A Potential Reduction Algorithm for Two-person Zero-sum Mean Payoff Stochastic Games *†

Endre Boros‡ Khaled Elbassioni§ Vladimir Gurvich†
Kazuhisa Makino¶

Abstract

We suggest a new algorithm for two-person zero-sum undiscounted stochastic games focusing on stationary strategies. Given a positive real ϵ, let us call a stochastic game ϵ-ergodic, if its values from any two initial positions differ by at most ϵ. The proposed new algorithm outputs for every $\epsilon > 0$ in finite time either a pair of stationary strategies for the two players guaranteeing that the values from any initial positions are within an ϵ-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 $\epsilon/24$ apart. In particular, the above result shows that if a stochastic game is ϵ-ergodic, then there are stationary strategies for the players proving 24ϵ-ergodicity. This result strengthens and provides a constructive version of an existential result by Vrieze (1980) claiming that if a stochastic game is 0-ergodic, then there are ϵ-optimal stationary strategies for every $\epsilon > 0$. 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

*This research was partially supported by DIMACS, Center for Discrete Mathematics and Theoretical Computer Science, Rutgers University, and by the Scientific Grant-in-Aid from Ministry of Education, Science, Sports and Culture of Japan. Part of this research was done at the Mathematisches Forschungsinstitut Oberwolfach during stays within the Research in Pairs Program. The first author also acknowledges the partial support of NSF Grant IIS-1161476.

†An extended abstract of this paper was published in the proceedings of the Combinatorial Optimization and Applications, 2014 [BEGM14].

‡MSIS Department and RUTCOR, Rutgers University, 100 Rockafellar Road, Livingston Campus Piscataway, NJ 08854, USA; ({boros,gurvich}@rutcor.rutgers.edu)

§Masdar Institute of Science and Technology, P.O.Box 54224, Abu Dhabi, UAE; (kelbassioni@masdar.ac.ae)

¶Research Institute for Mathematical Sciences (RIMS) Kyoto University, Kyoto 606-8502, Japan; (makino@kurims.kyoto-u.ac.jp)

1 Introduction

1.1 Basic Concepts and Notation

Stochastic games were introduced in 1953 by Shapley [Sha53] for the discounted case, and extended to the undiscounted case by Gillette [Gil57]. Each such game $\Gamma = (p^{vu}_{k\ell},\ r^{vu}_{k\ell} \ \mid\ k \in K^v,\ \ell \in L^v,\ u, v \in V)$ 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; $p^{vu}_{k\ell} \in [0,1]$ 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 $r^{vu}_{k\ell} \in \mathbb{R}$ is the reward player 1 (the maximizer) receives from player 2 (the minimizer), correpsonding to this transition. We assume that the game is non-stopping, that is, $\sum_{u \in V} p^{vu}_{k\ell} = 1$ 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 $p^{vu}_{k\ell}$, and associate in Γ a *local expected reward matrix* A^v to every $v \in V$ defined by

$$(A^v)_{k\ell} \ = \ \sum_{u \in V} p^{vu}_{k\ell} r^{vu}_{k\ell}. \tag{1}$$

In the game Γ, 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 $p^{vu}_{\alpha,\beta} \ = \ \alpha^v P^{vu} \beta^v$.

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

$$g^{v_0}(\Gamma) \ = \ \liminf_{N \to \infty} \frac{1}{N+1} \sum_{j=0}^{N} \mathbb{E}[b_j], \tag{2}$$

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 in their seminal paper [MN81] 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 paper [Gil57]; see also [BF68].

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 player 1 and player 2, 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 (1980) showed that if a stochastic game Γ 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 ϵ-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 Γ is said to be* ϵ-ergodic *if the game values from any two initial positions differ by at most ϵ, 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 Γ and $\epsilon > 0$, whether or not Γ is ϵ-ergodic, and provides a witness for its ϵ-ergodicity/non-ergodicity. As a corollary, we get a constructive proof of the above mentioned theorem of Vrieze [Vri80]. 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 [Gal58] 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 $r^{vu}_{k\ell}$ as follows:

$$r^{vu}_{k\ell}(x) = r^{vu}_{k\ell} + x^v - x^u.$$

Then the one step expected payoff amount changes to $\mathbb{E}[b_j(x)] = \mathbb{E}[b_j] + \mathbb{E}[x^{v_j}] - \mathbb{E}[x^{v_{j+1}}]$, 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:

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

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.

m^v is the value of the matrix game A^v at state v.

