A Potential Reduction Algorithm for Two-Person Zero-Sum Mean Payoff Stochastic Games

Endre Boros1 · Khaled Elbassioni2 · Vladimir Gurvich1,3 · Kazuhisa Makino4

Published online: 8 July 2016
© Springer Science+Business Media New York 2016

Abstract We suggest a new algorithm for two-person zero-sum undiscounted stochastic games focusing on stationary strategies. Given a positive real $\varepsilon$, let us call a stochastic game $\varepsilon$-ergodic, if its values from any two initial positions differ by at most $\varepsilon$. The proposed new algorithm outputs for every $\varepsilon > 0$ in finite time either a pair of stationary strategies for the two players guaranteeing that the values from any initial positions are within an $\varepsilon$-range, or identifies two initial positions $u$ and $v$ and corresponding stationary strategies for the players proving that the game values starting from $u$ and $v$ are at least $\varepsilon/24$ apart. In particular, the above result shows that if a stochastic game is $\varepsilon$-ergodic, then there are stationary strategies for the players proving $24\varepsilon$-ergodicity. This result strengthens and provides a constructive...
version of an existential result by Vrieze (Stochastic games with finite state and action spaces. PhD thesis, Centrum voor Wiskunde en Informatica, Amsterdam, 1980) claiming that if a stochastic game is 0-ergodic, then there are \( \varepsilon \)-optimal stationary strategies for every \( \varepsilon > 0 \).

The suggested algorithm is based on a potential transformation technique that changes the range of local values at all positions without changing the normal form of the game.

**Keywords** Undiscounted stochastic games · Limiting average payoff · Mean payoff · Local reward · Potential transformation · Computational game theory

1 Introduction

1.1 Basic Concepts and Notation

Stochastic games were introduced by Shapley [24] for the discounted case, and extended to the undiscounted case by Gillette [13]. Each such game \( \Gamma = (p_{k\ell}^{vu}, r_{k\ell}^{vu} \mid k \in K^v, \ell \in L^v, u, v \in V) \) is played by two players on a finite set \( V \) of vertices (states, or positions); \( K^v \) and \( L^v \) for \( v \in V \) are finite sets of actions (pure strategies) of the two players; \( p_{k\ell}^{vu} \in [0,1] \) is the transition probability from state \( v \) to state \( u \) if players chose actions \( k \in K^v \) and \( \ell \in L^v \) at state \( v \in V \); and \( r_{k\ell}^{vu} \in \mathbb{R} \) is the reward player 1 (the maximizer) receives from player 2 (the minimizer), corresponding to this transition. We assume that the game is non-stopping, that is, \( \sum_{u \in V} p_{k\ell}^{vu} = 1 \) for all \( v \in V \) and \( k \in K^v, \ell \in L^v \). To simplify later expressions, let us denote by \( P^{vu} \in [0,1]^{K^v \times L^v} \) the transition matrix, the elements of which are the probabilities \( p_{k\ell}^{vu} \), and associate in \( \Gamma \) a local expected reward matrix \( A^v \) to every \( v \in V \) defined by

\[
(A^v)_{k\ell} = \sum_{u \in V} p_{k\ell}^{vu} r_{k\ell}^{vu}.
\]

In the game \( \Gamma \), players first agree on an initial vertex \( v_0 \in V \) to start. Then, in a general step \( j = 0, 1, \ldots, \) when the game arrives to state \( v_j = v \in V \), they choose mixed strategies \( \alpha^v \in \Delta(K^v) := \{ y \in \mathbb{R}^{K^v} \mid \sum_{i \in K^v} y_i = 1, y_i \geq 0 \text{ for } i \in K^v \} \) and \( \beta^v \in \Delta(L^v) \), player 1 receives the amount of \( b_j = \alpha^v A^v \beta^v \) from player 2, and the game moves to the next state \( u \) chosen according to the transition probabilities \( p_{k\ell}^{vu} = \alpha^v P^{vu} \beta^v \).

The **undiscounted limiting average (effective) payoff** is the Cesaro average

\[
g^{v_0}(\Gamma) = \liminf_{N \to \infty} \frac{1}{N + 1} \sum_{j=0}^{N} \mathbb{E}\left[b_j\right],
\]

where the expectation is taken over all random choices made (according to mixed strategies and transition probabilities) up to step \( j \) of the play. The purpose of player 1 is to maximize \( g^{v_0}(\Gamma) \), while player 2 would like to minimize it.

In 1981, Mertens and Neymann [18] in their seminal paper proved that every stochastic game has a value from any initial position in terms of history-dependent strategies. An example (the so-called Big Match) showing that the same does not hold when restricted to stationary strategies was given in 1957 in Gillette’s [13] paper; see also [3].

In this paper, we shall restrict ourselves (and the players) to the so-called **stationary** strategies, that is, the mixed strategy chosen in a position \( v \in V \) can depend only on \( v \) but not on the preceding positions or moves before reaching \( v \) (i.e., not on the history of the play).
We will denote by \( \mathcal{K}(\Gamma) \) and \( \mathcal{L}(\Gamma) \) the sets of stationary strategies of WHITE and BLACK, respectively, that is,

\[
\mathcal{K}(\Gamma) = \bigotimes_{v \in V} \Delta (K^v) \quad \text{and} \quad \mathcal{L}(\Gamma) = \bigotimes_{v \in V} \Delta (L^v).
\]

Vrieze [25] showed that if a stochastic game \( \Gamma \) has a value \( g^{v_0}(\Gamma) = m \), which is a constant, independent of the initial state \( v_0 \in V \), then it has a value in \( \epsilon \)-optimal stationary strategies for any \( \epsilon > 0 \). We call such games ergodic and extend their definition as follows.

**Definition 1** For \( \epsilon > 0 \), a stochastic game \( \Gamma \) is said to be \( \epsilon \)-ergodic if the game values from any two initial positions differ by at most \( \epsilon \), that is, \(|g^v(\Gamma) - g^u(\Gamma)| \leq \epsilon\), for all \( u, v \in V \). A 0-ergodic game will be simply called ergodic.

Our main result in this paper is an algorithm that decides, for any given stochastic game \( \Gamma \), whether or not \( \Gamma \) is \( \epsilon \)-ergodic, and provides a certificate for its \( \epsilon \)-ergodicity/non-ergodicity. As a corollary, we get a constructive proof of the above-mentioned theorem of Vrieze [25]. A notion central to our algorithm is the concept of a potential transformation introduced in the following section.

### 1.2 Potential Transformations

In 1958, Gallai [12] suggested the following simple transformation. Let \( x : V \to \mathbb{R} \) be a mapping that assigns to each state \( v \in V \) a real number \( x^v \) called the potential of \( v \). For every transition \((v, u)\) and pair of actions \( k \in K^v \) and \( \ell \in L^v \), let us transform the payoff \( r_{k\ell}^v \) as follows:

\[
r_{k\ell}^v(x) = r_{k\ell}^v + x^v - x^u.
\]

Then, the one step expected payoff amount changes to \( E[b_j(x)] = E[b_j] + E[x^v] - E[x^{v+1}] \), where \( v_j \in V \) is the (random) position reached at step \( j \) of the play. However, as the sum of these expectations telescopes, the limiting average payoff remains the same for all finite potentials:

\[
g^{v_0}(\Gamma(x)) = g^{v_0}(\Gamma) + \lim_{N \to \infty} \frac{1}{N} E \left[ x^{v_0} - x^{v_N} \right] = g^{v_0}(\Gamma).
\]

Thus, the transformed game remains equivalent with the original one.

Using potential transformations, we may be able to obtain a proof for ergodicity/non-ergodicity. This is made more precise in the following section.

### 1.3 Local and Global Values and Concepts of Ergodicity

Let us consider an arbitrary potential \( x \in \mathbb{R}^V \) and define the local value \( m^v(x) \) at position \( v \in V \) as the value of the \( |K^v| \times |L^v| \) local reward matrix game \( A^v(x) \) with entries

\[
d_{k\ell}^v(x) = \sum_{a \in V} p_{k\ell}^v (r_{k\ell}^v + x^v - x^u), \quad \text{for all} \ k \in K^v, \ \ell \in L^v, \quad (3)
\]

that is,

\[
m^v(x) = \text{Val} (A^v(x)) := \max_{\alpha^v \in \Delta(K^v)} \min_{\beta^v \in \Delta(L^v)} \alpha^v A^v(x) \beta^v = \min_{\beta^v \in \Delta(L^v)} \max_{\alpha^v \in \Delta(K^v)} \alpha^v A^v(x) \beta^v.
\]

To a pair of stationary strategies \( \alpha = (\alpha^v|v \in V) \in \mathcal{K}(\Gamma) \) and \( \beta = (\beta^v|v \in V) \in \mathcal{L}(\Gamma) \), we associate a Markov chain \( M_{\alpha,\beta}(\Gamma) \) on states in \( V \), defined by the transition probabilities
Given an undiscounted zero-sum stochastic game, we try to reduce the range of its local values by a potential transformation \( x \in \mathbb{R}^V \). If they are equalized by some potential \( x \), that is, \( m^v(x) = m \) is a constant for all \( v \in V \), we say that the game is brought to its ergodic canonical form [4]. In this case, one can show that the values \( g^v \) exist and are equal to \( m \) for all initial positions \( v \in V \), and furthermore, locally optimal strategies are globally optimal [4]. Thus, the game is solved in uniformly optimal strategies. However, typically we are not that lucky.

