Noname manuscript No.
(will be inserted by the editor)

Extremal cases of poor approximation of undiscounted recursive games by discounted and time-bounded ones

Rasmus Ibsen-Jensen

the date of receipt and acceptance should be inserted later

Abstract A seminal result of Mertens and Neyman states that the value of a finite but infinite-horizon undiscounted two-player zero-sum stochastic game is the limits of the values of the corresponding discounted and time bounded games as the discount factor approaches one and the time bound approaches infinity. In this work, we are interested in games that are *extremal* in a class of stochastic games in the sense that the approximations for the value offered by the discounted and the time bounded versions are the worst possible. As our main result, we identify for each N and m such an extremal game among all games with N positions, m actions for each player in each position, all rewards either 0 or 1 with reward 1 only occurring in absorbing positions, and all positions having value 0 or 1. This extremal game is the following: Player II repeatedly selects and hides a number between 1 and m. Each time Player II hides such a number, Player I must try to guess which number it is. After the guess, the hidden number is revealed. If Player I ever guesses a number which is strictly higher than the one Player II is hiding, Player I loses the game. If Player I ever guesses correctly N times in a row, the game ends with Player I being the winner. If neither of these two events ever happen and the play thus continues forever, Player I loses.

Keywords Stochastic games, recursive games, extremal combinatorics, approximations

R. Ibsen-Jensen
IST Austria, E-mail: ribsen@ist.ac.at

1 Introduction

In this paper we consider classic two-player, zero-sum, infinite-horizon, finite-position stochastic games. Such games are played over turns and in turn i there is a current position k_i. Player p then selects an action a_i^p (without knowing the other player's action) and depending on their joint choice of actions and the position k_i, some reward r_i is given (formally, there is a *reward function* r for game G such that $r_i = r^G(k_i, a_i^1, a_i^2)$) and the next position k_{i+1} is defined (formally, there is a *position function* π for game G such that $k_{i+1} = \pi^G(k_i, a_i^1, a_i^2)$)[1]. The game continues indefinitely. For any T, by G_T we denote the finite-horizon game where the payoff to Player I is $\frac{\sum_{i=1}^T r_i}{T}$. By G we denote the game when the payoff to Player I is $\lim_{T\to\infty} \frac{\sum_{i=1}^T r_i}{T}$ (that is, the limit of the finite-horizon game G_T). For any $\lambda \in (0,1)$ we denote by G^λ the game when the payoffs are discounted by λ, that is where the payoff to Player I is $\lambda \sum_{i=0}^\infty (1-\lambda)^i r_i$. We will furthermore focus our attention on the special case of recursive games (Everett (1957)). In such games, all non-zero payoffs are in absorbing positions (absorbing positions are positions such that if v_i is an absorbing position, then $v_i = v_{i'}$ for all $i' > i$). For such games and any T, one can also consider an alternate finite-horizon game G'_T where the payoff to Player I is r_T, that is, the payoff is the reward of the last round. For recursive games, as shown by Everett (1957), one only needs to consider stationary strategies, i.e. strategies that only depends on the current position but not on how the game got there. For any game G, whether finite-horizon or infinite-horizon, discounted or undiscounted, we denote the vector of values of G by $\text{val}(G)$ (the vector has an entry for each possible starting position). If rewards are non-negative, then $\text{val}(G_T) \le \text{val}(G'_T)$. A seminal result of Mertens and Neyman (1981) states:

Theorem 1 (Mertens and Neyman)

$$\text{val}(G) = \lim_{T\to\infty} \text{val}(G_T) = \lim_{\lambda\to 0+} \text{val}(G^\lambda).$$

When G is a recursive game, we also have $\text{val}(G) = \lim_{T\to\infty} \text{val}(G'_T)$. *We are interested in the question of how good an approximation we get from Theorem 1.* More formally: For a given number of non-absorbing positions N, a given bound on the number of actions m per position and some $\epsilon > 0$, how large does T and λ^{-1} need to be before $|\text{val}(G) - \text{val}(G_T)| \le \epsilon$, $|\text{val}(G) - \text{val}(G'_T)| \le \epsilon$ and $|\text{val}(G) - \text{val}(G^\lambda)| \le \epsilon$ for every game G (with N non-absorbing positions and at most m actions per position)?

Main result and Purgatory. The main result of the present paper is to find, given N and m, the game G over some sub-set of games $\mathcal{G}_{N,m}$, to be defined next, that for any $\epsilon > 0$ maximizes the smallest value for T and λ^{-1} for which the above three inequalities are satisfied. (A priori, it is not clear that such a game exists at all). The set of games $\mathcal{G}_{N,m}$ is the sub-set of recursive games with N positions and at most m actions in each position for which $\text{val}(G)_k \in \{0,1\}$ for each position k.

For any N and m we next define the game of *Purgatory* $P_{N,m}$, as the following recursive game with N non-terminal positions and m actions in each position: *In each*

[1] For convenience, we consider only deterministic position functions

round, Player II hides a number between 1 and m and Player I attempts to guess this number. If Player I guesses correctly N times in a row he wins the game. If Player I even once guesses too low, he loses the game. If neither of those ever happens, and the game thus consists of an infinite number of rounds, Player I loses the game. We have that $P_{N,m} \in \mathcal{G}_{N,m}$ as shown by Hansen, Ibsen-Jensen, and Miltersen (2014) (indeed, except for the absorbing position corresponding to losing the game, Player 1 can win starting from any position with probability arbitrarily close to 1). We can now formulate our main result:

Theorem 2 *For arbitrary $N \geq 1, m \geq 2, T \geq 1$ and $\lambda \in (0, 1)$, the game $P_{N,m}$ maximizes each of the following three quantities among all games of $\mathcal{G}_{N,m}$:*

$$\mathrm{val}(G) - \mathrm{val}(G_T),\ \mathrm{val}(G) - \mathrm{val}(G'_T),\ \textit{and}\ \mathrm{val}(G) - \mathrm{val}(G^\lambda)$$

There is an illustration of the game $P_{2,2}$ in Figure 1a. The game is illustrated using the following convention: Each position corresponds to a matrix. The rows of the matrix corresponds to choices for Player I and the columns to choices for Player II. Each entry has an arrow, pointing to which position the game would advance to if the current position corresponded to that matrix and the players had played the defining row and column of the entry. Also, the entry contains the reward if it is non-zero (to make the figures more readable we omit zeros).

We conjecture that $P_{N,m}$ is also extremal over all other recursive games:

Conjecture 1 $P_{N,m}$ is extremal over all recursive games with N non-absorbing positions and at most m actions in each position, deterministic position function and rewards in $\{0, 1\}$.

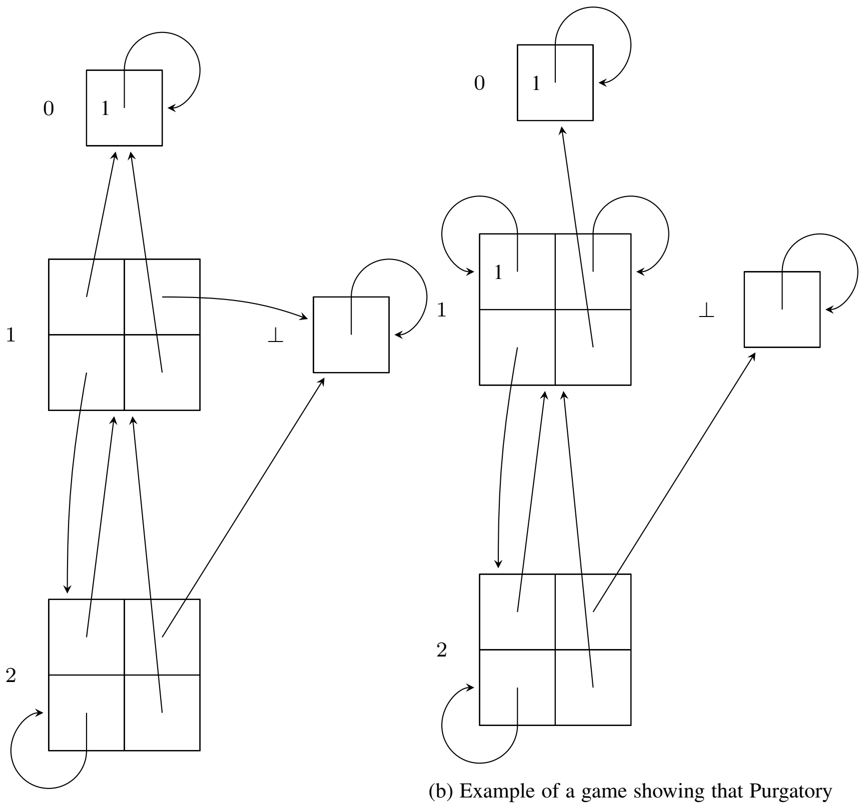
Note that while we conjecture that $P_{N,m}$ is extremal over all recursive games with rewards in $\{0, 1\}$, it is not extremal over all stochastic games with rewards in $\{0, 1\}$ and not even those where the value of each position is 0 or 1. Indeed, the stochastic game G in Figure 1b is such that except for the absorbing position $\perp$ with 0 reward, each other position k is such that $\mathrm{val}(G)_k = 1$, as shown in Ibsen-Jensen (2013, Chapter 3, Lemma 19). But we also have that $\mathrm{val}(G_4)_2 = 1/7 \approx 0.143$. Yet in $\tilde{G} = P_{2,2}$ (see Figure 1a), we have that $\mathrm{val}(\tilde{G}_4)_{k'} \geq 64/399 \approx 0.160$, for each non-absorbing position k' (it is equal for $k' = 2$).

