Oberwolfach Preprints

OWP 2010 - 22

ENDRE BOROS, KHALED ELBASSIONI, MAHMOUD FOUZ, VLADIMIR GURVICH, KAZUHISA MAKINO, BODO MANTHEY

Stochastic Mean Payoff Games: Smoothed Analysis and Approximation Schemes

Mathematisches Forschungsinstitut Oberwolfach gGmbH
Oberwolfach Preprints (OWP) ISSN 1864-7596

Oberwolfach Preprints (OWP)

Starting in 2007, the MFO publishes a preprint series which mainly contains research results related to a longer stay in Oberwolfach. In particular, this concerns the Research in Pairs-Programme (RiP) and the Oberwolfach-Leibniz-Fellows (OWLF), but this can also include an Oberwolfach Lecture, for example.

A preprint can have a size from 1 - 200 pages, and the MFO will publish it on its website as well as by hard copy. Every RiP group or Oberwolfach-Leibniz-Fellow may receive on request 30 free hard copies (DIN A4, black and white copy) by surface mail.

Of course, the full copy right is left to the authors. The MFO only needs the right to publish it on its website *www.mfo.de* as a documentation of the research work done at the MFO, which you are accepting by sending us your file.

In case of interest, please send a **pdf file** of your preprint by email to *rip@mfo.de* or *owlf@mfo.de*, respectively. The file should be sent to the MFO within 12 months after your stay as RiP or OWLF at the MFO.

There are no requirements for the format of the preprint, except that the introduction should contain a short appreciation and that the paper size (respectively format) should be DIN A4, "letter" or "article".

On the front page of the hard copies, which contains the logo of the MFO, title and authors, we shall add a running number (20XX - XX).

We cordially invite the researchers within the RiP or OWLF programme to make use of this offer and would like to thank you in advance for your cooperation.

Imprint:

Mathematisches Forschungsinstitut Oberwolfach gGmbH (MFO)
Schwarzwaldstrasse 9-11
77709 Oberwolfach-Walke
Germany

Tel +49 7834 979 50
Fax +49 7834 979 55
Email admin@mfo.de
URL www.mfo.de

The Oberwolfach Preprints (OWP, ISSN 1864-7596) are published by the MFO.

Stochastic Mean Payoff Games: Smoothed Analysis and Approximation Schemes *

Endre Boros[†] Khaled Elbassioni[‡] Mahmoud Fouz[§] Vladimir Gurvich[*]
Kazuhisa Makino[¶] Bodo Manthey[‖]

Abstract. We consider two-person zero-sum stochastic mean payoff games with perfect information modeled by a digraph with black, white, and random vertices. These *BWR-games* games are polynomially equivalent with the classical Gillette games, which include many well-known subclasses, such as cyclic games, simple stochastic games, stochastic parity games, and Markov decision processes. They can also be used to model parlor games such as Chess or Backgammon.

It is a long-standing open question whether a polynomial algorithm exists that solves BWR-games. In fact, a pseudo-polynomial algorithm for these games with an arbitrary number of random nodes would already imply their polynomial solvability. Currently, only two classes are known to have such a pseudo-polynomial algorithm: BW-games (the case with *no* random nodes) and *ergodic* BWR-games (i.e., in which the game's value does not depend on the initial position) with constant number of random nodes. In this paper, we show that the existence of a pseudo-polynomial algorithm for BWR-games with constant number of random vertices implies smoothed polynomial time complexity and the existence of absolute and relative polynomial-time approximation schemes. In particular, we obtain smoothed polynomial time complexity and derive absolute and relative approximation schemes for the above two classes.

1 Introduction

Stochastic games with perfect information and mean payoff were introduced in 1957 by Gillette [18]. In an equivalent formulation [20, 10, 7], which is called *mean stochastic payoff games* or *BWR-games*, we are given a directed graph $G = (V, E)$ whose vertex set V is partitioned into three subsets $V = V_B \cup V_W \cup V_R$ that correspond to black, white, and random positions, respectively. The black and white vertices are owned by two players: Black – the *minimizer* – owns the black vertices in V_B, and White – the *maximizer* – owns the white vertices in V_W. The vertices in V_R are owned by nature. Furthermore, we have a local reward $r_e \in \mathbb{R}$ for each arc $e \in E$. Finally, there are given probabilities p_{vu} for all arcs (v, u) going out of $v \in V_R$. Vertices $v \in V$ are also called *positions* and arcs $e \in E$ are also called *moves*. Starting from some vertex $v_0 \in V$, a token is moved along one arc e in every round of the game. If the token is on a black vertex, Black selects an outgoing arc e and moves the token along e. If the token is on a white vertex, then White selects an outgoing arc e. In a random position $v \in V_R$, a move $e = (v, u)$ is chosen according to the probabilities p_{vu} of the outgoing arcs of v. In all cases, Black pays White the reward r_e on the selected arc e.

From a given initial position $v_0 \in V$ the game produces an infinite walk $\{v_0, v_1, v_2, \ldots\}$ (called a *play*). Let b_i denote the reward $r_{v_i v_{i+1}}$ received by White in step $i \in \{0, 1, \ldots\}$. The *undiscounted limit average effective payoff* is defined as the *Cesàro average* $c = \liminf_{n\to\infty} \frac{\sum_{i=0}^{n} \mathbb{E}[b_i]}{n+1}$.

[*]This research was supported through the programme "Research in Pairs" by the Mathematisches Forschungsinstitut Oberwolfach in 2010.

[†]RUTCOR, Rutgers University, boros/gurvich@rutcor.rutgers.edu
[‡]Max-Planck-Institute for Informatics, elbassio@mpi-inf.mpg.de
[§]Saarland University, Department of Computer Science, mfouz@cs.uni-saarland.de
[¶]University of Tokyo, Graduate School of Information Science and Technology, makino@mist.i.u-tokyo.ac.jp
[‖]University of Twente, Department of Applied Mathematics, b.manthey@utwente.nl

WHITE's objective is to maximize c, while the objective of BLACK is the opposite, i.e., to minimize $\limsup_{n\to\infty} \frac{\sum_{i=0}^{n} \mathbb{E}[b_i]}{n+1}$.

The important special case of BWR-games without random vertices, i.e., $V_R = \emptyset$, is known as *cyclic* or *mean payoff* games [31, 17, 20]; for convenience, we will call these *BW-games*. A more special case was considered extensively in the literature under the name of *parity games* [11, 21, 22, 24], and later generalized also to include random vertices [10]. A BWR-game is reduced to a *minimum mean cycle problem* in case $V_W = V_R = \emptyset$ or $V_B = V_R = \emptyset$, which can be solved in polynomial time [25]. If one of the sets V_B or V_W is empty, we obtain a *Markov decision process* for which polynomial-time algorithms are also known [30]. Finally, if both sets are empty $V_B = V_W = \emptyset$, we get a *weighted Markov chain*.

In the special case of BWR-games when all rewards are zero except for m terminal loops we obtain the so-called *Backgammon-like* games [9]. In case of $m = 1$, we obtain *simple stochastic games* (SSGs), introduced by Condon [14, 15] and considered in several papers (e.g. [19, 21]). In these games, the objective of WHITE is to maximize the probability of reaching the terminal, while BLACK wants to minimize this probability. Recently, it was shown that Gillette games (and hence BWR-games [7]) are equivalent to SSGs under polynomial-time reductions [1]. Thus, by recent results of Halman [21], all these games can be solved in randomized strongly subexponential time $2^{O(\sqrt{n_d \log n_d})}\text{poly}(|V|)$, where $n_d = |V_B| + |V_W|$ is the number of *deterministic* vertices (see also [4, 29]). Let us note that several pseudo-polynomial and subexponential algorithms exists for BW-games [20, 26, 41, 33, 5, 21, 40]; see also Dhingra and Gaubert [16] for a policy iteration method, and Jurdzinski et al. [24] for parity games.

Besides their many applications (see, e.g., [28, 23]), all these games are of interest to complexity theory: Karzanov and Lebedev [26] (see also [41]) proved that the decision problem "whether the value of a BW-game is positive" is in the intersection of NP and co-NP. Yet, no polynomial algorithm is known even in this special case, see e.g., the recent survey by Vorobyov [40]. A similar complexity claim can be shown to hold for SSGs and BWR-games [1, 7]. On the other hand, there exists algorithms (see, e.g., [20]) that solve BW-games in practice very fast. The situation for these games is thus comparable to linear programming before the seminal discovery of the ellipsoid method, where the problem was also known to lie in the intersection of NP and co-NP and where the simplex algorithm proved to be a fast algorithm in practice. In [37], Spielman and Teng introduced smoothed analysis to explain the practical performance of the simplex method. We further enforce this analogy by showing a smoothed polynomial complexity for BWR-games. Recently, it was also shown that, in the *unit cost model*, a polynomial algorithm for linear programming would imply a polynomial algorithm for BW-games [36].

While there are numerous pseudo-polynomial algorithms known for the BW-case [20, 41, 33], pseudo-polynomiality for BWR-games (with no restriction on the number of random nodes) is in fact equivalent to polynomiality [1]. It was shown by Gimbert and Horn [19] that simple stochastic games on k random vertices can be solved in time $O(k!(|V||E| + L))$, where L is the maximum bit length of a transition probability. (Even though BWR-games are polynomially reducible to simple stochastic games, under this reduction the number of random vertices k becomes a polynomial in n, even if the original BWR-game has constantly many random vertices.) Recently, a pseudo-polynomial algorithm was given for BWR-games with a *constant* number of random vertices and polynomial common denominator of transition probabilities, but under the assumption that the game is *ergodic*, i.e., the game value does not depend on the initial position [8]. However, the existence of a similar algorithm for the non-ergodic or non-constant number of random vertices remains open, as the approach in [8] does not seem to generalize to these cases.

1.1 Our Results and Some Related Work

Approximation schemes for BWR-games. As for approximation schemes, the only result we are aware of is the observation made by Roth et al. [35] that the values of BW-games can be approximated within an *absolute* error of ε in polynomial-time, if all rewards are in the range $[-1, 1]$. This follows immediately from truncating the rewards and using any of the known pseudo-polynomial algorithms [20, 33, 41].

In this paper, we generalize this result in two directions. Let us say that a digraph $G = (V_B \cup V_W \cup V_R, E)$ admits a pseudo-polynomial algorithm, if there is an algorithm that solves any BWR-game $\mathcal{G}$ on G, with integral rewards and rational transition probabilities, in time polynomial in n, D, and R. Here, $n = n(\mathcal{G})$ is the total number of vertices, $R = R(\mathcal{G})$ is the size of the range of the rewards, and $D = D(\mathcal{G})$ is the common denominator of the transition probabilities. For instance, G admits a pseudo-polynomial algorithm if it has no random vertex (i.e., BW-games), or when it has a constant number of random nodes and is *structurally ergodic*, e.g., when G is a complete tripartite digraph; see [8] for more general sufficient conditions for structural ergodicity.

Let $p_{\min} = p_{min}(\mathcal{G})$ be the minimum positive transition probability in the game $\mathcal{G}$. Throughout the paper, we will assume that the number of random vertices k is bounded by a constant.

Theorem 1.1. *Let G be a digraph that admits a pseudo-polynomial algorithm. For any $\varepsilon > 0$, there is an algorithm that returns, for any given BWR-game on G with rewards in $[-1, 1]$, a pair of strategies that approximates the value, within an* absolute *error of ε, in time poly$(n, \frac{1}{p_{min}}, \frac{1}{\varepsilon})$.*

We also obtain an approximation scheme with a *relative* error.

Theorem 1.2. *Let G be a digraph that admits a pseudo-polynomial algorithm. For any $\varepsilon > 0$, there is an algorithm that returns, for any given BWR-game on G with non-negative integral rewards and rational transition probabilities, a pair of strategies that approximates the value, within a* relative *error of ε, in time* $\text{poly}(n, \log R, \frac{1}{\varepsilon})$.

Note that our reduction in Theorem 1.1, unlike Theorem 1.2, has the property that if the pseudo-polynomial algorithm returns optimal strategies that are *independent* of the starting vertex, the so-called *uniformly* optimal strategies, then so does the approximation scheme. For BW-games, i.e., the special case without random vertices, we can also strengthen the result of Theorem 1.2 to return a pair of strategies that is uniformly ε-optimal.

In deriving these approximation schemes from a pseudo-polynomial algorithm, we face two main technical challenges which distinguish the computation of ε-equilibria of BWR-games from similar standard techniques used in optimization: (i) the running time of the pseudo-polynomial algorithm depends polynomially both on the maximum reward and the common denominator D of the transition probabilities; thus to obtain an FPTAS with an absolute guarantee whose running time is independent of D, we need to truncate the probabilities and bound the change in the game value, which is a *non-linear* function of D, (ii) to obtain an FPTAS with a relative guarantee, one needs (as standard in optimization) a (trivial) lower/upper bound on the optimum value; this is not possible in the case of BWR-games, since the game value can be arbitrarily small; the situation becomes even more complicated, if we look for uniformly ε-optimal strategies, since we have to output one pair of strategies which guarantees ε-optimality from any starting position.

In order to resolve the first issue, we analyze the change in the game values and optimal strategies if the rewards or transition probabilities are changed. Roughly speaking, we use results from Markov chain perturbation theory to show that if the probabilities are perturbed by a small error δ, then the change in the game value is $O(\frac{\delta^2 n^3}{p_{\min}^{2k}})$; see Section 3.2. The second issue is resolved through repeated applications of the pseudo-polynomial algorithm on a truncated game; after each such application either the value of the game has already been approximated within the required accuracy, or it is guaranteed that the range of the rewards can be shrunk by a constant factor without changing the value of the game; see Sections 3.4 and 3.5.

Since BW-games and structurally ergodic BWR-games with constant number of random vertices admit pseudo-polynomial algorithms, we obtain the following results.

Corollary 1.3. *(i) There is an FPTAS that solves, within a relative error guarantee, in uniformly ε-optimal strategies, any BW-games with non-negative (rational) rewards.*

(ii) There is an FPTAS that solves, within an absolute error guarantee, in uniformly ε-optimal strategies, any structurally ergodic BWR-game with $1/p_{\min} = \text{poly}(n)$ and rewards in $[-1, 1]$.

(iii) There is an FPTAS that solves, within a relative error guarantee, in uniformly ε-optimal strategies, any structurally ergodic BWR-game with $1/p_{\min} = \text{poly}(n)$ and non-negative rational rewards.

Note that (i) strengthens the absolute FPTAS for BW-games [35], and (ii) and (iii) enlarge the class of games for which an FPTAS exists.

Smoothed Analysis for BWR-games. We further show that typical instances of digraphs that admit a pseudo-polynomial algorithm can be solved in polynomial time. Towards this end, we do a smoothed analysis using the one-step model introduced by Beier and Vöcking [3]: Given an upper bound ϕ for the densities, an adversary specifies a BWR game $\mathcal{G}$ together with density functions, one for each arc. Then the rewards for all arcs are drawn independently according to their respective density functions. We prove that in this setting, independent of the actual choices of the adversary, the resulting game can be solved in polynomial time with high probability, which shows that such BWR-games with a constant number of random vertices have smoothed polynomial complexity. This means that there exists a polynomial $P(n, \phi, 1/\varepsilon)$ such that the probability that the algorithm exceeds a running-time of $P(n, \phi, 1/\varepsilon)$ is at most ε.

Theorem 1.4. *Let G be a digraph that admits a pseudo-polynomial algorithm. There is an algorithm that solves any BWR-game on G with rational transition probabilities and $D = poly(n)$ in smoothed polynomial time.*

Our proof of Theorem 1.4 follows the general paradigm introduced by Beier and Vöcking [3] for using a pseudo-polynomial algorithm to analyze the smoothed complexity of integer programs. However, in the case of BWR-games, the situation becomes more complicated by the following two facts. First, we have to deal with two different objectives (of the two players), and, second, the coefficients of the objectives are not given explicitly, but correspond to the limiting distributions in the Markov chains corresponding to the different strategies. To prove that BWR-games can be solved in smoothed polynomial time, we first need a new isolation lemma. In contrast to the existing isolation lemmas used in smoothed analysis of optimization problems, our isolation lemma has to deal with two players who optimize the same objective function in two different directions. Second, our procedure for certifying that the solution found is indeed the optimal solution is considerably more involved. The reason is again that we have two competing players, which requires careful rounding of the coefficients in order to certify optimality.