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

$$a^v_{k\ell}(x) = \sum_{u \in V} p^{vu}_{k\ell}(r^{vu}_{k\ell} + x^v - x^u), \quad \text{for all } k \in K^v, \ell \in L^v, \tag{3}$$

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 $\mathcal{M}_{\alpha,\beta}(\Gamma)$ on states in V, defined by the transition probabilities $p^{vu}_{\alpha,\beta} = \alpha^v P^{vu} \beta^v$. Then, this Markov chain has unique limiting probability distributions $(q^{vu}_{\alpha,\beta} | \ u \in V)$, where $q^{vu}_{\alpha,\beta}$ is the probability of staying in state $u \in V$ when the initial vertex is $v \in V$. With this notation, The limiting average payoff (2) starting from vertex $v \in V$ can be computed as

$$g^v(\alpha, \beta) \ = \ \sum_{u \in V} q^{vu}_{\alpha,\beta} \ (\alpha^u A^u \beta^u) \, . \tag{4}$$

The game is said to be to be solvable in *uniformly optimal* stationary strategies, if there exist stationary strategies $\bar{\alpha} \in \mathcal{K}(\Gamma)$ and $\bar{\beta} \in \mathcal{L}(\Gamma)$, such that for all initial states $v \in V$

$$g^v(\bar{\alpha}, \bar{\beta}) \ = \ \max_{\alpha \in \mathcal{K}(\Gamma)} g^v(\alpha, \bar{\beta}) \ = \ \min_{\beta \in \mathcal{L}(\Gamma)} g^v(\bar{\alpha}, \beta). \tag{5}$$

This common quantity, if exists, is the value of the game with initial position $v \in V$, and will be simply denoted by $g^v = g^v(\Gamma)$.

1.4 Main Result

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* [BEGM13a]. 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 [BEGM13a]. 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.
$$0 \leq r^{vu}_{k\ell} \leq R \tag{6}$$
- $N = \max_{v \in V}\{\max\{|K^v|, |L^v|\}\}$.
- $n = |V|$
- $\eta = \max\{\log_2 R, \log_2 W\}$ (maximum "bit length")

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

The proof of Theorem 1 will be given in Section 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 [BPR96, GV88, Ren92] to reduce the range of potentials after each iteration; see the discussion preceding Lemma 9.

2 Related Work

The above definition of ergodicity follows Moulin's concept of the ergodic extension of a matrix game [Mou76] (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, [KS63].

The following four algorithms for undiscounted stochastic games are based on stronger "ergodicity type" conditions: the strategy iteration algorithm by Hoffman and Karp [HK66] 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 are based on similar but slightly weaker requirements; see [Fed80] for the definitions and more details; the recent algorithm of Chatterjee and Ibsen-Jensen [CIJ14] assumes a weaker requirement than the strong ergodicity required by Hoffman and Karp [HK66]: 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 [HK66]) 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 [HK66] and [CIJ14] 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 [Fed80] implicitly, as the differences of local values of positions, as well as in [HK66], 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 [BEGM10] and extended to the general (not necessarily ergodic) case in [BEGM13b]. This approach is also somewhat similar to the first of two value iteration algorithms suggested by Federgruen in [Fed80], though our approach has some distinct characteristics: It is assumed in [Fed80] that the values g^v exist and are equal for all v; in particular, this assumption implies the ϵ-ergodicity for every $\epsilon > 0$. For our approach we do not need such an assumption. We can verify ϵ-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 [Fed80] 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 [RF91]. 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 [CMH08, HKL$^+$11]: in [CMH08], 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 $(nN)^{nN} \operatorname{poly}(\eta, \log \frac{1}{\epsilon})$. In [HKL$^+$11], the authors give algorithms for discounted and recursive stochastic games that run in time $2^{N^{O(N^2)}} \operatorname{poly}(\eta, \log(\frac{1}{\epsilon}))$, and claim also that similar bounds can be obtained for general stochastic games, by reducing them to the discounted version using a discount factor of $\delta = \epsilon^{\eta N^{O(n^2)}}$ (and this bound on δ is almost tight [Mil11]). These results are based on quantifier elimination techniques and yield very complicated history-dependent strategies. For almost sure ergodic games, a variant of the algorithm of Hoffman and Karp [HK66] was given in [CIJ14]; this algorithm finds ϵ-optimal stationary strategies in time (roughly) $\left(\frac{Nn^2W^n}{\epsilon}\right)^{nN} \operatorname{poly}(N, \eta)$. This result is not comparable to ours, since the class of games they deal with are somewhat different (although both generalize the class of strongly ergodic games of [HK66]). Furthermore, the algorithm in Theorem 1 exhibits the additional feature that it either provides a solution in stationary strategies in the ergodic case, if one exists, or produces a pair of stationary strategies that witness the non-ergodicity.

