MIN}}$.

2.2. Existence of Optimal Strategies

Lemma 2.18 shows that, for any stopping SSG, there is a unique solution to the local optimality equations given by the operator F_G of Definition 2.12; this solution is also an optimal value vector of G. This lemma implies that there always exist optimal strategies and an optimal value vector of a stopping SSG. We refer to [Som05, Theorem 2.7] for a proof sketch of this lemma.

Lemma 2.18 ([Sha53, Con92]). *Let G be a* stopping *SSG. Then there is a* unique fixed point *$v_\star : V_{\mathrm{POS}} \to [0,1]$ of the operator F_G. Moreover, $v_\star$ is an optimal value vector of G, and* $v_\star$-greedy strategies *$\sigma_\star$ and $\tau_\star$ are optimal strategies of G.*

Lemma 2.19 states that any optimal value vector of a stopping SSG is a fixed point of the operator F_G. The proof of this lemma is similar to the proofs in [Con92, Lemmas 4 and 5].

Lemma 2.19 (see [Con92]). *Let G be a* stopping *SSG and let v_{opt} be an optimal value vector of G. Then v_{opt} is a* fixed point *of the operator F_G.*

Lemma 2.19 implies that there is a unique optimal value vector of a stopping SSG, since F_G has a unique fixed point by Lemma 2.18. Therefore, any pair of optimal strategies of a stopping SSG yields the same optimal value vector of the SSG. Henceforth, we refer to an optimal value vector of a stopping SSG as *the* optimal value vector of the game.

The proof of the following lemma can be found in [Con92, Lemma 6] and [Som05, Proposition 2.8].

Lemma 2.20 (see [Con92, Som05]). *Let G be a* stopping *SSG and let v_{opt} be* the optimal value vector *of G. Then the following statements are equivalent:*

1. *strategies σ and τ are optimal,*
2. $v_{\sigma,\tau}(x) = \max_{\sigma} \min_{\tau} v_{\sigma,\tau}(x) = \min_{\tau} \max_{\sigma} v_{\sigma,\tau}(x)$ *for every $x \in V$,*
3. $v_{\sigma,\tau} = v_{\mathrm{opt}}$,
4. *strategies σ and τ are $v_{\sigma,\tau}$-greedy for* MAX *and* MIN*, respectively,*
5. *strategies σ and τ are v_{opt}-greedy for* MAX *and* MIN*, respectively.*

As pointed out in [Som05], the following observations can be made from Lemma 2.20.

Observation 2.21. *For any* stopping *SSG, knowing the optimal value vector v_{opt} yields (in polynomial time) a pair of optimal strategies σ and τ, as σ and τ are v_{opt}-greedy by the equivalence of parts (1) and (5) in Lemma 2.20.*

Observation 2.22. *For any* stopping *SSG, a polynomial-time test for the optimality of strategies σ and τ can be done in two steps: First compute $v_{\sigma,\tau}$ given σ and τ, and then verify whether σ and τ are $v_{\sigma,\tau}$-greedy for* MAX *and* MIN*, respectively. Computing $v_{\sigma,\tau}$ in the first step takes polynomial time by Corollary 2.15. The correctness of the second step follows from the equivalence of parts (1) and (4) in Lemma 2.20.*

2.3. The SSG Value Problem

Definition 2.23 ([Con92]). *The value of a SSG and the problem* SSG-VALUE *are defined as follows:*

1. *The* value *of a SSG G is* $\max_{\sigma} \min_{\tau} v_{\sigma,\tau}(\text{start vertex})$.
2. *The problem* SSG-VALUE *is: Given a SSG G, is the value of $G > \frac{1}{2}$?*

The next lemma explains why it is sufficient to restrict attention to stopping SSGs for studying the problem SSG-VALUE.

Lemma 2.24 ([Con92]). *Given any SSG G, we can transform G to a* stopping *SSG G' in time polynomial in the size of G such that G' has the same number of* MAX *and* MIN *positions as G and the value of G' is greater than $1/2$ if and only if the value of G is greater than $1/2$.*

The notations used in this paper are summarized in Table 1.

Notation	**Meaning**
G	a simple stochastic game
MAX	player MAX
MIN	player MIN
V_{MAX}	the set of all MAX-positions
V_{MIN}	the set of all MIN-positions
V_{AVE}	the set of all AVE-positions
V_{SINK}	the set of all sink vertices
V_{POS}	the set of all game positions of G; equals to $V_{\text{MAX}} \cup V_{\text{MIN}} \cup V_{\text{AVE}}$
V	the vertex set of G; equals to $V_{\text{POS}} \cup \text{V}_{\text{SINK}}$
E	the edge set of G
start vertex	the start position of G
$q(x, y)$	the probability that the next move from an AVE-position x is to y
$p(x)$	the payoff associated with a sink vertex x
σ	a strategy of MAX
τ	a strategy of MIN
$v_{\sigma,\tau}$	the expected payoff vector corresponding to strategies σ and τ
v_{opt}	an optimal value vector of a SSG
$\overline{v}$	the extension of a value vector v to V_{SINK}
$\overline{v}_{\sigma,\tau}$	the extension of the expected payoff vector $v_{\sigma,\tau}$ to V_{SINK}
$F_{\sigma,\tau}$	an operator with respect to strategies σ and τ
F_G	an operator with respect to a SSG G
$\sigma(\tau)$	an optimal strategy of MAX w.r.t. a MIN-strategy τ
$\tau(\sigma)$	an optimal strategy of MIN w.r.t. a MAX-strategy σ
S_σ	the set of all $v_{\sigma,\tau(\sigma)}$-switchable MAX-positions
T_τ	the set of all $v_{\sigma(\tau),\tau}$-switchable MIN-positions
$\mathbf{0}$	any vector in which each component is zero

Table 1: Notations used in this paper

3. Results

The strategy improvement method is an iterative procedure for constructing optimal strategies within a finite number of iterations in a decision-making scenario (e.g., game). This technique was developed first in the context of Markov decision processes, which are SSGs that have only MAX- and AVE-positions, but no MIN-positions. In Section 3.2, we study a strategy improvement algorithm by Hoffman and Karp [HK66] (Algorithm 1) for SSG-VALUE.

The Hoffman-Karp algorithm (Algorithm 1) starts with an initial pair of player strategies. It then iteratively computes new player strategies until a pair of optimal strategies is found. W.l.o.g. assume that $|V_{\text{MAX}}| \leq |V_{\text{MIN}}|$ and let $n = \min\{|V_{\text{MAX}}|, |V_{\text{MIN}}|\}$. In each iteration, the current strategy σ of MAX is changed to σ' by switching the choices of all MAX-positions at which local optimality is not achieved, while the MIN-strategy is always the best response strategy $\tau(\sigma')$ w.r.t. the new MAX-strategy σ'. The algorithm terminates when all MAX-positions (and also MIN-positions) are locally optimal. At this point, the value vector $v_{\sigma,\tau(\sigma)}$ corresponding to the current MAX-strategy σ and MIN-strategy $\tau(\sigma)$ will satisfy the local optimality equations given by the operator F_G (Definition 2.12), and so it will be a fixed point of F_G.

Input : A stopping SSG G
Output: optimal strategies σ, τ and the optimal value vector v_{opt}

begin
  Let σ and τ be arbitrary strategies of MAX and MIN, respectively
  while $(F_G(v_{\sigma,\tau}) \neq v_{\sigma,\tau})$ **do**
    Let $S \leftarrow$ the set of all $v_{\sigma,\tau}$-switchable MAX-positions
    Let $\sigma' \leftarrow \text{switch}(\sigma, S)$
    Let τ' be an optimal strategy of MIN w.r.t. σ'
    Let $\sigma \leftarrow \sigma'$, $\tau \leftarrow \tau'$
  end
  Output σ, τ and the optimal value vector $v_{\sigma,\tau}$
end

Algorithm 1: The Hoffman-Karp Algorithm [HK66]

W.l.o.g. we may assume that Algorithm 1 begins with a strategy pair σ, τ such that τ is an optimal strategy of MIN w.r.t. σ. This assumption does not change the asymptotic run-time complexity of Algorithm 1 as the strategy pair σ, τ at the end of the first iteration of the while loop in fact constitutes such a pair. We will make this assumption in the analysis of Algorithm 1 presented in Section 3.2.

Each iteration of the while loop in Algorithm 1 requires (a) computing $v_{\sigma,\tau}$ given player strategies σ and τ, and (b) computing an optimal MIN-strategy τ' w.r.t. a MAX-strategy σ'. By Corollary 2.15, $v_{\sigma,\tau}$ can be computed in polynomial time, and as discussed after Lemma 2.17, so τ' can be as $\tau' = \tau(\sigma')$. It follows that every iteration of the while loop can be executed in polynomial time. Thus, it only remains to prove the correctness of the algorithm and bound the number of iterations of the while loop.