Chen et al. [12] have analyzed the smoothed complexity of the Lemke-Howson algorithm for computing equilibria in bimatrix games. They have shown that the Lemke-Howson algorithm – and also any other algorithm – has smoothed polynomial complexity only if all problems in PPAD can be solved in randomized polynomial time. In contrast to their negative result, our smoothed analysis shows that equilibria of BWR-games can typically be computed efficiently.

As a corollary, we obtain the following results.

Corollary 1.5. *(i) BW-games can be solved in smoothed polynomial time.*

(ii) Structurally ergodic BWR-games with $D = \text{poly}(n)$ *can be solved in smoothed polynomial time.*

Let us remark finally that removing the assumption that k is constant in the above results remains a challenging open problem that seems to require totally new ideas.

2 Preliminaries, Notation and Basic Properties

2.1 BWR-games and Markov Chains

A BWR-game is defined by a triple $\mathcal{G} = (G, P, r)$, where $G = (V = V_W \cup V_B \cup V_R, E)$ is a digraph that may have loops and multiple arcs, but no terminal vertices, i.e., vertices of out-degree 0; $P \in [0,1]^E$ is the vector of probability distributions for all $v \in V_R$ specifying the probability p_{vu} of a move from v to u; and $r \in \mathbb{R}^E$ is a local reward function. It is assumed that $\sum_{u:(v,u)\in E} p_{vu} = 1$ for all $v \in V_R$ and $p_{v,u} > 0$ whenever $(v,u) \in E$ and $v \in V_R$.

Standardly, we define a strategy $s_W \in S_W$ for White as a mapping that assigns a move $(v,u) \in E$ to each position $v \in V_W$. For simplicity, we may write $s_W(v) = u$ for $s_W(v) = (v,u)$. Strategies $s_B \in S_B$ for Black are analogously defined. A pair of strategies $s = (s_W, s_B)$ is called a *situation*. Given a BWR-game $\mathcal{G} = (G, P, r)$ and a situation $s = (s_B, s_W)$, we obtain a weighted Markov chain $\mathcal{G}(s) = (G(s) = (V, E(s)), P(s), r)$ with transition matrix $P(s)$ defined in the obvious way:

$$p_{vu}(s) = \begin{cases} 1 & \text{if } (v \in V_W \text{ and } u = s_W(v)) \text{ or } (v \in V_B \text{ and } u = s_B(v)); \\ 0 & \text{if } (v \in V_W \text{ and } u \neq s_W(v)) \text{ or } (v \in V_B \text{ and } u \neq s_B(v)); \\ p_{vu} & \text{if } v \in V_R. \end{cases}$$

Here, $E(s) = \{e \in E \mid p_e(s) > 0\}$ is the set of arcs with positive probability. Given an initial position $v_0 \in V$ from which the play starts, we define the limiting (mean) effective payoff $c_{v_0}(s)$ in $\mathcal{G}(s)$ as $c_{v_0}(s) = \rho(s)^T r = \sum_{e\in E} \rho_e(s) r_e$, where $\rho(s) = \rho(s, v_0) \in [0,1]^E$ is the arc-limiting distribution for $\mathcal{G}(s)$ starting from v_0. This means that for $(v,u) \in E$, $\rho_{vu}(s) = \pi_v(s) p_{vu}(s)$, where $\pi \in [0,1]^V$ is the limiting distribution in the Markov chain $\mathcal{G}(s)$ starting from v_0. In what follows, we will use $(\mathcal{G}, v_0)$ to denote the game starting from v_0. We will simply write $\rho(s)$ for $\rho(s, v_0)$, when v_0 is clear from the context. For rewards $r : E \to \mathbb{R}$, let $r^- = \min_e r_e$ and $r^+ = \max_e r_e$. Let $[r] = [r^-, r^+]$ be the range of r. Let $R = R(\mathcal{G}) = r^+ - r^-$.

2.2 Strategies and Saddle Points

If we consider $c_{v_0}(s)$ for all possible situations, we obtain a matrix game $C_{v_0} : S_W \times S_B \to \mathbb{R}$, with entries $C_{v_0}(s_W, s_B) = c_{v_0}(s_W, s_B)$. It is known that every such game has a *saddle point* in pure strategies [18, 27]. Such a saddle point defines an equilibrium state in which no player has an incentive to change her strategy, and as shown in [18, 27], the value at that state coincides with the limiting payoff in the corresponding BWR-game. We call a pair of strategies optimal if they correspond to a saddle point. It is well-known that there exists optimal strategies (s_W^*, s_B^*) that do not depend on the starting position v_0. Such strategies are called *uniformly optimal*. Although there might be several optimal strategies, it is easy to see that they all lead to the same value. We define this to be the value of the game and write $\mu_{v_0}(\mathcal{G}) := C_{v_0}(s_W^*, s_B^*)$ where (s_W^*, s_B^*) is any pair of optimal strategies. Note that $\mu_{v_0}(\mathcal{G})$ may depend on the starting node v_0. Note also that for a situation s, $\mu_u(\mathcal{G}(s))$ denotes the effective payoff $c_u(s)$ in the Markov chain $\mathcal{G}(s)$.

3 Approximation Schemes

3.1 Approximation and Approximate Equilibria

Given a BWR-game $\mathcal{G} = (G = (V, E), P, r)$, a constant $\varepsilon > 0$, and a starting position $v \in V$, an *ε-relative approximation* of the value of the game is determined by a situation (s_W^*, s_B^*) such that

$$\max_{s_W} \mu_v(\mathcal{G}(s_W, s_B^*)) \le (1+\varepsilon)\mu_v(\mathcal{G}) \quad \text{and} \quad \min_{s_B} \mu_v(\mathcal{G}(s_W^*, s_B)) \ge (1-\varepsilon)\mu_v(\mathcal{G}). \tag{1}$$

We may also consider *ε-relative equilibrium.* This is determined by a situation (s_W^*, s_B^*) such that

$$\max_{s_W} \mu_v(\mathcal{G}(s_W, s_B^*)) \le (1+\varepsilon)\mu_v(\mathcal{G}(s_W^*, s_B^*)) \quad \text{and} \quad \min_{s_B} \mu_v(\mathcal{G}(s_W^*, s_B)) \ge (1-\varepsilon)\mu_v(\mathcal{G}(s_W^*, s_B^*)). \tag{2}$$

Note that, for sufficiently small ε, an *ε-relative approximation* implies a *$\Theta(\varepsilon)$-relative equilibrium*, and vice versa. So, in what follows, we shall use these notions interchangeably.

An alternative to relative approximations is to look for an approximation with *absolute* error of ε. This is achieved by a situation (s_W^*, s_B^*) such that

$$\max_{s_W} \mu_v(\mathcal{G}(s_W, s_B^*)) \le \mu_v(\mathcal{G}) + \varepsilon \quad \text{and} \quad \min_{s_B} \mu_v(\mathcal{G}(s_W^*, s_B)) \ge \mu_v(\mathcal{G}) - \varepsilon, \tag{3}$$

or, equivalently, for an ε-absolute equilibrium:

$$\max_{s_W} \mu_v(\mathcal{G}(s_W, s_B^*)) \le \mu_v(\mathcal{G}(s_W^*, s_B^*)) + \varepsilon \quad \text{and} \quad \min_{s_B} \mu_v(\mathcal{G}(s_W^*, s_B)) \ge \mu_v(\mathcal{G}(s_W^*, s_B^*)) - \varepsilon. \tag{4}$$

Again, an *ε-absolute approximation* implies a *2ε-absolute equilibrium*, and vice versa.

A situation (s_W^*, s_B^*) satisfying (1) is called relative ε-optimal. If a situation satisfies (3), it is called absolute ε-optimal. In the following, we will drop the specification of absolute and relative, when it is clear from the context. If the pair (s_W^*, s_B^*) is ε-optimal for any starting position, it is called *uniformly* ε-optimal.

When considering relative errors, we assume that the rewards are non-negative. If we consider absolute errors, then we assume that the rewards lie in a certain range, say, $[-1, 1]$. Under such assumptions, the notion of relative approximation becomes stronger. Indeed, an ε-relative approximation of the game $\hat{\mathcal{G}}$ with local rewards given by $\hat{r} = r + \mathbf{1} \ge 0$, where $\mathbf{1}$ is the vector of all ones, implies strategies (s_W^*, s_B^*) satisfying

$$\begin{aligned}
\max_{s_W} \mu_v(\mathcal{G}(s_W, s_B^*)) &= \max_{s_W} \mu_v(\hat{\mathcal{G}}(s_W, s_B^*)) - 1 \le (1+\varepsilon)\mu_v(\hat{\mathcal{G}}) - 1 \\
&= \mu_v(\mathcal{G}) + \varepsilon\mu_v(\mathcal{G}) + \varepsilon \le \mu_v(\mathcal{G}) + 2\varepsilon \quad \text{and} \\
\min_{s_B} \mu_v(\mathcal{G}(s_W^*, s_B)) &= \min_{s_B} \mu_v(\hat{\mathcal{G}}(s_W^*, s_B)) - 1 \ge (1-\varepsilon)\mu_v(\hat{\mathcal{G}}) - 1 \\
&= \mu_v(\mathcal{G}) - \varepsilon\mu_v(\mathcal{G}) - \varepsilon \ge \mu_v(\mathcal{G}) - 2\varepsilon,
\end{aligned}$$

since $\mu_v(\hat{\mathcal{G}}(s)) = \mu_v(\mathcal{G}(s)) + 1$ for any situation s, and $\mu_v(\mathcal{G}) \le 1$. Thus, we obtain a 2ε-absolute approximation for the value of the original game.

An algorithm for approximating the values of the game is said to be a fully polynomial-time approximation scheme (FPTAS), if the running time depends polynomially on the input size and $1/\varepsilon$. In what follows, we will assume without loss of generality that that $\frac{1}{\varepsilon}$ is an integer.

3.2 The Effect of Perturbation

Our approximation schemes are based on the following three propositions. The first one (which is well-known) says that a linear change in the rewards will correspond to a linear change in the game value. In our approximation schemes, we truncate and scale the rewards to be able to run the pseudo-polynomial algorithm in polynomial-time. We will then need the proposition to bound the error in the game value resulting from the truncation.

Proposition 3.1. *Let $\mathcal{G} = (G = (V, E), P, r)$ be a BWR-game. Let $\theta_1, \gamma_1, \theta_2, \gamma_2$ be constants such that $\theta_1, \theta_2 > 0$. Let $\hat{\mathcal{G}}$ be a game $(G = (V, E), P, \hat{r})$ with $\theta_1 r + \gamma_1 \mathbf{1} \le \hat{r} \le \theta_2 r + \gamma_2 \mathbf{1}$. Then for any $v \in V$, we have $\theta_1 \mu_v(\mathcal{G}) + \gamma_1 \le \mu_v(\hat{\mathcal{G}}) \le \theta_2 \mu_v(\mathcal{G}) + \gamma_2$. Moreover, if $(\hat{s}_W, \hat{s}_B)$ is an absolute ε-optimal situation in $(\hat{\mathcal{G}}, v)$, then*

$$\max_{s_W} \mu_v(\mathcal{G}(s_W, \hat{s}_B)) \le \frac{\theta_2 \mu_v(\mathcal{G}) + \gamma_2 - \gamma_1 + \varepsilon}{\theta_1} \text{ and } \min_{s_B} \mu_v(\mathcal{G}(\hat{s}_W, s_B)) \ge \frac{\theta_1 \mu_v(\mathcal{G}) + \gamma_1 - \gamma_2 - \varepsilon}{\theta_2}. \quad (5)$$

Proof. This is somewhat standard. We give a proof for completeness. Let (s_W^*, s_B^*) and $(\hat{s}_W, \hat{s}_B)$ be pairs of optimal strategies for $(\mathcal{G}, v)$ and $(\hat{\mathcal{G}}, v)$, respectively. Denote by $\rho^*, \hat{\rho}, \rho'$, and ρ'' the (arc) limiting distributions, for the Markov chains starting from v, corresponding to pairs of strategies (s_W^*, s_B^*), $(\hat{s}_W, \hat{s}_B)$, $(s_W^*, \hat{s}_B)$, and $(\hat{s}_W, s_B^*)$, respectively. By definitions of optimal strategies and the facts that $\mathbf{1}^T \rho' = \mathbf{1}^T \rho'' = 1$, we have the following series of inequalities:

$$\begin{aligned}
\mu_v(\hat{\mathcal{G}}) &= (\hat{\rho})^T \hat{r} \ge (\rho')^T \hat{r} \ge \theta_1 (\rho')^T r + \gamma_1 \ge \theta_1 (\rho^*)^T r + \gamma_1 = \theta_1 \mu_v(\mathcal{G}) + \gamma_1, \\
\mu_v(\hat{\mathcal{G}}) &= (\hat{\rho})^T \hat{r} \le (\rho'')^T \hat{r} \le \theta_2 (\rho'')^T r + \gamma_2 \le \theta_2 (\rho^*)^T r + \gamma_2 = \theta_2 \mu_v(\mathcal{G}) + \gamma_2.
\end{aligned}$$

To see the first bound in (5), note that for any s_W, $\mu_v(\mathcal{G}(s_W, \hat{s}_B)) \le \frac{1}{\theta_1}(\mu_v(\hat{\mathcal{G}}(s_W, \hat{s}_B)) - \gamma_1)$. Also, by the ε-optimality of $\hat{s}_W$ in $(\hat{\mathcal{G}}, v)$, we have $\mu_v(\hat{\mathcal{G}}(s_W, \hat{s}_B)) \le \mu_v(\hat{\mathcal{G}}) + \varepsilon \le \theta_2 \mu_v(\mathcal{G}) + \gamma_2 + \varepsilon$. Thus the first bound in (5) follows. The second bound can be shown similarly. □

The second proposition, which is to the best of our knowledge new, states that if we truncate the transition probabilities within a small error ε, then the change in the game value is $O(\frac{\varepsilon^2 n^3}{p_{\min}^{2k}})$. More precisely, for a BWR-game $\mathcal{G}$ and a constant $\varepsilon > 0$, define

$$\delta(\mathcal{G}, \varepsilon) := \left(\frac{\varepsilon}{2} n^2 (\frac{1}{2} p_{\min})^{-k} [\varepsilon n k (k+1) (\frac{1}{2} p_{\min})^{-k} + 3k + 1] + \varepsilon n \right) r_*, \quad (6)$$

where $n = n(\mathcal{G})$, $p_{\min} = p_{\min}(\mathcal{G})$, $k = k(\mathcal{G})$, and $r_* = r_*(\mathcal{G}) := \max\{|r_+(\mathcal{G})|, |r_-(\mathcal{G})|\}$.