3 Pumping Algorithm

We begin by describing our procedure on an abstract level. Then we specialize it to stochastic games in Section 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 $\delta \geq 0$:

(i) $m^v(x - \delta e_S)$ is a monotone decreasing function of δ if $v \in S$;

(ii) $m^v(x - \delta e_S)$ is a monotone increasing function of δ if $v \notin S$;

(iii) $|m^v(x) - m^v(x - \delta e_S)| \leq \delta$ 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_τ at iteration τ, 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_τ or forth (bottom) set B_τ is empty, the procedure terminates; otherwise, the potentials of all the vertices in the first and second sets are reduced by the same amount δ, and the computation proceeds to the next iteration.

Algorithm 1 PUMP(x, S)

Input: a stochastic game Γ a subset S of states.
Output: a potential $x \in \mathbb{R}^S$.

1: Initialize $\tau := 0$, and $x_\tau := x$.
2: Set $m^+ := \max_{v \in S} m^v(x_\tau)$, $m^- := \min_{v \in S} m^v(x_\tau)$, and $\delta := (m^+ - m^-)/4$.
3: Define
$$\begin{aligned} T_\tau &:= \{v \in S \mid m^v(x_\tau) \geq m^- + 3\delta\} \\ B_\tau &:= \{v \in S \mid m^v(x_\tau) < m^- + \delta\} \\ M_\tau &:= S \setminus (T_\tau \cup B_\tau). \end{aligned}$$
4: **if** $T_\tau = \emptyset$ or $B_\tau = \emptyset$ **then**
5: **return** x_τ
6: **end if**
7: Otherwise, set $P_\tau := \{v \in S \mid m^v(x_\tau) \geq m^- + 2\delta\}$ and update
$$x^v_{\tau+1} := \begin{cases} x^v_\tau - \delta & \text{if } v \in P_\tau \\ x^v_\tau & \text{otherwise.} \end{cases}$$
8: Set $\tau := \tau + 1$ and Goto step 3.

We can show next that properties (i), (ii) and (iii) 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 (i) and (iii) we can conclude that $m^v(x_\tau) \geq m^- + \delta$ holds for all $v \in P_\tau$. Analogously, by (ii) and (iii) $m^v(x_\tau) < m^- + 3\delta$ follows for all $v \notin P_\tau$. □

Lemma 2 *Either $T_\tau = \emptyset$ or $B_\tau = \emptyset$ for some finite τ, or there are nonempty disjoint subsets $I, F \subseteq S$, $I \supseteq T_\tau$, $F \supseteq B_\tau$, and a threshold τ_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 $\tau \geq \tau_0$;

(b) $x_\tau^u - x_\tau^v \geq \Delta$ *for all* $v \in I$ *and* $u \notin I$*, and for all* $\tau \geq \tau(\Delta)$*;*

(c) $x_\tau^v - x_\tau^u \geq \Delta$ *for all* $v \in F$ *and* $u \notin F$*, and for all* $\tau \geq \tau(\Delta)$*.*

Proof By Lemma 1 sets T_τ and B_τ 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_τ and B_τ will never change and all positions in T_τ remain always pumped (that is, have their potentials reduced), while all positions in B_τ 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 τ_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_τ is always pumped by (iii) 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 δ is a fixed constant, for every Δ there must exist an iteration $\tau(\Delta) \geq \tau_0$ such that all positions not in I are not pumped by at least Δ/δ many more times than those in I, and all positions not in F are pumped by at least Δ/δ many more times than those in F, implying (b) and (c). $\square$

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:

Algorithm 2 $\textsc{RepeatedPumping}(\epsilon)$

```
1: Initialize h := 0, and x_h := 0 ∈ R^V.
2: Set m^+(h) := max_{v∈V} m^v(x_h) and m^-(h) := min_{v∈V} m^v(x_h).
3: If m^+(h) − m^-(h) ≤ ε then STOP.
4: x_{h+1} := PUMP(x_h, V); h := h + 1.
5: Goto step 2.
```

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 ϵ-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 those large potential differences can be used to prove that the game is not ϵ-ergodic.

4 Application of Pumping for Stochastic Games

We show in this section how to use $\textsc{RepeatedPumping}$ to find potential transformations verifying ϵ-ergodicity, or proving that the game is not ϵ-ergodic, thus establishing a proof of Theorem 1. Towards this end, we shall give some necessary and sufficient conditions for ϵ-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.