To state our main theorem, we need more notation.

- \( W > 0 \) is smallest integer s.t. either \( p^{vu}_{k\ell} = 0 \) or \( p^{vu}_{k\ell} \geq 1/W \)
- \( R \) is the smallest real s.t.

\[
0 \leq r^{vu}_{k\ell} \leq R
\]

- \( N = \max_{v \in V} \{ \max(|K^v|, |L^v|) \} \).
- \( n = |V| \).
- \( \eta = \max \{ \log_2 R, \log_2 W \} \) (maximum “bit length”)

**Theorem 1** For every stochastic game and \( \varepsilon > 0 \), we can find in \( O(2^n n N W R \varepsilon) \) time either a potential vector \( x \in \mathbb{R}^V \) proving that the game is \( (24\varepsilon) \)-ergodic, or stationary strategies for the players proving that it is not \( \varepsilon \)-ergodic.

Note that in the case when the game \( \Gamma \) is not \( 24\varepsilon \)-ergodic, the algorithm in Theorem 1 finds two states \( u \) and \( v \) and stationary strategies \( \alpha \) and \( \beta \) for the two players such that \( \alpha \) guarantees a value at least \( b \) for player 1 starting from state \( u \), while \( \beta \) guarantees a value at most \( a \) for player 2 starting from state \( v \), and such that \( b - a > \varepsilon \).

The proof of Theorem 1 will be given in Sect. 4. One major hurdle that we face is that the range of potentials can grow doubly exponentially as iterations proceed, leading to much worse bounds than those stated in the theorem. To deal with this issue, we use quantifier elimination techniques [2,14,23] to reduce the range of potentials after each iteration; see the discussion preceding Lemma 9.

**Remark 1** Vrieze [25] showed the existence of \( \varepsilon \)-optimal stationary strategies in ergodic stochastic games (but without algorithmic constructions) by considering the discounted version and comparing the undiscounted and discounted values, as the discount factor approaches zero.
1. It seems plausible that this existential result can be turned into an algorithm by analyzing how the difference between the discounted and undiscounted values change with the discount factor. This has been applied in the case of stochastic games with perfect information \[1, 26\]. However, typically the discount factor has to be exponentially close to 1; see, for example, \[5\]. Since our pumping algorithm is very simple and provides a self-contained proof of Theorem 1, we will not require any discounting in this paper.

2 Related Work

The above definition of ergodicity follows Moulin’s concept of the ergodic extension of a matrix game \[21\] (which is a very special example of a stochastic game with perfect information). Let us note that slightly different terminology is used in the Markov chain theory; see, for example, \[17\].

The following four algorithms for undiscounted stochastic games are based on stronger “ergodicity type” conditions: The strategy iteration algorithm by Hoffman and Karp \[16\] requires that for any pair of stationary strategies of the two players, the obtained Markov chain has to be irreducible; two value iteration algorithms by Federgruen \[11\] are based on similar but slightly weaker requirements; the recent algorithm of Chatterjee and Ibsen-Jensen \[10\] assumes a weaker requirement than the strong ergodicity required by Hoffman and Karp \[16\]: they call a stochastic game almost surely ergodic if for any pair of (not necessarily stationary) strategies of the two players, and any starting position, some strongly ergodic class (in the sense of \[16\]) is reached with probability 1.

While these restrictions apply to the structure of the game, our ergodicity definition only restricts the value. Moreover, the results in \[16\] and \[10\] apply to a game that already satisfies the ergodicity assumption, which seems to be hard to check. Our algorithm, on the other hand, always produces an answer, regardless whether the game is ergodic or not.

Interestingly, potentials appear in \[11\] implicitly, as the differences in local values of positions, as well as in \[16\], as the dual variables to linear programs corresponding to the controlled Markov processes, which appear when a player optimizes his strategy against a given strategy of the opponent. Yet, the potential transformation is not considered explicitly in these papers.

We prove Theorem 1 by an algorithm that extends the approach recently obtained for ergodic stochastic games with perfect information \[6\] and extended to the general (not necessarily ergodic) case in \[7\]. This approach is also somewhat similar to the first of two value iteration algorithms suggested by Federgruen \[11\], though our approach has some distinct characteristics: It is assumed in \[11\] that the values \(g^v\) exist and are equal for all \(v\); in particular, this assumption implies the \(\epsilon\)-ergodicity for every \(\epsilon > 0\). For our approach, we do not need such an assumption. We can verify \(\epsilon\)-ergodicity for an arbitrary given \(\epsilon > 0\), or provide a proof for non-ergodicity (with a small gap) in a finite time. Moreover, while the approach of \[11\] was only shown to converge, we provide a bound in terms of the input parameters for the number of steps.

Several other algorithms for solving undiscounted zero-sum stochastic games in stationary strategies are surveyed by Raghavan and Filar; see Sections 4 (B) and 5 in \[22\]. The only algorithmic results that we are aware of that provide bounds on the running time for approximating the value of general (undiscounted) stochastic games are those given in \[9, 15\]: in \[9\], the authors provide an algorithm that approximates, within any factor of \(\epsilon > 0\), the value of any stochastic game (in history-dependent strategies) in time \((nN)^nN\) \(\log \frac{1}{\epsilon}\). In
Lemma 1 We have $T_{\tau+1} \subseteq T_\tau$, $B_{\tau+1} \subseteq B_\tau$, and $M_{\tau+1} \supseteq M_\tau$ for all iterations $\tau = 0, 1, \ldots$

Proof Indeed, by (1) and (3) we can conclude that $m^v(x_\tau) \geq m^- + \delta$ holds for all $v \in P_\tau$. Analogously, by (2) and (3) $m^v(x_\tau) < m^- + 3\delta$ follows for all $v \notin P_\tau$. □

Lemma 2 Either $T_\tau = \emptyset$ or $B_\tau = \emptyset$ for some finite $\tau$, or there are nonempty disjoint subsets $I, F \subseteq S$, $I \supseteq T_\tau, F \supseteq B_\tau$, and a threshold $\tau_0$, such that for every real $\Delta \geq 0$, there exists a finite index $\tau(\Delta) \geq \tau_0$ such that

(a) $m^v(x_\tau) \geq m^- + 2\delta$ for all $v \in I$ and $m^v(x_\tau) < m^- + 2\delta$ for all $v \in F$, and for all $\tau \geq \tau_0$. 

3 Pumping Algorithm

We begin by describing our procedure on an abstract level. Then, we specialize it to stochastic games in Sect. 4.

Given a subset $S \subseteq V$, let us denote by $e_S \in \{0, 1\}^V$ the characteristic vector of $S$.

Let us further assume that $m^v(x)$ for $v \in V$ are functions depending on potentials $x \in \mathbb{R}^n$ (where $n = |V|$) and satisfying the following properties for all subsets $S \subseteq V$ and reals $\delta \geq 0$:

1. $m^v(x - \delta e_S)$ is a monotone decreasing function of $\delta$ if $v \in S$;
2. $m^v(x - \delta e_S)$ is a monotone increasing function of $\delta$ if $v \notin S$;
3. $|m^v(x) - m^v(x - \delta e_S)| \leq \delta$ for all $v \in V$.

We show in this section that under the above conditions, we can change iteratively the potentials to some $x' \in \mathbb{R}^n$ such that either all values $m^v(x')$, $v \in V$, are very close to one another or we can find a decomposition of the states $V$ into disjoint subsets proving that such convergence of the values is not possible.

Our main procedure is described in Algorithm 2 below. Given the current vector of potentials $x_\tau$ at iteration $\tau$, the procedure partitions the set of vertices into four sets according to the local value $m^v(x)$. If either the first (top) set $T_\tau$ or forth (bottom) set $B_\tau$ is empty, the procedure terminates; otherwise, the potentials of all the vertices in the first and second sets are reduced by the same amount $\delta$, and the computation proceeds to the next iteration.