Proposition 3.2. *Let $\mathcal{G} = (G = (V, E), P, r)$ be a BWR-game, with $r \in [-1, 1]^E$, and $\varepsilon \le \frac{1}{2} p_{\min} = \frac{1}{2} p_{\min}(\mathcal{G})$ be a positive constant. Let $\hat{\mathcal{G}}$ be a game $(G = (V, E), \hat{P}, r)$ with $\|P - \hat{P}\|_\infty \le \varepsilon$. Then for any $v \in V$, we have $|\mu_v(\mathcal{G}) - \mu_v(\hat{\mathcal{G}})| \le \delta(\mathcal{G}, \varepsilon)$. Moreover, if the pair $(\tilde{s}_W, \tilde{s}_B)$ is absolute ε'-optimal in $(\hat{\mathcal{G}}, v)$, then it is absolute $(\varepsilon' + 2\delta(\mathcal{G}, \varepsilon))$-optimal in $(\mathcal{G}, v)$.*

Proof. We make use of technical lemma A.3. Let (s_W^*, s_B^*) and $(\hat{s}_W, \hat{s}_B)$ be pairs of optimal strategies for $(\mathcal{G}, v)$ and $(\hat{\mathcal{G}}, v)$, respectively. Write $\delta = \delta(\mathcal{G}, \varepsilon)$. Then optimality and Lemma A.3 imply the following series of inequalities:

$$\begin{aligned}
\mu_v(\hat{\mathcal{G}}) &= \mu_v(\hat{\mathcal{G}}(\hat{s}_W, \hat{s}_B)) \ge \mu_v(\hat{\mathcal{G}}(s_W^*, \hat{s}_B)) \ge \mu_v(\mathcal{G}(s_W^*, \hat{s}_B)) - \delta \ge \mu_v(\mathcal{G}(s_W^*, s_B^*)) - \delta = \mu_v(\mathcal{G}) - \delta \\
\mu_v(\hat{\mathcal{G}}) &= \mu_v(\hat{\mathcal{G}}(\hat{s}_W, \hat{s}_B)) \le \mu_v(\hat{\mathcal{G}}(\hat{s}_W, s_B^*)) \le \mu_v(\mathcal{G}(\hat{s}_W, s_B^*)) + \delta \le \mu_v(\mathcal{G}(s_W^*, s_B^*)) + \delta = \mu_v(\mathcal{G}) + \delta.
\end{aligned}$$

To see the second claim, note that for any $s_W \in S_W$,

$$\mu_v(\mathcal{G}(s_W, \tilde{s}_B)) \le \mu_v(\hat{\mathcal{G}}(s_W, \tilde{s}_B)) + \delta \le \mu_v(\hat{\mathcal{G}}(\hat{s}_W, \hat{s}_B)) + \varepsilon' + \delta \le \mu_v(\mathcal{G}) + \varepsilon' + 2\delta.$$

Similarly, we can show that $\mu_v(\mathcal{G}(\tilde{s}_W, s_B)) \ge \mu_v(\mathcal{G}) - \varepsilon' - 2\delta$ for all $s_B \in S_B$. □

Since we assume that the running time of the pseudo-polynomial algorithm depends on the common denominator D of the transition probabilities, we need to truncate the probabilities to remove this dependence on D. By Proposition 3.2, the value of the game does not change too much after such truncation.

The third result we need concerns relative approximation. The main idea is to use the pseudo-polynomial algorithm to test whether the value of the game is larger than a certain threshold. If it is, we get already a good relative approximation. Otherwise, the next proposition says that we can somehow reduce all large rewards without changing the value of the game.

Proposition 3.3. *Let $\mathcal{G} = (G = (V, E), P, r)$ be a BWR-game with $r \geq 0$, and v be any vertex such that $\mu_v(\mathcal{G}) < t$. Suppose that for some $e \in E$, $r_e \geq t' = ntp_{\min}^{-(2k+1)}$. Let $\hat{\mathcal{G}} = (G = (V, E), P, \hat{r})$, where $\hat{r}_e \geq (1 + \varepsilon)t'$, for some $\varepsilon \geq 0$, and $\hat{r}_{e'} = r_{e'}$ for all $e' \neq e$. Then $\mu_v(\hat{\mathcal{G}}) = \mu_v(\mathcal{G})$, and any relative ε-optimal situation in $(\hat{\mathcal{G}}, v)$ is also ε-optimal in $(\mathcal{G}, v)$.*

Proof. Let $s^* = (s^*_W, s^*_B)$ be an optimal strategy for $(\mathcal{G}, v)$, that is, $\mu_v(\mathcal{G}) = \mu_v(\mathcal{G}(s^*)) = \rho(s^*)^T r < t$. By Lemma A.1, if $\rho_e(s^*) > 0$ then $\rho_e(s^*) \geq p_{\min}^{2k+1}/n$, and hence $r_e \rho_e(s^*) \leq \rho(s^*)^T r = \mu_v(\mathcal{G}) < t$ implies that $r_e < t'$. We conclude that $\rho_e(s^*) = 0$, and hence $\mu_v(\hat{\mathcal{G}}(s^*)) = \mu_v(\mathcal{G})$.

Since $\hat{r} \leq r$, $\mu_v(\hat{\mathcal{G}}(s)) \leq \mu_v(\mathcal{G}(s))$ for all situations s. In particular, for any $s_W \in S_W$,

$$\mu_v(\hat{\mathcal{G}}(s_W, s^*_B)) \leq \mu_v(\mathcal{G}(s_W, s^*_B)) \leq \mu_v(\mathcal{G}(s^*_W, s^*_B)) = \mu_v(\hat{\mathcal{G}}(s^*_W, s^*_B)).$$

We claim also that $\mu_v(\hat{\mathcal{G}}(s^*_W, s_B)) \geq \mu_v(\hat{\mathcal{G}}(s^*_W, s^*_B))$ for all $s_B \in S_B$. Indeed, if there is an s_B such that $\mu_v(\hat{\mathcal{G}}(s^*_W, s_B)) < \mu_v(\hat{\mathcal{G}}(s^*_W, s^*_B)) = \mu_v(\mathcal{G}) < t$, then by the same argument above, since $\rho_e(s^*_W, s_B)(1 + \varepsilon)t' \leq \rho_e(s^*_W, s_B)\hat{r}_e \leq \rho(s^*_W, s_B)^T \hat{r} = \mu_v(\hat{\mathcal{G}}(s^*_W, s_B)) < t$, we must have $\rho_e(s^*_W, s_B) = 0$, implying that $\mu_v(\mathcal{G}(s^*_W, s_B)) = \mu_v(\hat{\mathcal{G}}(s^*_W, s_B)) < \mu_v(\hat{\mathcal{G}}(s^*_W, s^*_B)) = \mu_v(\mathcal{G}(s^*_W, s^*_B))$, in contradiction to the optimality of s^* in $\mathcal{G}$. We conclude that (s^*_W, s^*_B) is also optimal in $\hat{\mathcal{G}}$ and hence $\mu_v(\hat{\mathcal{G}}) = \mu_v(\mathcal{G})$.

Suppose that $(\hat{s}_W, \hat{s}_B)$ is a relative ε-optimal situation in $(\hat{\mathcal{G}}, v)$. Then for any $s_W \in S_W$, $\rho_e(s_W, \hat{s}_B) = 0$. Indeed,

$$\begin{aligned}\rho_e(s_W, \hat{s}_B)(1 + \varepsilon)t' = \rho_e(s_W, \hat{s}_B)\hat{r}_e \leq \rho(s_W, \hat{s}_B)^T \hat{r} &= \mu_v(\hat{\mathcal{G}}(s_W, \hat{s}_B)) \\ &\leq (1 + \varepsilon)\mu_v(\hat{\mathcal{G}}) = (1 + \varepsilon)\mu_v(\mathcal{G}) < (1 + \varepsilon)t,\end{aligned}$$

gives a contradiction with Lemma A.1 if $\rho_e(s_W, \hat{s}_B) > 0$. It follows that, for any $s_W \in S_W$, $\mu_v(\mathcal{G}(s_W, \hat{s}_B)) = \mu_v(\hat{\mathcal{G}}(s_W, \hat{s}_B)) \leq (1 + \varepsilon)\mu_v(\mathcal{G})$. Furthermore, for any $s_B \in S_B$,

$$\mu_v(\mathcal{G}(\hat{s}_W, s_B)) \geq \mu_v(\hat{\mathcal{G}}(\hat{s}_W, s_B)) \geq (1 - \varepsilon)\mu_v(\hat{\mathcal{G}}) = (1 - \varepsilon)\mu_v(\mathcal{G}).$$

□

3.3 Absolute Approximation

Let $G = (V, E)$ be a graph that admits a pseudo-polynomial algorithm $\mathbb{A}$ and $\mathcal{G} = (G, P, r)$ be a BWR-game on G. In this section, we assume that $r^- = -1$ and $r^+ = 1$, i.e., all rewards are from the interval $[-1, 1]$. We apply the pseudo-polynomial algorithm $\mathbb{A}$ on a truncated game $\tilde{\mathcal{G}} = (G = (V, E), \tilde{P}, \tilde{r})$ defined by rounding the rewards to the nearest integer multiple of $\varepsilon/4$ (denoted $\tilde{r} := \lfloor r \rceil_{\frac{\varepsilon}{4}}$) and truncating the vector of probabilities $(p_{vu} : u \in V)$ for each random node $v \in V_R$, as described in the following lemma.

Lemma 3.4. *Let $\alpha \in [0, 1]^n$ with $\|\alpha\|_1 = 1$. Let $B \in \mathbb{Z}^+$ be an integer such that $\min_{i:\alpha_i > 0}\{\alpha_i\} > 2^{-B}$. Then there exists $\alpha' \in [0, 1]^n$ such that (i) $\|\alpha'\|_1 = 1$; (ii) for all $i = 1, \ldots, n$, $\alpha'_i = c_i/2^B$ where $c_i \in \mathbb{Z}^+$ is an integer; (iii) for all $i = 1, \ldots, n$, $\alpha'_i > 0$ if and only $\alpha_i > 0$, and (iv) $\|\alpha - \alpha'\|_\infty \leq 2^{-B}$.*

Proof. This is standard and easy. We include the proof for completeness. Without loss of generality $\alpha_i > 0$ for all i. Initialize $\varepsilon_0 = 0$ and iterate, for $i = 1, \ldots, n$: $\alpha'_i := \lfloor \alpha_i + \varepsilon_{i-1} \rceil_{2^{-B}}$; $\varepsilon_i := \alpha_i + \varepsilon_{i-1} - \alpha'_i$. Then $|\varepsilon_i| \leq 2^{-(B+1)}$ for all i, and $\varepsilon_n = \sum_i \alpha_i - \sum_i \alpha'_i$, implying (i). Furthermore, $|\alpha_i - \alpha'_i| = |\varepsilon_i - \varepsilon_{i-1}| \leq 2^{-B}$. Note finally that (iii) follows from (iv) since $\min_{i:\alpha_i > 0}\{\alpha_i\} > 2^{-B}$. □

Lemma 3.5. *Let $G = (V, E)$ be a graph admitting a pseudo-polynomial algorithm $\mathbb{A}$ that solves, in (uniformly) optimal strategies, any BWR-game on G in time $\tau(n, D, R)$. Then for any $\varepsilon > 0$, there is an algorithm that solves, in (uniformly) absolute ε-optimal strategies, any given BWR-game $\mathcal{G} = (G, P, r)$ in time bounded by $\tau(n, \frac{2^{2k+5}n^3k^2}{\varepsilon p_{\min}^{2k}}, \frac{4}{\varepsilon})$, where $p_{\min} = p_{\min}(\mathcal{G})$.*

Proof. We apply $\mathbb{A}$ to the game $\tilde{\mathcal{G}} = (G, \tilde{P}, \tilde{r})$, where $\tilde{r} := \frac{2}{\varepsilon}\lfloor r \rceil_{\frac{\varepsilon}{2}}$, and $\tilde{P}$ is obtained from P by applying Lemma 3.4 with $B = \lceil \log \frac{1}{\varepsilon'} \rceil$, where we select ε' such that $\delta(\mathcal{G}, \varepsilon') \le \frac{\varepsilon}{4}$ (as defined by (6)). Note that all rewards in $\tilde{\mathcal{G}}$ are integer in the range $[-\frac{4}{\varepsilon}, \frac{4}{\varepsilon}]$. Note also that $\delta(\mathcal{G}, \varepsilon') \le 4^{k+1}\varepsilon' n^3 k^2 p_{\min}^{-2k}$, and hence $\delta(\mathcal{G}, \varepsilon') \le \frac{\varepsilon}{4}$ for $\varepsilon' = \frac{\varepsilon p_{\min}^{2k}}{4^{k+2}n^3k^2}$. Since $D(\tilde{\mathcal{G}}) = 2^B$ and $R(\tilde{\mathcal{G}}) = \frac{4}{\varepsilon}$, the statement on the running time follows.

Let $\tilde{s}$ be the pair of (uniformly) optimal strategies returned by $\mathbb{A}$ (on $\tilde{\mathcal{G}}$). Let $\hat{\mathcal{G}}$ be the game $(G, \tilde{P}, r)$. Since $\|\tilde{r} - \frac{2}{\varepsilon}r\|_\infty \le 1$, we can apply Proposition 3.1 (with $\theta_1 = \theta_2 = \frac{2}{\varepsilon}$ and $\gamma_1 = -\gamma_2 = -\frac{1}{2}$) to conclude that $\tilde{s}$ is a (uniformly) $\frac{\varepsilon}{2}$-optimal pair for $\hat{\mathcal{G}}$. Now applying Proposition 3.2, we conclude further that $\tilde{s}$ is (uniformly) $(\frac{\varepsilon}{2} + 2\delta(\mathcal{G}, \varepsilon'))$-optimal for $\mathcal{G}$. □

Note that the above technique works (i.e., runs in polynomial time) only for the case $k = O(1)$, even if the given pseudo-algorithm $\mathbb{A}$ works for *any* k.

3.4 Relative approximation

Let $G = (V, E)$ be a graph that admits a pseudo-polynomial algorithm $\mathbb{A}$ and $\mathcal{G} = (G, P, r)$ be a BWR-game on G, with non-negative rational rewards (i.e., $r^- = 0$). Without loss of generality, we may assume that the rewards are integral with $\min_{e: r_e > 0} r_e = 1$. The algorithm is given as Algorithm 1. The main idea is to truncate the rewards, scaled by a certain factor $1/K$, and use the pseudo-polynomial algorithm on the truncated game $\hat{\mathcal{G}}$. If the value in the truncated game $\mu_w(\hat{\mathcal{G}})$, from the starting node w, is large enough (step 7) then we get a good relative approximation of the original value and we are done. Otherwise, the information that $\mu_w(\hat{\mathcal{G}})$ is small allows us to reduce the maximum reward by a factor of 2 in the original game (step 10); we invoke Proposition 3.3 for this. Thus the algorithm terminates in polynomial time (in the bit length of $R(\mathcal{G})$). To remove the dependence on D in the running time, we need also to truncate the transition probabilities. In the algorithm, we denote by $\tilde{P}$ the transition probabilities obtained from P by applying Lemma 3.4 with $B = \lceil \log \frac{1}{\varepsilon'} \rceil$, where we select $\varepsilon' = \frac{p_{\min}^{2k}}{2^{2k+3}n^3k^2\theta}$, where $\theta = \theta(\mathcal{G}) := \frac{2(1+\varepsilon)(3+2\varepsilon)n}{\varepsilon p_{\min}^{2k+1}}$, so that $2\delta(\mathcal{G}, \varepsilon') \le \frac{r_+(\mathcal{G})}{\theta(\mathcal{G})} := K(\mathcal{G})$.

Lemma 3.6. *Let $G = (V, E)$ be a graph admitting a pseudo-polynomial algorithm $\mathbb{A}$ that solves any BWR-game on G in time $\tau(n, D, R)$. Then for any $\varepsilon \in (0, 1)$, there is an algorithm that solves, in relative ε-optimal strategies, any BWR-game $(\mathcal{G} = (G, P, r), w)$, form any given starting position w, in time $(\tau(n, \frac{4^{k+2}n^4k^2(1+\varepsilon)(3+2\varepsilon)}{\varepsilon p_{\min}^{2k}}, \frac{2(1+\varepsilon)(3+2\varepsilon)n}{\varepsilon p_{\min}^{2k+1}}) + poly(n))(\lfloor \log R \rfloor + 1)$.*

Proof. The algorithm FPTAS-BWR$(\mathcal{G}, w, \varepsilon)$ is given as Algorithm 1. The bound on the running time is obvious since, by (7), each time we recurse on a game $\tilde{\mathcal{G}}$ with $r^+(\tilde{\mathcal{G}})$ reduced by a factor of at least half. Moreover, the reward in the truncated game $\hat{\mathcal{G}}$ is integral with maximum value $r^+(\hat{\mathcal{G}}) \le \theta$, and common denominator of transition probabilities at most . Thus the time taken by algorithm $\mathbb{A}$ for each recursive call is at most $\tau(n, D, \frac{2(1+\varepsilon)n}{\varepsilon p_{\min}^{2k+1}})$.