Previous work. Previous work have also considered our problem of interest (i.e. how good an approximation we get from Theorem 1). The problem refines a result by Milman (2002), who showed that for any stochastic game, there exists a number c, such that for all $\lambda > 0$, we have that $|\mathrm{val}(G^\lambda) - \mathrm{val}(G)| \leq \epsilon$ for $\lambda = \epsilon^c$, in that we are interested in how c depends on N and m. The problem has since then been considered in computer science literature, including, among others, by the author of the present paper. We will next provide a short review of these works. The game $P_{N,m}$ was previously considered in Hansen, Koucký, and Miltersen (2009) (for $m = 2$) and Hansen, Ibsen-Jensen, and Miltersen (2011a), where it was shown that:

Theorem 3 *Let $G = P_{N,m}$ for any m and any even N. If $T < 2^{m^{N/2}}$, then*

$$\mathrm{val}(G_T)_N \leq \mathrm{val}(G'_T)_N \leq 3m^{-N/2} \ ,$$

where position N is the start state.



(a) Illustration of Purgatory $P_{2,2}$.

(b) Example of a game showing that Purgatory is not extremal for all stochastic games.

Fig. 1: Illustration of games mentioned in the introduction

(The proof is elementary but somewhat involved.) This also easily implies that if $\lambda^{-1} < \frac{2^{m^{N/2}}}{N/2\log_2(m)}$, then val$(G^\lambda)_N \leq 4m^{-N/2}$. Since the papers also show that val$(G)_N$ is 1, this shows that T and λ^{-1} needs to be double exponential large before we get a good approximation.

To find a value for T and λ^{-1} that suffices techniques from semi-algebraic geometry and specifically first order theory of real numbers was used. Specifically, in Hansen, Koucký, Lauritzen, Miltersen, and Tsigaridas (2011b), it was shown that the value and strategies of a recursive game could be expressed in the first order theory of reals, and using results for such (especially the *sampling theorem* of Basu, Pollack, and Roy (2006, Theorem 13.11)), they showed that double exponential *patience* is sufficient. Patience is 1 divided by the smallest non-zero probability used in the strategy. Next, in Hansen, Ibsen-Jensen, and Miltersen (2011a), it was shown how lower bounds on patience of ϵ-strategies yield upper bounds on the value of its time bounded version, which again easily imply upper bounds on the value of its discounted version. Together, the two papers show that:

Theorem 4 (Hansen, Koucký, Lauritzen, Miltersen, and Tsigaridas (2011b) and Hansen, Ibsen-Jensen, and Miltersen (2011a)) *There is a constant c so that for any recursive game G with $N \geq 1$ positions and $m \geq 2$ actions for each player in each position and where the rewards are in $\{0,1\}$ and which is using a deterministic position function, for any $0 < \epsilon \leq \frac{1}{2}$ we have that:*

$$T \geq (1/\epsilon)^{m^{cN}} \Rightarrow |\text{val}(G) - \text{val}(G_T)| \leq \epsilon.$$

$$\lambda \leq \epsilon^{m^{cN}} \Rightarrow |\text{val}(G) - \text{val}(G^\lambda)| \leq \epsilon.$$

It is hard to get a bound on c using the literature on semi-algebraic geometry. However, if our Conjecture 1 is true, then c is close to 1.

For general stochastic games, the bounds are somewhat worse.

Theorem 5 (Hansen, Koucký, Lauritzen, Miltersen, and Tsigaridas (2011b)) *There is a constant c so that for any stochastic game G with $N \geq 1$ positions and $m \geq 2$ actions for each player in each position and where the rewards are in $\{0,1\}$ and which is using a deterministic position function, for any $0 < \epsilon \leq \frac{1}{2}$ we have that:*

$$T \geq (1/\epsilon)^{m^{c'N^2}} \Rightarrow |\text{val}(G) - \text{val}(G_T)| \leq \epsilon.$$

$$\lambda \leq \epsilon^{m^{c'N^2}} \Rightarrow |\text{val}(G) - \text{val}(G^\lambda)| \leq \epsilon.$$

The approach is roughly similar to before: If one express the proof of Theorem 1 in first order theory of the reals, one can use standard theorems of semi-algebraic geometry, including the aforementioned sampling theorem and the quantifier elimination theorem by Basu, Pollack, and Roy (2006, Theorem 14.16) to get a bound on how good the discounted/finite horizon games approximates the undiscounted and infinite-horizon game. Again, we know no explicit bound on c.

For a special case somewhere in between recursive and stochastic games, if one considers stochastic games, but where finite-memory, ϵ-optimal strategies exist and the value of each position is in $\{0,1\}$, then results by Chatterjee and Ibsen-Jensen (2015) show that ϵ-optimal stationary strategies exist and also bounds the patience for such. The proof is elementary, but not simple and relies on short characterization of such positions and gives an explicit construction of such strategies. Similarly to before, we can then apply the argument from Hansen, Ibsen-Jensen, and Miltersen (2011a), relating patience to values of time bounded games. Together, the two papers show that:

Theorem 6 (Chatterjee and Ibsen-Jensen (2015) and Hansen, Ibsen-Jensen, and Miltersen (2011a)) *For any stochastic game G with $N \geq 1$ positions and $m \geq 2$ actions for each player in each position and where the rewards and $\text{val}(G)_k$ for each position k are in $\{0,1\}$ and for which for any $\epsilon > 0$, there exists a finite memory strategy ensuring $\text{val}(G)_k - \epsilon$, we have that:*

$$T \geq (4/\epsilon)^{(2m)^N N} \Rightarrow |\text{val}(G) - \text{val}(G_T)| \leq \epsilon.$$

$$\lambda \leq (\epsilon/8)^{(2m)^N N} / \log(2/\epsilon) \Rightarrow |\text{val}(G) - \text{val}(G^\lambda)| \leq \epsilon.$$

2 Preliminaries

We consider recursive games where all absorbing positions have limiting average value 1 or 0. We assume without loss of generality that there is one position of value 0; we shall refer to this position as "$\perp$", and one position of value 1 ; we shall refer to this position as "0" (there are practical reasons for this choice of indexing related to statements of lemmas below). The non-absorbing positions are $\{1,\ldots,N\}=[N]$. As mentioned in the introduction, for a game G there is a position function π^G, such that if players play actions i,j in position k for some k,i,j, then play proceeds to position $k':=\pi^G(k,i,j)$. For convenience, we define the matrix $M:=\pi^G(k)$ such that $M_{i,j}=\pi^G(k,i,j)$.

Let $\mathcal{G}_{N,m}$ be the set of recursive games G defined in the introduction. For an absorbing position k, we will by v_k denote the reward of that position.

Let $R_{N,m}$ be the set of recursive games G, where G has N positions, each player has m actions in each position and all successors of position k are in $\{\perp,k-1,N\}$.

Five different recursive games. We will now define five different recursive games for a given set of positions and actions, like the games defined in the introduction. That is, the games differ only in their payoff. In each case, let r_i be the reward collected by Player I at stage i. In each case we will define the *value vector* $\mathbf{v}$. The value vector is such that $\mathbf{v}_k$ is the value of the game, when starting in position k.

1. We will by G denote the game with limiting average (undiscounted) payoffs, i.e, payoff $\liminf_{T\to\infty}(\sum_{i=0}^{T} r_i)/(T+1)$ to Player I. Such a game has a value vector as shown by Mertens and Neyman (1981). Let val(G) be this value vector.
2. For any integer $T\geq 0$, we will by G_T denote the finite-horizon game with T stages and payoff $(\sum_{i=0}^{T} r_i)/(T+1)$ to Player I. Such a game has a value vector by von Neumann's minmax theorem and backward induction. Let val(G_T) be this value vector.
3. For any real number $\lambda\in(0,1)$, we will by G^λ denote the game with payoffs discounted with a discount factor of $1-\lambda$, i.e., with payoff $\lambda\sum_{i=0}^{\infty}(1-\lambda)^i r_i$ to Player I. Such a game has a value vector as shown by Shapley (1953). Let val(G^λ) be this value vector.
4. For any integer $T\geq 0$, we will by G'_T denote the finite-horizon game with T stages and payoff r_T to Player I. Such a game has a value vector by von Neumann's minmax theorem and backward induction. Let val(G'_T) be this value vector.
5. For any integer $T\geq 0$ and any real number $\lambda\in(0,1)$, we will by G_T^λ denote the finite-horizon game with T stages and payoffs discounted with a discount factor of $1-\lambda$, i.e., with payoff $\lambda\sum_{i=0}^{T}(1-\lambda)^i r_i$ to Player I. Such a game has a value vector by von Neumann's theorem and backward induction. Let val(G_T^λ) be this value vector. The proof by Shapley (1953) of G^λ having a value consists of showing that $\lim_{T\to\infty}$ val(G_T^λ) exists and that val$(G^\lambda)=\lim_{T\to\infty}$ val(G_T^λ).

By convention, for all $\lambda\in(0,1)$ we have that val$(G_0)_k=$ val$(G_0^\lambda)_k=$ val$(G'_0)_k=0$ for $k\notin\{\perp,0\}$ and val$(G_0)_{k'}=$ val$(G'_0)_{k'}=v_{k'}$, for $k'\in\{\perp,0\}$. Also, val$(G_0^\lambda)_0=\lambda\cdot v_{k'}$, for $k'\in\{\perp,0\}$.