We can show next that properties (1), (2) and (3) above guarantee some simple properties for the above procedure.

Lemma 1 We have $T_{\tau+1} \subseteq T_\tau$, $B_{\tau+1} \subseteq B_\tau$, and $M_{\tau+1} \supseteq M_\tau$ for all iterations $\tau = 0, 1, \ldots$

Proof Indeed, by (1) and (3) we can conclude that $m^v(x_\tau) \geq m^- + \delta$ holds for all $v \in P_\tau$. Analogously, by (2) and (3) $m^v(x_\tau) < m^- + 3\delta$ follows for all $v \notin P_\tau$. □

Lemma 2 Either $T_\tau = \emptyset$ or $B_\tau = \emptyset$ for some finite $\tau$, or there are nonempty disjoint subsets $I, F \subseteq S$, $I \supseteq T_\tau, F \supseteq B_\tau$, and a threshold $\tau_0$, such that for every real $\Delta \geq 0$, there exists a finite index $\tau(\Delta) \geq \tau_0$ such that

(a) $m^v(x_\tau) \geq m^- + 2\delta$ for all $v \in I$ and $m^v(x_\tau) < m^- + 2\delta$ for all $v \in F$, and for all $\tau \geq \tau_0$. 

Birkhäuser
Proof By Lemma 1, sets $T_\tau$ and $B_\tau$ can change only monotonically, and hence only at most $|S|$ times. Thus, if \textsc{Pump}$(x, S)$ does not stop in a finite number of iterations, then after a finite number of iterations the sets $T_\tau$ and $B_\tau$ will never change and all positions in $T_\tau$ remain always pumped (that is, have their potentials reduced), while all positions in $B_\tau$ will be never pumped again.

Assuming now that the pumping algorithm \textsc{Pump}$(x, S)$ does not terminate, let us define the subset $I \subseteq S$ as the set of all those positions which are always pumped with the exception of a finite number of iterations. Analogously, let $F$ be the subset of all those positions that are never pumped with the exception of a finite number of iterations. Since $I$ and $F$ are finite sets, there must exist a finite $\tau_0$ such that for all $\tau \geq \tau_0$ we have $I \subseteq P_\tau$ and $F \cap P_\tau = \emptyset$, implying (a). Note that any vertex in $T_\tau$ is always pumped by (3) and hence $T_\tau \subseteq I$ for any $\tau \geq \tau_0$; similarly, $B_\tau \subseteq F$ for any $\tau \geq \tau_0$.

Let us next observe that all positions not in $I \cup F$ are both pumped and not pumped infinitely many times. Thus, since $\delta$ is a fixed constant, for every $\Delta$ there must exist an iteration $\tau(\Delta) \geq \tau_0$ such that all positions not in $I$ are not pumped by at least $\Delta/\delta$ many more times than those in $I$, and all positions not in $F$ are pumped by at least $\Delta/\delta$ many more times than those in $F$, implying (b) and (c).

Let us next describe the use of \textsc{Pump}$(x, S)$ for repeatedly shrinking the range of the $m^v$ values, or to produce some evidence that this is not possible. A simplest version is the following:

Note that by our above analysis, \textsc{RepeateDPumping} either returns a potential transformation for which all $m^v$, $v \in V$ values are within an $\varepsilon$-band, or returns the sets $I$ and $F$ as in Lemma 2 with arbitrary large potential differences from the other positions. In the next section, we use a modification of these procedures for stochastic games and show that these large potential differences can be used to prove that the game is not $\varepsilon$-ergodic.
Algorithm 2 REPEATEDPUMPING(ε)
1: Initialize \( h := 0 \), and \( x_h := 0 \in \mathbb{R}^V \).
2: Set \( m^+(h) := \max_{v \in V} m^v(x_h) \) and \( m^-(h) := \min_{v \in V} m^v(x_h) \).
3: If \( m^+(h) - m^-(h) \leq \varepsilon \) then STOP.
4: \( x_{h+1} := \text{PUMP}(x_h, V); h := h + 1 \).
5: Goto step 2.

Fig. 1  Stochastic game with two states \( V = \{u, v\} \) and with 2–2 actions for each of the players in both states. Arcs in the figure indicate transitions, and the corresponding probabilities are all 1, except in state \( u \) when both players choose their second actions, when the game remains in state \( u \) with probability \( p_{u,u}^{u,v} = \frac{1}{2} \) and moves to state \( v \) with probability \( p_{u,v}^{u,v} = \frac{1}{2} \). Local rewards are indicated next to the arcs. For instance, if player 1 chooses action 1 in state \( u \) and player 2 also chooses his action 1, then the game stays in state \( u \) with probability \( p_{1,1}^{u,u} = 1 \) and the local reward is \( r_{1,1}^{u,u} = 1 \). If player two changes his mind and chooses action 2, then the game moves to state \( v \) with probability \( p_{1,2}^{u,v} = 1 \), and the corresponding local reward is \( r_{1,2}^{u,v} = \frac{12}{2} \). The corresponding \( A^u \) and \( A^v \) matrix entries are indicated in the picture, as computed according to formula (3) (with \( x^u = x^v = 0 \))

4 Application of Pumping for Stochastic Games

We show in this section how to use REPEATEDPUMPING to find potential transformations verifying \( \varepsilon \)-ergodicity, or proving that the game is not \( \varepsilon \)-ergodic, thus establishing a proof of Theorem 1. Toward this end, we shall give some necessary and sufficient conditions for \( \varepsilon \)-non-ergodicity and consider a modified version of the pumping algorithm described in the previous section which will provide a constructive proof for the above theorem. It will be instructive to look first at some examples.

4.1 Examples for PUMP(\( x, S \))

We give two examples to illustrate the algorithm.

Example 1  Let us consider, as a simple example, the game in Fig. 1. Assuming that the initial potential vector is \( x_0 = (0, 0) \), we get
Fig. 2 Stochastic game of Fig. 1 with the updated local rewards after one iteration of PUMP($x, S$). The corresponding $A^u(x_1)$ and $A^v(x_1)$ matrix entries are indicated in the picture, as computed according to formula (3).

$$A^u(x_0) = \begin{pmatrix} 1 & 12 \\ 6 & 8 \end{pmatrix} \quad \text{and} \quad A^v(x_0) = \begin{pmatrix} 6 & 0 \\ 0 & 3 \end{pmatrix}.$$  

The optimal strategies for the players in state $u$ are $\alpha^u = (0, 1)$ and $\beta^u = (1, 0)$, while in state $v$, they are $\alpha^v = beta^v = \left( \frac{4}{11}, \frac{7}{11} \right)$. Thus, we get $m^u(x) = 6$ and $m^v(x) = 2$.

When running PUMP($x, S$) for this game, with $S = V$, we have $m^+ = 6$, $m^- = 2$, and consequently $\delta = 1$. Thus, in iteration $\tau = 0$, we get

$$T_0 = \{u\}, \quad B_0 = \{v\} \quad \text{and} \quad P_0 = \{u\}.$$  

Consequently, we update the potentials to $x^u_1 = -1$ and $x^v_1 = 0$, and proceed to iteration $\tau = 1$. The updated local rewards are shown in Fig. 2.

For the updated potential vector $x_1 = (-1, 0)$, we get

$$A^u(x_1) = \begin{pmatrix} 1 & 11 \\ 5 & 7.5 \end{pmatrix} \quad \text{and} \quad A^v(x_1) = \begin{pmatrix} 7 & 0 \\ 0 & 4 \end{pmatrix}.$$  

The optimal mixed strategies for the players in state $u$ are still $\alpha^u = (0, 1)$ and $\beta^u = (1, 0)$, while in state $v$ they are $\alpha^v = \beta^v = \left( \frac{4}{11}, \frac{7}{11} \right)$. Thus, we get $m^u(x_1) = 5$ and $m^v(x_1) = \frac{28}{11}$, and consequently we still have

$$T_1 = \{u\}, \quad B_1 = \{v\} \quad \text{and} \quad P_1 = \{u\}.$$  

Consequently, we update the potentials to $x^u_2 = -1 - \delta = -2$ and $x^v_2 = 0$, and proceed to iteration $\tau = 2$. The updated local rewards are shown in Fig. 3.