So it remains to argue by induction that the algorithm returns a pair of ε-optimal strategies $\tilde{s} = (\tilde{s}_W, \tilde{s}_B)$. For the base case, note that since $\|P - \tilde{P}\|_\infty \le \varepsilon'$ and $r_+(\mathcal{G}) = 1$, Proposition 3.2 implies that the pair $(\tilde{s}_W, \tilde{s}_B)$ returned in step 3 is absolute ε''-optimal, where $\varepsilon'' = 2\delta(\mathcal{G}, \varepsilon') < \frac{\varepsilon p_{\min}^{2k+1}}{n}$. Lemma A.1 and the integrality of the non-negative rewards imply that, for any situation s, $\mu_w(\mathcal{G}(s)) \ge \frac{p_{\min}^{2k+1}}{n}$ if $\mu_w(\mathcal{G}(s)) > 0$. Thus, if $\mu_w(\mathcal{G}) > 0$, then $\varepsilon'' \le \varepsilon \mu_w(\mathcal{G})$, and it follows that

Algorithm 1 FPTAS-BWR($\mathcal{G}, w, \varepsilon$)

Input: a BWR-game $\mathcal{G} = (G = (V, E), P, r)$, a starting vertex $w \in V$, and an accuracy ε.
Output: an ε-optimal pair $(\tilde{s}_W, \tilde{s}_B)$ for the game $(\mathcal{G}, w)$.
1: **if** $r^+(\mathcal{G}) = 1$ **then**
2: $\quad \hat{\mathcal{G}} := (G, \tilde{P}, r)$
3: $\quad$ **return** $\mathbb{A}(\hat{\mathcal{G}}, v)$
4: $K := \frac{r^+(\mathcal{G})}{\theta(\mathcal{G})}$
5: $\hat{r}_e = \lfloor \frac{r_e}{K} \rfloor$ for $e \in E$; $\hat{\mathcal{G}} = (G, \tilde{P}, \hat{r})$
6: $(\tilde{s}_W, \tilde{s}_B) := \mathbb{A}(\hat{\mathcal{G}}, w)$
7: **if** $\mu_w(\hat{\mathcal{G}}) \geq \frac{3}{\varepsilon}$ **then**
8: $\quad$ **return** $(\tilde{s}_W, \tilde{s}_B)$
9: **else**
10: $\quad$ for all $e \in E$, let

$$\tilde{r}_e = \begin{cases} \lceil \frac{r^+}{2} \rceil & \text{if } r_e > \frac{r^+}{2(1+\varepsilon)} \\ r_e & \text{otherwise} \end{cases} \quad (7)$$

11: $\quad \tilde{\mathcal{G}} := (G, P, \tilde{r})$
12: $\quad$ **return** FPTAS-BWR($\tilde{\mathcal{G}}, w, \varepsilon$)

Algorithm 2 FPTAS-BW($\mathcal{G}, \varepsilon$)

Input: a BW-game $\mathcal{G} = (G = (V = V_B \cup V_W, E), r)$, and an accuracy ε.
Output: a uniformly ε-optimal pair $(\tilde{s}_W, \tilde{s}_B)$ for $\mathcal{G}$.
1: **if** $r^+(\mathcal{G}) = 1$ **then**
2: $\quad$ **return** $\mathbb{A}(\mathcal{G})$
3: $K := \frac{\varepsilon' r^+}{2(1+\varepsilon')^2 n}$
4: $\hat{r}_e = \lfloor \frac{r_e}{K} \rfloor$ for $e \in E$; $\hat{\mathcal{G}} = (G, \hat{r})$
5: $(\hat{s}_W, \hat{s}_B) := \mathbb{A}(\hat{\mathcal{G}})$
6: $U := \{u \in V \mid \mu_u(\hat{\mathcal{G}}) \geq \frac{1}{\varepsilon'}\}$
7: **if** $U = V$ **then**
8: $\quad$ **return** $(\tilde{s}_W, \tilde{s}_B) = (\hat{s}_W, \hat{s}_B)$
9: **else**
10: $\quad \tilde{G} := G[V \setminus U]$
11: $\quad$ for all $e \in E(\tilde{G})$, let

$$\tilde{r}_e = \begin{cases} \lceil \frac{r^+}{2} \rceil & \text{if } r_e > \frac{r^+}{2(1+\varepsilon')} \\ r_e & \text{otherwise} \end{cases} \quad (8)$$

12: $\quad \tilde{\mathcal{G}} := (\tilde{G}, \tilde{r})$
13: $\quad (\tilde{s}_W, \tilde{s}_B) :=$FPTAS-BW($\tilde{\mathcal{G}}, \varepsilon$)
14: $\quad \tilde{s}(w) := \hat{s}(w)$ for all $w \in U$
15: $\quad$ **return** $\tilde{s} = (\tilde{s}_W, \tilde{s}_B)$

$(\tilde{s}_W, \tilde{s}_B)$ is relative ε-optimal. On the other hand, if $\mu_w(\mathcal{G}) = 0$, then $\mu_w(\mathcal{G}(\tilde{s})) \leq \mu_w(\mathcal{G}) + \varepsilon'' < \frac{p_{\min}^{2k+1}}{n}$, implying that $\mu_w(\mathcal{G}(\tilde{s})) = 0$. Thus, we get an ε-approximation in both cases.

Now let us consider the general case. Note that $\frac{1}{K}r - \mathbf{1} \leq \hat{r} \leq \frac{1}{K}r$, and $\|P - \tilde{P}\|_\infty \leq \varepsilon'$, and hence by Propositions 3.1 and 3.2, we have

$$K\mu_w(\hat{\mathcal{G}}) - 2\delta(\mathcal{G}, \varepsilon') \leq \mu_w(\mathcal{G}) \leq K\mu_w(\hat{\mathcal{G}}) + K + 2\delta(\mathcal{G}, \varepsilon'), \quad (9)$$

and the pair $(\tilde{s}_W, \tilde{s}_B)$ returned in step 8 is absolute $K + 2\delta(\mathcal{G}, \varepsilon') \leq 2K$-optimal for $\mathcal{G}$. Suppose that $\mathbb{A}$ determines that $\mu_w(\hat{\mathcal{G}}) \geq \frac{3}{\varepsilon}$ in step 7, and hence the algorithm returns $(\tilde{s}_W, \tilde{s}_B)$. Then (9) implies that

$$K \leq \frac{\mu_w(\mathcal{G})}{\mu_w(\hat{\mathcal{G}}) - 1} \leq \frac{\mu_w(\mathcal{G})}{3/\varepsilon - 1} \leq \frac{\varepsilon}{2}\mu(\mathcal{G}),$$

and we are done. On the other hand, if $\mu_w(\hat{\mathcal{G}}) < \frac{3}{\varepsilon}$ then, by (9), $\mu_w(\mathcal{G}) < \frac{K(3+2\varepsilon)}{\varepsilon} = \frac{p_{\min}^{2k+1} r^+}{2(1+\varepsilon)n}$. By Proposition 3.3, applied with $t = \frac{K(3+2\varepsilon)}{\varepsilon}$, we have that the game $\tilde{\mathcal{G}}$ defined in step 11 satisfies $\mu_v(\mathcal{G}) = \mu_v(\tilde{\mathcal{G}})$, and any (relative) ε-optimal strategy in $(\tilde{\mathcal{G}}, w)$ (in particular the one returned in step 12) is also ε-optimal for $(\mathcal{G}, w)$. □

Remark 3.7. *It is easy to see that, for structurally ergodic BWR-games, one can easily modify the above procedure to return uniformly ε-optimal strategies.*

3.5 Uniformly relative ε-approximation for BW-games

Note that FPTAS in Theorem 3.6 does not necessarily return a uniformly ε-optimal situation, even if the given pseudo-polynomial algorithm $\mathbb{A}$ provides a uniformly optimal situation. In case of BW-games, we can modify this FPTAS to return a situation which is ε-optimal for all $v \in V$. The algorithm is given as Algorithm 2. The main difference is that when we recurse on a game with reduced rewards (step 13), we have also to delete all nodes that have large values $\mu(\tilde{\mathcal{G}}, v)$ in the truncated game. This is similar to the approach used to decompose a BW-game into ergodic classes [20]. However, the main technical difficulty is that, with approximate equilibria, White (respectively, Black) might still have some incentive to move from a higher-value (respectively, lower-value) class to a lower-value (respectively, higher-value) class, since the values are just estimated approximately. We show that such a move will not be very profitable for White (respectively, Black). As before, we assume that the rewards are integral with $\min_{e:r_e>0} r_e = 1$.

Lemma 3.8. *Let $\mathbb{A}$ be a pseudo-polynomial algorithm that solves, in uniformly optimal strategies, any BW-game $\mathcal{G}$ in time $\tau(n, R)$. Then for any $\varepsilon > 0$, there is an algorithm that solves, in uniformly relative ε-optimal strategies, any BW-game $\mathcal{G}$, in time $(\tau(n, \frac{2(1+\varepsilon')^2 n}{\varepsilon'}) + poly(n))h$, where $h = \lfloor \log R \rfloor + 1$, and $\varepsilon' = \frac{\ln(1+\varepsilon)}{4h-2}$.*

Proof. The algorithm FPTAS-BW($\mathcal{G}, \varepsilon$) is given as Algorithm 2. The bound on the running time is obvious since, by (8), each time we recurse on a game $\tilde{\mathcal{G}}$ with $r^+(\tilde{\mathcal{G}})$ reduced by a factor of at least half. Moreover, the reward in the truncated game $\hat{\mathcal{G}}$ is integral with maximum value $r^+(\hat{\mathcal{G}}) \le \frac{r^+}{K} \le \frac{2(1+\varepsilon')^2 n}{\varepsilon'}$. Thus the time taken by algorithm $\mathbb{A}$ for each recursive call is at most $\tau(n, \frac{2(1+\varepsilon')n}{\varepsilon'})$.

So it remains to argue (by induction) that the algorithm returns ε-optimal strategies $(\tilde{s}_W, \tilde{s}_B)$. Let us index the different recursive calls of the algorithm by $i = 1, 2, \ldots, h' \le h$ and denote by $\mathcal{G}^{(i)} = (G^{(i)} = (V, E^{(i)}), r^{(i)})$ the game input to the ith recursive call of the algorithm (so $\mathcal{G}^{(1)} = \mathcal{G}$) and by $\hat{s}^{(i)} = (\hat{s}_W^{(i)}, \hat{s}_B^{(i)})$, $\tilde{s}^{(i)} = (\tilde{s}_W^{(i)}, \tilde{s}_B^{(i)})$ the pair of strategies returned either in steps 2, 5, 8, or 15. Similarly, we denote respectively by $V^{(i)} = B_W^{(i)} \cup V_B^{(i)}$, $U^{(i)}$, $r^{(i)}$, $K^{(i)}$ $\hat{r}^{(i)}$, $\hat{\mathcal{G}}^{(i)}$, $\tilde{\mathcal{G}}^{(i)}$ the instantiations of V, U, r, $\hat{r}$, K, $\hat{\mathcal{G}}$, $\tilde{\mathcal{G}}$ in the ith call of the algorithm. We denote by $S_W^{(i)}$ and $S_B^{(i)}$ the set of strategies in $\mathcal{G}^{(i)}$ for White and Black, respectively. For a set U, game $\mathcal{G}$ and situation s, denote respectively by $\mathcal{G}[U] = (G[U], r)$ and $s[U]$ the game and situation induced on U.

Observation 3.9. *(i)* $\nexists (v, u) \in E : v \in V_B^{(i)} \cap U^{(i)},\ u \in V^{(i)} \setminus U^{(i)}$;

(ii) $\forall v \in V_W^{(i)} \cap U^{(i)}\ \ \exists u \in U^{(i)} :\ (v, u) \in E$;

(i′) $\nexists (v, u) \in E : v \in V_W^{(i)} \setminus U^{(i)},\ u \in U^{(i)}$;

(ii′) $\forall v \in V_B^{(i)} \setminus U^{(i)}\ \ \exists u \in V^{(i)} \setminus U^{(i)} :\ (v, u) \in E$;

(iii) *Let $\hat{s}^{(i)} = (\hat{s}_W^{(i)}, \hat{s}_B^{(i)})$ be the situation returned in step 5. Then $\forall v \in U^{(i)} :\ \hat{s}^{(i)}(v) \in U^{(i)}$ and $\forall v \in V^{(i)} \setminus U^{(i)} :\ \hat{s}^{(i)}(v) \in V^{(i)} \setminus U^{(i)}$.*

Proof. By the optimality conditions in $\hat{\mathcal{G}}^{(i)}$ (see e.g. [20]), we have

(I) $\mu_v(\hat{\mathcal{G}}^{(i)}) = \min\{\mu_u(\hat{\mathcal{G}}^{(i)}) \mid u \in V^{(i)}$ such that $(v, u) \in E\}$, for $v \in V_B^{(i)}$, and

(II) $\mu_v(\hat{\mathcal{G}}^{(i)}) = \max\{\mu_u(\hat{\mathcal{G}}^{(i)}) \mid u \in V^{(i)}$ such that $(v, u) \in E\}$, for any $v \in V_W^{(i)}$.

(I) and (II), together with the definition of $U^{(i)}$, imply (i) and (ii), respectively,. Similarly (i') and (ii') can be shown. The optimality conditions also imply that $\mu_v(\hat{\mathcal{G}}^{(i)}) = \mu_{\hat{s}^{(i)}(v)}(\hat{\mathcal{G}}^{(i)})$, which in turn implies (iii). □

Note that Observation 3.9 implies that the game $\mathcal{G}^{(i)}[V^{(i)} \setminus U^{(i)}]$ is well-defined, since the graph $G[V^{(i)} \setminus U^{(i)}]$ has no sinks.

For a strategy s_W (similarly for a strategy S_B) and a subset $V' \subseteq V$, we write $S_W(V') = \{s_W(u) : u \in V'\}$.

Observation 3.10. *Let $\hat{s}^{(i)}$ be the situation returned in step 5. Then for all $i = 1, \ldots, h'$ and any $w \in U^{(i)}$, we have*

$$\max_{s_W:\ s_W(U^{(i)} \cap V_W) \subseteq U^{(i)}} \mu_w(\mathcal{G}^{(i)}(s_W, \hat{s}_B^{(i)})) \leq (1 + \varepsilon')\mu_w(\mathcal{G}^{(i)}),$$

$$\min_{s_B:\ s_B(U^{(i)} \cap V_B) \subseteq U^{(i)}} \mu_w(\mathcal{G}^{(i)}(\hat{s}_W^{(i)}, s_B)) \geq (1 - \varepsilon')\mu_w(\mathcal{G}^{(i)}).$$

Proof. This follows from Proposition 3.1 by the uniformly optimality of $\hat{s}^{(i)}$ in $\hat{\mathcal{G}}^{(i)}$ and the fact that $\mu_w(\hat{\mathcal{G}}^{(i)}) \geq \frac{1}{\varepsilon'}$, for every $w \in U^{(i)}$. □

Observation 3.11. $\forall u \in U^{(i)},\ v \in V^{(i)} \setminus U^{(i)} :\ (1 + \varepsilon')\mu_u(\mathcal{G}^{(i)}) > \mu_v(\mathcal{G}^{(i)})$.

Proof. For $u \in U^{(i)},\ v^{(i)} \in V^{(i)} \setminus U^{(i)}$, we have $\mu_u(\hat{\mathcal{G}}^{(i)}) \geq \frac{1}{\varepsilon'}$ and $\mu_v(\hat{\mathcal{G}}^{(i)}) < \frac{1}{\varepsilon'}$. Thus,

$$\mu_v(\mathcal{G}^{(i)}) \leq K^{(i)}\mu_v(\hat{\mathcal{G}}^{(i)}) + K^{(i)} < \frac{K^{(i)}}{\varepsilon'}(1 + \varepsilon') \leq K^{(i)}\mu_u(\hat{\mathcal{G}}^{(i)})(1 + \varepsilon') \leq \mu_u(\mathcal{G}^{(i)})(1 + \varepsilon').$$

□

Observe that the strategy $\tilde{s}^{(i)}$, returned by the ith call to the algorithm, is determined as follows (c.f. steps 13 and 14): for $w \in U^{(i)}$, $\tilde{s}^{(i)}(w) = \hat{s}^{(i)}(w)$ is chosen by the solution of the game $\hat{\mathcal{G}}^{(i)}$, and for $w \notin U^{(i)}$, $\tilde{s}^{(i)}(w)$ is determined by the (recursive) solution on the residual game $\tilde{\mathcal{G}}^{(i)} = \mathcal{G}^{(i+1)}$. The following claim states that the value of any vertex $u \in V^{(i)} \setminus U^{(i)}$ in the residual game is a good approximation of the value in the original game $\mathcal{G}^{(i)}$.