The proof of correctness is based on a property of value vectors that is formally stated in Lemma 3.5. Using this property, it can be shown that in every iteration, the new value vector $v_{\sigma',\tau'}$ improves upon the original vector $v_{\sigma,\tau}$ in the following sense: For all positions x, $v_{\sigma,\tau}(x) \leq v_{\sigma',\tau'}(x)$, and for all MAX-switchable positions y, this inequality is strict. Thus, it follows that no MAX-strategy can repeat over all the iterations of the algorithm. Since there can be at most 2^n distinct MAX-strategies in a (binary) SSG with n MAX-positions, this algorithm requires at most 2^n iterations of the while loop in the worst case.

It is important to study the convergence time (i.e., the number of iterations) of strategy improvement algorithms for solving SSGs (e.g., the Hoffman-Karp algorithm and its many variants). These algorithms are not complicated from implementation perspective, and so a non-trivial upper bound on their convergence time might have a practical value. Melekopoglou and Condon [MC94] showed that many variations of the Hoffman-Karp algorithm require $\Omega(2^n)$ iterations in the worst case. In these variations, *only one* MAX-switchable position is switched at every iteration as opposed to *all* MAX-switchable positions in the Hoffman-Karp algorithm. In a recent breakthrough, Friedmann [Fri09] presented a super-polynomial lower bound for the discrete strategy improvement algorithm of Vöge and Jurdziński [VJ00] for solving parity games. The paper [And09] reports that Friedmann's lower bound can be extended to the Hoffman-Karp algorithm for solving SSGs.

In Section 3.2, we prove that the Hoffman-Karp algorithm requires $O(2^n/n)$ iterations in the worst case. This is the first non-trivial upper bound on the worst-case convergence time of this algorithm. In Section 3.3, we propose a randomized variant of the Hoffman-Karp algorithm and prove that with probability almost one this randomized strategy improvement algorithm requires $O(2^{0.78n})$ iterations in the worst case. Our analyses in these sections extend those of Mansour and Singh [MS99] for policy improvement algorithms for Markov decision processes.

We now present some definitions that will be used in these sections. Definition 3.1 shows how a new MAX-strategy $\mathrm{switch}(\sigma, S)$ is constructed from a MAX-strategy σ by switching the choices at any set $S \subseteq V_{\mathrm{MAX}}$.

Definition 3.1. *For every* MAX-*strategy σ and $S \subseteq V_{\mathrm{MAX}}$, let* $\mathrm{switch}(\sigma, S) : V_{\mathrm{MAX}} \to V$ *be a* MAX-*strategy obtained from σ by switching the choices of the positions of S only. That is, for every $x \in V_{\mathrm{MAX}}$ such that (x, y), $(x, z) \in E$ and $y = \sigma(x)$, we have*

$$\mathrm{switch}(\sigma, S)(x) = \begin{cases} y & \text{if } x \notin S, \text{ and} \\ z & \text{if } x \in S. \end{cases}$$

The following definition associates partial orders ($\prec$ and $\preceq$) on the set of value vectors.

Definition 3.2. *Let u, $v \in [0, 1]^N$, for some $N \in \mathbb{N}^+$. We say that*

- *$u \preceq v$ if for each $x \in [N]$, it holds that $u(x) \leq v(x)$.*
- *$u \prec v$ if $u \preceq v$ and there is an $x \in [N]$ such that $u(x) < v(x)$.*
- *u and v are incomparable if there are $x, y \in [N]$ such that $u(x) < v(x)$ and $v(y) < u(y)$.*
- *$u \npreceq v$ if either $v \prec u$, or u and v are incomparable.*
- *$u \nprec v$ if either $v \preceq u$, or u and v are incomparable.*

We will need Fact 3.3 in the proofs of Theorems 3.12 and 3.17 for bounding the number of iterations of some strategy improvement algorithms for solving SSGs.

Fact 3.3. (see [Juk01]) *Let $H(x) =_{df} -x \log_2 x - (1 - x) \log_2(1 - x)$ for $0 < x < 1$ and $H(0) = H(1) = 0$. Then, for every integer $0 \leq s \leq t/2$, we have $\sum_{k=0}^{s} \binom{t}{k} \leq 2^{t \cdot H(s/t)}$.*

3.1. Some Properties of Value Vectors

We establish some properties of value vectors that are crucial in the analyses presented in Sections 3.2 and 3.3. The first property says that optimal strategies are simultaneously best for each player assuming that the other player always plays optimally.

Lemma 3.4. *Let σ, τ be a pair of* optimal strategies *of a* stopping *SSG. Then, for any* MAX-*strategy σ', it holds that $v_{\sigma', \tau(\sigma')} \preceq v_{\sigma, \tau}$. Similarly, for any* MIN-*strategy τ', it holds that $v_{\sigma, \tau} \preceq v_{\sigma(\tau'), \tau'}$.*

Proof Since $\tau(\sigma')$ is optimal w.r.t. σ', we know from Definition 2.6 that

$$\forall x \in V_{\mathrm{POS}}, v_{\sigma', \tau(\sigma')}(x) = \min_{\tau} v_{\sigma', \tau}(x).$$

Thus, it holds that

$$\forall x \in V_{\mathrm{POS}}, v_{\sigma', \tau(\sigma')}(x) \leq \max_{\sigma} \min_{\tau} v_{\sigma, \tau}(x). \quad (1)$$

Since σ, τ are optimal strategies of G, we know from Lemma 2.20 that

$$\forall x \in V_{\text{POS}}, v_{\sigma,\tau}(x) = \max_{\sigma} \min_{\tau} v_{\sigma,\tau}(x). \tag{2}$$

Combining (1) and (2), we get that

$$\forall x \in V_{\text{POS}}, v_{\sigma',\tau(\sigma')}(x) \leq v_{\sigma,\tau}(x).$$

This completes the proof that $v_{\sigma',\tau(\sigma')} \preceq v_{\sigma,\tau}$.

The proof of the second part, namely, $v_{\sigma,\tau} \preceq v_{\sigma(\tau'),\tau'}$, is similar. So, we omit this proof. □ (Lemma 3.4)

Lemma 3.5 is a generalization of [Lud95, Lemma 6]. Its proof is based on proofs by Howard [How60] and Condon [Con92]. The lemma says that switching the choices of any nonempty subset of $v_{\sigma,\tau(\sigma)}$-switchable MAX-positions results in an improved value vector.

Lemma 3.5. *For any* stopping *SSG G, let σ be any strategy of* MAX *such that S_σ is nonempty. Let S be any nonempty subset of S_σ. Let $\sigma' =_{df}$* switch(σ, S) *be a* MAX*-strategy in G. Then, it holds that $v_{\sigma,\tau(\sigma)} \prec v_{\sigma',\tau(\sigma')}$.*

Proof Recall that $\tau(\sigma)$ ($\tau(\sigma')$) denotes an optimal MIN-strategy w.r.t. σ (respectively, σ'). For brevity, we write τ for $\tau(\sigma)$ and τ' for $\tau(\sigma')$.

Since G is a stopping SSG, we know from Lemma 2.14 that $v_{\sigma,\tau}$ and $v_{\sigma',\tau'}$ are the unique solutions to the matrix equations

$$v_{\sigma,\tau} = Q_{\sigma,\tau} v_{\sigma,\tau} + b_{\sigma,\tau} \tag{3}$$
$$v_{\sigma',\tau'} = Q_{\sigma',\tau'} v_{\sigma',\tau'} + b_{\sigma',\tau'}, \tag{4}$$

for some $|V_{\text{POS}}| \times |V_{\text{POS}}|$ matrices $Q_{\sigma,\tau}$ and $Q_{\sigma',\tau'}$, and $|V_{\text{POS}}|$-vectors $b_{\sigma,\tau}$ and $b_{\sigma',\tau'}$. Let $\vartheta = v_{\sigma',\tau'} - v_{\sigma,\tau}$. We show that $\mathbf{0} \leq \vartheta$ and that some entry of ϑ is positive. By Definition 3.2, this would imply that $v_{\sigma,\tau} \prec v_{\sigma',\tau'}$.