For the updated potential vector $x_2 = (-2, 0)$ we get

$$A^u(x_2) = \begin{pmatrix} 1 & 10 \\ 4 & 7 \end{pmatrix} \quad \text{and} \quad A^v(x_2) = \begin{pmatrix} 8 & 0 \\ 0 & 5 \end{pmatrix}.$$
The optimal mixed strategies for the players in state $u$ are still $\alpha^u = (0, 1)$ and $\beta^u = (1, 0)$, while in state $v$ they are $\alpha^v = \beta^v = \left(\frac{5}{13}, \frac{8}{13}\right)$. Thus, we get $m^u(x_2) = 4$ and $m^v(x_2) = \frac{40}{13}$, and consequently, we have

$$T_2 = \emptyset, \quad B_2 = \emptyset \quad \text{and} \quad P_2 = \{u\}.$$ 

Thus, $\text{PUMP}(x, S)$ stops after two iterations and outputs a game equivalent with the game of Fig. 1, with the reduced $m$-range of $4 - \frac{40}{13} = \frac{12}{13} < 1$.

**Example 2** Gilette [13] introduced the so-called Big Match game, see Fig. 4, to illustrate that the value in stationary strategies from some initial state may not exist in a stochastic game. In this game $\Gamma$, one can show that with stationary strategies player 1 can guarantee only 0 by choosing, for instance $(1, 0)$ as his strategy in state $u$, i.e.,

$$0 = \max_{\alpha \in \mathcal{X}(\Gamma)} \min_{\beta \in \mathcal{L}(\Gamma)} g^u(\alpha, \beta),$$

while player 2 can only guarantee $1/2$ by choosing $(1/2, 1/2)$ as his strategy in state $u$, implying that $1/2 = \min_{\beta \in \mathcal{L}(\Gamma)} \max_{\alpha \in \mathcal{X}(\Gamma)} g^v(\alpha, \beta)$.

Let us see how $\text{PUMP}(x, S)$ behaves when it is applied to this game. Assuming that the initial potential vector is $x_0 = (0, 0)$, we have

$$A^u(x_0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

while for all potentials $x$, we have

$$A^v(x) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad A^w(x) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad m^w(x) = 1 \quad \text{and} \quad m^u(x) = 0.$$ 

The optimal strategies for the players in state $u$ are $\alpha^u = \beta^u = \left(\frac{1}{2}, \frac{1}{2}\right)$. Thus, we get $m^u(x) = \frac{1}{4}$.

When running $\text{PUMP}(x, S)$ for this game, with $S = V$, we have $m^+ = 1$, $m^- = 0$, and consequently $\delta = \frac{1}{4}$. Thus, in iteration $\tau = 0$ we get

$$T_0 = \{v\}, \quad B_0 = \{w\} \quad \text{and} \quad P_0 = \{u, v\}.$$
Fig. 4  This game (Big Match) has three states \( V = \{1, 2, 3\} \), and zero-one transition probabilities. State 1 has a \( 2 \times 2 \) payoff matrix, the other two states are absorbing with payoffs equal to 1 and 0, respectively. Arrows in the picture indicate the nonzero transitions, e.g., if player 1 chooses the first row in state \( u \), and player 2 chooses the second column, then player 1 earns a local payoff of 0 and then the game returns to state \( u \); while if player 1 chooses the second row and player 2 chooses the first column then player 1 earns 0 and then the game moves to state \( w \), where it gets stuck forever, providing an effective payoff of 0 for player 1.

Consequently, we update the potentials to \( x_1^u = x_1^v = -\frac{1}{4} \) and \( x_1^w = 0 \), and proceed to iteration \( \tau = 1 \).

For the updated potential vector \( x_1 = (\frac{1}{4}, \frac{1}{4}, 0) \) we get

\[
A^u(x_1) = \begin{pmatrix}
0 & 1 \\
-\frac{1}{4} & 1 \\
\end{pmatrix}.
\]

The optimal mixed strategies for the players in state \( u \) are now \( \alpha^u = (\frac{5}{9}, \frac{4}{9}) \) and \( \beta^u = (\frac{4}{9}, \frac{5}{9}) \). Thus, we get \( m^u(x_1) = \frac{4}{9} \), and consequently, we have

\( T_1 = \{v\}, \quad B_1 = \{w\} \quad \text{and} \quad P_1 = \{v\} \).

Next, we update the potential of \( v \) to \( x_2^v = -\frac{1}{2} \), while \( x_2^u = -\frac{1}{4} \) and \( x_2^w = 0 \) stay unchanged, and proceed to iteration \( \tau = 2 \).

For the updated potential vector \( x_2 = (\frac{1}{4}, \frac{1}{4}, 0) \) we get

\[
A^u(x_2) = \begin{pmatrix}
0 & 1 \\
-\frac{1}{4} & \frac{5}{4} \\
\end{pmatrix}.
\]

The optimal mixed strategies for the players in state \( u \) are now \( \alpha^u = (\frac{3}{5}, \frac{2}{5}) \) and \( \beta^u = (\frac{1}{5}, \frac{1}{5}) \), and we get \( m^u(x_2) = \frac{1}{2} \). Consequently, we have again, as in iteration \( \tau = 0 \),

\( T_2 = \{v\}, \quad B_2 = \{w\} \quad \text{and} \quad P_2 = \{u, v\} \).

In general, it is easy to verify by induction that we will have

\[
x_\tau = \left( \frac{1}{4} \left\lfloor \frac{\tau + 1}{2} \right\rfloor, \frac{1}{4} \tau, 0 \right)
\]

\[
A^u(x_\tau) = \begin{pmatrix}
0 & 1 \\
-\frac{1}{4} \left\lfloor \frac{\tau + 1}{2} \right\rfloor & 1 + \frac{1}{4} \left\lfloor \frac{\tau}{2} \right\rfloor \\
\end{pmatrix}.
\]
and the corresponding optimal mixed strategies for the players in state $u$ are

$$
\alpha^u = \left( \frac{1 + \frac{1}{4} \tau}{2 + \frac{1}{4} \tau}, \frac{1}{2 + \frac{1}{4} \tau} \right) \\
\beta^u = \left( \frac{1 + \frac{1}{4} \tau}{2 + \frac{1}{4} \tau}, \frac{1 + \frac{1}{4} \tau}{2 + \frac{1}{4} \tau} \right),
$$

resulting in the value $m^u(x_\tau) = \frac{1 + \frac{1}{4} \tau}{2 + \frac{1}{4} \tau} \in (\frac{1}{4}, \frac{1}{2}]$. Consequently, we will have

$$
T_\tau = \{v\}, \quad B_\tau = \{w\} \quad \text{and} \quad P_\tau = \{u, v\} \quad \text{if} \ \tau \ \text{is even}
$$
$$
T_\tau = \{v\}, \quad B_\tau = \{w\} \quad \text{and} \quad P_\tau = \{v\} \quad \text{if} \ \tau \ \text{is odd}.
$$

Thus, as $\tau$ increases, $u$ keeps oscillating between the middle level of the range $m^- + 2\delta = \frac{1}{2}$ and some level below it that approaches $\frac{1}{2}$ as $\tau \to \infty$. However, the $m$-range does not shrink as we always have $m^u(x_\tau) = 0$ and $m^v(x_\tau) = 1$. After sufficiently large $\tau$, we obtain stationary strategies $\alpha^u \approx (1, 0)$ and $\beta^u \approx (\frac{1}{2}, \frac{1}{2})$ that are not optimal for the game $\Gamma$.

4.2 Some Useful Lemmas

Let us first observe that the local value function of stochastic games satisfies the properties required to run the pumping algorithm described in the previous section.

**Lemma 3** For every subset $S \subseteq V$ and $\delta \geq 0$ and for all $v \in V$, we have

$$
\begin{align*}
m^v(x) &\geq m^v(x - \delta e_S) \geq m^v(x) - \delta \max_{k, \ell} \sum_{u \notin S} p_{k\ell}^{vu} \text{ if } v \in S, \\
m^v(x) &\leq m^v(x - \delta e_S) \leq m^v(x) + \delta \max_{k, \ell} \sum_{u \in S} p_{k\ell}^{vu} \text{ if } v \notin S.
\end{align*}
$$

Furthermore, the value functions $m^v(x)$ for $v \in V$ satisfy properties (1), (2) and (3) stated in Sect. 3.

**Proof** According to (3), we must have for all $\delta \geq 0$ that $A^v(x) \geq A^v(x - \delta e_S)$ for all $v \in S$ and $A^v(x) \leq A^v(x - \delta e_S)$ for all $v \notin S$ proving properties (1) and (2) (Indeed, $A^v(x - \delta e_S) = A^v(x) - \delta (E^v - \sum_{u \in S} P^{vu})$ for $v \in S$ and $A^v(x - \delta e_S) = A^v(x) + \delta \sum_{u \in S} P^{vu}$ for $v \notin S$, where $E^v$ is the $|K^v| \times |L^v|$-matrix of all ones. Since the operator $\text{Val}(B)$ is monotone increasing in $B$, inequalities (7) follow. Property (3) follows directly from (7). \hfill \Box

The above lemma implies that procedures PUMP and REPEATEDPUMPING could, in principle, be used to find a potential transformation yielding an $\epsilon$-ergodic solution. It does not offer, however, a way to discover $\epsilon$-non-ergodicity. Toward this end, we need to find some sufficient and algorithmically achievable conditions for $\epsilon$-non-ergodicity.