Claim 3.12. *For all $i = 1, \ldots, h'$ and any $u \in V^{(i)} \setminus U^{(i)}$, we have*

$$\mu_u(\mathcal{G}^{(i)}) \leq \mu_u(\mathcal{G}^{(i)}[V^{(i)} \setminus U^{(i)}]) \leq (1 + 2\varepsilon')\mu_u(\mathcal{G}^{(i)}). \tag{10}$$

Proof. Fix $u \in V^{(i)} \setminus U^{(i)}$. Let $s^* = (s_W^*, s_B^*)$ and $(\bar{s}_W, \bar{s}_B)$ be optimal situations in $(\mathcal{G}^{(i)}, u)$ and $(\bar{\mathcal{G}}^{(i)}, u) := (\mathcal{G}^{(i)}[V^{(i)} \setminus U^{(i)}], u)$, respectively. Let us extend $\bar{s}$ to a situation in $\mathcal{G}^{(i)}$ by setting $\bar{s}(v) = \hat{s}^{(i)}(v)$ for all $v \in U^{(i)}$, where $\hat{s}$ is the situation returned in step 5. Then, by Observation 3.9-(i'), White has no way to escape to $U^{(i)}$, or in other words, $s_W^*(u') \in V^{(i)} \setminus U^{(i)}$ for all $u' \in V_W^{(i)} \setminus U^{(i)}$. Hence

$$\begin{aligned} \mu_u(\mathcal{G}^{(i)}) &= \mu_u(\mathcal{G}^{(i)}(s_W^*, s_B^*)) \leq \mu_u(\mathcal{G}^{(i)}(s_W^*, \bar{s}_B)) \\ &= \mu_u(\bar{\mathcal{G}}^{(i)}(s_W^*, \bar{s}_B)) \leq \mu_u(\bar{\mathcal{G}}^{(i)}(\bar{s}_W, \bar{s}_B)) = \mu_u(\bar{\mathcal{G}}^{(i)}). \end{aligned}$$

For similar reasons, $\mu_u(\mathcal{G}^{(i)}) \geq \mu_u(\bar{\mathcal{G}}^{(i)})$, if $s_B^*(v) \in V^{(i)} \setminus U^{(i)}$ for all $v \in V_B^{(i)} \setminus U^{(i)}$ such that v is reachable from u in the graph $G(s_W^*, s_B^*)$. Suppose, on the other hand, that there is a $v \in V_B^{(i)} \setminus U^{(i)}$ such that $u' = s_B^*(v) \in U^{(i)}$, and v is reachable from u in the graph $G(s_W^*, s_B^*)$. Then $\mu_u(\mathcal{G}^{(i)}) = \mu_{u'}(\mathcal{G}^{(i)}) \geq K^{(i)}\mu_{u'}(\hat{\mathcal{G}}^{(i)}) \geq \frac{K^{(i)}}{\varepsilon'}$. Moreover, the optimality of $(\hat{s}_W, \hat{s}_B)$ in $\hat{\mathcal{G}}^{(i)}$, and the fact that $\frac{1}{K^{(i)}}r^{(i)} - \mathbf{1} \leq \hat{r}^{(i)} \leq \frac{1}{K^{(i)}}r^{(i)}$, imply that

$$\begin{aligned} \forall s_W \in S_W^{(i)} :\ \mu_u(\mathcal{G}^{(i)}(\hat{s}_W, \hat{s}_B)) &\geq K^{(i)}\mu_u(\hat{\mathcal{G}}^{(i)}(\hat{s}_W, \hat{s}_B)) \geq K^{(i)}\mu_u(\hat{\mathcal{G}}^{(i)}(s_W, \hat{s}_B)) \\ &\geq \mu_u(\mathcal{G}^{(i)}(s_W, \hat{s}_B)) - K^{(i)} \geq \mu_u(\mathcal{G}^{(i)}(s_W, \hat{s}_B)) - \varepsilon'\mu_u(\mathcal{G}^{(i)}) \\ \forall s_B \in S_B^{(i)} :\ \mu_u(\mathcal{G}^{(i)}(\hat{s}_W, \hat{s}_B)) &\leq K^{(i)}\mu_u(\hat{\mathcal{G}}^{(i)}(\hat{s}_W, \hat{s}_B)) + K^{(i)} \leq K^{(i)}\mu_u(\hat{\mathcal{G}}^{(i)}(\hat{s}_W, s_B)) + K^{(i)} \\ &\leq \mu_u(\mathcal{G}^{(i)}(\hat{s}_W, s_B)) + K^{(i)} \leq \mu_u(\mathcal{G}^{(i)}(\hat{s}_W, s_B)) + \varepsilon'\mu_u(\mathcal{G}^{(i)}). \end{aligned}$$

In particular,

$$\begin{aligned}\mu_u(\mathcal{G}^{(i)}) &= \mu_u(\mathcal{G}^{(i)}(s^*_W, s^*_B)) \geq \mu_u(\mathcal{G}^{(i)}(\hat{s}_W, s^*_B)) \geq \mu_u(\mathcal{G}^{(i)}(\hat{s}_W, \hat{s}_B)) - \varepsilon' \mu_u(\mathcal{G}^{(i)}) \\ &\geq \mu_u(\mathcal{G}^{(i)}(\bar{s}_W, \hat{s}_B)) - 2\varepsilon' \mu_u(\mathcal{G}^{(i)}) = \mu_u(\bar{\mathcal{G}}^{(i)}(\bar{s}_W, \hat{s}_B)) - 2\varepsilon' \mu_u(\mathcal{G}^{(i)}) \\ &\geq \mu_u(\bar{\mathcal{G}}^{(i)}(\bar{s}_W, \bar{s}_B)) - 2\varepsilon' \mu_u(\mathcal{G}^{(i)}) = \mu_u(\bar{\mathcal{G}}^{(i)}) - 2\varepsilon' \mu_u(\mathcal{G}^{(i)}),\end{aligned}$$

where $\mu_u(\mathcal{G}^{(i)}(\bar{s}_W, \hat{s}_B)) = \mu_u(\bar{\mathcal{G}}^{(i)}(\bar{s}_W, \hat{s}_B))$ follows by Observation 3.9 (since $(\bar{s}_W, \hat{s}_B)(v) \in V^{(i)} \setminus U^{(i)}$). It follows that $\mu_u(\mathcal{G}^{(i)}) \geq \frac{1}{1+2\varepsilon'} \mu_u(\bar{\mathcal{G}}^{(i)})$. □

Let us fix $\varepsilon_{h'} = \varepsilon'$, and for $i = 1, 2, \ldots, h'-1$, let us choose ε_i such that $1+\varepsilon_i \geq (1+\varepsilon')(1+2\varepsilon')(1+\varepsilon_{i+1})$. We next claim that the strategies $(\tilde{s}^{(i)}_W, \tilde{s}^{(i)}_B)$ returned by the ith call are relative ε_i-optimal in $\mathcal{G}^{(i)}$.

Claim 3.13. *For all $i = 1, \ldots, h'$ and any $w \in V^{(i)}$, we have*

$$\max_{s_W \in S^{(i)}_W} \mu_w(\mathcal{G}^{(i)}(s_W, \tilde{s}^{(i)}_B)) \leq (1+\varepsilon_i)\mu_w(\mathcal{G}^{(i)}) \tag{11}$$

$$\min_{s_B \in S^{(i)}_B} \mu_w(\mathcal{G}^{(i)}(\tilde{s}^{(i)}_W, s_B)) \geq (1-\varepsilon_i)\mu_w(\mathcal{G}^{(i)}). \tag{12}$$

Proof. By induction on $i = h', h'-1, \ldots, 1$. For $i = h'$, the statement follows directly from Proposition 3.1 since $\mu_w(\hat{\mathcal{G}}^{(i)}) \geq \frac{1}{\varepsilon'}$ for all $w \in V^{(i)}$. So suppose that $i < h'$.

(I) Consider an arbitrary strategy $s_W \in S^{(i)}_W$ for White. Suppose first that $w \in U^{(i)}$. Note that, by Observation 3.9-(*iii*), $\tilde{s}^{(i)}_B(v) \in U^{(i)}$ for all $v \in V_B \cap U^{(i)}$. If also $s_W(v) \in U^{(i)}$ for all $v \in V_W \cap U^{(i)}$, such that v is reachable from w in the graph $G(s_W, \tilde{s}^{(i)}_B)$, then $\mu_w(\mathcal{G}^{(i)}(s_W, \tilde{s}^{(i)}_B)) \leq (1+\varepsilon')\mu_w(\mathcal{G}^{(i)}) \leq (1+\varepsilon_i)\mu_w(\mathcal{G}^{(i)})$ follows from Observation 3.10.

Suppose that $u = s_W(v) \notin U^{(i)}$ for some $v \in V_W \cap U^{(i)}$ such that v is reachable from w in the graph $G(s_W, \tilde{s}^{(i)}_B)$.

By induction, $\bar{s}^{(i)} = (\bar{s}^{(i)}_W, \bar{s}^{(i)}_B) := (\tilde{s}^{(i)}_W, \tilde{s}^{(i)}_B)[V^{(i)} \setminus U^{(i)}]$ is ε_{i+1}-optimal in $\mathcal{G}^{(i+1)} = \tilde{\mathcal{G}}^{(i)}$. Recall that the game $\tilde{\mathcal{G}}^{(i)}$ is obtained from $\bar{\mathcal{G}}^{(i)} := \mathcal{G}^{(i)}[V^{(i)} \setminus U^{(i)}]$ by reducing the rewards according to (8). Thus, we get by Proposition 3.3 that $\mu_u(\bar{\mathcal{G}}^{(i)}) = \mu_u(\tilde{\mathcal{G}}^{(i)})$, and

$$\max_{s'_W \in S^{(i+1)}_W} \mu_u(\mathcal{G}^{(i)}(s'_W, \bar{s}^{(i)}_B)) \leq (1+\varepsilon_{i+1})\mu_u(\bar{\mathcal{G}}^{(i)}) \tag{13}$$

$$\min_{s'_B \in S^{(i+1)}_B} \mu_u(\mathcal{G}^{(i)}(\bar{s}^{(i)}_W, s'_B)) \geq (1-\varepsilon_{i+1})\mu_u(\bar{\mathcal{G}}^{(i)}). \tag{14}$$

Note that $\tilde{s}^{(i)}_B(u') \in V^{(i)} \setminus U^{(i)}$ for all $u' \in V^{(i)}_B \setminus U^{(i)}$, and by Observation 3.9-(i'), $S^{(i+1)}_W$ is the restriction of $S^{(i)}_W$ to $V^{(i)} \setminus U^{(i)}$. Thus, we get the following series of inequalities

$$\begin{aligned}\mu_w(\mathcal{G}^{(i)}(s_W, \tilde{s}^{(i)}_B)) &= \mu_u(\mathcal{G}^{(i)}(s_W, \tilde{s}^{(i)}_B)) \\ &\leq (1+\varepsilon_{i+1})\mu_u(\mathcal{G}^{(i)}[V^{(i)} \setminus U^{(i)}]) && \text{(by (13))} \\ &\leq (1+\varepsilon_{i+1})(1+2\varepsilon')\mu_u(\mathcal{G}^{(i)}) && \text{(by (10))} \\ &< (1+\varepsilon_{i+1})(1+2\varepsilon')(1+\varepsilon')\mu_w(\mathcal{G}^{(i)}) && \text{(by Observation 3.11)} \\ &\leq (1+\varepsilon_i)\mu_w(\mathcal{G}^{(i)}) && \text{(since } (1+\varepsilon_{i+1})(1+2\varepsilon')(1+\varepsilon') \leq (1+\varepsilon_i)\text{)}.\end{aligned}$$

If $w \in V^{(i)} \setminus U^{(i)}$, then the above argument also shows that $\mu_w(\mathcal{G}^{(i)}(s_W, \tilde{s}^{(i)}_B)) \leq (1+\varepsilon_{i+1})(1+2\varepsilon')\mu_w(\mathcal{G}^{(i)}) \leq (1+\varepsilon_i)\mu_w(\mathcal{G}^{(i)})$. Thus (11) follows.

(II) Consider an arbitrary strategy $s_B \in S^{(i)}_B$ for Black. If $w \in U^{(i)}$, then $\mu_w(\mathcal{G}^{(i)}(\tilde{s}^{(i)}_W, s_B)) \geq (1-\varepsilon')\mu_w(\mathcal{G}^{(i)}) \geq (1-\varepsilon_i)\mu_w(\mathcal{G}^{(i)}w)$ follows from Observations 3.9-(i), 3.9-(iii), and 3.10.

Suppose now that $w \in V^{(i)} \setminus U^{(i)}$. If $s_B \in S_B^{(i+1)}$, then we get by (14) and (10) that $\mu_w(\mathcal{G}^{(i)}(\tilde{s}_W^{(i)}, s_B)) \geq (1-\varepsilon_{i+1})\mu_w(\mathcal{G}^{(i)}) \geq (1-\varepsilon_i)\mu_w(\mathcal{G}^{(i)})$. A similar situation holds if $s_B(v) \in V^{(i)} \setminus U^{(i)}$ for all $v \in V_B^{(i)} \setminus U^{(i)}$, such that v is reachable from w in the graph $G(\tilde{s}_W^{(i)}, s_B)$. So it remains to consider the case when there is a $v \in V_B^{(i)} \setminus U^{(i)}$ such that $u = s_B(v) \in U^{(i)}$, and v is reachable from w in the graph $G(\tilde{s}_W^{(i)}, s_B)$. Since, in this case, Black has no escape from $U^{(i)}$ (by Observation 3.9-(i)), we get from Observations 3.10 and 3.11 that

$$\mu_w(\mathcal{G}^{(i)}(\tilde{s}_W^{(i)}, s_B)) = \mu_u(\mathcal{G}^{(i)}(\tilde{s}_W^{(i)}, s_B)) \geq (1-\varepsilon')\mu_u(\mathcal{G}^{(i)}) > (1-\varepsilon')^2\mu_w(\mathcal{G}^{(i)}) \geq (1-\varepsilon_i)\mu_w(\mathcal{G}^{(i)}).$$

□

Finally, we set the ε's such that $\varepsilon_1 = [(1+2\varepsilon')(1+\varepsilon')]^{h'-1}(1+\varepsilon') - 1 \leq \varepsilon$. □

4 Smoothed Analysis

4.1 The notion of smoothed complexity

For many applications, the classic worst-case analysis is far too pessimistic. Worst-case analysis is often dominated by artificially constructed worst-case instances, that do not reflect typical instances. On the other hand, one drawback of average-case analysis is that random instances usually have very special properties with high probability, and these properties are often not shared with typical instances.

To overcome the drawbacks of worst-case and average-case complexity, Spielman and Teng introduced smoothed analysis [37]. Smoothed analysis is a hybrid of worst-case and average-case analysis: An adversary specifies an instance, which is then subject to a small amount of random noise. The smoothed complexity is the maximum expected running-time that the adversary can achieve by his choice of an instance. Smoothed analysis rules out the drawbacks of both average-case and worst-case analysis and often allows more realistic conclusions about the performance of an algorithm or the complexity of a problem. It takes into account that realistic data is often subject to a small amount of random noise, be it from measurement or rounding errors or from some arbitrary unknown circumstances. Smoothed analysis has originally been invented to explain the practical performance of the simplex method [37]. Since then, smoothed analysis has been applied successfully to a variety of algorithms and problems [2, 39, 34]. We refer to Spielman and Teng for a survey [38].

A generalization of smoothed analysis has been introduced by Beier and Vöcking [3]: They went from the two-step model (an adversary chooses the instance, which is then perturbed) to a one-step model. In the one-step model, the adversary specifies the density functions according to which the numbers of the instance are drawn. In this case, the perturbation parameter ϕ restricts the adversary to density functions that are bounded by ϕ: The larger ϕ, the more powerful the adversary. For instance, in case of Gaussian perturbation, ϕ is proportional to the inverse of the standard deviation.