Subtracting Eq. (3) from Eq. (4), we get

$$\vartheta = (Q_{\sigma',\tau'} v_{\sigma',\tau'} + b_{\sigma',\tau'}) - (Q_{\sigma,\tau} v_{\sigma,\tau} + b_{\sigma,\tau}).$$

Adding and subtracting $(Q_{\sigma',\tau'} v_{\sigma,\tau} + b_{\sigma',\tau'})$ to the r.h.s. of the above equation, we get

$$\vartheta = (Q_{\sigma',\tau'} v_{\sigma',\tau'} + b_{\sigma',\tau'}) - (Q_{\sigma',\tau'} v_{\sigma,\tau} + b_{\sigma',\tau'}) + (Q_{\sigma',\tau'} v_{\sigma,\tau} + b_{\sigma',\tau'}) - (Q_{\sigma,\tau} v_{\sigma,\tau} + b_{\sigma,\tau}), \text{ or}$$
$$\vartheta = Q_{\sigma',\tau'} \vartheta + d,$$

where $d = (Q_{\sigma',\tau'} v_{\sigma,\tau} + b_{\sigma',\tau'}) - (Q_{\sigma,\tau} v_{\sigma,\tau} + b_{\sigma,\tau})$. From Lemma 2.14, we know that $I - Q_{\sigma',\tau'}$ is invertible since G is a stopping SSG. Therefore, there is a unique solution to ϑ given by $\vartheta = (I - Q_{\sigma',\tau'})^{-1} d$. Lemma 2.14 also implies that $(I - Q_{\sigma',\tau'})^{-1}$ has all entries non-negative and only positive entries along the diagonal. So, it suffices to show that $\mathbf{0} \leq d$ and that some entry of d is positive.

Notice that the vector $Q_{\sigma',\tau'} v_{\sigma,\tau} + b_{\sigma',\tau'}$ equals to $F_{\sigma',\tau'}(v_{\sigma,\tau})$ and the vector $Q_{\sigma,\tau} v_{\sigma,\tau} + b_{\sigma,\tau}$ equals to $F_{\sigma,\tau}(v_{\sigma,\tau})$, where the operators $F_{\sigma',\tau'}$ and $F_{\sigma,\tau}$ are defined as in Definition 2.13. Thus, it follows that

$$d = F_{\sigma',\tau'}(v_{\sigma,\tau}) - F_{\sigma,\tau}(v_{\sigma,\tau}). \tag{5}$$

Consider any $x \in S$. In this case, $F_{\sigma',\tau'}(v_{\sigma,\tau})(x) = \overline{v}_{\sigma,\tau}(\sigma'(x))$ and $F_{\sigma,\tau}(v_{\sigma,\tau})(x) = \overline{v}_{\sigma,\tau}(\sigma(x))$. Thus, by Eq. (5), $d(x) = \overline{v}_{\sigma,\tau}(\sigma'(x)) - \overline{v}_{\sigma,\tau}(\sigma(x))$. Since $x \in S$ and $S \subseteq S_\sigma$, x is a $v_{\sigma,\tau}$-switchable MAX-position. Therefore, by Definition 2.10, we must have $v_{\sigma,\tau}(x) < \max\{\overline{v}_{\sigma,\tau}(y) \mid (x,y) \in E\} = \overline{v}_{\sigma,\tau}(\sigma'(x))$. Since $v_{\sigma,\tau}(x) = \overline{v}_{\sigma,\tau}(\sigma(x))$, we get that $\overline{v}_{\sigma,\tau}(\sigma(x)) < \overline{v}_{\sigma,\tau}(\sigma'(x))$. It follows that $0 < d(x)$.

Next, consider any $x \in V_{\mathrm{MIN}}$. In this case, $F_{\sigma',\tau'}(v_{\sigma,\tau})(x) = \overline{v}_{\sigma,\tau}(\tau'(x))$ and $F_{\sigma,\tau}(v_{\sigma,\tau})(x) = \overline{v}_{\sigma,\tau}(\tau(x))$. Thus, by Eq. (5), $d(x) = \overline{v}_{\sigma,\tau}(\tau'(x)) - \overline{v}_{\sigma,\tau}(\tau(x))$. Since τ is optimal w.r.t. σ, by Lemma 2.16 we must have $v_{\sigma,\tau}(x) = \min\{\overline{v}_{\sigma,\tau}(y) \mid (x,y) \in E\} \leq \overline{v}_{\sigma,\tau}(\tau'(x))$. Since $v_{\sigma,\tau}(x) = \overline{v}_{\sigma,\tau}(\tau(x))$, we get that $\overline{v}_{\sigma,\tau}(\tau(x)) \leq \overline{v}_{\sigma,\tau}(\tau'(x))$. It follows that $0 \leq d(x)$.

Finally, consider any $x \in (V_{\mathrm{MAX}} - S) \cup V_{\mathrm{AVE}}$. In this case, it is easy to see from Definition 2.13 that when restricted to position x, the actions of $F_{\sigma',\tau'}$ and $F_{\sigma,\tau}$ on any value vector v are the same. It follows from Eq. (5) that $d(x) = 0$.

Thus, we have shown that for every $x \in V_{\mathrm{POS}}$, $v_{\sigma,\tau}(x) \leq v_{\sigma',\tau'}(x)$, and for every $x \in S$, $v_{\sigma,\tau}(x) < v_{\sigma',\tau'}(x)$. This proves that $v_{\sigma,\tau} \prec v_{\sigma',\tau'}$. ☐ (Lemma 3.5)

Lemma 3.6 says that switching the choice of a single MAX-position would either result in an improved value vector or result in a value vector that is no better than the original, if MIN always plays optimally. Notice that in general switching the choices of a subset of MAX-positions may result in an incomparable value vector. (See Figure 2 for an example and refer to the discussion made after the proof of Lemma 3.6 on this example.) However, this lemma states that if the choice of only *one* such position is switched and if MIN always plays optimally, then the value vectors must be comparable.

Lemma 3.6. *Let G be any* stopping *SSG. Let σ and σ' be* MAX*-strategies in G such that σ' is obtained from σ by switching the choice of a single* MAX*-position x, i.e., $\sigma' = \mathrm{switch}(\sigma, \{x\})$. Then, $v_{\sigma,\tau(\sigma)}$ and $v_{\sigma',\tau(\sigma')}$ are not incomparable, i.e., either $v_{\sigma,\tau(\sigma)} \prec v_{\sigma',\tau(\sigma')}$ or $v_{\sigma',\tau(\sigma')} \preceq v_{\sigma,\tau(\sigma)}$. Moreover, we have $v_{\sigma,\tau(\sigma)} \prec v_{\sigma',\tau(\sigma')}$ if and only if x is $v_{\sigma,\tau(\sigma)}$*-switchable.

Proof For brevity, we use τ to denote $\tau(\sigma)$, an optimal MIN-strategy w.r.t. σ, and use τ' to denote $\tau(\sigma')$, an optimal MIN-strategy w.r.t. σ'.

If x is a $v_{\sigma,\tau}$-switchable MAX-position, then, by Lemma 3.5, it holds that $v_{\sigma,\tau} \prec v_{\sigma',\tau'}$ since $\sigma' = \mathrm{switch}(\sigma, \{x\})$. So, we suppose next that x is a MAX-position such that $v_{\sigma,\tau}(x) = \max\{\overline{v}_{\sigma,\tau}(y) \mid (x,y) \in E\}$.

We can break down the construction of the pair of strategies σ', τ' from σ, τ into two steps: (i) the construction of the pair of strategies σ', τ from σ, τ in the first step, and (ii) the construction of the pair σ', τ' from σ', τ in the second step. In the first step σ' is obtained from σ as a result of switching the choice of a position $x \in V_{\mathrm{MAX}}$, i.e., $\sigma' = \mathrm{switch}(\sigma, \{x\})$. In the second step, τ' is an optimal strategy of MIN w.r.t. σ'. We will show that $v_{\sigma',\tau} \preceq v_{\sigma,\tau}$ and $v_{\sigma',\tau'} \preceq v_{\sigma',\tau}$. By transitivity of $\preceq$, it would immediately imply that $v_{\sigma',\tau'} \preceq v_{\sigma,\tau}$.

The proof of $v_{\sigma',\tau} \preceq v_{\sigma,\tau}$ is similar to that of Lemma 3.5. Since G is a stopping SSG, we know from Lemma 2.14 that $v_{\sigma,\tau}$ and $v_{\sigma',\tau}$ are the unique solutions to the following matrix equations:

$$v_{\sigma,\tau} = Q_{\sigma,\tau} v_{\sigma,\tau} + b_{\sigma,\tau} \tag{6}$$

$$v_{\sigma',\tau} = Q_{\sigma',\tau} v_{\sigma',\tau} + b_{\sigma',\tau}, \tag{7}$$

for some $|V_{\mathrm{POS}}| \times |V_{\mathrm{POS}}|$ matrices $Q_{\sigma,\tau}$ and $Q_{\sigma',\tau}$, and $|V_{\mathrm{POS}}|$-vectors $b_{\sigma,\tau}$ and $b_{\sigma',\tau}$. Let $\vartheta = v_{\sigma,\tau} - v_{\sigma',\tau}$. We show that $\mathbf{0} \preceq \vartheta$. By Definition 3.2, this would imply that $v_{\sigma',\tau} \preceq v_{\sigma,\tau}$.