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*

$$\begin{array}{lllll} m^v(x) & \geq & m^v(x-\delta e_S) & \geq & m^v(x) - \delta \max_{k,\ell} \sum_{u \notin S} p^{vu}_{k\ell} \quad \text{if } v \in S, \\ m^v(x) & \leq & m^v(x-\delta e_S) & \leq & m^v(x) + \delta \max_{k,\ell} \sum_{u \in S} p^{vu}_{k\ell} \quad \text{if } v \notin S. \end{array} \tag{7}$$

Furthermore, the value functions $m^v(x)$ for $v \in V$ satisfy properties (i), (ii) and (iii) stated in Section 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 (i) and (ii) (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 Val (B) is monotone increasing in B, inequalities (7) follow). Property (iii) follows directly from (7). □

The above lemma implies that procedures Pump and RepeatedPumping could, in principle, be used to find a potential transformation yielding an ϵ-ergodic solution. It does not offer, however, a way to discover ϵ-non-ergodicity. Towards this end, we need to find some sufficient and algorithmically achievable conditions for ϵ-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 ϵ-non-ergodic for some positive ϵ.*

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

Lemma 5 *A stochastic game Γ is ϵ-non-ergodic if there exist disjoint non-empty subsets of the positions $I, F \subseteq V$, reals a, b with $b - a \geq \epsilon$, stationary strategies α^v, $v \in I$, for player 1, and β^u, $u \in F$, for player 2, and a vector of potentials $x \in \mathbb{R}^V$, such that*

(N1) $\alpha^v_k p^{vu}_{k\ell} = 0$ *for all* $v \in I$, $u \notin I$, $k \in K^v$ *and* $\ell \in L^v$,

(N2) $\beta^u_\ell p^{uw}_{k\ell} = 0$ *for all* $u \in F$, $w \notin F$, $\ell \in L^u$ *and* $k \in K^u$, *and*

(N3) *for all* $v \in I$ *and* $u \in F$:

$$\min_{\widetilde{\beta}^v \in \Delta(L^v)} (\alpha^v)^T A^v(x) \widetilde{\beta}^v \ \geq \ b \quad \text{and} \quad \max_{\widetilde{\alpha}^u \in \Delta(K^u)} (\widetilde{\alpha}^u)^T A^u(x) \beta^u \ < \ a.$$

Proof Let us note that (N1) and (N3) imply that for all strategies $\beta' \in \mathcal{L}(\Gamma)$ of player 2, the pair of strategies $(\bar{\alpha}, \beta')$, where $\bar{\alpha}^v := \alpha^v$ for $v \in I$ and $\bar{\alpha}^v \in \Delta(K^v)$ is chosen arbitrarily for $v \notin I$, results in a Markov chain in which subset I induces one or more absorbing sets (that is, $p^{vu}_{\bar{\alpha}\beta'} = 0$), 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 $(\alpha', \bar{\beta})$, where α' is any strategy in $\mathcal{K}(\Gamma)$, $\bar{\beta}^v := \beta^v$ for $v \in F$ and $\bar{\beta}^v \in \Delta(L^v)$ is chosen arbitrarily, for $v \notin F$. Hence choosing any positions $v \in I$ and $u \in F$ and strategies $\bar{\alpha}$ and $\bar{\beta}$ provides a witness for the ϵ-nonergodicity of Γ. (Here, we use the well-known fact [MO70] that, to each player's stationary strategy, there is a best response of the opponent which is also stationary.) $\square$

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

$$\begin{aligned} \widetilde{a}^v_{k\ell}(x) \quad &= \sum_{u \in V} p^{vu}_{k\ell} r^{vu}_{k\ell} \ + \sum_{u \in V,\ x^u \leq x^v} p^{vu}_{k\ell}(x^v - x^u) \\ \widetilde{b}^v_{k\ell}(x) \quad &= m^+(x) - \sum_{u \in V} p^{vu}_{k\ell} r^{vu}_{k\ell} \ - \sum_{u \in V,\ x^u \geq x^v} p^{vu}_{k\ell}(x^v - x^u) \end{aligned} \tag{8}$$

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

$$\begin{aligned} R^v(x) \quad &= \max_{k \in K^v, \ell \in L^v} (\widetilde{a}^v_{k\ell}(x)) \qquad \text{if } m^v(x) \geq \tfrac{m^+(x)+m^-(x)}{2}, \\ R^v(x) \quad &= \max_{k \in K^v, \ell \in L^v} \left(\widetilde{b}^v_{k\ell}(x)\right) \qquad \text{otherwise.} \end{aligned} \tag{9}$$

Note that

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

which implies

$$\begin{array}{lll} m^v(x) & \leq \ R^v(x) & \text{if } m^v(x) \geq \frac{m^+(x)+m^-(x)}{2}, \\ m^v(x) & \geq \ m^+(x) - R^v(x) & \text{otherwise.} \end{array} \quad \text{for all } v \in V \text{ and } x \in \mathbb{R}^V, \tag{10}$$

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* $I, F \subseteq V$*, a vector of potentials* $x \in \mathbb{R}^V$*, and reals* $a', b' \in [0, m^+(x)]$ *with* $b' - a' \geq 3\epsilon$*,* $a' < \frac{m^+(x)+m^-(x)}{2}$*,* $b' \geq \frac{m^+(x)+m^-(x)}{2}$*,such that*