To characterize integer programs that can typically be solved in polynomial time, Beier and Vöcking introduced the notion of polynomial smoothed complexity [3]. Their notion is inspired by the notion of average polynomial time [6]. A problem is said to have smoothed polynomial complexity if and only if there exists an algorithm $\mathcal{A}$ with running-time T and a constant α such that

$$\forall \phi \geq 1, \forall n \in \mathbb{N} : \max_{\vec{f} \in \mathcal{D}_N(\phi)} \mathbb{E}_{X \sim \vec{f}}(T(X)^\alpha) = O(N\phi) . \tag{15}$$

Here, $\mathcal{D}_N(\phi)$ denotes all possible vectors of density functions bounded by ϕ for instances of size N, and X is an instance drawn according to $\vec{f}$. Another way to phrase smoothed polynomial running-time is that there exists a polynomial $P(N, \phi, 1/\varepsilon)$ such that the probability that $\mathcal{A}$

exceeds a running-time of $P(N, \phi, 1/\varepsilon)$ is at most ε. Note that smoothed polynomial running-time does not always give polynomial expected running-time. To get the latter, the α must be placed outside the expectation. The reason for defining smoothed polynomial running-time in this way is that it makes the notion robust against, e.g., simulation on a slower machine.

4.2 Smoothed complexity of BWR-games

Let $G = (V, E)$ be a graph that admits a pseudo-polynomial algorithm $\mathbb{A}$ and $\mathcal{G} = (G, P, r)$ be a BWR-game on G. In this section, we show that any such game can be solved in smoothed polynomial time. For this, we assume that an adversary specifies a game together with density functions for the rewards (one for each arc), and these density functions are bounded by ϕ, and show a bound as in (15). One (technical) issue is that the perturbed rewards are of course real, non-rational numbers with probability 1. Thus, we cannot really use existing algorithms as sub-routine, and we cannot even compute anything with these numbers on an ordinary RAM. To cope with this problem, we use Beier and Vöcking's [3] approach and assume that the rewards are in $[-1, 1]$ and that we can access the bits of the rewards one-by-one.

To state our results in a bit more general setting, we will assume that $\mathbb{A}$ solves any BWR-game on G in uniformly optimal strategies. If this was not the case, then it is easy to modify the procedure and analysis in this section to solve the game starting from a given vertex.

Before describing the procedure (Algorithm 3), we need to introduce some notation. Let us write $\lfloor x \rfloor_b$ for the largest integer smaller than or equal to x that has b bits (i.e., we basically cut off all bits after the b-th bit). Let $\gamma = \gamma(\mathcal{G}) := (kn)^{-2}(2D)^{-2(k+2)}$ and $\varepsilon > 0$. Given the game $\mathcal{G} = (G = (V, E), P, r)$, define, for each $e \in E$, the game $\mathcal{G}_{e,\varepsilon} = (G, P, r(e))$, where

$$r_{e'}(e) = \begin{cases} r_e + 2\gamma^{-1}\varepsilon & \text{if } e' = e, \\ r_{e'} & \text{otherwise.} \end{cases} \tag{16}$$

The basic idea behind our smoothed analysis is as follows: We use a certain number of bits for each reward. Then we run the pseudo-polynomial algorithm to solve the resulting game with the rewards rounded down (and scaled to integers) because we do not have more bits at that point (Step 5). This can be done in polynomial-time as long as we have at most roughly $O(\log n)$ bits. Then we try to certify that the solution obtained is also a solution for the true rewards (Step 7). If this succeeds, then we are done. If this fails, then we use one more bit and repeat the process.

To prove a smoothed polynomial running time, we need to show that with high probability a logarithmic number of bits suffices to compute an equilibrium for the original (untruncated) game. Furthermore, we have to devise a certificate proving that the computed equilibrium is indeed an equilibrium for the original game (we will show that such a certificate is given in Step 7). Both results are based on a sensitivity analysis of the game: we show that by changing the rewards slightly, an optimal strategy remains optimal for the changed game.

A key ingredient for our smoothed analysis is an adaption of the isolation lemma [32] to our setting. An adaption of the isolation lemma has already been used successfully in smoothed analysis of integer programs [3, 34]. It basically says the following: Of course there are exponentially many alternative strategies for each player. But if we consider the optimal pair of strategies, then choosing an alternative strategy makes the payoff for the respective player significantly worse (with high probability, at least).

As usual, we assume that the maximum (respectively, the minimum) over an empty set is $-\infty$ (respectively, $+\infty$).

Lemma 4.1 (Isolation Lemma)**.** *Let E be a finite set, and $\mathcal{F} \subset \mathbb{R}_+^E$ be a family of (distinct) vectors, such that for any distinct $\rho, \rho' \in \mathcal{F}$, there exists an $e \in E$ with $|\rho_e - \rho'_e| \geq \gamma$. Let $\{w_e\}_{e \in E}$*

Algorithm 3 Solve($\mathcal{G}$)

Input: a BWR-game $\mathcal{G} = (G = (V, E), P, r)$.
Output: an optimal pair $(\tilde{s}_W, \tilde{s}_B)$ for the game $\mathcal{G}$.
1: $\ell_0 \leftarrow \log((nD)^{c_0}\phi)$ $\{c_0$ is a constant to be specified later$\}$
2: $i \leftarrow 0$
3: **repeat**
4: $\quad \ell := \ell_0 + i$; $\varepsilon \leftarrow 2^{-\ell}$; $i := i + 1$
5: $\quad \tilde{r} := \lfloor r \rfloor_\ell$; $\tilde{\mathcal{G}} := (G, P, \tilde{r})$; $\tilde{\mathcal{G}}' := (G, P, 2^\ell \tilde{r})$
6: $\quad (\tilde{s}_W, \tilde{s}_B) := \mathbb{A}(\tilde{\mathcal{G}}')$
7: **until** $\tilde{s}$ is optimal in $\tilde{\mathcal{G}}_{e,\varepsilon}$ for all $e \in E$

be independent continuous random variables with maximum density ϕ. Define $\text{gap}(w) := w^T\rho^* - w^T\rho^{**}$*, where* $\rho^* = \text{argmax}_{\rho\in\mathcal{F}}\, w^T\rho$ *and* $\rho^{**} = \text{argmax}_{\rho\in\mathcal{F},\ \rho\neq\rho^*}\, w^T\rho$*. Then* $\Pr(\text{gap}(w) \leq \varepsilon) \leq |E|\varepsilon\phi\frac{\kappa^2}{\gamma}$*, where* $\kappa = \max_{e\in E}|\mathcal{F}_e|$*, and* $\mathcal{F}_e = \{x \mid \rho_e = x \text{ for some } \rho \in \mathcal{F}\}$*.*

Proof. For $e \in E$ and $x, y \in \mathbb{R}_+$, define

$$\Delta_{e,x,y} = \max_{\rho\in\mathcal{F},\ \rho_e=x}(w^T\rho - w_e x) - \max_{\rho\in\mathcal{F},\ \rho_e=y}(w^T\rho - w_e y).$$

It is crucial to note that $\Delta_{e,x,y}$ is *independent* of w_e. With probability 1, there exist unique ρ^* and ρ^{**}, as defined in the lemma above. Since $\rho^* \neq \rho^{**}$, there exists an $e \in E$ such that $|\rho^*_e - \rho^{**}_e| \geq \gamma$. Suppose that $\rho^*_e = x$ and $\rho^{**}_e = y$. Then $w^T\rho^* - w_e x = \max_{\rho\in\mathcal{F},\ \rho_e=x}(w^T\rho - w_e x)$ and $w^T\rho^{**} - w_e y = \max_{\rho\in\mathcal{F},\ \rho_e=y}(w^T\rho - w_e y)$ imply that $w^T\rho^* - w^T\rho^{**} = \Delta_{e,x,y} + w_e(x - y)$. Thus

$$\begin{aligned}
\Pr(\text{gap}(w) < \varepsilon) &\leq \Pr(\exists e \in E,\ x, y:\ 0 \leq w^T\rho^* - w^T\rho^{**} \leq \varepsilon,\ \rho^*_e = x,\ \rho^{**}_e = y,\ |x - y| \geq \gamma) \\
&= \sum_{e\in E}\sum_{x,y\in\mathcal{F}_e:\ |x-y|\geq\gamma} \Pr(\ 0 \leq w^T\rho^* - w^T\rho^{**} \leq \varepsilon,\ \rho^{**}_e = x,\ \rho^*_e = y) \\
&\leq \sum_{e\in E}\sum_{x,y\in\mathcal{F}_e:\ |x-y|\geq\gamma} \Pr(\ 0 \leq \Delta_{e,x,y} + w_e(x - y) \leq \varepsilon) \\
&= \sum_{e\in E}\sum_{x,y\in\mathcal{F}_e:\ x-y\geq\gamma} \Pr\left(\frac{-\Delta_{e,x,y}}{x - y} \leq w_e \leq \frac{\varepsilon - \Delta_{e,x,y}}{x - y}\right) + \\
&\quad \sum_{e\in E}\sum_{x,y\in\mathcal{F}_e:\ y-x\geq\gamma} \Pr\left(\frac{-\varepsilon + \Delta_{e,x,y}}{y - x} \leq w_e \leq \frac{\Delta_{e,x,y}}{y - x}\right) \\
&\leq \sum_{e\in E}\sum_{x,y\in\mathcal{F}_e:\ |x-y|\geq\gamma} \frac{\phi\varepsilon}{|x - y|} \\
&\leq |E|\kappa^2\frac{\phi\varepsilon}{\gamma}.
\end{aligned}$$

□

We will use the above lemma with the set $\mathcal{F}$ representing a set of arc-limiting distributions, corresponding to a set of situations in the game starting from a certain vertex. For that we need bounds for κ and γ, given by the following lemma.

Lemma 4.2. *Let $\mathcal{G} = (G = (V, E), P, r)$ be a BWR-game, $u \in V$ be any vertex, and s be an arbitrary any situation. Then*

(i) every entry of the arc-limiting distribution $\rho(s)$ for the Markov chain $(\mathcal{G}(s), u)$ can be written as rational numbers of the form $\frac{a}{b}$, where $a, b \in \mathbb{Z}_+$ and $a, b \leq kn(2D)^{k+2}$. Hence,

(ii) *the number of possible entries in* $\rho(s)$ *is bounded by* $\kappa = (kn)^2(2D)^{2(k+2)}$*, and*

(iii) *for any situation* s' *such that* $\rho(s') \neq \rho(s)$*, there is an arc* e *such that* $\rho_e(s) - \rho_e(s') \geq \gamma = \gamma(\mathcal{G})$*.*

Proof. The first claim was proved in [8], and the second claim is immediate from it. It is also immediate that if $\rho(s) \neq \rho(s')$, then there is an arc e such that $|\rho_e(s) - \rho_e(s')| \geq \gamma$. To see the stronger claim in *(iii)*, we assume without loss of generality that there is no arc e such that $\rho_e(s) > 0$ and $\rho_e(s') = 0$. Then for all e, $\rho_e(s) > 0$ if and only if $\rho_e(s') > 0$ and hence, $P(s)$ and $P(s')$ have the same absorbing classes. (Indeed, if the absorbing classes of $P(s)$ and $P(s')$ are respectively $C_1, \ldots, C_\ell$ and $C'_1, \ldots, C'_{\ell'}$, then for all $i \in [\ell]$, there is a $j \in [\ell']$ such that $C_i \subseteq C'_j$. Suppose that for some i and j, $C_i \subset C'_j$. Then there should be vertices $v \in V_W \cup V_B$ and $u \in C'_j \setminus C_i$, such that $\rho_{(v,u)}(s) = 0$ while $\rho_{(v,u)}(s') > 0$. But then there should also exist a $u' \in C_i$ (corresponding to the strategy $s(v)$) such that $\rho_{(v,u')}(s) > 0$ and $\rho_{(v,u')}(s') = 0$, which contradicts our assumption. Hence, $\ell = \ell'$ and $C_i = C'_i$ for all i.) Let $\pi_{C_i}(s)$ and $\pi_{C_i}(s')$ be the absorption probabilities into class C_i in $P(s)$ and $P(s')$, respectively. Since $\rho(s) \neq \rho(s')$ and $\sum_i \pi_{C_i}(s) = \sum_i \pi_{C_i}(s') = 1$, there must exist an $i \in [\ell]$ such $\pi_{C_i}(s) > \pi_{C_i}(s')$. Hence any arc $e = (v, u)$ with $u, v \in C_i$ satisfies $\rho_e(s) > \rho_e(s')$ and hence the claim. □

To use the given pseudo-polynomial algorithm, we have to truncate the (perturbed) rewards after a certain number of bits. The following lemma assures that this is possible (with high probability) without changing the optimal strategies, as long as the rounded rewards and the true rewards are close enough. Before we state the lemma, it is useful to observe that, if the rewards are continuous, independently distributed random variables, then, for any two distinct situations s and s', we have $\Pr(\mu_u(\mathcal{G}(s)) = \mu_u(\mathcal{G}(s'))) = 0$ if and only if $\rho(s) \neq \rho(s')$. Thus for the structurally ergodic case, with probability one, two distinct situations result in two distinct values. On the other hand, in the general case, there might be many optimal situations, but all of them must lead to the same limiting distribution.

Given a strategy $s_W \in S_W$ of White we call a *uniform best response* (UBR) of Black any strategy $s^*_B \in S_B$, such that $\mu_u(\mathcal{G}(s_W, s^*_B)) \leq \mu_u(\mathcal{G}(s_W, s_B))$ for all $s_B \in S_B$. Similarly, a UBR of White is defined. (Note that the existence of such a UBR is an immediate corollary of the existence of uniformly optimal situations in BWR-games.) We denote by $\mathrm{UBR}_{\mathcal{G}}(s_W)$ and $\mathrm{UBR}_{\mathcal{G}}(s_B)$ the sets of uniform best responses in $\mathcal{G}$, corresponding to strategies s_W and s_B, respectively.

Lemma 4.3. *Let* $\mathcal{G} = (G = (V, E), P, r)$*,* $\mathcal{G}' = (G = (V, E), P, r')$ *be two BWR-games such that* $r = (r_e)_{e \in E}$ *is a vector of independent continuous random variables with maximum density* ϕ*, and* $\|r' - r\|_\infty \leq \varepsilon$*, for some given* $\varepsilon > 0$*. Let* $\theta := \frac{2n^3\varepsilon\phi}{\gamma(\mathcal{G})^3}$*. Then, the following holds for any situation* s*:*

(i) $\Pr($*s is not uniformly optimal in* $\mathcal{G}'$ $\mid$ *s is uniformly optimal in* $\mathcal{G}) \leq 2\theta$*;*

(ii) $\Pr($*s is not uniformly optimal in* $\mathcal{G}$ $\mid$ *s is uniformly optimal in* $\mathcal{G}') \leq 2\theta$*.*

Proof. It will be enough to prove the following claim for any $s_B \in S_B$ (and the analogous claim for any $s_W \in S_W$).

Claim 4.4. *Let* $s_B \in S_B$ *be an arbitrary strategy of* Black*, and* $s^*_W \in \mathrm{UBR}_{\mathcal{G}}(s_B)$*. Then*

$$\Pr(\exists u \in V,\ s_W \in S_W : \mu_u(\mathcal{G}'(s_W, s_B)) \geq \mu_u(\mathcal{G}'(s^*_W, s_B)) \text{ and } \rho((s_W, s_B), u) \neq \rho((s^*_W, s_B), u)) \leq \theta.$$

Proof. For a starting vertex u, let $A(u)$ be the event that there exists a strategy $s_W \in S_W$ of White such that $\mu_u(\mathcal{G}'(s_W, s_B)) \geq \mu_u(\mathcal{G}'(s^*_W, s_B))$ and $\rho((s_W, s_B), u) \neq \rho((s^*_W, s_B), u))$. The probability we want to bound is $\Pr(\bigcup_{u\in V} A(u))$. Thus it is enough to bound $\Pr(A(u))$ for a fixed vertex u and then apply a union bound.