Subtracting Eq. (7) from Eq. (6), we get

$$\vartheta = (Q_{\sigma,\tau}v_{\sigma,\tau} + b_{\sigma,\tau}) - (Q_{\sigma',\tau}v_{\sigma',\tau} + b_{\sigma',\tau}).$$

Adding and subtracting $(Q_{\sigma',\tau}v_{\sigma,\tau} + b_{\sigma',\tau})$ to the r.h.s. of the above equation, we get

$$\vartheta = (Q_{\sigma',\tau}v_{\sigma,\tau} + b_{\sigma',\tau}) - (Q_{\sigma',\tau}v_{\sigma',\tau} + b_{\sigma',\tau}) + (Q_{\sigma,\tau}v_{\sigma,\tau} + b_{\sigma,\tau}) - (Q_{\sigma',\tau}v_{\sigma,\tau} + b_{\sigma',\tau}), \text{ or}$$
$$\vartheta = Q_{\sigma',\tau}\vartheta + d,$$

where $d = (Q_{\sigma,\tau}v_{\sigma,\tau} + b_{\sigma,\tau}) - (Q_{\sigma',\tau}v_{\sigma,\tau} + b_{\sigma',\tau})$. From Lemma 2.14, we know that $I - Q_{\sigma',\tau}$ is invertible since G is a stopping SSG. Therefore, there is a unique solution to ϑ given by $\vartheta = (I - Q_{\sigma',\tau})^{-1}d$. Lemma 2.14 also implies that $(I - Q_{\sigma',\tau})^{-1}$ has all entries non-negative. So, it suffices to show that $\mathbf{0} \leq d$.

Notice that the vector $Q_{\sigma,\tau}v_{\sigma,\tau} + b_{\sigma,\tau}$ equals to $F_{\sigma,\tau}(v_{\sigma,\tau})$ and the vector $Q_{\sigma',\tau}v_{\sigma,\tau} + b_{\sigma',\tau}$ equals to $F_{\sigma',\tau}(v_{\sigma,\tau})$, where the operators $F_{\sigma,\tau}$ and $F_{\sigma',\tau}$ are defined as in Definition 2.13. Thus, it follows that

$$d = F_{\sigma,\tau}(v_{\sigma,\tau}) - F_{\sigma',\tau}(v_{\sigma,\tau}). \tag{8}$$

Consider the MAX-position x such that $\sigma' = \text{switch}(\sigma, \{x\})$. In this case, $F_{\sigma,\tau}(v_{\sigma,\tau})(x) = \overline{v}_{\sigma,\tau}(\sigma(x))$ and $F_{\sigma',\tau}(v_{\sigma,\tau})(x) = \overline{v}_{\sigma,\tau}(\sigma'(x))$. Thus, by Eq. (8), $d(x) = \overline{v}_{\sigma,\tau}(\sigma(x)) - \overline{v}_{\sigma,\tau}(\sigma'(x))$. Next recall the assumption that $v_{\sigma,\tau}(x) = \max\{\overline{v}_{\sigma,\tau}(y) \mid (x, y) \in E\}$. Since $v_{\sigma,\tau}(x) = \overline{v}_{\sigma,\tau}(\sigma(x))$, our assumption implies that $\overline{v}_{\sigma,\tau}(\sigma(x)) = \max\{\overline{v}_{\sigma,\tau}(y) \mid (x, y) \in E\}$; so, we get that $\overline{v}_{\sigma,\tau}(\sigma'(x)) \leq \overline{v}_{\sigma,\tau}(\sigma(x))$. It follows that $0 \leq d(x)$.

Next, consider any position $x' \neq x$. In this case, it is easy to see from Definition 2.13 that when restricted to position x', the actions of $F_{\sigma,\tau}$ and $F_{\sigma',\tau}$ on any value vector v are the same. It follows from Eq. (8) that $d(x') = 0$. Thus, we have shown that $v_{\sigma',\tau} \leq v_{\sigma,\tau}$.

The other implication, $v_{\sigma',\tau'} \leq v_{\sigma',\tau}$, follows from the fact that τ' is an optimal MIN-strategy w.r.t. σ'. Specifically, since τ' is optimal w.r.t. σ', we have, for every $x \in V_{\text{POS}}$, $v_{\sigma',\tau'} = \min_\tau v_{\sigma',\tau}(x)$ by Definition 2.6. It follows that, for any strategy τ and for every $x \in V_{\text{POS}}$, we have $v_{\sigma',\tau'}(x) \leq v_{\sigma',\tau}(x)$. This gives $v_{\sigma',\tau'} \leq v_{\sigma',\tau}$ by Definition 3.2. □ (Lemma 3.6)

Lemma 3.6 is no longer true if we require the strategies σ and σ' to differ in the choices of more than one MAX-positions. Figure 2 shows a stopping SSG in which switching the choices of two MAX-positions results in an incomparable value vector when MIN always plays optimally. In this SSG, there are the following four possibilities of a MAX-strategy:

Case 1: $\boldsymbol{\sigma = \{(1, 2), (3, 2)\}}$. In this case, it can be shown that an optimal MIN-strategy $\tau(\sigma)$ is $\{(2, 4)\}$ and $v_{\sigma,\tau(\sigma)} = (1/2, 1/2, 1/2, 1/2, 1/2)$. Thus, we have $S_\sigma = T_{\tau(\sigma)} = \emptyset$. Let this strategy pair σ, $\tau(\sigma)$ be denoted by σ_1, τ_1.

Case 2: $\boldsymbol{\sigma = \{(1, 7), (3, 2)\}}$. In this case, it can be shown that $\tau(\sigma)$ must be $\{(2, 5)\}$ and $v_{\sigma,\tau(\sigma)} = (1/4, 1/3, 1/3, 5/12, 1/3)$. Thus, we have $S_\sigma = \{1\}$ and $T_{\tau(\sigma)} = \emptyset$. Let this strategy pair σ, $\tau(\sigma)$ be denoted by σ_2, τ_2.

Case 3: $\boldsymbol{\sigma = \{(1, 2), (3, 7)\}}$. In this case, it can be shown that an optimal MIN-strategy $\tau(\sigma)$ is $\{(2, 4)\}$ and $v_{\sigma,\tau(\sigma)} = (3/8, 3/8, 1/4, 3/8, 3/8)$. Thus, we have $S_\sigma = \{3\}$ and $T_{\tau(\sigma)} = \emptyset$. Let this strategy pair σ, $\tau(\sigma)$ be denoted by σ_3, τ_3.

Case 4: $\sigma = \{(1,7),(3,7)\}$. In this case, it can be shown that $\tau(\sigma)$ must be $\{(2,5)\}$ and $v_{\sigma,\tau(\sigma)} = (1/4, 5/16, 1/4, 3/8, 5/16)$. Thus, we have $S_\sigma = \{1,3\}$ and $T_{\tau(\sigma)} = \emptyset$. Let this strategy pair σ, $\tau(\sigma)$ be denoted by σ_4, τ_4.

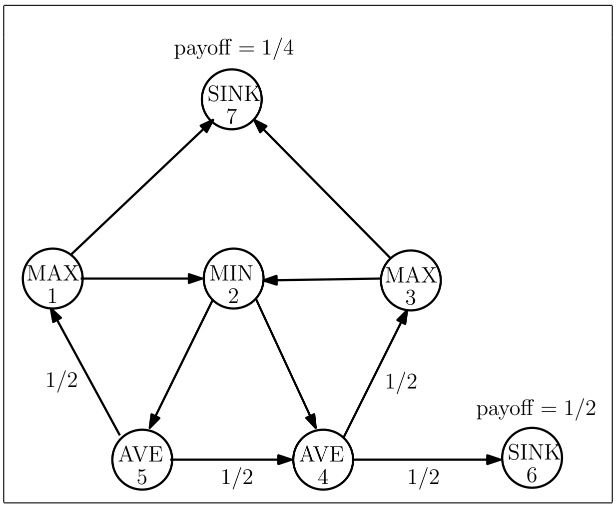

Figure 2: A *stopping* SSG G with five game positions and two sink vertices. The game positions are vertices 1 and 3 (MAX-positions), vertex 2 (MIN-position), and vertices 4 and 5 (AVE-positions). The sink vertices, 6 and 7, have payoffs 1/2 and 1/4, respectively. The probability associated with each edge leaving some AVE-position is 1/2.