Let us first analyze (0-)non-ergodicity of stochastic games (in stationary strategies).

**Lemma 4** A stochastic game is non-ergodic if and only if it is $\epsilon$-non-ergodic for some positive $\epsilon$.

**Proof** A stochastic game is non-ergodic by definition if there exists a threshold $\sigma$, positions $v, u \in V$, and stationary strategies $\alpha$ and $\beta$ for the players, such that no matter what other strategy $\beta'$ player 2 chooses the Markov chain resulting by fixing $(\alpha, \beta')$ has a value $> \sigma$ when using initial position $v_0 = v$ (guaranteeing for player 1 more than $\sigma$ from $v$), and the Markov chain obtained by fixing $(\alpha', \beta)$ has a value $< \sigma$ when using initial position $v_0 = u$.\hfill \Box
(guaranteeing for player 2 > σ from u). Since strategies α' and β' are chosen from a compact space, the above implies that there are σ' > σ > σ'' such that α guarantees for player 1 at least σ' from the initial position v, and β guarantees for player 2 at most σ'' from initial position u. Hence, the game is ε-non-ergodic for any ε < σ' − σ''.

**Lemma 5** A stochastic game Γ is ε-non-ergodic if there exist disjoint non-empty subsets of the positions I, F ⊆ V, reals a, b with b − a ≥ ε, stationary strategies αν, v ∈ I, for player 1, and βν, u ∈ F, for player 2, and a vector of potentials x ∈ ℝV, such that

\[
\begin{align*}
\text{(N1)} & \quad \alpha \uparrow \nu \nu \beta \downarrow \nu \uparrow 0 \quad \text{for all } v \in I, u \notin I, k \in K^v \text{ and } \ell \in L^v, \\
\text{(N2)} & \quad \beta \uparrow \nu \nu \beta \downarrow \nu \uparrow 0 \quad \text{for all } u \in F, w \notin F, \ell \in L^u \text{ and } k \in K^u, \text{ and} \\
\text{(N3)} & \quad \text{for all } v \in I \text{ and } u \in F:
\end{align*}
\]

\[
- \min_{\beta \in \Delta(L^v)} (\alpha \uparrow \nu \nu \beta) \geq b \quad \text{and} \quad \max_{\alpha \nu \in \Delta(K^v)} (\alpha \uparrow \nu \nu \beta) \leq a.
\]

**Proof** Let us note that (N1) and (N3) imply that for all strategies β′ ∈ \(\mathcal{L}(\Gamma)\) of player 2, the pair of strategies (\(\hat{\alpha}, \hat{\beta}\)), where \(\hat{\alpha} \uparrow \nu \nu \hat{\beta} \uparrow \nu \downarrow \nu \) for v ∈ I and \(\hat{\alpha} \nu \in \Delta(K^v)\) is chosen arbitrarily for v ∈ I, results in a Markov chain in which subset I induces one or more absorbing sets (that is, \(p_{\hat{\alpha} \nu}^{\hat{\beta} \nu} = 0\)), and in which all positions have values at least b. Analogously, (N2) and (N3) imply that F will always induce an absorbing set with values less than a, if we fix any pair of strategies (α′, β′), where α′ is any strategy in \(\mathcal{X}(\Gamma)\), \(\hat{\beta} \uparrow \nu \nu \hat{\beta} \downarrow \nu \uparrow \) for v ∈ F and \(\hat{\beta} \nu \in \Delta(L^v)\) is chosen arbitrarily, for v ∈ F. Hence, choosing any positions v ∈ I and u ∈ F and strategies \(\hat{\alpha} \) and \(\hat{\beta} \) provides a certificate for the ε-nongenericity of \(\Gamma\). (Here, we use the well-known fact [20] that, to each player’s stationary strategy, there is a best response of the opponent which is also stationary.)

4.3 Discovering ε-Ergodicity Using Repeated Pumping

Let us introduce a notation for denoting upper bounds on the entries of the matrices, more precisely on the part of these entries which do not depend on negative potential differences. Specifically, define

\[
\begin{align*}
\tilde{\alpha}^v_{k\ell}(x) & = \sum_{u \in V} p^\nu_{k\ell}^{vu} r^\nu_{k\ell} + \sum_{u \in V, x \leq u} p^\nu_{k\ell}(x - u) \\
\tilde{b}^v_{k\ell}(x) & = m^+(x) - \sum_{u \in V} p^\nu_{k\ell}^{vu} r^\nu_{k\ell} - \sum_{u \in V, x \geq u} p^\nu_{k\ell}(x - u)
\end{align*}
\]

where, as before, \(m^+(x) := \max_v m^+_v(x)\), \(m^-(x) := \min_v m^+_v(x)\). Define further

\[
\begin{align*}
R^v(x) & = \max_{k \in K^v, \ell \in L^v} \left( \tilde{\alpha}^v_{k\ell}(x) \right) \quad \text{if } m^v(x) \geq \frac{m^+(x) + m^-(x)}{2}, \\
R^v(x) & = \max_{k \in K^v, \ell \in L^v} \left( \tilde{b}^v_{k\ell}(x) \right) \quad \text{otherwise}.
\end{align*}
\]

Note that

\[
m^+(x) - \tilde{b}^v_{k\ell}(x) \leq a^v_{k\ell}(x) \leq \tilde{\alpha}^v_{k\ell}(x) \quad \text{for all } v \in V, k \in K^v, \ell \in L^v \text{ and } x \in \mathbb{R}^V,
\]

which implies

\[
\begin{align*}
m^v(x) & \leq R^v(x) \quad \text{if } m^v(x) \geq \frac{m^+(x) + m^-(x)}{2}, \quad \text{for all } v \in V \text{ and } x \in \mathbb{R}^V, \\
m^v(x) & \geq m^+(x) - R^v(x) \quad \text{otherwise}.
\end{align*}
\]

With this notation we can state a more constructive version of Lemma 5.

**Lemma 6** A stochastic game Γ satisfying (6) is ε-non-ergodic if there exist disjoint non-empty subsets I, F ⊆ V, a vector of potentials x ∈ ℝV, and reals a′, b′ ∈ [0, m^+(x)] with

\[
b' - a' \geq 3\varepsilon, a' < \frac{m^+(x) + m^-(x)}{2}, \quad b' \geq \frac{m^+(x) + m^-(x)}{2},
\]

such that
Proof. We first show that (N4)–(N5) imply the existence of strategies \( \alpha^v \), for \( v \in I \), satisfying (N1) and (N3). We shall then observe that a similar argument can be applied to (N4) and (N6) to show the existence of strategies \( \beta^u \), for \( u \in F \), such that those satisfy (N2) and (N3). Consequently, our claim will follow by Lemma 5.

Let us now fix a position \( v \in I \) and denote, respectively, by \( \tilde{\alpha}^v \) and \( \tilde{\beta}^v \) the optimal strategies of players with respect to the payoff matrix \( A^v(x) \). Denote further by \( \hat{\beta}^v = v/L(x) = (1, 1, \ldots, 1) \) the uniform strategy for player 2, and set \( \tilde{K}^v = \{ k \in K^v | \sum_{u \notin I} \sum_{\ell \in L^u} p_{k\ell}^{vu} = 0 \} \).

Let us then note that we have

\[
(A^v(x)\hat{\beta}^v)_k \leq \begin{cases} R^v(x) & \text{if } k \in \tilde{K}^v, \\ R^v(x) - \frac{R^v(x)^2}{\epsilon} & \text{otherwise}, \end{cases}
\]

since at least one of the entries of (N5) has at least \( \frac{R^v(x)}{W} \) as a coefficient in rows which are not in \( \tilde{K}^v \).