Suppose that there exists a strategy s_W that causes the event $A(u)$ to occur (this means that $\mu_u(\mathcal{G}'(s_W, s_B)) \geq \mu_u(\mathcal{G}'(s_W^*, s_B)))$. By the optimality of s_W^* (in $\mathcal{G}$), we have $\mu_u(\mathcal{G}(s_W^*, s_B)) \geq \mu_u(\mathcal{G}(s_W, s_B))$ with a probability of 1, with equality holding if an only if $\rho((s_W^*, s_B), u) = \rho((s_W, s_B), u))$. Since $\|r' - r\|_\infty \leq \varepsilon$, Proposition 3.1 implies that

$$\mu_u(\mathcal{G}(s_W, s_B)) \geq \mu_u(\mathcal{G}'(s_W, s_B)) - \varepsilon \geq \mu_u(\mathcal{G}'(s_W^*, s_B)) - \varepsilon.$$

Furthermore, $\mu_u(\mathcal{G}'(s_W^*, s_B)) \geq \mu_u(\mathcal{G}(s_W^*, s_B)) - \varepsilon$. This yields

$$\mu_u(\mathcal{G}(s_W^*, s_B)) \geq \mu_u(\mathcal{G}(s_W, s_B)) \geq \mu_u(\mathcal{G}(s_W^*, s_B)) - 2\varepsilon. \tag{17}$$

Now we show that the existence of such an s_W is unlikely using Lemma 4.1.

Let $\mathcal{F} = \{\rho((s_W, s_B), u) \mid s_W \in S_W\}$. Then the elements of $\mathcal{F}$ satisfy the conditions of Lemma 4.1, with γ and κ as defined in Lemma 4.2, and

$$\text{gap}(r) = \mu_u(\mathcal{G}(s_W^*, s_B)) - \max_{s'_W : \rho((s'_W, s_B), u) \neq \rho((s_W^*, s_B), u)} \mu_u(\mathcal{G}(s'_W, s_B)).$$

Note that (17) implies that $\text{gap}(r) \leq 2\varepsilon$, and by Lemmas 4.1 and 4.2, the probability of this happening is at most $\frac{2n^2\varepsilon\phi}{\gamma(\mathcal{G})^3}$. The claim follows. □

Now we proceed to prove the lemma.

(i). Suppose that $s = (s_W, s_B)$ is uniformly optimal in $\mathcal{G}$. Then, $s_W \in \text{UBR}_{\mathcal{G}}(s_B)$ and $s_B \in \text{UBR}_{\mathcal{G}}(s_W)$, and the claim implies immediately that $\Pr(s \text{ is not uniformly optimal in } \mathcal{G}') \leq 2\theta$.

(ii). Suppose that $s = (s_W, s_B)$ is uniformly optimal in $\mathcal{G}'$. Pick $s_W^* \in \text{UBR}_{\mathcal{G}}(s_B)$ and $s_B^* \in \text{UBR}_{\mathcal{G}}(s_W)$. Then, by the claim, with probability at least $1 - \theta$, for all $u \in V$, either $\mu_u(\mathcal{G}'(s_W, s_B)) < \mu_u(\mathcal{G}'(s_W^*, s_B))$ or $\rho((s_W, s_B), u) = \rho((s_W^*, s_B), u)$. The former condition contradicts the optimality of s_W in $\mathcal{G}'$, so we must have the latter condition, which in turn implies that $\mu_u(\mathcal{G}(s_W, s_B)) = \mu_u(\mathcal{G}(s_W^*, s_B))$, i.e., $s_W \in \text{UBR}_{\mathcal{G}}(s_B)$. Similarly, we can show that with probability at least $1 - \theta$, $s_B \in \text{UBR}_{\mathcal{G}}(s_W)$, and the result follows. □

Still, it can happen that rounding results in different optimal strategies. How can we be sure that the solution obtained from the rounded rewards is also optimal for the game with the true rewards? Step 7 in Algorithm 3 is one way to do this. The basic idea is as follows: Let $\tilde{s}$ be a uniformly optimal situation in the rounded game. Lemma 4.3 says that with high probability $\tilde{s}$ is a uniformly optimal situation in $\mathcal{G}$, and hence it is also uniformly optimal in any game on the same graph and transition matrix, but with rewards lying in a small interval around the rounded rewards. Thus, we create $|E|$ copies of the truncated game; in each copy the reward on a single arc is perturbed by a certain amount within this small interval. If $\tilde{s}$ is uniformly optimal in all these games, then it is also uniformly optimal for all rewards from that small interval. The following lemma justifies the correctness of this certificate.

Lemma 4.5. *Let $\tilde{\mathcal{G}} = (G = (V, E), P, \tilde{r})$ be a BWR-game and u be an arbitrary vertex. Consider a situation $\tilde{s} = (\tilde{s}_W, \tilde{s}_B)$ such that, for all $e \in E$, $\tilde{s}$ is optimal in the game $(\tilde{\mathcal{G}}_{e,\varepsilon}, u)$ (defined in (16)). Then $\tilde{s}$ is also optimal in $(\mathcal{G} = (G, P, r), u)$, for any r such that $\|r - \tilde{r}\|_\infty \leq \varepsilon$.*

Proof. Fix an r such that $\|r - \tilde{r}\|_\infty \leq \varepsilon$. Let $s_W \in S_W$ be any strategy of White. If $\mu_u(\mathcal{G}(s_W, \tilde{s}_B)) \neq \mu_u(\mathcal{G}(\tilde{s}_W, \tilde{s}_B))$, then according to Lemma 4.2, there exists an arc e with $\rho_e(s_W, \tilde{s}_B) - \rho_e(\tilde{s}_W, \tilde{s}_B) \geq \gamma$. Note that for any situation s, $\mu_u(\tilde{\mathcal{G}}_{e,\varepsilon}(s)) = \mu_u(\tilde{\mathcal{G}}(s)) + \frac{2\varepsilon}{\gamma}\rho_e(s)$.

Thus, we have

$$
\begin{aligned}
\mu_u(\mathcal{G}(s_W,\tilde{s}_B)) &\le \mu_u(\tilde{\mathcal{G}}(s_W,\tilde{s}_B))+\varepsilon && \text{by Proposition 3.1}\\
&= \mu_u(\tilde{\mathcal{G}}_{e,\varepsilon}(s_W,\tilde{s}_B))-\frac{2\varepsilon}{\gamma}\rho_e(s_W,\tilde{s}_B)+\varepsilon\\
&\le \mu_u(\tilde{\mathcal{G}}_{e,\varepsilon}(\tilde{s}_W,\tilde{s}_B))-\frac{2\varepsilon}{\gamma}\rho_e(s_W,\tilde{s}_B)+\varepsilon && \text{since } \tilde{s} \text{ is optimal in } \tilde{\mathcal{G}}_{e,\varepsilon}\\
&= \mu_u(\tilde{\mathcal{G}}(\tilde{s}_W,\tilde{s}_B))-\frac{2\varepsilon}{\gamma}(\rho_e(s_W,\tilde{s}_B)-\rho_e(\tilde{s}_W,\tilde{s}_B))+\varepsilon\\
&\le \mu_u(\tilde{\mathcal{G}}(\tilde{s}_W,\tilde{s}_B))-\varepsilon\\
&\le \mu_u(\mathcal{G}(\tilde{s}_W,\tilde{s}_B)). && \text{by Proposition 3.1}
\end{aligned}
$$

This shows that White cannot improve by switching to a different strategy. By a similar argument, $\mu_u(\mathcal{G}(\tilde{s}_W,s_B))\ge\mu_u(\mathcal{G}(\tilde{s}_W,\tilde{s}_B))$ for all $s_B\in S_B$, which completes the proof. □

Now, we have all ingredients to prove that BWR-games, on graphs which admit a pseudo-polynomial algorithm and have a constant number of random vertices, can be solved in smoothed polynomial time.

Theorem 4.6. *Algorithm 3 solves (in uniformly optimal strategies) any BWR-game $\mathcal{G}=(G,P,r)$ in smoothed polynomial time, given that G admits a pseudo-polynomial algorithm $\mathbb{A}$ (that solves any such $\mathcal{G}$ in uniformly optimal strategies), the number k of random vertices is constant, and $D=poly(n)$.*

Proof. For the correctness of Algorithm 3, it suffices to show the correctness of the certificate step (Step 7). Note that, in the iteration corresponding to $i=\ell$, $\|\tilde{r}-r\|_\infty\le\varepsilon=2^{-\ell}$. Thus, by Lemma 4.5, since the situation $\tilde{s}$, found in Step 6, is uniformly optimal in $(\tilde{\mathcal{G}}_{e,\varepsilon},u)$ for all $e\in E$, then it is also uniformly optimal in $\mathcal{G}$.

It remains to argue about the running time of the algorithm. For this, we bound from above the probability that we fail to certify that $\tilde{s}$ is uniformly optimal in $\tilde{\mathcal{G}}_{e,\varepsilon}$ for some $e\in E$. Note that, for any $e\in E$,

$$\|\tilde{r}(e)-r\|_\infty\le 2^{-\ell}+2\gamma^{-1}\cdot 2^{-\ell}\le 3\cdot 2^{-\ell}\gamma^{-1}. \tag{18}$$

Now conditioned on the event that $\tilde{s}$ is uniformly optimal in $\tilde{\mathcal{G}}$, let A be the event that $\tilde{s}$ is uniformly optimal in $\mathcal{G}$, and let, for $e\in E$, A_e be the event that $\tilde{s}$ is not uniformly optimal for $\tilde{\mathcal{G}}_{e,\varepsilon}$. By Lemma 4.3(i), applied to $\mathcal{G}$ and $\mathcal{G}'=\tilde{\mathcal{G}}_{e,\varepsilon}$, where the difference in rewards satisfies (18), we have $\Pr(A_e\mid A)\le\frac{12n^3\phi 2^{-\ell}}{\gamma(\mathcal{G})^4}$. By Lemma 4.3(i), applied to $\mathcal{G}$ and $\mathcal{G}'=\tilde{\mathcal{G}}$, we have $\Pr(\overline{A})\le\frac{4n^3\phi 2^{-\ell}}{\gamma(\mathcal{G})^3}$. Thus, the probability that $\tilde{s}$ is not uniformly optimal in $\tilde{\mathcal{G}}_{e,\varepsilon}$, for some $e\in E$ is

$$
\begin{aligned}
\Pr(\exists e\in E:\ A_e) &\le \sum_{e\in E}\Pr(A_e)\le\sum_{e\in E}(\Pr(A_e\mid A)+\Pr(\overline{A}))\\
&\le |E|(\frac{12n^3\phi 2^{-\ell}}{\gamma(\mathcal{G})^4}+\frac{4n^3\phi 2^{-\ell}}{\gamma(\mathcal{G})^3})\le\frac{16n^5\phi 2^{-\ell}}{\gamma(\mathcal{G})^4}\le(nD)^{c_0}2^{-\ell}\phi,
\end{aligned}
$$

for some constant $c_0>1$.

Note that Step 7 of the procedure can be implemented in polynomial-time (in n and $\log D$) since it amounts to solving a number of Markov decision processes (which can be solved by linear programming [30]) with polynomially many bits. Let $c_4n^{c_1}2^{\ell c_2}D^{c_3}$ be an upper bound on the running time of each iteration of the Algorithm 3 which is dominated by the time required by the pseudo-polynomial algorithm on an instance of size n where the maximum weight is 2^ℓ, and such that all probabilities are integer multiples of $\frac{1}{D}$, where c_1,c_2,c_3 and c_4 are non-negative constants. By our assumption, there exists a non-negative constant c_5, such that $D \leq n^{c_5}$. Define $\alpha = \min\{\frac{1}{c_1+c_3c_5+c_0c_2+c_0c_2c_5}, \frac{1}{2c_2}\}$. Then,

$$\begin{aligned}\mathbb{E}(\text{running-time}^{\alpha}) &\leq \sum_{\ell \geq \ell_0} (nD)^{c_0} 2^{-\ell} \phi \big(c_4 n^{c_1} 2^{\ell c_2} D^{c_3}\big)^{\alpha} \\ &= c_4^{\alpha} n^{\alpha(c_1+c_3c_5+c_0c_2+c_0c_2c_5)} \phi^{\alpha c_2} \cdot \sum_{i=0}^{\infty} 2^{-i(1-\alpha c_2)} = O(n\phi).\end{aligned}$$

This shows that Algorithm 3 runs in smoothed polynomial time. □

References

[1] D. Andersson and P. B. Miltersen. The complexity of solving stochastic games on graphs. In *Proc. 20th ISAAC*, volume 5878 of *LNCS*, pages 112–121, 2009.

[2] D. Arthur, B. Manthey, and H. Röglin. k-means has smoothed polynomial complexity. In *Proc. 50th FOCS*, pages 405–414, 2009.

[3] R. Beier and B. Vöcking. Typical properties of winners and losers in discrete optimization. *SIAM J. Comput.*, 35(4):855–881, 2006.

[4] H. Björklund and S. Vorobyov. Combinatorial structure and randomized sub-exponential algorithm for infinite games. *Theoret. Comput. Sci.*, 349(3):347–360, 2005.

[5] H. Björklund and S. Vorobyov. A combinatorial strongly sub-exponential strategy improvement algorithm for mean payoff games. *Discrete Appl. Math.*, 155(2):210–229, 2007.

[6] A. Bogdanov and L. Trevisan. Average-case complexity. *Foundations and Trends in Theoretical Computer Science*, 2(1):1–106, 2006.

[7] E. Boros, K. Elbassioni, V. Gurvich, and K. Makino. Every stochastic game with perfect information admits a canonical form. RRR-09-2009, RUTCOR, Rutgers University, 2009.

[8] E. Boros, K. M. Elbassioni, V. Gurvich, and K. Makino. A pumping algorithm for ergodic stochastic mean payoff games with perfect information. In *Proc. 14th IPCO*, volume 6080 of *LNCS*, pages 341–354. Springer, 2010.

[9] E. Boros and V. Gurvich. Why chess and backgammon can be solved in pure positional uniformly optimal strategies. RRR-21-2009, RUTCOR, Rutgers University, 2009.

[10] K. Chatterjee and T. A. Henzinger. Reduction of stochastic parity to stochastic mean-payoff games. *Inf. Process. Lett.*, 106(1):1–7, 2008.

[11] K. Chatterjee, M. Jurdziński, and T. A. Henzinger. Quantitative stochastic parity games. In *Proc. 15th SODA*, pages 121–130, 2004.

[12] X. Chen, X. Deng, and S.-H. Teng. Settling the complexity of computing two-player nash equilibria. *J. ACM*, 56(3), 2009.

[13] G. E. Cho and C. D. Meyer. Markov chain sensitivity measured by mean first passage times. *Linear Algebra and its Applications*, 316(1–3):21–28, 2000.

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

[15] A. Condon. An algorithm for simple stochastic games. In *Advances in computational complexity theory, volume 13 of DIMACS series in discrete mathematics and theoretical computer science*, 1993.

[16] V. Dhingra and S. Gaubert. How to solve large scale deterministic games with mean payoff by policy iteration. In *valuetools '06: Proc. 1st Int. Conf. on Performance evaluation methodolgies and tools*, page 12. ACM, 2006.

[17] A. Eherenfeucht and J. Mycielski. Positional strategies for mean payoff games. *International Journal of Game Theory*, 8:109–113, 1979.

[18] 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.

[19] H. Gimbert and F. Horn. Simple stochastic games with few random vertices are easy to solve. In *Proc. 11th FoSSaCS*, volume 4962 of *LNCS*, pages 5–19, 2008.

[20] V. Gurvich, A. Karzanov, and L. Khachiyan. Cyclic games and an algorithm to find minimax cycle means in directed graphs. *USSR Computational Mathematics and Mathematical Physics*, 28:85–91, 1988.

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

[22] M. Jurdziński. Deciding the winner in parity games is in UP $\cap$ co-UP. *Inf. Process. Lett.*, 68(3):119–124, 1998.

[23] M. Jurdziński. *Games for Verification: Algorithmic Issues.* PhD thesis, Faculty of Science, University of Aarhus, USA, 2000.

[24] M. Jurdzinski, M. Paterson, and U. Zwick. A deterministic subexponential algorithm for solving parity games. In *Proc. 17th SODA*, pages 117–123, 2006.