Extended recursive games. For technical reasons, we will extend the definitions above to a more general, slightly artificial extension of the class of finite position recursive games. We call this class *extended recursive games*. An extended recursive game is defined as a recursive game, except that there is a special position called $\mathscr{M}$, such that if the current position is about to become $\mathscr{M}$ at stage t, player Min instead chooses some other non-terminal position v as the current position at stage t. In an extended recursive game G, we let N denote the number of positions but *not* counting the three special positions ($\perp$, 0, $\mathscr{M}$). It is easy to see that an extended recursive game has a limiting average value, by a reduction to the case of recursive games. It is also straightforward that

$$\mathrm{val}(G)_{\mathscr{M}} = \min_{k\in([N]\setminus\{\perp,0,\mathscr{M}\})} \mathrm{val}(G)_k.$$

Value iteration operator. For a fixed extended recursive game, its *value iteration operator* $\mathcal{T}^G : \mathbb{R}^{N+3} \to \mathbb{R}^{N+3}$ is defined as follows: The operator is such that $(\mathcal{T}^G v)_\perp = v_\perp$ and $(\mathcal{T}^G v)_0 = v_0$. If $k \in ([N] \setminus \{\mathscr{M}, \perp, 0\})$ then $(\mathcal{T}^G v)_k = \mathrm{val}(M_k(v))$, where $M_k(v)$ is the $m \times m$ matrix game, with $M_k(v)_{i,j} = v_{\pi^G(k,i,j)}$. Also, the operator is such that $(\mathcal{T}^G v)_{\mathscr{M}} = \min_{k\in([N]\setminus\{\mathscr{M},\perp,0\})}(\mathcal{T}^G v)_k$. The value of a matrix game is monotonically increasing in the value of the entries. Hence, the operator $\mathcal{T}^G$ is monotonically increasing.

For $T = 0$ and all $\lambda \in (0,1)$, the value $\mathrm{val}(G_0)_k = \mathrm{val}(G'_0)_k = \mathrm{val}(G^\lambda_0)_k = 0$ for all $k \neq 0$. Also, $\mathrm{val}(G_0)_0 = \mathrm{val}(G'_0)_0 = 1$ and $\mathrm{val}(G^\lambda_0)_0 = \lambda$. For $T > 0$, we can use the operator $\mathcal{T}^G$ to describe the value of G_T given the value of G_{T-1} as follows.

1. We have that $\mathrm{val}(G_T)_k = (\mathrm{val}(G_{T-1}))_k$ for $k \in \{\perp, 0\}$ and that $\mathrm{val}(G_T)_{k'} = \frac{T-1}{T}(\mathcal{T}^G \mathrm{val}(G_{T-1}))_{k'}$ for $k' \in \{1, \ldots, N, \mathscr{M}\}$.
2. Also similarly, for $T \geq 0$ we get that $\mathrm{val}(G'_T) = (\mathcal{T}^G \mathrm{val}(G'_{T-1}))$.
3. Again, similarly, for $T \geq 0$ and $\lambda \in (0,1)$ we get that $\mathrm{val}(G^\lambda_T)_0 = \lambda + (1-\lambda)\cdot(\mathcal{T}^G \mathrm{val}(G^\lambda_{T-1}))_0$ and that $\mathrm{val}(G^\lambda_T)_k = (1-\lambda)\cdot(\mathcal{T}^G \mathrm{val}(G^\lambda_{T-1}))_k$ for $k \neq 0$.

Less valuable. For T some integer and $\lambda \in (0,1)$ and some (extended) recursive games with positive payoffs G and $\tilde{G}$ we say that a position k in G is *(T,λ)-less valuable* than a position k' in $\tilde{G}$ if

$$\mathrm{val}(G'_T)_k \leq \mathrm{val}(\tilde{G}'_T)_{k'} \text{ and } \mathrm{val}(G_T)_k \leq \mathrm{val}(\tilde{G}_T)_{k'} \text{ and } \mathrm{val}(G^\lambda_T)_k \leq \mathrm{val}(\tilde{G}^\lambda_T)_{k'} \ .$$

We say that k is *less valuable* than k' if for all T and $\lambda \in (0,1)$ the position k is (T,λ)-less valuable than k'. We say that an (extended) recursive game with positive payoffs G is *less valuable* when another recursive game $\tilde{G}$ with the same set of positions if position k of G is less valuable than position k of $\tilde{G}$.

Slower and extremal. For T some integer and $\lambda \in (0,1)$, we say that an (extended) recursive game with positive payoffs G is *(T,λ)-slower* than another (extended) recursive game $\tilde{G}$ if

$$\max_{k\in[N]}(\mathrm{val}(\tilde{G})_k - \mathrm{val}(\tilde{G}'_T)_k) \leq \max_{k\in[N]}(\mathrm{val}(G)_k - \mathrm{val}(G'_T)_k)$$

and

$$\max_{k\in[N]}(\mathrm{val}(\tilde{G})_k - \mathrm{val}(\tilde{G}_T)_k) \leq \max_{v\in[N]}(\mathrm{val}(G)_k - \mathrm{val}(G_T)_k)$$

and

$$\max_{k\in[N]}(\mathrm{val}(\tilde{G})_k - \mathrm{val}(\tilde{G}^{\lambda}_T)_v) \leq \max_{v\in[N]}(\mathrm{val}(G)_k - \mathrm{val}(G^{\lambda}_T)_k)$$

We say that an (extended) recursive game G is *slower* than an (extended) recursive game $\tilde{G}$ if for all $T \in \mathbb{Z}$ and $\lambda \in (0,1)$ G is (T,λ)-slower than $\tilde{G}$. In practice, we are going to show that a game G is slower than a game $\tilde{G}$ (over the same set of positions) by showing that G is less valuable than $\tilde{G}$ and $\mathrm{val}(G) = \mathrm{val}(\tilde{G})$.

A recursive game G is *extremal within* S for S being a set of recursive games iff $G \in S$ and G is slower than each game $\tilde{G} \in S$.

3 The game $P(N, m)$ is extremal

We have that $P(N,m)$ is in $\mathcal{G}_{N,m}$ by Hansen, Ibsen-Jensen, and Miltersen (2011a). The remainder of the paper will focus on showing the following key lemma.

Lemma 1 *For any N, m, the game $P(N,m)$ is extremal within $\mathcal{G}_{N,m}$.*

Our key lemma implies our main result, Theorem 2. This can be seen as follows: We directly get that $P(N,m)$ maximizes $\mathrm{val}(G) - \mathrm{val}(G_T)$ and $\mathrm{val}(G) - \mathrm{val}(G'_T)$ over all games in $\mathcal{G}_{N,m}$. Furthermore, by letting T go to infinity, we get from that $P(N,m)$ maximizes $\max_{v\in[N]}(\mathrm{val}(G)_k - \mathrm{val}(G^{\lambda}_T)_k)$ over $\mathcal{G}_{N,m}$ and that $\mathrm{val}(G^{\lambda}) = \lim_{T\to\infty} \mathrm{val}(G^{\lambda}_T)$, as shown by Shapley (1953), that $P(N,m)$ maximizes $\mathrm{val}(G) - \mathrm{val}(G^{\lambda})$ over $\mathcal{G}_{N,m}$.

To prove Lemma 1 we will make use of a standard technique from extremal combinatorics pioneered by Moon and Moser (1965), by starting with an arbitrary game in $\mathcal{G}_{N,m}$ and iteratively creating a sequence of games each of which is slower than the previous game in the sequence, with the last game in the sequence being $P(N,m)$. The technique was also used in Ibsen-Jensen and Miltersen (2012). In essence, the proof goes as follows: We first show that for any game in $\mathcal{G}_{N,m}$, there is a slower game in $R_{N,m} \cap \mathcal{G}_{N,m}$ (Lemma 5) and then we show that we in such games can replace a position with the corresponding Purgatory position and get a slower game (Lemma 8). But first we will show how to order positions so that they satisfy a useful property for our proof.

For a stationary strategy profile (σ,τ), a non-terminal position $k \in [N]$ and some position $\ell \in [N]$, consider the probability measure on plays $\Omega^{\sigma,\tau}_k$ defined from Player I following σ and Player II following τ and the game having starting position k. Under that probability space we will let $F^{\sigma,\tau}_{k,\ell}$ be the event that the position of the next stage of play is position ℓ and $E^{\sigma,\tau}_{k,\ell}$ be the event that the position at some future stage is position ℓ.

Lemma 2 *Let G be a game in $\mathcal{G}_{N,m}$. Let V be some non-empty subset of the non-terminal positions. For all $1 > \epsilon > 0$, there exists some position $k \in V$ and some stationary strategy in position v for Player I, σ, such that for all stationary strategies τ for Player II*

$$\Pr(\bigcup_{\ell\in([N]\setminus V)} F^{\sigma,\tau}_{k,\ell}) > 0$$

and

$$\Pr(F_{k,\perp}^{\sigma,\tau}) / \Pr(\bigcup_{\ell \in ([N] \setminus V)} F_{k,\ell}^{\sigma,\tau}) \leq \epsilon.$$

Proof Assume not. Let σ be a some ϵ-optimal stationary strategy for Player I.

For a strategy τ for Player II and position k, let

$$h_k^\tau := \Pr(F_{k,\perp}^{\sigma,\tau}) / \Pr(\bigcup_{\ell \in ([N] \setminus V)} F_{k,\ell}^{\sigma,\tau}.)$$

To make it well-defined let h_k^τ be ∞ if $\Pr(\bigcup_{\ell \in ([N] \setminus V)} F_{k,\ell}^{\sigma,\tau}) = 0$. The number h_k^τ only depends on the strategy used when the play starts in k and we can hence find a strategy for Player II that maximizes h_k^τ for all $k \in V$. Therefore let τ^* be some strategy that maximizes h_k^τ. Let h be the smallest $h_k^{\tau^*}$ for all $k \in V$. Our assumption then implies that $h > \epsilon$.

Look at some position $\ell \in V$ and consider again the probability measure $\Omega_\ell^{\sigma,\tau^*}$. Let P be the event that a play ends in 0. Each play p that ends in 0 must have gone from some position inside V to a position outside V, since p starts inside and ends outside.

There are two cases. Either $h = \infty$ or not.