Out of the $\binom{4}{2} = 6$ pairs of MAX-strategies (σ_i, σ_j), where $1 \le i < j \le 4$, the following four pairs of MAX-strategies differ only in the choice of a single MAX-position: (σ_1, σ_2), (σ_1, σ_3), (σ_2, σ_4), and (σ_3, σ_4). For these pairs of MAX-strategies, the corresponding expected payoff vectors satisfy the following relationships: $v_{\sigma_2,\tau_2} \prec v_{\sigma_1,\tau_1}$, $v_{\sigma_3,\tau_3} \prec v_{\sigma_1,\tau_1}$, $v_{\sigma_4,\tau_4} \prec v_{\sigma_2,\tau_2}$, and $v_{\sigma_4,\tau_4} \prec v_{\sigma_3,\tau_3}$. The other two pairs of MAX-strategies, namely (σ_1, σ_4) and (σ_2, σ_3), differ in the choices of two MAX-positions as $\sigma_4 = \text{switch}(\sigma_1, \{1,3\})$ and $\sigma_3 = \text{switch}(\sigma_2, \{1,3\})$. Although, for the pair (σ_1, σ_4), the corresponding expected payoff vector satisfies $v_{\sigma_4,\tau_4} \prec v_{\sigma_1,\tau_1}$, the other pair (σ_2, σ_3) has no such relationship. In fact, v_{σ_2,τ_2} and v_{σ_3,τ_3} are incomparable as $v_{\sigma_2,\tau_2}(1) < v_{\sigma_3,\tau_3}(1)$ and $v_{\sigma_2,\tau_2}(3) > v_{\sigma_3,\tau_3}(3)$.

Lemma 3.7 says that if a MAX-strategy σ' is such that all $v_{\sigma,\tau(\sigma)}$-switchable MAX-positions of another MAX-strategy σ are also $v_{\sigma',\tau(\sigma')}$-switchable, then $v_{\sigma',\tau(\sigma')}$ cannot be better than $v_{\sigma,\tau(\sigma)}$.

Lemma 3.7. *Let G be any* stopping *SSG. If σ and σ' are two* MAX*-strategies in G such that they both agree on positions in S_σ (i.e., $\forall x \in S_\sigma$, $\sigma(x) = \sigma'(x)$), then $v_{\sigma',\tau(\sigma')} \le v_{\sigma,\tau(\sigma)}$.*

Proof Consider a new game H obtained from G as follows: The vertex set $V(H)$ of H is the same as the vertex set $V(G)$ of G and the edge set $E(H)$ of H is given by

$$E(H) = \{(x,y) \in E(G) \mid x \notin S_\sigma\} \biguplus \left(\biguplus_{x \in S_\sigma} \{(x,\sigma(x)),(x,\sigma(x))\} \right).$$

Here, '$\uplus$' is the union operation on *multisets*. Thus, $E(H)$ contains all edges $(x, y) \in E(G)$ such that $x \notin S_\sigma$. On the other hand, for each $x \in S_\sigma$, $E(H)$ contains two copies of $(x, \sigma(x))$, but does not contain any edge (x, z) such that $z \neq \sigma(x)$.

Notice that every MAX- (MIN-) strategy in H is also a MAX- (respectively, MIN-) strategy in G. From this observation, it follows that (similar to G) H is a stopping SSG. Since σ and σ' agree on vertices in S_σ, both $\sigma, \tau(\sigma)$ and $\sigma', \tau(\sigma')$ are strategies in H by the definition of $E(H)$. Let $v_{\sigma,\tau(\sigma)}[G]$ and $v_{\sigma,\tau(\sigma)}[H]$ be the value vectors corresponding to the strategies σ, $\tau(\sigma)$ in G and H, respectively. In a similar way, we define $v_{\sigma',\tau(\sigma')}[G]$ and $v_{\sigma',\tau(\sigma')}[H]$ as the value vectors corresponding to the strategies σ', $\tau(\sigma')$ in G and H, respectively. It is easy to see that

$$v_{\sigma,\tau(\sigma)}[G] = v_{\sigma,\tau(\sigma)}[H] \qquad \text{and} \qquad v_{\sigma',\tau(\sigma')}[G] = v_{\sigma',\tau(\sigma')}[H]. \tag{9}$$

Claim 1. *σ, $\tau(\sigma)$ are optimal strategies of H.*

Proof of Claim 1. By the definition of S_σ, we know that every position $x \notin S_\sigma$ is $v_{\sigma,\tau(\sigma)}[G]$-stable. The equality of $v_{\sigma,\tau(\sigma)}[G]$ and $v_{\sigma,\tau(\sigma)}[H]$ in (9) implies that every $x \notin S_\sigma$ is also $v_{\sigma,\tau(\sigma)}[H]$-stable. For every $x \in S_\sigma$, the edge $(x, \sigma(x))$ is duplicated in H. It follows that every position $x \in S_\sigma$ is also $v_{\sigma,\tau(\sigma)}[H]$-stable. This shows that every position of H is $v_{\sigma,\tau(\sigma)}[H]$-stable. By Definition 2.11, we get that σ and $\tau(\sigma)$ are $v_{\sigma,\tau(\sigma)}[H]$-greedy strategies for MAX and MIN, respectively. Since H is a stopping SSG, it follows from Lemma 2.20 (1) and (4) that σ, $\tau(\sigma)$ are optimal strategies of H. □ (Claim 1)

Claim 2. *$\tau(\sigma')$ is an optimal* MIN-*strategy w.r.t. σ' in H.*

Proof of Claim 2. Since $\tau(\sigma')$ is optimal w.r.t. σ' in G, Lemma 2.16(2) implies that, for every $x \in V_{\mathrm{MIN}}$,

$$v_{\sigma',\tau(\sigma')}[G](x) = \min\{\bar{v}_{\sigma',\tau(\sigma')}[G](y) \mid (x, y) \in E(G)\}.$$

The equality of $v_{\sigma',\tau(\sigma')}[G]$ and $v_{\sigma',\tau(\sigma')}[H]$ in (9) and the definition of $E(H)$ imply that, for every $x \in V_{\mathrm{MIN}}$,

$$v_{\sigma',\tau(\sigma')}[H](x) = \min\{\bar{v}_{\sigma',\tau(\sigma')}[H](y) \mid (x, y) \in E(H)\}.$$

Thus, it follows from Lemma 2.16(2) that $\tau(\sigma')$ is optimal w.r.t. σ' in H. □ (Claim 2)

As a consequence of Claim 1, Claim 2, and Lemma 3.4, we get that $v_{\sigma',\tau(\sigma')}[H] \leq v_{\sigma,\tau(\sigma)}[H]$. The equality in (9) now implies that $v_{\sigma',\tau(\sigma')} \leq v_{\sigma,\tau(\sigma)}$. □ (Lemma 3.7)

3.2. An Improved Analysis of the Hoffman-Karp Algorithm

In this section, we prove that the number of iterations in the worst case required by the Hoffman-Karp strategy improvement algorithm is $O(2^n/n)$. This improves upon the previously known, trivial, worst-case upper bound 2^n for this algorithm.

Lemma 3.8 shows that in the Hoffman-Karp algorithm the value vectors monotonically increase with the number of iterations.

Lemma 3.8. *Let $\sigma_i, \tau(\sigma_i)$ and $\sigma_j, \tau(\sigma_j)$ be pairs of player strategies at the start of iterations i and j, respectively, for $1 \leq i < j$, of the* Hoffman-Karp *algorithm. Then, it holds that $v_{\sigma_i,\tau(\sigma_i)} \prec v_{\sigma_j,\tau(\sigma_j)}$.*

Proof The proof is by induction on $j - i$. Note that at the start of the $(i+1)$'th iteration of the Hoffman-Karp algorithm, we have $\sigma_{i+1} =_{df}$ switch(σ_i, S_{σ_i}). So, Lemma 3.5 implies that $v_{\sigma_i,\tau(\sigma_i)} \prec v_{\sigma_{i+1},\tau(\sigma_{i+1})}$. This completes the proof when $j = i+1$. If $i+1 < j$, then by induction hypothesis we get that $v_{\sigma_{i+1},\tau(\sigma_{i+1})} \prec v_{\sigma_j,\tau(\sigma_j)}$. Using this relation and $v_{\sigma_i,\tau(\sigma_i)} \prec v_{\sigma_{i+1},\tau(\sigma_{i+1})}$, it follows by the transitivity of $\prec$ that $v_{\sigma_i,\tau(\sigma_i)} \prec v_{\sigma_j,\tau(\sigma_j)}$. □ (Lemma 3.8)

An easy consequence of Lemma 3.8 is Corollary 3.9. This corollary shows that if a MAX-strategy σ' does not yield an improved value vector $v_{\sigma',\tau(\sigma')}$ compared to the value vector $v_{\sigma_{i+1},\tau(\sigma_{i+1})}$ at the start of iteration $i+1$, then σ' cannot appear after iteration $i+1$.