[25] R. M. Karp. A characterization of the minimum cycle mean in a digraph. *Discrete Math.*, 23:309–311, 1978.

[26] A. V. Karzanov and V. N. Lebedev. Cyclical games with prohibition. *Mathematical Programming*, 60:277–293, 1993.

[27] T. M. Liggett and S. A. Lippman. Stochastic games with perfect information and time-average payoff. *SIAM Review*, 4:604–607, 1969.

[28] M. L. Littman. *Algorithm for sequential decision making, CS-96-09.* PhD thesis, Dept. of Computer Science, Brown Univ., USA, 1996.

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

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

[31] H. Moulin. Extension of two person zero sum games. *Journal of Mathematical Analysis and Application*, 5(2):490–507, 1976.

[32] K. Mulmuley, U. V. Vazirani, and V. V. Vazirani. Matching is as easy as matrix inversion. *Combinatorica*, 7(1):105–113, 1987.

[33] N. N. Pisaruk. Mean cost cyclical games. *Mathematics of Operations Research*, 24(4):817–828, 1999.

[34] H. Röglin and S.-H. Teng. Smoothed analysis of multiobjective optimization. In *Proc. 50th FOCS*, pages 681–690, 2009.

[35] A. Roth, M.-F. Balcan, A. Kalai, and Y. Mansour. On the equilibria of alternating move games. In *Proc. 21st SODA*, pages 805–816. SIAM, 2010.

[36] S. Schewe. From parity and payoff games to linear programming. In *Proc. 34th MFCS*, volume 5734 of *LNCS*, pages 675–686. Springer, 2009.

[37] D. A. Spielman and S.-H. Teng. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *J. ACM*, 51(3):385–463, 2004.

[38] D. A. Spielman and S.-H. Teng. Smoothed analysis: An attempt to explain the behavior of algorithms in practice. *C. ACM*, 52(10):76–84, 2009.

[39] R. Vershynin. Beyond hirsch conjecture: Walks on random polytopes and smoothed complexity of the simplex method. *SIAM J. Comput.*, 39(2):646–678, 2009.

[40] S. Vorobyov. Cyclic games and linear programming. *Discrete Appl. Math.*, 156(11):2195–2231, 2008.

[41] U. Zwick and M. Paterson. The complexity of mean payoff games on graphs. *Theoret. Comput. Sci.*, 158(1-2):343–359, 1996.

A Some Technical lemmas

For a situation s, we denote by $d_{G(s)}(u,v)$ the *stochastic distance* from u to v in $G(s)$, which is the shortest (directed) distance between vertices u and v in the graph obtained from $G(s)$ by setting the length of every deterministic arc to 0 and of every stochastic arc to 1. Let $\lambda = \lambda(\mathcal{G}) = \max_s \max\{d_{G(s)}(v,u) \mid v,u \in V, d_{G(s)}(v,u) \text{ is finite}\}$ be the *stochastic diameter* of $\mathcal{G}$. Clearly, $\lambda(\mathcal{G}) \le k(\mathcal{G})$. Some of our bounds will be given in terms of λ, implying stronger bounds on the running times of some of the approximation schemes.

A set of vertices $U \subseteq V$ is called an absorbing class of the Markov chain $\mathcal{M}$ if there is no arc with positive probability from U to $V \setminus U$, i.e., U can never be left once it is entered, and U is strongly connected, i.e., any vertex of U is reachable from any other vertex of U.

Lemma A.1. *Let $\mathcal{M} = (G = (V,E), P)$ be a Markov chain on n vertices with starting vertex u. Then the limiting probability of any vertex $v \in V$ is either 0 or at least $p_{\min}^{2\lambda}/n$ and the limiting probability of any arc $(u,v) \in E$ is either 0 or $p_{\min}^{2\lambda+1}/n$.*

Proof. Let π and ρ denote the limiting vertex- and arc-distributions, respectively. Let $C_1,\ldots,C_\ell$ denote the absorbing classes of $\mathcal{M}$ reachable from u. We deal with π first. Clearly, for any v that does not lie in any of these absorbing classes, we have $\pi_v = 0$. It remains to show that for all i and every $v' \in C_i$ we have $\pi_{v'} \ge p_{\min}^{2\lambda}/n$. Denote by $\pi_{C_i} = \sum_{v\in C_i} \pi_v$ the total limiting probability of C_i. Note that π_{C_i} is equal to the probability that we reach some vertex $v \in C_i$ starting from u. Since there is a simple path in G from u to C_i with at most λ stochastic vertices, this probability is at least $p_{\min}^{\lambda}$. Furthermore, there exists a vertex $v \in C_i$ with $\pi_v \ge \pi_{C_i}/|C_i| \ge p_{\min}^{\lambda}/n$. Now for any $v' \in C_i$, there exists again a simple path in G from v to v' with at most λ stochastic nodes, so the probability that we reach v' starting from v is at least $p_{\min}^{\lambda}$. It follows that $\pi_{v'} \ge p_{\min}^{2\lambda}/n$.

Now for ρ, note that $\rho_{uv} \ge \pi_u p_{\min}$, if $(u,v) \in E$. Since π_u is either 0 or at least $p_{\min}^{2\lambda}/n$, the claim follows. □

A Markov chain is said to be *irreducible* if its state space is a single absorbing class. For an irreducible Markov chain, let m_{uv} denote the *mean first passage time* from vertex u to vertex v, and m_{vv} denotes the *mean return time* to vertex v: m_{uv} is the expected number of steps to reach vertex v for the first time, starting from vertex u, and m_{vv} is the expected number of steps to return to vertex v for the first time, starting from vertex v. The following lemma by Cho any Meyer [13] relates these values to the sensitivity of the limiting probabilities of a Markov chain.

Lemma A.2 ([13])**.** *Let $\varepsilon > 0$. Let $\mathcal{M} = (G = (V,E), P)$ be an irreducible Markov chain. For any transition probabilities $\tilde{P}$ with $\|\tilde{P} - P\|_\infty \le \varepsilon$ such that the corresponding Markov chain $\tilde{\mathcal{M}}$ is also irreducible, we have $\|\tilde{\pi} - \pi\|_\infty \le \frac{1}{2}\varepsilon \max_v \frac{\max_{u\ne v} m_{uv}}{m_{vv}}$, where m_{vu} are the mean values defined w.r.t $\mathcal{M}$.*

Let $\mathcal{M} = (G = (V,E), P, r)$ be a weighted Markov chain. We denote by $\mu_u(\mathcal{M}) := \sum_{(v,u)\in E} \pi_v p_{vu} r_{vu}$ the limiting average weight, where π is the limiting distribution when u is the starting position. We will write μ_u when $\mathcal{M}$ is understood from the context.

Lemma A.3. *Let $\mathcal{M} = (G = (V,E), P, r)$ be a weighted Markov chain with arc weights in $[r_-, r_+]$, and $\varepsilon \le \frac{1}{2}p_{\min} = \frac{1}{2}p_{\min}(\mathcal{M})$ be a positive constant. Let $\tilde{\mathcal{M}} = (G = (V,E), \tilde{P}, r)$ be the weighted Markov chain with transition probabilities $\tilde{P}$ such that $\|\tilde{P} - P\|_\infty \le \varepsilon$. Then, for any $u \in V$, $|\mu_u(\tilde{\mathcal{M}}) - \mu_u(\mathcal{M})| \le \delta(\mathcal{M},\varepsilon)$ (defined as in (6)).*

Proof. Fix the starting vertex $u_0 \in V$. Let π and $\tilde{\pi}$ denote the limiting distributions corresponding to $\mathcal{M}$ and $\tilde{\mathcal{M}}$, respectively. We first bound $|\pi - \tilde{\pi}|_\infty$. Since $\varepsilon < p_{\min}$, we have $\tilde{p}_{uv} = 0$ if and only if $p_{uv} = 0$. It follows that $\mathcal{M}$ and $\tilde{\mathcal{M}}$ have the same absorbing classes.

Let $C_1, \ldots, C_\ell$ denote these classes. Denote by $\pi_{C_i} = \sum_{v \in C_i} \pi_v$ and $\tilde{\pi}_{C_i} = \sum_{v \in C_i} \tilde{\pi}_v$ the total limiting probability of C_i with respect to π and $\tilde{\pi}$, respectively. Furthermore, let $\pi^{|i}$ and $\tilde{\pi}^{|i}$ be the limiting distributions, corresponding respectively to $\mathcal{M}$ and $\tilde{\mathcal{M}}$, conditioned on the event that the Markov process is started in C_i (i.e., $u_0 \in C_i$). Note that these conditional limiting distributions describe the limiting distributions for the irreducible Markov chains restricted to C_i. By Lemma A.2, we have $|\pi^{|i} - \tilde{\pi}^{|i}|_\infty \le \frac{1}{2}\varepsilon \max_{v \in C_i} \max_{\substack{u \in C_i \\ u \neq v}} \frac{m_{uv}}{m_{vv}}$.

Claim A.4. *For any $u, v \in C_i$, $m_{uv} \le \frac{(\lambda+1)|C_i|}{p_{\min}^\lambda}$.*

Proof. Fix $v \in C_i$. Note that, for any $u \in C_i$, we have

$$m_{uv} = \sum_{w \neq v} p_{uw}(1 + m_{wv}) + p_{uv}. \tag{19}$$

Let $h = \max\{d_G(u, v) : \ u \in C_i\}$. For $l = 0, 1, \ldots, h$, let $X_l = \max\{m_{uv} : \ u \in C_i, \ d_G(u, v) = l\}$. Let $\ell = \operatorname{argmax}\{X_l : \ l = 1 \ldots, h\}$. Then $X_0 \le |C_i|$ and, for $l = 1, \ldots, h$, (19) implies that

$$X_l \le |C_i| + p_{\min} X_{l-1} + (1 - p_{\min}) X_\ell. \tag{20}$$

(Indeed, for a vertex for $u \in V$ such that $d_G(u, v) = l$, there is a path Q from u to v with l stochastic arcs. Let u' be the vertex closest to u on Q such that $d_G(u', v) = l - 1$, and let u'' be the vertex on Q preceding u'. Then u'' is stochastic, and hence by (19)

$$\begin{aligned} m_{u''v} &\le p_{u''u'}(1 + X_{l-1}) + \sum_{w \neq u'} p_{u''w}(1 + X_\ell) \\ &= p_{u''u'}(1 + X_{l-1}) + (1 - p_{u''u'})(1 + X_\ell) \\ &\le p_{\min}(1 + X_{l-1}) + (1 - p_{\min})(1 + X_\ell), \end{aligned}$$

using the fact that $X_l \le X_\ell$ for all l and $p_{u''u'} \ge p_{\min}$. Finally, $m_{uv} \le |C_i| - 1 + m_{u''v}$ implies (20).)

Iterating (20), for $l = 1, \ldots, \ell$, we get

$$X_\ell \le |C_i| \frac{1 - p_{\min}^{\ell+1}}{1 - p_{\min}} + X_\ell(1 - p_{\min}^\ell),$$

implying that $X_\ell \le |C_i| \frac{1 - p_{\min}^{\ell+1}}{1 - p_{\min}} p_{\min}^{-\ell} \le |C_i|(\lambda + 1) p_{\min}^{-\lambda}$. ☐

It follows that $\|\pi^{|i} - \tilde{\pi}^{|i}\|_\infty \le \frac{\varepsilon(\lambda+1)|C_i|}{2p_{\min}^\lambda}$.

Claim A.5. $|\pi_{C_i} - \tilde{\pi}_{C_i}| \le \varepsilon n \lambda p_{\min}^{-\lambda}$.

Proof. Without loss of generality we assume that $u_0 \notin C_i$. For a transient vertex v, let y_v (resp., $\tilde{y}_v$) be the absorption probability into class C_i, in $\mathcal{M}$, (resp., $\tilde{\mathcal{M}}$). In particular $y_{u_0} = \pi_{C_i}$. Note that

$$y_v = \sum_{u \notin C_i} p_{vu} y_u + p_{vC_i}, \text{ where } p_{vC_i} = \sum_{u \in C_i} p_{vu}. \tag{21}$$

Similarly,

$$\tilde{y}_v = \sum_{u \notin C_i} \tilde{p}_{vu} \tilde{y}_u + \tilde{p}_{vC_i} = \sum_{u \notin C_i} p_{vu} \tilde{y}_u + \sum_{u \notin C_i} (\tilde{p}_{vu} - p_{vu}) \tilde{y}_u + \tilde{p}_{vC_i}. \tag{22}$$

Write $\Delta_v := |\tilde{y}_v - y_v|$. Subtracting (21) from (22) and bounding, we get:

$$\begin{aligned} \Delta_v &\le \sum_{u \notin C_i} p_{vu} \Delta_u + \sum_{u \notin C_i} |\tilde{p}_{vu} - p_{vu}| \tilde{y}_u + |\tilde{p}_{vC_i} - p_{vC_i}| \\ &\le \sum_{u \notin C_i} p_{vu} \Delta_u + (n - |C_i|)\varepsilon + |C_i|\varepsilon = \sum_{u \notin C_i} p_{vu} \Delta_u + \varepsilon n. \end{aligned} \tag{23}$$

Let $h=\max\{d_G(u,C_i):\ u\notin C_i,\ d_G(u,C_i)<\infty\}$, where $d_G(u,C_i)=\min\{d_G(u,v):\ v\in C_i\}$ is the stochastic distance in G from u to C_i. For $l=0,1,\ldots,h$, let $X_l=\max\{\Delta_u:\ u\notin C_i,\ d(u,C_i)=l\}$, and let $\ell=\operatorname{argmax}\{X_l:\ l=1\ldots,h\}$. Then $X_0=0$ (since deterministic vertices in $\mathcal{M}$ remain deterministic in $\tilde{\mathcal{M}}$) and, for $l=1,\ldots,h$, (23) implies that

$$X_l\le\varepsilon n+p_{\min}X_{l-1}+(1-p_{\min})X_\ell. \tag{24}$$

Iterating we get that $X_\ell\le\varepsilon n\frac{1-p_{\min}^{\ell}}{1-p_{\min}}p_{\min}^{-\ell}\le\varepsilon n\lambda p_{\min}^{-\lambda}$. □

Let $v\in V$ be an arbitrary vertex. If v does not lie in any absorbing class, then $\pi_v=\tilde{\pi}_v=0$. Otherwise, let $v\in C_i$. By the above claims, we have

$$\begin{aligned}\pi_v &= \pi_{C_i}\pi_v^{|i}\le(\tilde{\pi}_{C_i}+\varepsilon n\lambda p_{\min}^{-\lambda})(\tilde{\pi}_v^{|i}+\frac{\varepsilon}{2}(\lambda+1)|C_i|p_{\min}^{-\lambda})\\ &\le \tilde{\pi}_v+\frac{\varepsilon}{2}np_{\min}^{-\lambda}[\varepsilon n\lambda(\lambda+1)p_{\min}^{-\lambda}+3\lambda+1]:=\tilde{\pi}_v+\delta'(\mathcal{M},\varepsilon).\end{aligned}$$

Similarly, we can conclude that $\tilde{\pi}_v\le\pi_v+\delta'(\tilde{\mathcal{M}},\varepsilon)$. Note that $p_{\min}(\tilde{\mathcal{M}})\ge p_{\min}(\mathcal{M})/2$, since $\varepsilon\le p_{\min}(\mathcal{M})/2$. It follows that

$$\begin{aligned}|\mu_{u_0}(\mathcal{M})-\mu_{u_0}(\tilde{\mathcal{M}})| &\le \sum_{(u,v)\in E}|\pi_u p_{uv}-\tilde{\pi}_u\tilde{p}_{uv}||r_{uv}|\\ &\le \sum_{(u,v)\in E}(|\pi_u-\tilde{\pi}_u|p_{uv}+\tilde{\pi}_u|\tilde{p}_{uv}-p_{uv}|)r_*\\ &\le \sum_{(u,v)\in M}(\delta'(\tilde{\mathcal{M}},\varepsilon)p_{uv}+\tilde{\pi}_u\varepsilon)r_*\\ &\le (\delta'(\tilde{\mathcal{M}},\varepsilon)+\varepsilon)n\le\delta(\mathcal{M},\varepsilon)r_*,\end{aligned}$$

and the result follows. □