(N4) $m^v(x) \geq b'$ *for all* $v \in I$*, and* $m^u(x) < a'$ *for all* $u \in F$*;*

(N5) $x^u - x^v \geq |L^v| W R^v(x)^2/\epsilon$ *for all* $u \notin I$*, and* $v \in I$*;*

(N6) $x^u - x^v \geq |K^v| W R^v(x)^2/\epsilon$ *for all* $u \in F$*, and* $v \notin F$*.*

Proof We first show that (N4)-(N5) imply the existence of strategies α^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 β^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 $\bar{\alpha}^v$ and $\bar{\beta}^v$ the optimal strategies of players with respect to the payoff matrix $A^v(x)$. Denote further by $\widehat{\beta}^v = \frac{1}{|L^v|}(1, 1, \ldots, 1)$ the uniform strategy for player 2, and set $\bar{K}^v = \{k \in K^v \mid \sum_{u \notin I} \sum_{\ell \in L^v} p_{k\ell}^{vu} = 0\}$.

Let us then note that we have

$$\left(A^v(x)\widehat{\beta}^v\right)_k \leq \begin{cases} R^v(x) & \text{if } k \in \bar{K}^v, \\ R^v(x) - \frac{R^v(x)^2}{\epsilon} & \text{otherwise,} \end{cases}$$

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

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

$$0 < b' \leq m^v(x) \leq \bar{\alpha}^v A^v(x) \widehat{\beta}^v \leq R^v(x) - \left(\sum_{k \notin \bar{K}^v} \bar{\alpha}_k^v\right) \frac{R^v(x)^2}{\epsilon}$$

implying that $\sum_{k \notin \bar{K}^v} \bar{\alpha}_k^v < \frac{\epsilon}{R^v(x)}$. 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 $\bar{K}^v$ is not empty. Let us then denote by $\widetilde{\alpha}^v$ the truncated strategy defined by

$$\widetilde{\alpha}_k^v = \begin{cases} \dfrac{\bar{\alpha}_k^v}{\sum_{k \in \bar{K}^v} \bar{\alpha}_k^v} & \text{if } \quad k \in \bar{K}^v, \\ 0 & \text{if } \quad k \notin \bar{K}^v. \end{cases}$$

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

$$\begin{aligned} b' \leq m^v(x) \quad & \leq (\bar{\alpha}^v A^v(x)\widetilde{\beta}^v \\ & = \left(\widetilde{\alpha}^v A^v(x)\widetilde{\beta}^v\right)\left(\sum_{k\in\bar{K}^v}\bar{\alpha}^v_k\right) + \sum_{k\notin\bar{K}^v}\bar{\alpha}^v_k\left(\sum_{\ell\in L^v} a^v_{k\ell}(x)\widetilde{\beta}^v_\ell\right) \\ & \leq \left(\widetilde{\alpha}^v A^v(x)\widetilde{\beta}^v\right) + \left(\sum_{k\notin\bar{K}^v}\bar{\alpha}^v_k\right) R^v(x) \\ & < \left(\widetilde{\alpha}^v A^v(x)\widetilde{\beta}^v\right) + \epsilon. \end{aligned}$$

Let us then define $\alpha^v = \widetilde{\alpha}^v$ 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 β^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 2 in procedure Pump by

2: Set $\delta := (m^+ - m^-)/4$.

Next, Let us replace in procedure Pump, line 7 by the following lines, where $\epsilon > 0$ is a prespecified parameter, and call the new procedure with these modifications ModifiedPump$(\epsilon, x, S, m_-, m_+)$:

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

$$\begin{aligned} R^v_\tau \quad &:= \max_{k\in K^v,\ell\in L^v} (\widetilde{a}^v_{k\ell}(x_\tau)) \qquad \text{if} \quad v \in P_\tau, \\ R^v_\tau \quad &:= \max_{k\in K^v,\ell\in L^v} \left(\widetilde{b}^v_{k\ell}(x_\tau)\right) \qquad \text{if} \quad v \notin P_\tau, \end{aligned}$$

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

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

$$\begin{aligned} x^u_\tau - x^v_\tau < \tfrac{|L^v|W(R^v_\tau)^2}{\epsilon} \qquad \text{if} \quad v \in P_\tau, \\ x^v_\tau - x^u_\tau < \tfrac{|K^v|W(R^v_\tau)^2}{\epsilon} \qquad \text{if} \quad v \notin P_\tau. \end{aligned}$$

7c: Find subsets I_τ and F_τ 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 int he proof of Theorem 1).

7d: if such sets are found STOP and output these sets, otherwise continue with step 8.