If $h = \infty$, then τ^* prevents play from leaving V. In that case $\Pr(P) = 0 < 1 - \epsilon$.

If $h < \infty$, then at the last stage, where the current position is inside V, the probability that the next position is Position $\perp$ is at least h. Hence, $\Pr(P) \leq 1 - h < 1 - \epsilon$.

In both cases $\Pr(P) < 1 - \epsilon$. But that contradicts that σ was ϵ-optimal or that v had value one. □

The quantifiers in Lemma 2 appear in the wrong order for our applications, so we will now prove a corollary, Corollary 1, where the quantifiers appear in another order.

Corollary 1 *Let G be some game in $\mathcal{G}_{N,m}$. Let V be some non-empty subset of the non-terminal positions. There exists some position $k \in V$, such that for all $1 > \epsilon > 0$, there exists some stationary strategy σ for Player I in position k such that for all stationary strategies τ for Player II*

$$\Pr(\bigcup_{\ell \in ([N] \setminus V)} F_{k,\ell}^{\sigma,\tau}) > 0$$

and

$$\Pr(F_{k,\perp}^{\sigma,\tau}) / \Pr(\bigcup_{\ell \in ([N] \setminus V)} F_{k,\ell}^{\sigma,\tau}) \leq \epsilon.$$

Proof Let $\epsilon_n = 1/n$ for $n \in \mathbb{Z}_+$. Apply Lemma 2 with ϵ_n. Let the position and strategy found be k_n and σ_n respectively.

Since the sequence has unbounded length and V is a finite set, we must have that for some $k \in V$, that $k = k_n$ for infinitely many distinct n in $\mathbb{Z}_+$.

Let $\epsilon > 0$ be fixed. Since $k = k_n$ for infinitely many distinct n in $\mathbb{Z}_+$, there must be a $n' > 1/\epsilon$, for which $k_{n'} = k$. Therefore, for the stationary strategy $\sigma_{n'}$ we have that

$$\Pr(\bigcup_{\ell \in ([N]\setminus V)} F_{k,\ell}^{\sigma_{n'},\tau}) > 0$$

and

$$\Pr(F_{k,\perp}^{\sigma_{n'},\tau}) / \Pr(\bigcup_{\ell \in ([N]\setminus V)} F_{k,\ell}^{\sigma_{n'},\tau}) \leq 1/n < \epsilon$$

for all stationary strategies τ for Player II, and we are done. □

For a game G, a pair of positions k, k' in G and a pair of actions i, j in position k, let $G[(k,i,j) \leftarrow k']$ be the game similar to G, except such that $\pi^{G[(k,i,j)\leftarrow k']}(k,i,j) = k'$.

The next lemma will be used heavily in the reminder of the paper to find less valuable games.

Lemma 3 *Let G be an (extended) recursive game. Consider some positions ℓ, ℓ' in G. Assume that $\pi^G(k,i,j) = \ell$ for some $k \in [N]$, $i \in \{1,\dots,m\}$ and $j \in \{1,\dots,m\}$. Let $\tilde{G} := G[(k,i,j) \leftarrow \ell']$.*

If ℓ' in G is less valuable than ℓ in G, then $\tilde{G}$ is less valuable than G.

Proof By definition of less valuable we can just show that:

1. If $\forall T \in \mathbb{N} : \text{val}(G'_T)_{\ell'} \leq \text{val}(G'_T)_\ell$, then
$$\forall k' \in [N], T \in \mathbb{N} : \text{val}(\tilde{G}'_T)_{k'} \leq \text{val}(G'_T)_{k'} \quad .$$
2. If $\forall T \in \mathbb{N} : \text{val}(G_T)_{\ell'} \leq \text{val}(G_T)_\ell$, then
$$\forall k' \in [N], T \in \mathbb{N} : \text{val}(\tilde{G}_T)_{k'} \leq \text{val}(G_T)_{k'} \quad .$$
3. If $\forall T \in \mathbb{N}, \lambda \in (0,1) : \text{val}(G_T^\lambda)_{\ell'} \leq \text{val}(G_T^\lambda)_\ell$, then
$$\forall k' \in [N], T \in \mathbb{N}, \lambda \in (0,1) : \text{val}(\tilde{G}_T)_{k'} \leq \text{val}(G_T)_{k'} \quad .$$

We will refer to them as Item 1 to 3 respectively.

Consider first Item 1. We will show the statement using induction in T.

For $T = 0$, we clearly have

$$\text{val}(\tilde{G}'_0) = \text{val}(G'_0).$$

For $T > 0$, we have by the induction hypothesis that $\forall k' : \text{val}(\tilde{G}'_T)_{k'} \leq \text{val}(\tilde{G}'_T)_{k'}$ and we need to show that $\forall k' : \text{val}(\tilde{G}'_{T+1})_{k'} \leq \text{val}(G'_{T+1})_{k'}$. The statement is trivial for the terminals and if we show it for all other positions it will follow for the $\mathscr{M}$ position. Thus, let k' be some position in $([N] \setminus \{\perp, 0, \mathscr{M}\})$.

There are now two cases. Either $k = k'$ or not. If not, we have that

$$\begin{aligned}\text{val}(\tilde{G}'_{T+1})_{k'} &= (\mathcal{T}^{\tilde{G}}\text{val}(\tilde{G}'_T))_{k'} \leq (\mathcal{T}^{\tilde{G}}\text{val}(G'_T))_{k'} \\ &= (\mathcal{T}^{G}\text{val}(G'_T))_{k'} = \text{val}(G'_{T+1})_{k'},\end{aligned}$$

where the inequality is by induction and by the fact that the value iteration operator is monotonically increasing. The equality just after the inequality comes from the fact that $\pi^G(k',i',j') = \pi^{\tilde{G}}(k',i',j')$ for all $i' \in \{1,\ldots,m\}, j' \in \{1,\ldots,m\}$.

For $k = k'$, we will use a similar argument.

First we will show that

$$(\mathcal{T}^{\tilde{G}} \text{val}(G'_T))_k \leq (\mathcal{T}^G \text{val}(G'_T))_k.$$

For $v \in \mathbb{R}^N$, we have that $(\mathcal{T}^G v)_k$ (resp. $(\mathcal{T}^{\tilde{G}} v)_k$) is the value of the matrix game, with matrix A (resp. A'), where $a_{i',j'} = v_{\pi^G(k,i',j')}$ (resp. $a'_{i',j'} = v_{\pi^{\tilde{G}}(k,i',j')}$) for all $i' \in \{1,\ldots,m\}, j' \in \{1,\ldots,m\}$. But $\pi^G(k,i',j') = \pi^{\tilde{G}}(k,i',j')$ for $i' \neq i$ or $j' \neq j$. Hence, the matrix is entry-wise the same except possible for the entry (i,j). But

$$\text{val}(G'_T)_{\pi^{\tilde{G}}(k,i,j)} = \text{val}(G'_T)_\ell \leq \text{val}(G'_T)_{\ell'} = \text{val}(G'_T)_{\pi^G(k,i,j)}.$$

A matrix game which is entry-wise smaller than another matrix game has smaller value. Hence,

$$(\mathcal{T}^{\tilde{G}} \text{val}(G'_T))_k \leq (\mathcal{T}^G \text{val}(G'_T))_k.$$

Now we are ready to show the remainder of the statement.

$$\text{val}(\tilde{G}'_{T+1})_k = (\mathcal{T}^{\tilde{G}} \text{val}(\tilde{G}'_T))_k \leq (\mathcal{T}^{\tilde{G}} \text{val}(G'_T))_k \leq (\mathcal{T}^G \text{val}(G'_T))_k = \text{val}(G'_{T+1})_k,$$

where the first inequality is by induction and that the value iteration operator is monotonically increasing.

Item 2 can be shown similar to Item 1, except we use that for all $k' \notin \{\perp, 0, \mathscr{M}\}$, we have that,

$$\text{val}(\tilde{G}_{T+1})_{k'} = \frac{T-1}{T} (\mathcal{T}^{\tilde{G}} \text{val}(\tilde{G}_T))_{k'},$$

instead of

$$\text{val}(\tilde{G}'_{T+1})_{k'} = (\mathcal{T}^{\tilde{G}} \text{val}(\tilde{G}'_T))_{k'}.$$

Item 3 can also be shown similar to Item 1, except we use that for all $k' \neq \{\perp, 0, \mathscr{M}\}$, we have that

$$\text{val}(\tilde{G}^\lambda_{T+1})_{k'} = (1-\lambda) \cdot (\mathcal{T}^{\tilde{G}} \text{val}(\tilde{G}^\lambda_T))_{k'},$$

instead of

$$\text{val}(\tilde{G}'_{T+1})_{k'} = (\mathcal{T}^{\tilde{G}} \text{val}(\tilde{G}'_T))_{k'}.$$

Note that the argument for the terminals is still trivial, even though the formula is different. □

Remark 1 Consider some game G, some positions ℓ, ℓ' and let $\tilde{G} := G[(k,i,j) \leftarrow \ell']$ for some k,i,j. If ℓ' in G is less valuable than ℓ' in G and $\text{val}(G)_k = \text{val}(\tilde{G})_k$, then $\tilde{G}$ is slower than G.

Alternately, consider a sequence of games $G^1, G^2, \ldots, G^\ell$, such that for all x, we have that G^{x+1} is $G^x[(k^x,i^x,j^x) \leftarrow (k^x)']$ (for some k^x, i^x, j^x and $(k^x)'$ such that $(k^x)'$ in G^x is less valuable than $\pi^{G^x}(k,i,j)$ in G^x) then G^x is less valuable than G^1 (by transitivity) and in particular, if $\text{val}(G^1)_k = \text{val}(G^x)_k$ for all k, then G^x is slower than G^1.