Corollary 3.9. *For every integer $1 \leq i$, let $\sigma_i, \tau(\sigma_i)$ denote a pair of player strategies at the start of iteration i of the* Hoffman-Karp *algorithm. Let σ' be any* MAX*-strategy. If $v_{\sigma_{i+1},\tau(\sigma_{i+1})} \not\prec v_{\sigma',\tau(\sigma')}$, then for any $i+2 \leq j$, we have $\sigma' \neq \sigma_j$.*

Proof Assume to the contrary that, for some $i+2 \leq j$, we have $\sigma' = \sigma_j$. Then Lemma 3.8 implies that $v_{\sigma_{i+1},\tau(\sigma_{i+1})} \prec v_{\sigma',\tau(\sigma')}$, a contradiction. □ (Corollary 3.9)

Lemma 3.10 shows that the set of all switchable MAX-positions differ from one iteration to another in the Hoffman-Karp strategy improvement algorithm.

Lemma 3.10. *Let $\sigma_i, \tau(\sigma_i)$ and $\sigma_j, \tau(\sigma_j)$ be pairs of player strategies at the start of iterations i and j, respectively, for $1 \leq i < j$, of the* Hoffman-Karp *algorithm. Then, it holds that $S_{\sigma_i} \nsubseteq S_{\sigma_j}$.*

Proof. Assume to the contrary that for some $1 \leq i < j$, we have $S_{\sigma_i} \subseteq S_{\sigma_j}$. Let S be the (possibly empty) subset of S_{σ_i} containing all positions x on which σ_i and σ_j disagree. (That is, $S =_{df} \{x \in S_{\sigma_i} \mid \sigma_i(x) \neq \sigma_j(x)\}$.) We define a new strategy σ' of MAX as follows: $\sigma' =_{df}$ switch(σ_j, S). Then, σ_i and σ' agree on positions in S_{σ_i}. Application of Lemma 3.7 gives $v_{\sigma',\tau(\sigma')} \preceq v_{\sigma_i,\tau(\sigma_i)}$. On the other hand, $v_{\sigma_j,\tau(\sigma_j)} \preceq v_{\sigma',\tau(\sigma')}$ by Lemma 3.5, since $\sigma' =$ switch(σ_j, S) and $S \subseteq S_{\sigma_i} \subseteq S_{\sigma_j}$. By transitivity of $\preceq$, we get that $v_{\sigma_j,\tau(\sigma_j)} \preceq v_{\sigma_i,\tau(\sigma_i)}$. However, this gives a contradiction with Lemma 3.8. □ (Lemma 3.10)

Lemma 3.11 allows to show that the Hoffman-Karp algorithm rules out, at the end of every iteration, a number of MAX-strategies that is at least *linear* in the number of switchable MAX-positions. To see this, notice that by this lemma, in every iteration, there are at least $|S_\sigma| - 1$ strategy pairs σ_i, $\tau(\sigma_i)$ such that $v_{\sigma,\tau(\sigma)} \prec v_{\sigma_i,\tau(\sigma_i)} \preceq v_{\sigma',\tau(\sigma')}$. Thus, Lemma 3.8 implies that none of these strategies can appear in any earlier iteration and Corollary 3.9 implies that none of them can appear in any later iteration.

Lemma 3.11. *Let σ, $\tau(\sigma)$ and $\sigma' =_{df}$* switch(σ, S_σ)*, $\tau' =_{df} \tau(\sigma')$ be pairs of player strategies at the start and at the end, respectively, of an iteration of the* Hoffman-Karp *algorithm. Then, there are at least $|S_\sigma| - 1$ strategy pairs σ_i, $\tau(\sigma_i)$ such that $v_{\sigma,\tau} \prec v_{\sigma_i,\tau(\sigma_i)} \preceq v_{\sigma',\tau'}$.*

Proof. Let the elements of S_σ be denoted by $1, 2, \ldots, |S_\sigma|$. For every $S \subseteq S_\sigma$, let $\sigma(S) =_{df}$ switch(σ, S) be a MAX-strategy. For notational convenience, let $\tau(S)$ denote $\tau(\sigma(S))$, which is an optimal MIN-strategy w.r.t. $\sigma(S)$.

Assume w.l.o.g. that $v_{\sigma(\{1\}),\tau(\{1\})}$ is a minimal value vector among the set of value vectors $v_{\sigma(\{i\}),\tau(\{i\})}$ for $1 \leq i \leq |S_\sigma|$. In other words, we assume that for every $2 \leq i \leq |S_\sigma|$, we have either $v_{\sigma(\{1\}),\tau(\{1\})} \preceq v_{\sigma(\{i\}),\tau(\{i\})}$ or $v_{\sigma(\{i\}),\tau(\{i\})}$ is incomparable to $v_{\sigma(\{1\}),\tau(\{1\})}$. By Lemma 3.5, we know that $v_{\sigma,\tau} \prec v_{\sigma(\{1\}),\tau(\{1\})}$.

Claim 3. *For every* $2 \le i \le |S_\sigma|$*, it holds that* $v_{\sigma(\{1\}),\tau(\{1\})} \preceq v_{\sigma(\{1,i\}),\tau(\{1,i\})}$.

For now assume that Claim 3 is true. (We prove this claim at the end of the proof of Lemma 3.11.) Pick a minimal value vector, say $v_{\sigma(\{1,2\}),\tau(\{1,2\})}$, among the set of value vectors $v_{\sigma(\{1,i\}),\tau(\{1,i\})}$ for $2 \le i \le |S_\sigma|$. This gives the sequence

$$v_{\sigma,\tau} \prec v_{\sigma(\{1\}),\tau(\{1\})} \preceq v_{\sigma(\{1,2\}),\tau(\{1,2\})}.$$

Repeating the arguments of Claim 3 with $v_{\sigma(\{1,2\}),\tau(\{1,2\})}$ in place of $v_{\sigma(\{1\}),\tau(\{1\})}$ gives the following statement: For every $3 \le i \le |S_\sigma|$, it holds that $v_{\sigma(\{1,2\}),\tau(\{1,2\})} \preceq v_{\sigma(\{1,2,i\}),\tau(\{1,2,i\})}$. Next, we pick a minimal value vector, say $v_{\sigma(\{1,2,3\}),\tau(\{1,2,3\})}$, among the set of value vectors $v_{\sigma(\{1,2,i\}),\tau(\{1,2,i\})}$ for $3 \le i \le |S_\sigma|$. This gives the sequence

$$v_{\sigma,\tau} \prec v_{\sigma(\{1\}),\tau(\{1\})} \preceq v_{\sigma(\{1,2\}),\tau(\{1,2\})} \preceq v_{\sigma(\{1,2,3\}),\tau(\{1,2,3\})}.$$

By proceeding in the above manner, the monotonically increasing sequence

$$v_{\sigma,\tau} \prec v_{\sigma(\{1\}),\tau(\{1\})} \preceq v_{\sigma(\{1,2\}),\tau(\{1,2\})} \preceq v_{\sigma(\{1,2,3\}),\tau(\{1,2,3\})} \preceq \cdots \preceq v_{\sigma(S_\sigma),\tau(S_\sigma)} = v_{\sigma',\tau'}$$

is obtained. This completes the proof of the lemma. □ (Lemma 3.11)

We now give a proof of Claim 3.

Proof of Claim 3. Assume to the contrary that, for some $2 \le i \le |S_\sigma|$, we have $v_{\sigma(\{1\}),\tau(\{1\})} \npreceq v_{\sigma(\{1,i\}),\tau(\{1,i\})}$. Then, it must be the case that either $v_{\sigma(\{1\}),\tau(\{1\})}$ and $v_{\sigma(\{1,i\}),\tau(\{1,i\})}$ are incomparable or $v_{\sigma(\{1,i\}),\tau(\{1,i\})} \prec v_{\sigma(\{1\}),\tau(\{1\})}$. Since $\sigma(\{1,i\}) = \text{switch}(\sigma(\{1\}),\{i\})$, Lemma 3.6 implies that $v_{\sigma(\{1\}),\tau(\{1\})}$ and $v_{\sigma(\{1,i\}),\tau(\{1,i\})}$ are not incomparable. Therefore, we must have $v_{\sigma(\{1,i\}),\tau(\{1,i\})} \prec v_{\sigma(\{1\}),\tau(\{1\})}$.

We next show that $v_{\sigma(\{1,i\}),\tau(\{1,i\})} \prec v_{\sigma(\{i\}),\tau(\{i\})}$. Notice that $\sigma(\{1,i\}) = \text{switch}(\sigma(\{i\}),\{1\})$. By Lemma 3.6, we must have either $v_{\sigma(\{i\}),\tau(\{i\})} \preceq v_{\sigma(\{1,i\}),\tau(\{1,i\})}$ or $v_{\sigma(\{1,i\}),\tau(\{1,i\})} \prec v_{\sigma(\{i\}),\tau(\{i\})}$. If the former holds, then by the transitivity of $\preceq$, we get that $v_{\sigma(\{i\}),\tau(\{i\})} \preceq v_{\sigma(\{1,i\}),\tau(\{1,i\})} \prec v_{\sigma(\{1\}),\tau(\{1\})}$. However, the relation $v_{\sigma(\{i\}),\tau(\{i\})} \prec v_{\sigma(\{1\}),\tau(\{1\})}$ contradicts the minimality of $v_{\sigma(\{1\}),\tau(\{1\})}$. Therefore, we must have $v_{\sigma(\{1,i\}),\tau(\{1,i\})} \prec v_{\sigma(\{i\}),\tau(\{i\})}$.