Note that \( b' > 0 \) implies by (10) that \( R^v(x) > 0 \). Thus, by the optimality of \( \tilde{\alpha} \) and by the above inequalities, we have

\[
0 < b' \leq m^v(x) \leq \tilde{\alpha}^v A^v(x)\hat{\beta}^v \leq R^v(x) - \left( \sum_{k \notin \tilde{K}^v} \tilde{\alpha}^v \right) \frac{R^v(x)^2}{\epsilon}
\]

implying that \( \sum_{k \notin \tilde{K}^v} \tilde{\alpha}^v < \frac{\epsilon}{R^v(x)} \). Since by (N4), we have \( 0 < a' \), inequalities \( \epsilon < a' + 3 \epsilon \leq b' < m^v(x) \leq R^v(x) \) follow, and hence, \( \frac{3 \epsilon}{R^v(x)} < 1 \) must hold, implying that the set \( \tilde{K}^v \) is not empty. Let us then denote by \( \tilde{\alpha}^v \) the truncated strategy defined by

\[
\tilde{\alpha}^v = \begin{cases} \tilde{\alpha}^v & \text{if } k \in \tilde{K}^v, \\ 0 & \text{if } k \notin \tilde{K}^v. \end{cases}
\]

With this, we have for any \( \tilde{\beta}^v \in \Delta(L^v) \)

\[
b' \leq m^v(x) \leq (\tilde{\alpha}^v A^v(x)\tilde{\beta}^v) = (\tilde{\alpha}^v A^v(x)\tilde{\beta}^v) \left( \sum_{k \in \tilde{K}^v} \tilde{\alpha}^v_k \right) + \sum_{k \notin \tilde{K}^v} \tilde{\alpha}^v_k \left( \sum_{\ell \in L^v} a^v_{k\ell}(x)\tilde{\beta}^v_\ell \right)
\]

\[
\leq (\tilde{\alpha}^v A^v(x)\tilde{\beta}^v) + \left( \sum_{k \notin \tilde{K}^v} \tilde{\alpha}^v_k \right) R^v(x) + \epsilon.
\]

Let us then define \( \alpha^v = \tilde{\alpha}^v \) and repeat the same for all \( v \in I \). Then, these strategies satisfy (N1) and (N3) with \( b = b' - \epsilon \).

Let us next note that by adding a constant to a matrix game, it changes its value with exactly the same constant. Furthermore, multiplying all entries by \(-1\) and transposing it, changes its value by a factor of \(-1\), interchanges the roles of row and column players, but leaves otherwise optimal strategies still optimal. Thus, we can repeat the above arguments for the matrices \( B^u(x) = m^+(x)E^u - A^u(x)^T \), where \( E \) is the \( |L^u| \times |K^u| \)-matrix of all ones, and obtain the same way strategies \( \beta^u, u \in F \) satisfying (N2) and (N3) with \( a = a' + \epsilon \). This completes the proof of the lemma.
To create a finite algorithm to find sets $I$ and $F$ and potentials satisfying (N4)-(N6), we need to do some modifications in our procedures.

First, we allow a more flexible partitioning of the $m$-range by allowing the $m$-range boundaries to be passed as parameters and replacing line 1 in procedure PUMP by

1: Initialize $\delta := (m^+ - m^-)/4$, $x_0 = x$, and $\tau = 0$.

Next, let us replace in procedure PUMP, line 6 by the following lines, where $\varepsilon > 0$ is a prespecified parameter, and call the new procedure with these modifications ModifiedPump($\varepsilon$, $x$, $S$, $m_-, m_+$):

7a: Otherwise, set $P_\tau := \{v \in S \mid m^v(x_\tau) \geq m^- + 2\delta\}$ and compute

$$R^v_\tau := \max_{k \in K^v, t \in L^v} \left(\tilde{a}_{k, t}(x_\tau)\right) \text{ if } v \in P_\tau,$$

$$R^v_\tau := \max_{k \in K^v, t \in L^v} \left(\tilde{b}_{k, t}(x_\tau)\right) \text{ if } v \notin P_\tau,$$

where $\tilde{a}$ and $\tilde{b}$ are defined by (8).

7b: Create an auxiliary directed graph $G = (V, E)$ on vertex set $V$ such that $(v, u) \in E$ iff

$$x^v_\tau - x^u_\tau < \frac{|L^v|W(R^v_\tau)^2}{\varepsilon} \text{ if } v \in P_\tau,$$

$$x^v_\tau - x^u_\tau < \frac{|K^v|W(R^v_\tau)^2}{\varepsilon} \text{ if } v \notin P_\tau.$$

7c: Find subsets $I_\tau$ and $F_\tau$ of $V$ such that $T_\tau \subseteq I_\tau \subseteq P_\tau$, $B_\tau \subseteq F_\tau \subseteq V \setminus P_\tau$, and no arcs are leaving these sets in $G$ (this can be done by a finding the strong components of $G$, or by the method described in the proof of Theorem 1).

7d: If such sets are found STOP and output these sets, otherwise continue with step 7.

Before starting to analyze this modified pumping algorithm, let us observe that we have for all iterations

$$m^- < m^- + \frac{\varepsilon}{2} < \frac{m^- + m^+}{2} \leq m^v(x_\tau) \leq R^v_\tau \text{ for all } v \in P_\tau \quad (11)$$

as long as $m^+ - m^- > \varepsilon$.

**Lemma 7** Procedure ModifiedPump($\varepsilon$, $x$, $S$) terminates in a finite number of steps.

**Proof** Let us observe that by Lemma 2 procedure PUMP would either terminate with $T_\tau = B_\tau = \emptyset$ for some finite $\tau \geq \tau_0$, or there exist sets $I = I_\tau$, and $F = F_\tau$ satisfying conditions (b) and (c) of the lemma, for $\Delta = NWQ^2/\varepsilon$, where $N = \max\{\max\{|K^v|, |L^v|\} : v \in I \cup F\}$, and $Q = \max\{R^v_\tau(\Delta) : v \in I \cup F\}$. Thus, in the latter case, ModifiedPump will indeed find some sets $I_\tau$ and $F_\tau$, and hence terminate for some finite $\tau$.

**Lemma 8** Procedure ModifiedPump($\varepsilon$, $x$, $F$) either shrinks the $m$-range by a factor of $3/4$ or outputs potentials $x = x_\tau$ and sets $I = I_\tau$ and $F = F_\tau$ which satisfy conditions (N4)-(N6) with $\alpha' < b'$.\[\Box\]

**Proof** When the procedure terminates without shrinking the $m$-range, then it outputs sets $I = I_\tau$ and $F = F_\tau$ such that in the auxiliary graph $G$, there are no arcs leaving these sets. Since $I \subseteq P_\tau$ and $F \subseteq V \setminus P_\tau$, condition (N4) holds with $\alpha' = \max_{v \notin P_\tau} m^v(x_\tau) < b' = (m^+ - m^-)/2$. Furthermore, the lack of leaving arcs in $G$ implies that for all $(v, u)$, $v \in I$ and $u \notin I$ and also for all $(u, v)$ with $u \in F$ and $v \notin F$ we must have the reverse inequalities in step 7b above, implying that conditions (N5) and (N6) hold.\[\Box\]
Let us observe that the bounds and strategies obtained by Lemmas 7 and 8 do not necessarily imply the \( \varepsilon \)-non-ergodicity of the game since those positions in \( I_\tau \) and \( F_\tau \) may not have enough separation in \( m \)-values (i.e., the condition \( b' - a' \geq 3\varepsilon \) in Lemma 6 is not satisfied). To fix this, we need to make one more use of the pumping algorithm, as described in the ModifiedRepeatedPumping procedure below. After each range-shrinking in this algorithm, we use a routine called REDUCEPOTENTIAL which takes the current potential vector \( x \) and range \([m_-, m_+]\) and produces another potential vector \( y \) such that \( \|y\|_\infty \leq 2^{\text{poly}(n,N,n)} \). We need to this because, as the algorithm proceeds, the potentials, and hence the transformed rewards, might grow doubly exponentially high.

The potential reduction can be done as follows. We write the following quadratic program in the variables \( x' \in \mathbb{R}^V, \alpha = (\alpha^v \mid v \in V) \in \mathcal{K}(\Gamma) \), and \( \beta = (\beta^v \mid v \in V) \in \mathcal{L}(\Gamma) \):  
\[
\begin{align*}
\alpha^v A^v(x') &\geq m_- \cdot e, \\
\alpha^v e &\leq 1, \\
\alpha^v e &\geq 0, \\
A^v(x') \beta^v &\leq m_+ \cdot e, \\
\beta^v e &\geq 1, \\
\beta^v e &\leq 0, \\
\end{align*}
\]
for all \( v \in V \), where \( e \) denotes the vector of all ones of appropriate dimension. This is a quadratic system of at most \( 6N \) (in)equalities on at most \((2N + 1)n\) variables. Moreover, the system is feasible since the original potential vector \( x \) satisfies it. Thus, a rational approximation to the solution to within an additive accuracy of \( \delta \) can be computed, using quantifier elimination algorithms, in time \( \text{poly}(\eta, N^{O(nN)}, \log \frac{1}{\delta}) \); see [2, 14, 23]. Note that the resulting solution will satisfy (12) but within the approximate range \([m_- - \delta, m_+ + \delta]\). By choosing \( \delta \) sufficiently smaller than the desired accuracy \( \varepsilon \), we can ignore the effect of such approximation.