We next give a simple lemma that allows us to apply the idea in Remark 1.

Lemma 4 *Let G be an (extended) recursive game. Let V be some non-empty subset, not including position 0. Assume that there is a position $k \in V$, such that for all other positions (except $\mathscr{M}$ and $\perp$) $k' \in (V \setminus \{k, \mathscr{M}, \perp\})$, and all $i, j \in [m]$, we have that*

$$\pi^G(k', i, j) \in V \ \ ,$$

then for all $k' \in (V \setminus \{k\})$ (i.e. possible including $\mathscr{M}$ and $\perp$) we have that k' is less valuable than k.

Proof We will show the statement using induction in T. Hence, consider some set V not including 0 and some position k satisfying the properties required in the lemma statement. The lemma statement is true for $T = 0$, since

$$\text{val}(G'_0)_{k'} = \text{val}(G'_0)_k = \text{val}(G_0)_{k'} = \text{val}(G_0)_k = \text{val}(G^\lambda_0)_{k'} = \text{val}(G^\lambda_0)_k = 0 \ \ .$$

For $T > 0$, note that the statement is trivially true if $k' = \mathscr{M}$ by definition of $\mathscr{M}$. We thus consider that $k' \neq \mathscr{M}$. Consider the operator $\mathcal{T}^G$. We have that in the matrix game defining $(\mathcal{T}^G \text{val}(G_{T-1}))_{k'}$, the greatest entry is $\text{val}(G_{T-1})_k$ by induction. Hence, $(\mathcal{T}^G \text{val}(G_{T-1}))_{k'} \leq \text{val}(G_{T-1})_k$ The operator $\mathcal{T}^G$ is monotone increasing, so $\text{val}(G_{T-1})_k \leq \text{val}(G_T)_k$ and hence, $(\mathcal{T}^G \text{val}(G_{T-1}))_{k'} \leq \text{val}(G_T)_k$. Also, similarly, $(\mathcal{T}^G \text{val}(G^\lambda_{T-1}))_{k'} \leq \text{val}(G^\lambda_T)_k$ and $(\mathcal{T}^G \text{val}(G'_{T-1}))_{k'} \leq \text{val}(G'_T)_k$.

Therefore, for $k' \in (V \setminus \{\mathscr{M}\})$, we have

$$\text{val}(G_T)_{k'} = \frac{T}{T+1}(\mathcal{T}^G \text{val}(G_{T-1}))_{k'} \leq \frac{T}{T+1}\text{val}(G_T)_k \leq \text{val}(G_T)_k \ \ .$$

Also,

$$\text{val}(G^\lambda_T)_{k'} = (1-\lambda)(\mathcal{T}^G \text{val}(G_{T-1}))_{k'} \leq (1-\lambda)\text{val}(G_T)_k \leq \text{val}(G_T)_k$$

and lastly

$$\text{val}(G'_T)_{k'} = (\mathcal{T}^G \text{val}(G'_{T-1}))_{k'} \leq \text{val}(G'_T)_k \ \ . \quad \square$$

We will use the idea expressed in Remark 1 together with Lemma 4 to show Lemma 5 which, for any game in $\mathcal{G}_{N,m}$, finds a slower game in $\mathcal{G}_{N,m} \cap R_{N,m}$.

Lemma 5 *Let $G \in \mathcal{G}_{N,m}$. There exists a game $\tilde{G} \in \mathcal{G}_{N,m} \cap R_{N,m}$ so that $\tilde{G}$ is slower than G.*

Proof We will first renumber the positions in G. We will inductively define the k'th element after the renumbering as follows: Apply Corollary 1 with $V = ([N] \setminus \{1, \ldots, k-1\})$. Let the position identified be k and let the strategy identified be ${\sigma'}^k_\epsilon$.

We will create a graphical representation of the game as follows: Each position will correspond to a node and for all positions k and l, there is a transition from the node corresponding to k to the node corresponding to l, if $p^{kl}_{ij} > 0$ for some i, j. To illustrate this proof we will be doing a running example starting with an example of a game G on four positions, where all (non-absorbing) positions has all positions as successors. Figure 2a illustrates this G.

Let G^1 be a copy of G, except that G^1 is an extended recursive game. (That is, the action maps are identical in the two games.) There is an illustration of the game that corresponds to G^1 in Figure 2b.

This proof will construct a sequence of different recursive games found using Lemma 3 such that each game is less valuable than the last.

For all $k \geq 1$, let

$$U_k = \{\ell \mid k \leq \ell\} \cup \{\mathscr{M}\}$$

and

$$L_k = \{\ell \neq \perp | k > \ell\}.$$

The only possible successor of position k in any game G with N non-terminal positions which is in neither U_k nor L_k is then $\perp$.

We will next construct a sequence S of games, eventually ending in $\tilde{G}$, such that for each game $\widehat{G}$ in S and each position k and pair of actions i, j, we have that

$$\pi^G(k,i,j) \in U_k \Leftrightarrow \pi^{\widehat{G}}(k,i,j) \in U_k$$

and that

$$\pi^G(k,i,j) \in L_k \Leftrightarrow \pi^{\widehat{G}}(k,i,j) \in L_k$$

and that

$$\pi^G(k,i,j) = \perp \Leftrightarrow \pi^{\widehat{G}}(k,i,j) = \perp \ .$$

We will now inductively construct a sequence S^1 of games,

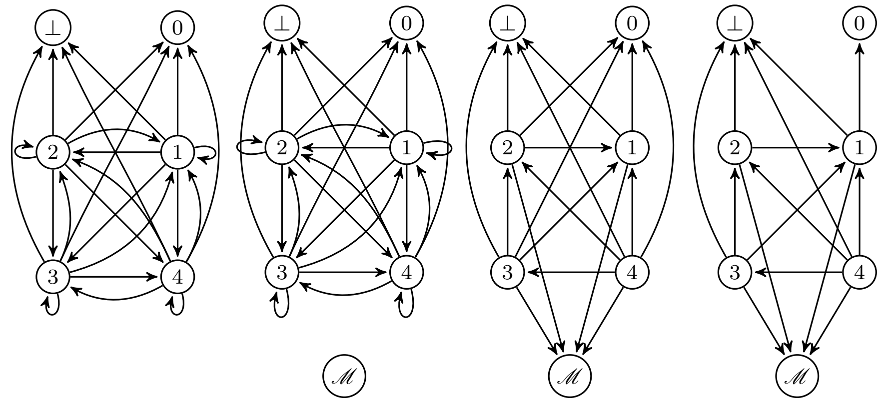
$$G^{1,0} = G^1, G^{1,1}, \ldots, G^{1,s^1} = G^2 \ .$$

We will define game $G^{1,r+1}$ from $G^{1,r}$. If there does not exist $k \leq \ell, i, j$ where $\pi^{G^{1,r}}(k,i,j) = \ell$, then $r = s^1$. Otherwise, let $G^{1,r+1}$ be the game $G^{1,r}[(k,i,j) \leftarrow \mathscr{M}]$. The position $\mathscr{M}$ is less valuable than all other non-terminal positions by definition (and hence less valuable than ℓ) and thus $G^{1,r+1}$ is less valuable than $G^{1,r}$ by Lemma 3. Note that $\ell, \mathscr{M} \in U_k$ (because $k \leq \ell$). Also, the successors of position k in game G^2, which is in U_k is $\{\mathscr{M}\}$. There is an illustration of G^2 in Figure 2c.

We will now inductively construct a sequence S^2 of games,

$$G^2 = G^{2,0}, G^{2,1}, \ldots, G^{2,s^2} = G^3 \ .$$

We will define game $G^{2,r+1}$ from $G^{2,r}$. If there does not exist $k \geq 2, i, j$ where $\pi^{G^{2,r}}(k,i,j') = 0$, then $r = s^2$. Otherwise, let $G^{2,r+1}$ be the game $G^{2,r}[(k,i,j) \leftarrow 1]$. Any position k (especially position 1) is less valuable than position 0 by definition and thus $G^{2,r+1}$ is less valuable than $G^{2,r}$ by Lemma 3. Note that $0, 1 \in L_k$ (because $k \geq 2$). Also, $k \geq 2$ does not have 0 as a successor in G^3. There is a illustration of G^3 in Figure 2d.

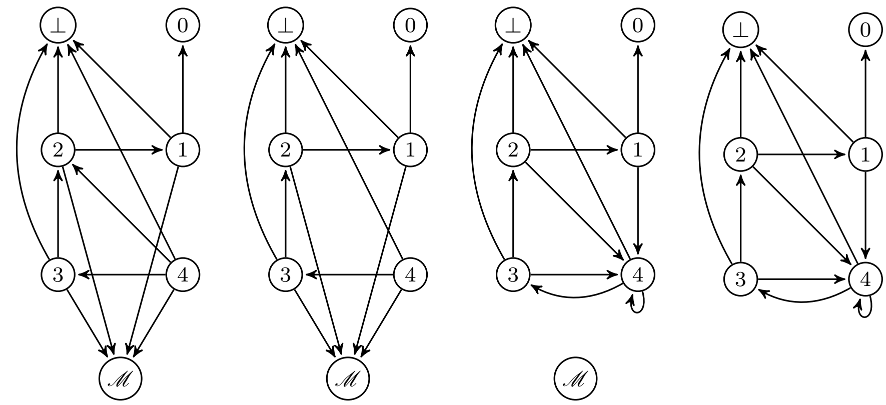


(a) The game G.

(b) The game G^1. The artificial position $\mathscr{M}$ has been added.

(c) The game G^2. We have changed each successor $\ell \in U_k$ of position k to $\mathscr{M}$.

(d) The game G^3. The terminal position 0 is not a successor for $i > 1$.