Thus, we have shown that $v_{\sigma(\{1,i\}),\tau(\{1,i\})} \prec v_{\sigma(\{1\}),\tau(\{1\})}$ and $v_{\sigma(\{1,i\}),\tau(\{1,i\})} \prec v_{\sigma(\{i\}),\tau(\{i\})}$. We claim that both $1, i \in S_{\sigma(\{1,i\})}$. To see this, note that $\sigma(\{1\}) = \text{switch}(\sigma(\{1,i\}),\{i\})$ and $v_{\sigma(\{1,i\}),\tau(\{1,i\})} \prec v_{\sigma(\{1\}),\tau(\{1\})}$. So, by Lemma 3.6, we must have $i \in S_{\sigma(\{1,i\})}$. In the same way, we must have $1 \in S_{\sigma(\{1,i\})}$.

We now know that $\sigma = \text{switch}(\sigma(\{1,i\}),\{1,i\})$, where $\{1,i\} \subseteq S_{\sigma(\{1,i\})}$. Lemma 3.5 implies that we must have $v_{\sigma(\{1,i\}),\tau(\{1,i\})} \prec v_{\sigma,\tau}$. However, we also have $\{1,i\} \subseteq S_\sigma$ and $\sigma(\{1,i\}) = \text{switch}(\sigma,\{1,i\})$. So, Lemma 3.5 also implies that $v_{\sigma,\tau} \prec v_{\sigma(\{1,i\}),\tau(\{1,i\})}$. This gives a contradiction. □ (Claim 3)

We are now ready to prove the main result of Section 3.2.

Theorem 3.12. *The* Hoffman-Karp *algorithm requires at most* $O(2^n/n)$ *iterations in the worst case, where* $n = \min\{|V_{\text{MAX}}|, |V_{\text{MIN}}|\}$.

Proof. W.l.o.g. assume that $n = |V_{\text{MAX}}| \le |V_{\text{MIN}}|$. We partition the analysis of the number of iterations into two cases: (1) iterations in which $|S_\sigma| \le n/3$ and (2) iterations in which $|S_\sigma| > n/3$.

By Lemma 3.10, the set of all switchable MAX-positions cannot repeat throughout the iterations of the Hoffman-Karp algorithm. Therefore, the number of iterations in which $|S_\sigma| \leq n/3$ is bounded by $\sum_{k=0}^{n/3} \binom{n}{k}$, which is at most $2^{n \cdot H(1/3)}$ by Fact 3.3. In case (2), since $|S_\sigma| > n/3$ in each such iteration, by Lemma 3.11 the Hoffman-Karp algorithm rules out at least $n/3$ strategies σ_i such that $v_{\sigma,\tau} \prec v_{\sigma_i,\tau(\sigma_i)} \preceq v_{\sigma',\tau'}$. (Note that σ, τ refers to the strategy pair at the start of the current iteration and σ', τ' refers to the strategy pair at the start of the next iteration.) Therefore, the number of iterations in which $|S_\sigma| > n/3$ is bounded by $2^n/(n/3) = 3 \cdot 2^n/n$.

It follows that the Hoffman-Karp algorithm requires at most $2^{n \cdot H(1/3)} + 3 \cdot 2^n/n \leq 4 \cdot 2^n/n$ iterations in the worst case. □ (Theorem 3.12)

3.3. *A Randomized Variant of the Hoffman-Karp Algorithm*

We propose a randomized strategy improvement algorithm (Algorithm 2) for the SSG value problem. This algorithm can be seen as a variation of the Hoffman-Karp algorithm in that, instead of deterministically choosing all switchable MAX-positions, the randomized algorithm chooses a uniformly random subset of the switchable MAX-positions in each iteration. Similar to the results in Section 3.2, our results in this section extend those of Mansour and Singh [MS99] for policy iteration algorithms for Markov decision processes. We mention that Condon [Con93] also presented a randomized variant of the Hoffman-Karp algorithm that is different from ours. The expected number of iterations of Condon's algorithm is at most $2^{n-f(n)}+2^{o(n)}$, for any function $f(n) = o(n)$, where $n = \min\{|V_{\text{MAX}}|, |V_{\text{MIN}}|\}$.

Input : A stopping SSG G
Output: Optimal strategies σ, τ and the optimal value vector v_{opt}

```
begin
    Let σ be an arbitrary MAX-strategy and τ = τ(σ) be a MIN-strategy
    while (F_G(v_{σ,τ}) ≠ v_{σ,τ}) do
        S ← a uniformly random subset of v_{σ,τ}-switchable MAX-positions
        Let σ' ← switch(σ, S)
        Let τ' be an optimal strategy of MIN w.r.t. σ'
        Let σ ← σ', τ ← τ'
    end
    Output σ, τ and the optimal value vector v_{σ,τ}
end
```

Algorithm 2: A Randomized Variant of the Hoffman-Karp Algorithm

It can be verified that analogs of Lemma 3.8, Corollary 3.9, and Lemma 3.10 hold for Algorithm 2. Therefore, we state the following results without giving their proofs.

Lemma 3.13. *Let $\sigma_i, \tau(\sigma_i)$ and $\sigma_j, \tau(\sigma_j)$ be pairs of player strategies at the start of iterations i and j, respectively, for $1 \leq i < j$, of Algorithm 2. Then, it holds that $v_{\sigma_i,\tau(\sigma_i)} \prec v_{\sigma_j,\tau(\sigma_j)}$.*

Corollary 3.14. *For every integer $1 \leq i$, let $\sigma_i, \tau(\sigma_i)$ denote a pair of player strategies at the start of iteration i of Algorithm 2. Let σ' be any* MAX*-strategy. If $v_{\sigma_{i+1},\tau(\sigma_{i+1})} \not\prec v_{\sigma',\tau(\sigma')}$, then for any $i + 2 \leq j$, we have $\sigma' \neq \sigma_j$.*

Lemma 3.15. *Let $\sigma_i, \tau(\sigma_i)$ and $\sigma_j, \tau(\sigma_j)$ be pairs of player strategies at the start of iterations i and j, respectively, for $1 \leq i < j$, of Algorithm 2. Then, it holds that $S_{\sigma_i} \not\subseteq S_{\sigma_j}$.*

Lemma 3.16 allows to show that Algorithm 2 rules out, at the end of every iteration, a number of MAX strategies that on expectation is at least *exponential* in the number of switchable MAX-positions. To see this, notice that by this lemma, in every iteration, the expected number of MAX-strategies σ' such that $v_{\sigma_i,\tau(\sigma_i)} \preceq v_{\sigma',\tau(\sigma')}$ and $v_{\sigma_{i+1},\tau(\sigma_{i+1})} \not\preceq v_{\sigma',\tau(\sigma')}$ is at least $2^{|S_{\sigma_i}|-1}$. Thus, Lemma 3.13 implies that none of these strategies can appear in any earlier iteration and Corollary 3.14 implies that none of them can appear in any later iteration.