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

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

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

Lemma 7 *Procedure* ModifiedPump(ϵ, 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 $\Delta = NWQ^2/\epsilon$, 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_τ and F_τ, and hence terminate for some finite τ. □

Lemma 8 *Procedure* ModifiedPump(ϵ, x, V) *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 $a' < b'$.*

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 $a' = \max_{v \notin P_\tau} m^v(x_\tau) < b' = (m^+ + m^-)/2$. 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 (7b), implying that conditions (N5) and (N6) hold. □

Let us observe that the bounds and strategies obtained by Lemmas 7 and 8 do not necessarily imply the ϵ-non-ergodicity of the game since those positions in I_τ and F_τ may not have enough separation in m-values (i.e. the condition $b' - a' \geq 3\epsilon$ 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(Γ, x, m_-, m_+) which takes the current potential vector x and range $[m_-, m_+]$ and produces another potential vector y such that

$\|y\|_\infty \leq 2^{\mathrm{poly}(n,N,\eta)}$. 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 $\beta = (\beta^v \mid v \in V)] \in \mathcal{L}(\Gamma)$:

$$\begin{aligned} \alpha^v A^v(x') &\geq m_- \cdot \mathbf{e}, & A^v(x')\beta^v &\leq m_+ \cdot \mathbf{e}, \\ \alpha^v \mathbf{e} &= 1, & \mathbf{e}\beta^v &= 1, \\ \alpha^v &\geq 0, & \beta^v &\geq 0, \end{aligned} \tag{12}$$

for all $v \in V$, where $\mathbf{e}$ 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 δ can be computed,using quantifier elimination algorithms, in time $\mathrm{poly}(\eta, N^{O(nN)}, \log\frac{1}{\delta})$; see [BPR96, GV88, Ren92]. Note that the resulting solution will satisfy (12) but within the approximate range $[m_- - \delta, m_+ + \delta]$. By choosing δ sufficiently smaller than the desired accuracy ϵ, we can ignore the effect of such approximation.

Lemma 9 ModifiedRepeatedPumping(ϵ) *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 24ϵ-ergodic, or outputs two nonempty subsets I and F and strategies α^v, $v \in I$, for player 1 and β^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 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 ModifiedPump, then we would also have $m^v(x_\tau) \geq \frac{5}{8}m^+ + \frac{3}{8}m^- = b'$ 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_τ, the potentials of these positions may go down, while those of the positions outside $\mathcal{I}_\tau$ 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_τ, 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 ModifiedPump.

Algorithm 3 ModifiedRepeatedPumping(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** $m^+(h) - m^-(h) \leq 24\epsilon$ **then**
4:     **return** x_h.
5: **end if**
6: x_{h+1} :=ModifiedPump($\epsilon, x_h, V, m_-, m_+$) 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} :=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} :=ModifiedPump($\epsilon, x_h, I_\tau, m_-, m_+$) and let T_τ, B_τ be the sets obtained from this call of ModifiedPump.
13: **if** $T_\tau = \emptyset$ **then**
14:     x_{h+1} :=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.

Let us note that as long as $m^+ - m^- > 24\epsilon$ we pump the upper half P_τ by exactly $\delta \geq 6\epsilon$. Let $\mathcal{P}_\tau(v)$ (resp., $\mathcal{N}_\tau(v)$) denote the number of iterations, among the first τ, 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 $j = 1, 2, ..., n-1$. Note that $\mathcal{P}_\tau(v_1) = \tau$ and $\mathcal{N}_\tau(v_n) = \tau$.

Let i_τ 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

$$0 \leq \widetilde{a}_{k\ell}^{v_{i+1}}(x_\tau) \leq R + \sum_{j=1}^{i} \Delta_j, \tag{13}$$

where the sum over the empty sum is zero by definition. Similarly, for $i = i_\tau + 1, \ldots, n$, we have

$$-R \leq \widetilde{b}_{k\ell}^{v_i}(x_\tau) \leq R + \sum_{j=1}^{n-i} \Delta_{n-j}. \tag{14}$$

From (13) and (14), it follows that

$$|R_\tau^{v_{i+1}}| \leq \begin{cases} R + \sum_{j=1}^{i} \Delta_i, & \text{for } i = 0, 1, 2, \ldots, i_\tau - 1 \\ R + \sum_{j=1}^{n-i-1} \Delta_{n-j}, & \text{for } i = i_\tau, i_\tau + 1, \ldots, n-1. \end{cases} \tag{15}$$

Let $\widetilde{i}_\tau$ be the smallest index i such that

$$\Delta_i > \frac{NW(R + \sum_{j=1}^{i-1} \Delta_j)^2}{\epsilon}, \tag{16}$$

and let $\widehat{i}_\tau$ be the largest index $i \leq n-1$ such that

$$\Delta_i > \frac{NW(R + \sum_{j=1}^{n-i-1} \Delta_{n-j})^2}{\epsilon}. \tag{17}$$