(e) The game G^4. The position 1 is not a successor for position k, for $k > 2$.

(f) The game G^{N+1}. The position k is only a successor for position $k+1$ for $k \geq 0$.

(g) The game G^*. The position N is a successor to each position k instead of $\mathscr{M}$.

(h) The game G'. The position $\mathscr{M}$ has been removed. For each $i > 0$, the position i's successors are in $\{i-1, N, \perp\}$, as desired.

Fig. 2

We will now inductively construct a sequence S^3 of games,

$$G^{3,0} = G^3, G^{3,1}, \dots, G^{3,s^3} = G^4 = G^{4,0},$$
$$G^{4,1}, \dots, G^{4,s^4} = G^5 = G^{5,0},$$
$$\dots, G^{N,s^N} = G^{N+1} \quad ,$$

as follows:

We will ensure inductively that for all $k \geq x$, we have that position k in G^x does not have a successor $< x-1$ and that we do not change position $k' < x$ again in the sequence S^3. Note that for all $k \geq 2$, we have that position k does not have 0 as a successor in G^3, as already noted, settling the base case.

We will define game $G^{x,r+1}$ from $G^{x,r}$ for all $x \geq 3$ and $r \geq 1$ and $G^{x,0} = G^{x-1,s^{x-1}}$. If there does not exist $k \geq x, i, j$ where $\pi^{G^{x,r}}(k,i,j) = x-2$, then $r = s^x$. Otherwise, let $G^{x,r+1}$ be the game $G^{x,r}[(k,i,j) \leftarrow x-3]$. The choice of $k \geq x$ ensures that no position $k' < x$ changes after game G^x in the sequence S^3. We have that position $x-2$ is less valuable than position $x-3$ by Lemma 4 (with $V = \{x-2, x-1, x, \dots, N, \mathscr{M}\}$ and $k = x-2$, using the induction hypothesis that for any $k \geq x-1$ the set $L_k^{G^{x,0}}$ does not contain any position $< x-2$ and thus also thus $L_k^{G^{x,r}}$ does not contain any position $< x-2$, since no entries has been changed to a position $< x-2$, since $G^{x,0}$). After the sequence of modifications up to game G^{x+1}, we have that for any $k \geq x$ position k does not have $x-2$ as a successor (because all such successors have been replaced by $x-3$). Since position k in G^x did not have any successors $< x-2$ (and only the successor $x-3$ has been added since then), this shows the induction hypothesis. Note that $x-2, x-3 \in L_k$ (because $k \geq x$). Also, the induction hypothesis shows that successors of position k in game G^{N+1}, which is in L_k is $\{k-1\}$. There is a illustration of G^4 in Figure 2e and of G^{N+1} in Figure 2f.

Note that G^{N+1} then ensures that the only successor in L_k for position k is position $k-1$ for all k.

We will now inductively construct a sequence S' of games,

$$G^{N+1} = G^{N+1,0}, G^{N+1,1}, \dots, G^{N+1,s^{N+1}} = G^* \quad .$$

We will define game $G^{N+1,r+1}$ from $G^{N+1,r}$. If there does not exist k, i, j where $\pi^{G^{N+1,r}}(k,i,j) = \mathscr{M}$, then $r = s^{N+1}$. Otherwise, let $G^{N+1,r+1}$ be the game $G^{N+1,r}[(k,i,j) \leftarrow N]$. Note that $\mathscr{M}, N \in L_k$. Hence, (because of S' and S^1), k only has $\{N\}$ as a successor in L_k in G^*. There is an illustration of G^* in Figure 2g.

We have that N is less valuable than $\mathscr{M}$ in $G^{N+1,r}$ for all r, as can be be seen as follows: Consider some $x \in [N-1]$. Applying Lemma 4 (with $V = \{x-1, x, \dots, N, \mathscr{M}\}$ and $k = x-1$) to $G^{N+1,r}$, we see that position $x+1$ is less valuable than position x in $G^{N+1,r}$. Thus, N is less valuable than all other non-terminal positions and thus is less valuable than $\mathscr{M}$ by definition (they are each less valuable than the other).

In game G^* we have the set of successors of position k is $\{k-1, N, \bot\}$ as wanted and thus $G^* \in R_{N,m}$. Hence, the recursive game $\tilde{G}$ (that is, without the $\mathscr{M}$ position) with the same action map as G^*, is such that, for all $T \in \mathbb{N}$ and $i \in \{1, \dots, N\}$ and

$\lambda \in (0,1)$, we have that $\text{val}((G^*)'_T)_v = \text{val}(\tilde{G}'_T)_v$ and that $\text{val}((G^*)_T)_v = \text{val}(\tilde{G}_T)_v$ and that $\text{val}((G^*)^\lambda_T)_v = \text{val}(\tilde{G}^\lambda_T)_v$. This $\tilde{G}$ is the game mentioned in the statement of the lemma. There is a illustration of $\tilde{G}$ in Figure 2h.

We have that $\tilde{G} \in R_{N,m}$ (because it was the case for G^*). Thus, we need to show that $\tilde{G}$ is slower than G and that $\tilde{G} \in \mathcal{G}_{N,m}$. We see that we have found $\tilde{G}$ from G using a sequence of games $G_1 = G, \dots, G_\ell$, such that $G_{x+1} = G_x[(k^x, i^x, j^x) \to (k^x)']$ and such that (k^x) is less valuable than $\pi^{G_x}(k,i,j)$. By Remark 1, to show that $\tilde{G}$ is slower than G we thus just need to show that $\text{val}(G) = \text{val}(\tilde{G})$. This also shows that $\tilde{G} \in \mathcal{G}_{N,m}$. Any play that reaches terminal position 0 from position N in $\tilde{G}$ has passed through all other positions in $\tilde{G}$. Therefore, position N must have the smallest value of all positions in $\tilde{G}$. Hence, for any given $\epsilon > 0$, we just need to find a strategy, σ, for Player I that ensures probability at least $1-\epsilon$ against any strategy for Player II in $\tilde{G}$ to reach terminal position 0 from position N.

We will do so by constructing a sequence of parametrized strategies $(\sigma^\epsilon_k)_{0 \le k < N}$ in $\tilde{G}$, each strategy parametrized with $\epsilon > 0$, for Player I. Each strategy σ^ϵ_k ensures that with probability at least $1-\epsilon$ that we eventually get from position N to v_k. Thus, especially, σ^ϵ_0 ensured that we get from position N to 0.

For any k and $\epsilon > 0$, remember that the strategy ${\sigma'}^k_\epsilon$ defined at the very beginning of this proof guaranteed against any strategy, τ, for Player II, that

$$\Pr(F^{{\sigma'_k}^\epsilon,\tau}_{k,\perp}) / \Pr(\bigcup_{\ell \in L_k \cup \{\perp\}} F^{{\sigma'_N}^\epsilon,\tau}_{k,\ell}) \le \epsilon.$$

The strategy ${\sigma'_i}^\epsilon$ can also be interpreted as a strategy for $\tilde{G}$, because $\tilde{G}$ has the same set of positions and actions as G. But if for a given i,j we have that $\pi^G(k,i,j) = \ell$ for some ℓ in L_k (resp. $\{\perp\}$ or U_k) then we have that $\pi^{\tilde{G}}(k,i,j) = k-1$ (resp. $\{\perp\}$ or N), because we carefully noted that if $\pi^G(k,i,j) \in L_k$ (resp. in $\{\perp\}$ and U_k), then so too for all other games in the sequence. Hence, ${\sigma'_k}^\epsilon$ must guarantee that

$$\Pr(F^{{\sigma'}^\epsilon_k,\tau}_{k,\perp}) / \Pr(\bigcup_{\ell' \in L^\cup_k \{\perp\}} F^{{\sigma'}^\epsilon_k,\tau}_{k,\ell'})) \le \epsilon.$$

But ${\sigma'_N}^\epsilon$ is precisely the type of strategy we want to be our σ^ϵ_{N-1}. Hence, we extend ${\sigma'_N}^\epsilon$ with arbitrary stationary strategies in the other positions and let the resulting strategy be σ^ϵ_{N-1}.

Therefore, assume that we know $\forall \epsilon > 0 : \sigma^\epsilon_{k+1}$ and want to find $\forall \epsilon : \sigma^\epsilon_k$ for $k \ge 0$.

Look at strategy ${\sigma'_k}^{\epsilon/2}$. For that strategy, let

$$h = \Pr(F^{{\sigma'_i}^{\epsilon/2},\tau}_{k,\perp}) \quad \wedge \quad u = \Pr(F^{{\sigma'_i}^{\epsilon/2},\tau}_{k,k-1}) > 0.$$

The expected number of stages where the position is position k before the first stage where the position is either position $\perp$ or position $k-1$ is thus $\frac{1}{h+u}$. Whenever the position of the next stage is neither position $\perp$ nor position $k-1$, the position of the next stage is position N (by construction, the successors are always in $\{k-$

$1, N, \perp\}$). Let σ_k^ϵ be the same strategy as $\sigma_{k-1}^{\epsilon/2\cdot(h+u)}$, except that distribution over the actions in k is ${\sigma'_k}^{\epsilon/2}$.