---

**Algorithm 3 MODIFIEDREPEATEDPUMPING(\( \varepsilon \))**

1: Initialize \( h := 0 \), and \( x_h := 0 \in \mathbb{R}^V \).
2: Set \( m^+(h) := \max_{v \in V} m^v(x_h) \) and \( m^-(h) := \min_{v \in V} m^v(x_h) \).
3: if \( m^+(h) - m^-(h) \leq 24\varepsilon \) then return \( x_h \).
   end if
4: \( x_{h+1} := \text{MODIFIEDPUMP}(\varepsilon, x_h, V, m_-, m_+) \) and let \( F_\tau, I_\tau, T_\tau, B_\tau, P_\tau \) be the sets obtained from MODIFIEDPUMP.
7: if \( T_\tau = \emptyset \) or \( B_\tau = \emptyset \) then
8: \( x_{h+1} := \text{REDUCEPOTENTIAL}(\Gamma, x_\tau, m^-(h), m^+(h)) \)
9: Set \( h := h + 1 \) and Goto step 2
10: end if
11: Otherwise, set \( F = F_\tau \) and \( I = I_\tau \).
12: \( x_{h+1} := \text{MODIFIEDPUMP}(\varepsilon, x_h, I_\tau, m_-, m_+) \) and let \( T_\tau, B_\tau \) be the sets obtained from this call of MODIFIEDPUMP.
13: if \( T_\tau = \emptyset \) then
14: \( x_{h+1} := \text{REDUCEPOTENTIAL}(\Gamma, x_\tau, m^-(h), m^+(h)) \)
15: Set \( h := h + 1 \) and Goto step 2.
16: end if
17: if \( B_\tau = \emptyset \) then
18: Goto step 21
19: end if
20: Otherwise, update \( I := I_\tau \).
21: return \( x_{h+1} \) and the sets \( I \) and \( F \).
Lemma 9 \textsc{ModifiedRepeatedPumping}(\epsilon) terminates in a finite number \( h \leq \log \frac{R}{24\epsilon} / \log \frac{7}{8} \), of iterations and either provides a potential transformation proving that the game is 24\epsilon-ergodic, or outputs two nonempty subsets \( I \) and \( F \) and strategies \( \alpha^v, v \in I \), for player 1 and \( \beta^v, v \in F \), for player 2 such that conditions (N4), (N5) and (N6) hold with \( b', a' \) satisfying the condition in Lemma 6.

Proof Let us note that if \( T_\tau = \emptyset \) after the second \textsc{ModifiedPump} call, then the range of the \( m \)-values has shrunk by a factor of \( \frac{7}{8} \) (at least), while if this happens in the first stage, the \( m \)-range has shrunk by a factor of 3/4.

On the other hand, if the \( m \)-range is not shrinking, and we have \( B_\tau = \emptyset \) after the second call of \textsc{ModifiedPump}, then we would also have \( m^u(x_\tau) \geq \frac{5}{8}m^+ + \frac{3}{8}m^- = b' \) for all \( v \in I \), while \( m^u(x_\tau) < (m^+ + m^-)/2 = a' \) for all \( u \in F \), and hence, (N4)–(N6) hold with these \( a' \) and \( b' \) values. Since the \( m \)-range has not shrunk, we must have \( m^+ - m^- > 24\epsilon \), and hence, \( b' - a' = \frac{1}{8}(m_+ - m_-) > 3\epsilon \) follows. [Note that, since in the second stage we pump only positions in \( I_\tau \), the potentials of these positions may go down, while those of the positions outside \( \mathcal{F}_\tau \) remain unchanged, and hence, condition (N5) remains satisfied.]

Finally, if the \( m \)-range is not shrinking, and the second call returns a new set \( I_\tau \), then all \( m \)-values of this set are at least \( \frac{3}{4}m^+ + \frac{1}{4}m^- > \frac{5}{8}m^+ + \frac{3}{8}m^- = b' \), and with the same set \( F \) we can conclude again that conditions (N4)–(N6) hold. \( \square \)

To complete the proof of Theorem 1, we need to analyze the time complexity of the above procedure, in particular, bounding the number of pumping steps performed in \textsc{ModifiedPump}.

Let us note that as long as \( m^+ - m^- > 24\epsilon \) we pump the upper half \( P_\tau \) by exactly \( \delta \geq 6\epsilon \). Let \( \mathcal{P}_\tau(v) \) [resp., \( \mathcal{N}_\tau(v) \)] denote the number of iterations, among the first \( \tau \), in which position \( v \) was pumped, that is, \( v \in P_\tau \) (resp., not pumped, that is, \( v \notin P_\tau \)).

Let us next sort the positions \( v \in V \) such that we have
\[
x_\tau^v_1 \leq x_\tau^v_2 \leq \cdots \leq x_\tau^v_n,
\]
and write \( \Delta_j = x_\tau^{v_{j+1}} - x_\tau^{v_j} \) for \( j = 1, 2, \ldots, n - 1 \). Note that \( \mathcal{P}_\tau(v_1) = \tau \) and \( \mathcal{N}_\tau(v_n) = \tau \).

Let \( i_\tau \) be the largest index in \{1, 2, \ldots, n\}, such that \( v_{i_\tau} \in P_\tau \). Then, by (8), we have for \( i = 0, 1, 2, \ldots, i_\tau - 1 \) that
\[
0 \leq \tilde{\alpha}_{k\ell}^{v_{i+1}}(x_\tau) \leq R + \sum_{j=1}^i \Delta_j,
\]
where the sum over the empty sum is zero by definition. Similarly, for \( i = i_\tau + 1, \ldots, n \), we have
\[
-R \leq \tilde{\alpha}_{k\ell}^{v_{i}}(x_\tau) \leq R + \sum_{j=1}^{n-i} \Delta_{n-j}.
\]

From (13) and (14), it follows that
\[
|R_\tau^{v_{i+1}}| \leq \begin{cases} R + \sum_{j=1}^i \Delta_j, & \text{for } i = 0, 1, 2, \ldots, i_\tau - 1 \\ R + \sum_{j=1}^{n-i-1} \Delta_{n-j}, & \text{for } i = i_\tau, i_\tau + 1, \ldots, n - 1. \end{cases}
\]

Let \( \tilde{i}_\tau \) be the smallest index \( i \) such that
\[
\Delta_i > \frac{\sqrt{NW} \left( R + \sum_{j=1}^{i-1} \Delta_j \right)^2}{\epsilon},
\]
and let \( \hat{\iota} \) be the largest index \( i \leq n - 1 \) such that
\[
\Delta_i > \frac{NW \left( R + \sum_{j=1}^{n-i} \Delta_{n-j} \right)^2}{\varepsilon}.
\]
(17)

From the definition of \( \hat{\iota} \), we know that
\[
\Delta_i \leq \frac{NW \left( R + \sum_{j=1}^{i-1} \Delta_j \right)^2}{\varepsilon}, \text{ for all } i = 1, \ldots, \hat{\iota} - 1.
\]

Solving this recurrence, we get
\[
x^{\nu_{\hat{\iota}}} - x^{\nu_{\hat{\iota}+1}} = \sum_{i=1}^{\hat{\iota}-1} \Delta_i \leq \left( \frac{(\hat{\iota} - 1) NW R}{\varepsilon} \right)^{2^{\hat{\iota}+1}} \left( \hat{\iota} - 1 \right)^2 R \leq \left( \frac{n NW R}{\varepsilon} \right)^{2^{\hat{\iota}-1}} n^2 R.
\]
Similarly, the definition of \( \hat{\iota} \) gives
\[
\Delta_i \leq \frac{NW \left( R + \sum_{j=1}^{n-i} \Delta_{n-j} \right)^2}{\varepsilon}, \text{ for all } i = \hat{\iota}, \ldots, n - 1,
\]
from which follows
\[
x^{\nu_{\hat{\iota}}} - x^{\nu_{\hat{\iota}+1}} \leq \left( \frac{n NW R}{\varepsilon} \right)^{2^{\hat{\iota}+1}} n^2 R.
\]
(18)

Note that if \( \hat{\iota} \leq \iota \) then (15) implies that taking \( I_\iota = \{ v_1, \ldots, v_{\hat{\iota}} \} \) would satisfy condition (N5) and guarantee that \( I_\iota \subseteq P_\tau \).