From the definition of $\widetilde{i}_\tau$, we know that

$$\Delta_i \leq \frac{NW(R + \sum_{j=1}^{i-1} \Delta_j)^2}{\epsilon}, \text{ for all } i = 1, \ldots, \widetilde{i}_\tau - 1.$$

Solving this recurrence, we get

$$x_\tau^{v_{\widetilde{i}_\tau}} - x_\tau^{v_1} = \sum_{i=1}^{\widetilde{i}_\tau - 1} \Delta_i \leq \left(\frac{(\widetilde{i}_\tau - 1)NWR}{\epsilon}\right)^{2^{\widetilde{i}_\tau - 1} - 1} (\widetilde{i}_\tau - 1)^2 R \leq \left(\frac{nNWR}{\epsilon}\right)^{2^n - 1} n^2 R. \tag{18}$$

Similarly, the definition of $\widehat{i}_\tau$ gives

$$\Delta_i \le \frac{NW(R+\sum_{j=1}^{n-i-1}\Delta_{n-i})^2}{\epsilon}, \text{ for all } i=\widehat{i}_\tau+1,\ldots,n-1,$$

from which follows

$$x_\tau^{v_n} - x_\tau^{v_{\widehat{i}_\tau+1}} \le \left(\frac{nNWR}{\epsilon}\right)^{2^n-1} n^2 R. \tag{19}$$

Note that if $\widetilde{i}_\tau \le i_\tau$ then (15) implies that taking $I_\tau = \{v_1,\ldots,v_{\widetilde{i}_\tau}\}$ would satisfy condition (N5) and guarantee that $I_\tau \subseteq P_\tau$.

Indeed, for all $i \le \widetilde{i}_\tau$ and $u \notin I_\tau$, we have

$$x_\tau^u - x_\tau^{v_i} \ge \Delta_{\widetilde{i}_\tau} > \frac{NW(R+\sum_{j=1}^{i-1}\Delta_j)^2}{\epsilon} \ge \frac{|L^{v_i}|W\left(R^{v_i}(x_\tau)\right)^2}{\epsilon}.$$

Similarly, having $\widehat{i}_\tau \ge i_\tau$ guarantees that taking $F_\tau = \{v_{\widehat{i}_\tau+1},\ldots,v_n\}$ would satisfy (N6) and $F_\tau \cap P_\tau = \emptyset$.

, since for all $i \ge \widehat{i}_\tau+1$ and $u \notin F_\tau$, we have

$$x_\tau^{v_i} - x_\tau^u \ge \Delta_{\widehat{i}_\tau} > \frac{NW(R+\sum_{j=1}^{n-i-1}\Delta_{n-j})^2}{\epsilon} \ge \frac{|K^{v_i}|W\left(R^{v_i}(x_\tau)\right)^2}{\epsilon}.$$

On the other hand, if $\widetilde{i}_\tau \ge i_\tau+1$, then (18) implies that $v_{i_\tau+1}$ was always pumped except for at most $\kappa(R) := \left(\frac{nNWR}{\epsilon}\right)^{2^n-1}\frac{n^2R}{\delta}$ iterations, that is, $\mathcal{N}_\tau(v_{i_\tau+1}) \le \kappa(R)$. Also, since $v_{i_\tau+1} \notin P_\tau$, then at time τ, $v_{i_\tau+1}$ is not pumped. Similarly, if $\widehat{i}_\tau < i_\tau$, then (19) implies that v_{i_τ} was never pumped except for at most $\kappa(R)$ iterations, that is, $\mathcal{P}_\tau(v_{i_\tau}) \le \kappa(R)$, while it is pumped at time τ. Since we have at most n candidates for each of v_{i_τ} and $v_{i_\tau+1}$, it follows that after $\tau = 2n\kappa(R)+1$, neither of these events ($\widetilde{i}_\tau \ge i_\tau+1$ and $\widehat{i}_\tau < i_\tau$) can happen, which by our earlier observations implies that the algorithm constructs the sets I_τ and F_τ. We can conclude that ModifiedPump(ϵ, x, V) must terminate in at most $2n\kappa(R)+1$ iterations, either producing $m^+ - m^- \le 24\epsilon$ or outputting the subsets I_τ and F_τ proving ϵ-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 $2^{\text{poly}(n,N,\eta)}$).

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 $\delta(2n\kappa(R)+1)$, it follows that the maximum absolute value of the transformed rewards at time τ is $r_{k\ell}^{vu}(x_\tau) \le R_2 := R + \delta(2n\kappa(R)+1)$

(note that the non-negativity 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 $2n\kappa(R_2) + 1 = \left(\frac{nNWR}{\epsilon}\right)^{O(2^{2n})}$.