The expected probability to eventually reach $k-1$ from N if Player I follows σ_k^ϵ and Player II follows the stationary strategy τ is then at least

$$1 - \Pr(\text{not reach } k \ \frac{1}{h+u} \text{ times from position } N) - \Pr(F_{k,\perp}^{{\sigma'_k}^{\epsilon/2},\tau} | \bigcup_{\ell \in \{\perp, k-1\}} F_{k,\ell}^{{\sigma'_k}^{\epsilon/2},\tau}).$$

By the choice of parameters the expected probability to eventually reach $k-1$ from N is then at least

$$\begin{aligned}
&\Pr(E_{N,k-1}^{\sigma_k^\epsilon,\tau}) \\
&= 1 - (1 - \Pr(E_{N,k}^{\sigma_k^\epsilon,\tau})) \cdot \frac{1}{h+u} - \Pr(F_{k,\perp}^{{\sigma'_k}^{\epsilon/2},\tau} | \bigcup_{\ell \in \{\perp, k-1\})} F_{k,\ell}^{{\sigma'_k}^{\epsilon/2},\tau} \\
&\geq 1 - \frac{\epsilon/2 \cdot (h+u)}{h+u} - \Pr(F_{k,\perp}^{{\sigma'_k}^{\epsilon/2},\tau} | \bigcup_{k \in \{\perp, k-1\}} F_{v_k,\ell}^{{\sigma'_k}^{\epsilon/2},\tau}) \\
&= 1 - \epsilon/2 - \frac{\Pr(F_{k,\perp}^{{\sigma'_k}^{\epsilon/2},\tau})}{\Pr(\bigcup_{\ell \in \{\perp, k-1\}} F_{k,\ell}^{{\sigma'_k}^{\epsilon/2},\tau})} \\
&\geq 1 - \epsilon.
\end{aligned}$$

Therefore, for $\epsilon > 0$, we have that σ_1^ϵ assures Player I of expected payoff $1-\epsilon$ if the game starts in position N in $\tilde{G}$. □

We next show some properties for games in $R_{N,m}$ or in $R_{N,m} \cap \mathcal{G}_{N,m}$, which will allow us to show that each position can be changed to a position from Purgatory and the resulting game is less valuable (which is in essence all we need to show our result).

Lemma 6 *For any recursive game G in $R_{N,m}$ and any positions $k < k'$, we have that k' is less valuable than k*

Proof Any play that reaches 0, the unique position that gives non-zero reward, from k' passes through k. □

In a game $G \in R_{N,m}$ and for a position k, we say that an total ordering $\prec$ of Player I's actions in position k is *rising* if

$$\forall j \in \{1, \dots m\} \exists i \in \{1, \dots m\} : \pi^G(k, i, j) = k-1 \wedge (\forall i' \prec i : \pi^G(k, i', j) = N)$$

Note that, by the properties of $R_{N,m}$, if $\prec$ is rising and for some position k and action i for Player I and j for Player II we have $\pi^G(k, i, j) = \perp$, then there is some action $i' \prec i$ for Player I such that $\pi^G(k, i', j) = k-1$.

Lemma 7 *Let $G \in R_{N,m} \cap \mathcal{G}_{N,m}$ and let k be some position in G. There exists a rising total order $\prec$ on Player I's actions in position k.*

Proof Consider some game $G \in R_{N,m} \cap \mathcal{G}_{N,m}$. The proof will be by contradiction. Consider a position k in G for which no rising total order exists.

Consider some strategy σ ensuring outcome $1-\epsilon$ for play starting in N, for $\epsilon < 1/m$. Since $G \in \mathcal{G}_{N,m}$ such a strategy exists. Let $\prec'$ be some arbitrary total ordering on the actions in position k, such that $i \prec' i'$ if $\sigma(k)(i) > \sigma(k)(i')$. The total ordering $\prec'$ is not rising (since no rising total order exists in position k) and thus there exists some actions i, j such that $\pi^G(k,i,j) = \bot$ and such that $\pi^G(k,i',j) = N$ for all $i' \prec' i$. Especially, i is an action played with highest probability for which $\pi^G(k,i',j) \in \{\bot, k-1\}$. Consider an arbitrary strategy τ for Player II that plays j in k. Consider a play starting in k that reaches a terminal, while the players follows σ and τ respectively. For the play to reach 0, it must have eventually gotten to $k-1$ from k, i.e. event $E^{\sigma,\tau}_{k,k-1}$ happens. We have that

$$\Pr[E^{\sigma,\tau}_{k,k-1}] \leq \Pr[F^{\sigma,\tau}_{k,k-1} \mid F^{\sigma,\tau}_{k,k-1} \cup F^{\sigma,\tau}_{k,\bot}] \ ,$$

because play cannot leave $\bot$. We have that

$$\Pr[F^{\sigma,\tau}_{k,k-1} \mid F^{\sigma,\tau}_{k,k-1} \cup F^{\sigma,\tau}_{k,\bot}] = 1 - \Pr[F^{\sigma,\tau}_{k,\bot} \mid F^{\sigma,\tau}_{k,k-1} \cup F^{\sigma,\tau}_{k,\bot}] = 1 - \frac{a}{b}$$

where $a = \Pr[F^{\sigma,\tau}_{k,\bot}] \geq \sigma(k)(i)$ and

$$\begin{aligned} b = \Pr[F^{\sigma,\tau}_{k,k-1} \cup F^{\sigma,\tau}_{k,\bot}] &= \sum_{i' \mid \pi^G(k,i',j) \in \{\bot,k-1\}} \sigma(k)(i') \\ \leq \sum_{i' \mid \pi^G(k,i',j) \in \{\bot,k-1\}} \sigma(k)(i) &\leq m \cdot \sigma(k)(i) \end{aligned}$$

(we have that $\sigma(k)(i') \leq \sigma(k)(i)$, because i is an action played with highest probability for which $\pi^G(k,i',j) \in \{\bot, k-1\}$ as already argued). Thus $\frac{a}{b} \geq \frac{1}{m}$ and thus the strategy σ ensures at most $1 - 1/m$, contradicting that it ensured $1-\epsilon$ for $\epsilon < 1/m$. □

For a game G, a position k with m actions for each player and an matrix $M \in \{0, 1, \ldots, N, \bot\}^{m \times m}$, let $G[k \leftarrow M]$ be the game similar to G except that

$$\pi^{G[k \leftarrow M]}(k,i,j) = M_{i,j} \ .$$

The next lemma will in essence allow us to change each position, one at a time, to the corresponding Purgatory position.

Lemma 8 *Let G be a game in $R_{N,m}$ and let k be a position in G with a rising total order. Let $\tilde{G}$ be the game $G[k \leftarrow \pi^{P(N,m)}(k)]$. Then $\tilde{G}$ is less valuable than G.*

Proof Let $\prec^k$ be a rising total order for position k. For any recursive game $\bar{G}$ and any action j for Player II, let $f^{\bar{G}}(j)$ be the smallest action i such that $\pi^{\bar{G}}(k,i,j) = k-1$ (or ∞ if such an i does not exists). Let A' be the subset of Player I's actions $\{f^{\tilde{G}}(j) | j \in \{1, \ldots, m\}\}$. Let $\prec$ be the same ordering as the total ordering $\prec^k$, except that for two actions i and i', if $i \in A'$ and $i' \notin A'$, then $i \prec i'$.

$$\begin{bmatrix} N & k-1 & k-1 & k-1 & N & N \\ N & k-1 & \bot & k-1 & N & N \\ N & N & \bot & k-1 & k-1 & N \\ k-1 & k-1 & \bot & \bot & \bot & k-1 \\ N & N & N & N & N & N \\ N & \bot & N & N & N & N \end{bmatrix}$$

(a) The matrix $\pi^G(k)$.

$$\begin{bmatrix} k-1 & k-1 & k-1 & N & N & N \\ N & \bot & k-1 & k-1 & N & N \\ k-1 & \bot & \bot & \bot & k-1 & k-1 \\ k-1 & \bot & k-1 & N & N & N \\ N & N & N & N & N & N \\ \bot & N & N & N & N & N \end{bmatrix}$$

(b) The matrix $\pi^H(k)$.

$$\begin{bmatrix} k-1 & k-1 & k-1 & N & N & N \\ \bot & \bot & \bot & k-1 & N & N \\ \bot & \bot & \bot & \bot & k-1 & k-1 \\ \bot & \bot & \bot & \bot & \bot & \bot \\ \bot & \bot & \bot & \bot & \bot & \bot \\ \bot & \bot & \bot & \bot & \bot & \bot \end{bmatrix}$$

(c) The matrix $\pi^{G^1}(k)$.

$$\begin{bmatrix} k-1 & k-1 & k-1 & N & N & N \\ k-1 & k-1 & k-1 & N & N & N \\ k-1 & k-1 & k-1 & N & N & N \\ \bot & \bot & \bot & k-1 & N & N \\ \bot & \bot & \bot & \bot & k-1 & k-1 \\ \bot & \bot & \bot & \bot & k-1 & k-1 \end{bmatrix}$$

(d) The matrix $\pi^{G^2}(k)$.

$$\begin{bmatrix} k-1 & k-1 & k-1 & N & N & N \\ \bot & k-1 & k-1 & N & N & N \\ \bot & \bot & k-1 & N & N & N \\ \bot & \bot & \bot & k-1 & N & N \\ \bot & \bot & \bot & \bot & k-1 & k-1 \\ \bot & \bot & \bot & \bot & \bot & k-1 \end{bmatrix}$$

(e) The matrix $\pi^{G^3}(k)$.

$$\begin{bmatrix} k-1 & N & N & N & N & N \\ \bot & k-1 & N & N & N & N \\ \bot & \bot & k-1 & N & N & N \\ \bot & \bot & \bot & k-1 & N & N \\ \bot & \bot & \bot & \bot & k-1 & N \\ \bot & \bot & \bot & \bot & \bot & k-1 \end{bmatrix}$$

(f) The matrix $\pi^{G^4}(k)$.

Fig. 3

Let H be the same game as G, except that Player I's actions are renumbered such that $\prec$ is the standard ordering and that Player II actions are renumbered such that $f^{H,k}$ is (weakly) increasing. (The change from $\tilde{G}$ to H ensures that $f^{H,k}(j) + 1 \geq f^{H,k}(j+1)$).