Indeed, for all \( i \leq \hat{\iota} \) and \( u \notin I_\iota \), we have
\[
x^{\nu_i} - x^{\nu_{\hat{\iota}+1}} \geq \frac{NW \left( R + \sum_{j=1}^{\hat{\iota}-1} \Delta_j \right)^2}{\varepsilon} \geq \frac{|L^{\nu_i} W (R^{\nu_i})^2}{\varepsilon}.
\]

Similarly, having \( \hat{\iota} \geq \iota \) guarantees that taking \( F_\iota = \{ v_{\hat{\iota}+1}, \ldots, v_n \} \) would satisfy (N6) and \( F_\iota \cap P_\tau = \emptyset \).

Since for all \( i \geq \hat{\iota} + 1 \) and \( u \notin F_\iota \), we have
\[
x^{\nu_i} - x^{\nu_{\hat{\iota}+1}} \geq \frac{NW \left( R + \sum_{j=1}^{n-\hat{\iota}} \Delta_{n-j} \right)^2}{\varepsilon} \geq \frac{|K^{\nu_i} W (R^{\nu_i})^2}{\varepsilon}.
\]

On the other hand, if \( \hat{\iota} \geq \iota + 1 \), then (18) implies that \( v_{\hat{\iota}+1} \) was always pumped except for at most \( \kappa (R) := \left( \frac{n NW R}{\varepsilon} \right)^{2^{\hat{\iota}-1}} n^2 R \) iterations, that is, \( \mathcal{N} (v_{\hat{\iota}+1}) \leq \kappa (R) \). Also, since \( v_{\hat{\iota}+1} \notin P_\tau \), then at time \( \tau \), \( v_{\hat{\iota}+1} \) is not pumped. Similarly, if \( \hat{\iota} < \iota \), then (19) implies that \( v_i \) was never pumped except for at most \( \kappa (R) \) iterations, that is, \( \mathcal{P}_\tau (v_i) \leq \kappa (R) \), while it is pumped at time \( \tau \). Since we have at most \( n \) candidates for each of \( v_i \) and \( v_{\hat{\iota}+1} \), it follows that after \( \tau = 2n \kappa (R) + 1 \), neither of these events \( \hat{\iota} \geq \iota + 1 \) and \( \hat{\iota} < \iota \) can happen, which by our earlier observations implies that the algorithm constructs the sets \( I_\iota \) and \( F_\iota \). We can conclude that \( \text{MODIFIEDPUMP}(\varepsilon, x, V) \) must terminate in at most \( 2n \kappa (R) + 1 \) iterations, either producing \( m^+ - m^- \leq 24 \varepsilon \) or outputting the subsets \( I_\iota \) and \( F_\iota \) proving \( \varepsilon \)-non-ergodicity.
One can similarly bound the running time for the second call of ModifiedPump (line 12), and the running time for each iteration of ModifiedRepeatedPumping(ε) (but with R replaced by 2poly(n,n,η)).

It remains now to bound the running time for the second call of ModifiedPump (line 12), and the running time for each iteration of ModifiedRepeatedPumping(ε). We can repeat essentially the same analysis as above, assuming that we modify the rewards with the potential vector obtained up to this point in time. Since, by the above argument, the maximum potential difference between any vertices before at the time τ, when we make the second call to ModifiedPump is at most δ(2nk(R) + 1), it follows that the maximum absolute value of the transformed rewards at time τ is \( r^{\nu}_{k}\ell(x_\tau) \leq R_2 := R + \delta(2nk(R) + 1) \) (note that the nonnegativity of the rewards was only needed to bound \( m_\geq 0 \) initially). It follows by the same argument as above that the second call ModifiedPump terminates in time 2nk(R2) + 1 = \( \frac{\nu NWR}{\epsilon} O(2^{2n}) \).

After shrinking the \( m \)-range, we apply potential reductions which guarantees that the bit length of each entry in potential vector is bounded by a polynomial in the original bit length \( \eta \). It follows that the new transformed rewards will have absolute value bounded by \( R_3 = 2^{\text{poly}(n,N,\eta)} \). We repeat the same argument for the different phases of ModifiedRepeatedPumping(ε) to arrive at the running time claimed in Theorem 1.

This completes the proof of the theorem.

References

1. Andersson D, Miltersen PB (2009) The complexity of solving stochastic games on graphs. In: Proceedings of the 20th ISAAC, LNCS, vol 5878, pp 112–121
2. Basu S, Pollack R, Roy M (1996) On the combinatorial and algebraic complexity of quantifier elimination. J ACM 43(6):1002–1045 (Preliminary version in FOCS 1994)
3. Blackwell D, Ferguson TS (1968) The big match. Ann Math Stat 39(1):159–163
4. Boros E, Elbassioni K, Gurvich V, Makino K (2013) On canonical forms for zero-sum stochastic mean payoff games. Dyn Games Appl 3(2):128–161
5. Boros E, Elbassioni K, Gurvich V, Makino K (2013) On discounted approximations of undiscounted stochastic games and markov decision processes with limited randomness. Oper Res Lett 41(4):357–362
6. Boros E, Elbassioni K, Gurvich V, Makino K (2010) A pumping algorithm for ergodic stochastic mean payoff games with perfect information. In: Proceedings of the 14th IPCO, LNCS, vol 6080. Springer, pp 341–354
7. Boros E, Elbassioni K, Gurvich V, Makino K (2013) A pseudo-polynomial algorithm for mean payoff stochastic games with perfect information and a few random positions. In: Fomin FV, Freivalds R, Kwiatkowska M, Peleg D (eds) Automata, languages, and programming, LNCS, vol 7965. Springer, Berlin Heidelberg, pp 220–231
8. Boros E, Elbassioni K, Gurvich V, Makino K (2014) A potential reduction algorithm for ergodic two-person zero-sum limiting average payoff stochastic games. In: Zhang Z, Wu L, Xu W, Du DZ (eds) Combinatorial optimization and applications, LNCS, vol 8881. Springer, pp 694–709
9. Chatterjee K, Majumdar R, Henzinger TA (2008) Stochastic limit-average games are in exptime. Int J Game Theory 37:219–234
10. Chatterjee K, Ibsen-Jensen R (2014) The complexity of ergodic mean-payoff games. In: Proceedings of the 41st ICALP: 2, LNCS, vol 8573, pp 122–133
11. Federgruen A (1980) Successive approximation methods in undiscounted stochastic games. Oper Res 1:794–810
12. Gallai T (1958) Maximum–minimum Sätze über Graphen. Acta Math Acad Sci Hung 9:395–434
13. Gillette D (1957) Stochastic games with zero stop probabilities. In: Dresher M, Tucker AW, Wolfe P (eds) Contribution to the theory of games III. Annals of mathematical studies, vol 39. Princeton University Press, Princeton, pp 179–187
14. Grigoriev D, Vorobjov N (1988) Solving systems of polynomial inequalities in subexponential time. J Symb Comput 5(1/2):37–64
15. Hansen KA, Koucky M, Lauritzen N, Miltersen PB, Tsigaridas EP (2011) Exact algorithms for solving stochastic games: extended abstract. In: Proceedings of the 43rd annual ACM symposium on theory of computing, STOC ’11. ACM, New York, pp 205–214
16. Hoffman AJ, Karp RM (1966) On nonterminating stochastic games. Manag Sci A 12(5):359–370
17. Kemeny JG, Snell JL (1963) Finite Markov chains. Springer, New York
18. Mertens JF, Neyman A (1981) Stochastic games. Int J Game Theory 10:53–66
19. Miltersen PB (2011) Discounted stochastic games poorly approximate undiscounted ones, manuscript. Technical report
20. Mine H, Osaki S (1970) Markovian decision process. American Elsevier Publishing Co., New York
21. Moulin H (1976) Prolongement des jeux à deux joueurs de somme nulle. Bull Soc Math Fr Mem 45:5–111
22. Raghavan TES, Filar JA (1991) Algorithms for stochastic games: a survey. Math Methods Oper Res 35(6):437–472
23. Renegar J (1992) On the computational complexity and geometry of the first-order theory of the reals. J Symb Comput 13(3):255–352
24. Shapley LS (1953) Stochastic games. Proc Natl Acad Sci USA 39:1095–1100
25. Vrieze OJ (1980) Stochastic games with finite state and action spaces. PhD thesis, Centrum voor Wiskunde en Informatica, Amsterdam
26. Zwick U, Paterson M (1996) The complexity of mean payoff games on graphs. Theor Comput Sci 158(1–2):343–359