Lemma 3.16. *In Algorithm 2, let σ_i, $\tau(\sigma_i)$ be a pair of player strategies at the start of an iteration in which S_{σ_i} is nonempty, and let σ_{i+1}, $\tau(\sigma_{i+1})$ be a pair of player strategies at the end of this iteration. Then, the expected number of* MAX*-strategies σ' such that $v_{\sigma_i,\tau(\sigma_i)} \preceq v_{\sigma',\tau(\sigma')}$ and $v_{\sigma_{i+1},\tau(\sigma_{i+1})} \not\preceq v_{\sigma',\tau(\sigma')}$ is at least $2^{|S_{\sigma_i}|-1}$.*

Proof. Consider an iteration in which σ_i, $\tau(\sigma_i)$ is a pair of player strategies such that S_{σ_i} is nonempty. Let U denote the set of all MAX-strategies obtained from σ_i by switching some subset of S_{σ_i}, i.e., $U = \{\sigma \mid (\exists S \subseteq S_{\sigma_i})[\sigma = \text{switch}(\sigma_i, S)]\}$. Clearly, $|U| = 2^{|S_{\sigma_i}|}$. For each strategy $\sigma \in U$, we associate sets U^+_σ and U^-_σ that are defined as follows:

$$U^+_\sigma =_{df} \{\sigma' \in U \mid v_{\sigma,\tau(\sigma)} \prec v_{\sigma',\tau(\sigma')}\} \text{ and } U^-_\sigma =_{df} \{\sigma' \in U \mid v_{\sigma',\tau(\sigma')} \prec v_{\sigma,\tau(\sigma)}\}.$$

Notice that, for any pair σ, $\sigma' \in U$, we have $\sigma' \in U^+_\sigma$ if and only if $\sigma \in U^-_{\sigma'}$. From this equivalence, it follows that

$$\sum_{\sigma\in U} |U^+_\sigma| = \sum_{\sigma\in U} |U^-_\sigma| \leq \frac{|U|^2}{2}.$$

Thus, for a strategy σ_{i+1} chosen uniformly at random from U in the current iteration, the expected number of MAX-strategies $\sigma' \in U$ such that $v_{\sigma_{i+1},\tau(\sigma_{i+1})} \not\preceq v_{\sigma',\tau(\sigma')}$ is

$$\begin{aligned} &= |U| - \frac{1}{|U|} \cdot \sum_{\sigma\in U} |U^+_\sigma| \\ &\geq \frac{|U|}{2} = 2^{|S_{\sigma_i}|-1}. \end{aligned}$$

Moreover, by Lemma 3.5, every such MAX-strategy σ' satisfies $v_{\sigma_i,\tau(\sigma_i)} \preceq v_{\sigma',\tau(\sigma')}$. □ (Lemma 3.16)

We are now ready to prove the main result of Section 3.3.

Theorem 3.17. *With probability at least $1 - 2^{-2^{\Omega(n)}}$, Algorithm 2 requires at most $O(2^{0.78n})$ iterations in the worst case, where $n = \min\{|V_{\text{MAX}}|, |V_{\text{MIN}}|\}$.*

Proof. Let $c \in (0, 1/2)$ that we will fix later in the proof. W.l.o.g. assume that $n = |V_{\text{MAX}}| \leq |V_{\text{MIN}}|$. We partition the analysis of the number of iterations i into two cases: (1) iterations in which $|S_{\sigma_i}| \leq cn$ and (2) iterations in which $|S_{\sigma_i}| > cn$.

By Lemma 3.15, the set of all switchable MAX-positions cannot repeat throughout the iterations of Algorithm 2. Therefore, as in the proof of Theorem 3.12, the number of iterations i in which $\sigma_i, \tau(\sigma_i)$ is a pair of player strategies at the start of it and $|S_{\sigma_i}| \leq cn$ is bounded by $\sum_{k=0}^{cn} \binom{n}{k}$, which is at most $2^{n \cdot H(c)}$ by Fact 3.3.

We next bound the number t^* of iterations i in which $\sigma_i, \tau(\sigma_i)$ is a pair of player strategies at the start of it and $|S_{\sigma_i}| > cn$. By Lemma 3.16, the expected number of MAX-strategies σ' that

Algorithm 2 rules out in each such iteration is at least $2^{|S_{\sigma_i}|-1} \geq 2^{cn}$. It follows from Markov's inequality that, with probability at least $1/2$, Algorithm 2 rules out at least 2^{cn-1} MAX-strategies in each such iteration.

We say that an iteration i in which $|S_{\sigma_i}| > cn$ is *good* if Algorithm 2 rules out at least 2^{cn-1} MAX-strategies at the end of it. We know from above that the probability that an iteration i in which $|S_{\sigma_i}| > cn$ is good is at least $1/2$. By Chernoff bounds, for any $t > 0$, at least $1/4$ of the t iterations i in which $|S_{\sigma_i}| > cn$ will be good with probability at least $1 - e^{-t/16}$. Thus, with probability at least $1 - e^{-t^*/16}$, there are at least $t^*/4$ iterations such that Algorithm 2 rules out at least 2^{cn-1} MAX-strategies at the end of each one of them; so, this gives $2^{cn-1} \cdot t^*/4 < 2^n$ or $t^* < 2^{n(1-c)+3}$. In other words, with probability at least $1 - e^{-t^*/16}$, the number t^* of iterations i in which $|S_{\sigma_i}| > cn$ is at most $2^{n(1-c)+3}$.

Thus, the total number of iterations is at most $2^{n \cdot H(c)} + 2^{n \cdot (1-c)+3}$ with probability at least $1 - e^{-t^*/16}$. Choosing $c \in (0, 1/2)$ such that $H(c) = 1 - c$ gives $c \approx 0.227$. For $c = 0.227$, this number of iterations is $O(2^{0.78n})$. Also, when the number of iterations (and so t^*) is $O(2^{0.78n})$, then the probability of success is at least $1 - 2^{-2^{\Omega(n)}}$. □ (Theorem 3.17)

Acknowledgment: V.S. Anil Kumar was supported partially by the following grants: NSF Nets Grant CNS-0626964, NSF HSD Grant SES-0729441, NIH MIDAS project 2U01GM070694-7, DTRA CNIMS Grant HDTRA1-07-C-0113, NSF NETS CNS-0831633, DOE DE-SC0003957, NSF NetSE CNS-1011769.

References

[And09] D. Andersson. Extending Friedmann's lower bound to the Hoffman-Karp algorithm. Preprint, June 2009.

[Con92] A. Condon. The complexity of stochastic games. *Information and Computation*, 96(2):203–224, 1992.

[Con93] A. Condon. On algorithms for simple stochastic games. In J. Cai, editor, *Advances in Computational Complexity Theory*, volume 13, pages 51–73. DIMACS series in Discrete Mathematics and Theoretical Computer Science, American Mathematical Society, 1993.

[Der70] C. Derman. *Finite State Markov Decision Processes*, volume 67 of *Mathematics in Science and Engineering*. Academic Press, New York, 1970.

[Fri09] O. Friedmann. An exponential lower bound for the parity game strategy improvement algorithm as we know it. In *Proceedings of the 24th Annual IEEE Symposium on Logic in Computer Science*, pages 145–156. IEEE Computer Society, 2009.

[GH09] H. Gimbert and F. Horn. Solving simple stochastic games with few random vertices. *Logical Methods in Computer Science*, 5(2):1–17, 2009.

[GW02] E. Grädel and T. Wolfgang. Automata, logics, and infinite games. *Lecture Notes in Computer Science*, 2500, 2002.

[Hal07] N. Halman. Simple stochastic games, parity games, mean payoff games and discounted payoff games are all LP-type problems. *Algorithmica*, 49(1):37–50, 2007.

[HK66] A. Hoffman and R. Karp. On nonterminating stochastic games. *Management Science*, 12:359–370, 1966.

[How60] R. Howard. *Dynamic Programming and Markov Processes*. M.I.T. Press, Cambridge, MA, 1960.

[Jub05] B. Juba. On the hardness of simple stochastic games. Master's thesis, Carnegie Mellon University, 2005.

[Juk01] S. Jukna. *Extremal Combinatorics*. Springer, 2001.

[Lud95] W. Ludwig. A subexponential randomized algorithm for the simple stochastic game problem. *Information and Computation*, 117(1):151–155, 1995.

[MC94] M. Melekopoglou and A. Condon. On the complexity of the policy improvement algorithm for Markov decision processes. *ORSA Journal of Computing*, 6(2):188–192, 1994.

[MS99] Y. Mansour and S. Singh. On the complexity of Policy Iteration. In *Proceedings of the 15th Conference on Uncertainity in Artificial Intelligence, Stockholm, Sweden*, pages 401–408. Morgan Kaufmann, July 1999.

[Poo03] C. Poon. Verifying minimal stable circuit values. *Information Processing Letters*, 86(1):27–32, 2003.

[Pur95] A. Puri. *Theory of Hybrid Systems and Discrete Event Systems*. PhD thesis, EECS, University of Berkley, 1995.

[Sha53] L. Shapley. Stochastic games. In *Proceedings of National Academy of Sciences (U.S.A.)*, volume 39, pages 1095–1100, 1953.

[Som05] R. Somla. New algorithms for solving simple stochastic games. *Electronic Notes in Theoretical Computer Science*, 119(1):51–65, 2005.

[TVK10] R. Tripathi, E. Valkanova, and V. Kumar. On strategy improvement algorithms for simple stochastic games. In *Proceedings of the 7th International Conference on Algorithms and Complexity*, pages 240–251. Springer Verlag *Lecture Notes in Computer Science #6078*, 2010.

[VJ00] J. Vöge and M. Jurdziński. A discrete strategy improvement algorithm for solving parity games. In *Proceedings of the 12th International Conference on Computer Aided Verification*, pages 202–215. Springer Verlag *Lecture Notes in Computer Science #1855*, 2000.

[ZP96] U. Zwick and M. Paterson. The complexity of mean payoff games on graphs. *Theoretical Computer Science*, 158:343–359, 1996.