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 η. 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

[BEGM10] E. Boros, K. Elbassioni, V. Gurvich, and K. Makino. A pumping algorithm for ergodic stochastic mean payoff games with perfect information. In *Proc. 14th IPCO*, pages 341–354, 2010.

[BEGM13a] E. Boros, K. Elbassioni, V. Gurvich, and K. Makino. On canonical forms for zero-sum stochastic mean payoff games. *Dynamic Games and Applications*, 3(2):128–161, 2013.

[BEGM13b] Endre Boros, Khaled Elbassioni, Vladimir Gurvich, and Kazuhisa Makino. A pseudo-polynomial algorithm for mean payoff stochastic games with perfect information and a few random positions. In FedorV. Fomin, Rsi Freivalds, Marta Kwiatkowska, and David Peleg, editors, *Automata, Languages, and Programming*, volume 7965 of *Lecture Notes in Computer Science*, pages 220–231. Springer Berlin Heidelberg, 2013.

[BEGM14] Endre Boros, Khaled Elbassioni, Vladimir Gurvich, and Kazuhisa Makino. A potential reduction algorithm for ergodic two-person zero-sum limiting average payoff stochastic games. In Zhao Zhang, Lidong Wu, Wen Xu, and Ding-Zhu Du, editors, *Combinatorial Optimization and Applications*, volume 8881 of *Lecture Notes in Computer Science*, pages 694–709. Springer International Publishing, 2014.

[BF68] D. Blackwell and T.S. Ferguson. The big match. *Ann. Math. Statist.*, 39(1):159–163, 1968.

[BPR96] S. Basu, R. Pollack, and M. Roy. On the combinatorial and algebraic complexity of quantifier elimination. *J. ACM*, 43(6):1002–1045, 1996. Preliminary version in *FOCS* 1994.

[CIJ14] K. Chatterjee and R. Ibsen-Jensen. The complexity of ergodic mean-payoff games. In *ICALP (2)*, pages 122–133, 2014.

[CMH08] K. Chatterjee, R. Majumdar, and T. A. Henzinger. Stochastic limit-average games are in exptime. *Int. J. Game Theory*, 37:219–234, 2008.

[Fed80] A. Federgruen. Successive approximation methods in undiscounted stochastic games. *Operations Research*, 1:794–810, 1980.

[Gal58] T. Gallai. Maximum-minimum Sätze über Graphen. *Acta Mathematica Academiae Scientiarum Hungaricae*, 9:395–434, 1958.

[Gil57] D. Gillette. Stochastic games with zero stop probabilities. In M. Dresher, A. W. Tucker, and P. Wolfe, editors, *Contribution to the Theory of Games III*, volume 39 of *Annals of Mathematics Studies*, pages 179–187. Princeton University Press, 1957.

[GV88] D. Grigoriev and N. Vorobjov. Solving systems of polynomial inequalities in subexponential time. *J. Symb. Comput.*, 5(1/2):37–64, 1988.

[HK66] A. J. Hoffman and R. M. Karp. On nonterminating stochastic games. *Management Science, Series A*, 12(5):359–370, 1966.

[HKL+11] K. A. Hansen, M. Koucky, N. Lauritzen, P. B. Miltersen, and E. P. Tsigaridas. Exact algorithms for solving stochastic games: extended abstract. In *Proceedings of the 43rd annual ACM symposium on Theory of computing*, STOC '11, pages 205–214, New York, NY, USA, 2011. ACM.

[KS63] J. G. Kemeny and J. L. Snell. *Finite Markov chains.* Springer, 1963.

[Mil11] P. B. Miltersen. Discounted stochastic games poorly approximate undiscounted ones, manuscript. Technical report, 2011.

[MN81] J.F. Mertens and A. Neyman. Stochastic games. *Int. J. Game Theory*, 10:53–66, 1981.

[MO70] H. Mine and S. Osaki. *Markovian decision process.* American Elsevier Publishing Co., New York, 1970.

[Mou76] H. Moulin. Prolongement des jeux à deux joueurs de somme nulle. *Bull. Soc. Math. France, Memoire*, 45, 1976.

[Ren92] J. Renegar. On the computational complexity and geometry of the first-order theory of the reals. *J. Symb. Comput.*, 13(3):255–352, 1992.

[RF91] T. E. S. Raghavan and J. A. Filar. Algorithms for stochastic games: A survey. *Mathematical Methods of Operations Research*, 35(6):437–472, 1991.

[Sha53] L. S. Shapley. Stochastic games. *Proc. Nat. Acad. Science, USA*, 39:1095–1100, 1953.

[Vri80] O. J. Vrieze. *Stochastic games with finite state and action spaces.* PhD thesis, Centrum voor Wiskunde en Informatica, Amsterdam, The Netherlands, 1980.