For a particular G, there is an illustration of $\pi^G(k, i, j)$ in Figure 3a and an illustration of $\pi^H(k, i, j)$ in Figure 3b for the corresponding game H.

We will now inductively construct a sequence S of games

$$G^{0,0} = H, G^{0,1}, \ldots, G^{0,\ell^0} = G^1 \quad .$$

We will define game $G^{0,r+1}$ from $G^{0,r}$. If there does not exist i, j, such that $i > f^{G^{0,r}}(j)$ and $\pi^{G^{0,r}}(k, i, j) \neq \bot$, then $r = \ell^0$. Otherwise, let $G^{0,r+1}$ be the game $G^{0,r}[(k, i, j) \leftarrow \bot]$. Each position k' is less valuable than position $\bot$ and thus $G^{0,r+1}$ is less valuable than $G^{0,r}$ by Lemma 3. There is an illustration of $\pi^{G^1}(k, i, j)$ in Figure 3c, when π^G is the map illustrated in Figure 3a. Note that $f^H(j) = f^{G^1}(j)$ for all j, because all changes are for $i > f^{G^{0,r}}(j)$.

For a recursive game $\bar{G}$, a position k and some action i for Player I in position k, let $S_i^{\bar{G},k}$ be the set $\{j|\pi^{\bar{G}}(k,i,j)=k-1\}$. Let G^2 be the same game as G^1, except that Player I's action set in position k consists of $\left|S^i(\bar{G},k)\right|$ instances of action i for all i. The actions for Player I are ordered such that all instances of action i for Player I are before any instance of action $i+1$. There is an illustration of $\pi^{G^2}(k,i,j)$ in Figure 3d, when $\pi^{\tilde{G}}$ is the map illustrated in Figure 3a.

Clearly, G^2 is less valuable than G^1, because Player I has a subset of the actions in position k in G^2 compared to in G^1 (he has lost each action i for which $\left|S^i(G^1,k)\right|=0$) and Player I is trying to maximize the values in the definition of less valuable.

We see the following:

1. Player I has m actions in position k of G^2, because $\sum_i \left|S_i^{G^1}\right|=m$ (which comes from there being exactly one $k-1$ in each column of the matrix $\pi^{G^1}(k)$)
2. $\pi^{G^2}(k,i,i)=k-1$ for all i, because $f^{H,k}$ is weakly increasing (and hence $f^{G^1,k}(j)$) and because there is exactly one $k-1$ in each column
3. $\pi^{G^2}(k,i,j)\in\{k-1,N\}$ for all $j<i$, because, (1) if $\pi^{G^1}(k,i,j)=\bot$ then so too is $\pi^{G^1}(k,i,j')=\bot$ for all $j'>j$, because of the changes over the sequence S and (2) $\pi^{G^2}(k,i,i)=k-1$ by item 2 and the actions before action i of G^2 was instances of actions in G^1 before the action in G^1 of which i is an instance of in G^2

We will refer to these as Property 1 to 3 of G^2.

We will now inductively construct a sequence S' of games

$$G^{2,0}=G^2,G^{2,1},\ldots,G^{2,\ell^2}=G^3 \ .$$

We will define game $G^{2,r+1}$ from $G^{2,r}$. If there does not exist i,j, such that $j>i$ and $\pi^{G^{2,r}}(k,i,j)\neq\bot$, then $r=\ell^2$. Otherwise, let $G^{2,r+1}$ be the game $G^{2,r}[(k,i,j)\leftarrow\bot]$. (It is not difficult to see that $\pi^{G^{2,r}}(k,i,j)\in\{\bot,k-1\}$ for such i,j, but it is not important for the proof.) Each position k' is less valuable than position $\bot$ and thus $G^{2,r+1}$ is less valuable than $G^{2,r}$ by Lemma 3. There is an illustration of $\pi^{G^3}(k,i,j)$ in Figure 3e, when $\pi^{\tilde{G}}$ is the map illustrated in Figure 3a.

We will next inductively construct another sequence S'' of games

$$G^{3,0}=G^3,G^{3,1},\ldots,G^{3,\ell^3}=G^4 \ .$$

We will define game $G^{3,r+1}$ from $G^{3,r}$. If there does not exist i,j, such that $j<i$ and $\pi^{G^{3,r}}(k,i,j)\neq N$, then $r=\ell^2$. Otherwise, let $G^{3,r+1}$ be the game $G^{3,r}[(k,i,j)\leftarrow N]$. We have that $\pi^{G^2}(k,i,j)\in\{k-1,N\}$ by Property 3 of G^2 and since we only change entries for which $i<j$ to N it is also true in $G^{3,r}$ and thus $G^{3,r+1}$ is less valuable than $G^{3,r}$ by Lemma 6. There is an illustration of $\pi^{G^4}(k,i,j)$ in Figure 3f, when $\pi^{\tilde{G}}$ is the map illustrated in Figure 3a.

We thus have that

$$\pi^{G^4}(k,i,j)=\begin{cases}k-1 & \text{if } i=j\\ \bot & \text{if } i>j\\ N & \text{if } i<j \ .\end{cases}$$

(The first case by Property 2 of G^2 since we never changed such entries and the other two by the changes over S' and S'' respectively.) Hence, $G^4 = G[k \leftarrow \pi^{P(N,m)}(k)] = \tilde{G}$. We also see using transitivity of less valuable that $G^4 = \tilde{G}$ is less valuable than G, which completes the proof. □

We are now ready to show the key lemma, Lemma 1.

Lemma 1 *For all N, m, the game $P(N,m)$ is extremal within $\mathcal{G}_{N,m}$.*

Proof Let G be a game in $\mathcal{G}_{N,m}$. By Lemma 5, there is a slower game $\tilde{G}$ in $R_{N,m} \cap \mathcal{G}_{N,m}$. By Lemma 7, there exists a rising total order for each position k in $\tilde{G}$. We can then construct the sequence of games $\tilde{G} = G^0, G^1, \ldots, G^N$, such that $G^{x+1} = G^x[x+1 \leftarrow \pi^{P(N,m)}(k)]$ for all x. We have that G^{x+1} is less valuable than G^x by Lemma 8 (because game G^x has only modified the positions $\leq x$ as compared to $\tilde{G}$, there is still a rising total order for position $x+1$ in G^x). Also, G^N is $P(N,m)$. Hence, $P(N,m)$ is less valuable than $\tilde{G}$. As shown by Hansen, Ibsen-Jensen, and Miltersen (2011a) $P(N,m)$ is in $\mathcal{G}_{N,m}$. Thus, $P(N,m)$ is slower than $\tilde{G}$ and hence by transitivity slower than G. Therefore, $P(N,m)$ is slower than all games in $\mathcal{G}_{N,m}$ and thus extremal within $\mathcal{G}_{N,m}$. □

Acknowledgements The author would like to thank Peter Bro Miltersen for promoting the paper and for long discussions about it.

The paper was supported by the Sino-Danish Center for the Theory of Interactive Computation, funded by the Danish National Research Foundation and the National Science Foundation of China (under the grant 61061130540); and by the Center for research in the Foundations of Electronic Markets (CFEM), supported by the Danish Strategic Research Council. Also supported by the Austrian Science Fund (FWF) NFN Grant no. S11407-N23 (RiSE/SHiNE) and ERC Starting grant (279307: Graph Games).

References

S. Basu, R. Pollack, and M. Roy. *Algorithms in Real Algebraic Geometry*. Springer, 2nd edition, 2006.

Krishnendu Chatterjee and Rasmus Ibsen-Jensen. The value 1 problem under finite-memory strategies for concurrent mean-payoff games. In *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2015, San Diego, CA, USA, January 4-6, 2015*, pages 1018–1029, 2015.

H. Everett. Recursive games. In H. W. Kuhn and A. W. Tucker, editors, *Contributions to the Theory of Games Vol. III*, volume 39 of *Annals of Mathematical Studies*. Princeton University Press, 1957.

K.A. Hansen, M. Koucký, and P.B. Miltersen. Winning concurrent reachability games requires doubly exponential patience. In *Proc. of IEEE Symp. on Logic in Comp. Sci., LICS*, pages 332–341, 2009.

K.A. Hansen, R. Ibsen-Jensen, and P.B. Miltersen. The complexity of solving reachability games using value and strategy iteration. In *6th Int. Comp. Sci. Symp. in Russia, CSR*, LNCS. Springer, 2011a.

K.A. Hansen, M. Koucký, N. Lauritzen, P.B. Miltersen, and E.P. Tsigaridas. Exact algorithms for solving stochastic games. In *Proc. 43rd Annual ACM Symp. Theory of Computing (STOC)*, 2011b.

Kristoffer Arnsfelt Hansen, Rasmus Ibsen-Jensen, and Peter Bro Miltersen. The complexity of solving reachability games using value and strategy iteration. *Theor. Comp. Sys.*, 55(2):380–403, August 2014. ISSN 1432-4350. doi: 10.1007/s00224-013-9524-6. URL `http://dx.doi.org/10.1007/s00224-013-9524-6`.

R. Ibsen-Jensen. *Strategy complexity of two-player, zero-sum games*. PhD thesis, Aarhus University, 2013.

Rasmus Ibsen-Jensen and Peter Bro Miltersen. Solving simple stochastic games with few coin toss positions. In *ESA*, pages 636–647, 2012.

J. F. Mertens and A. Neyman. Stochastic games. *International Journal of Game Theory*, 10:53–66, 1981.

Emanuel Milman. The semi-algebraic theory of stochastic games. *Mathematics of Operations Research*, 27(2):401–418, 2002.

J. Moon and L. Moser. On cliques in graphs. *Israel Journal of Mathematics*, 3:23–28, 1965. ISSN 0021-2172.

L. S. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences, U. S. A.*, 39:1095–1100, 1953.