Approximation Schemes for Stochastic Mean Payoff Games with Perfect Information and Few Random Positions

Endre Boros · **Khaled Elbassioni** ·
Mahmoud Fouz · **Vladimir Gurvich** ·
Kazuhisa Makino · **Bodo Manthey**

June 4, 2020

Abstract We consider two-player zero-sum stochastic mean payoff games with perfect information. We show that any such game, with a constant number of random positions and polynomially bounded positive transition proba-

E. Boros
MSIS Department and RUTCOR, Rutgers University, 100 Rockafellar Road, Livingston Campus, Piscataway, NJ 08854, USA;
E-mail: Endre.Boros@rutgers.edu

K. Elbassioni
Masdar Institute, Khalifa University of Science and Technology, P.O. Box 54224, Abu Dhabi, UAE
E-mail: kelbassioni@masdar.ac.ae

M. Fouz
Department of Computer Science, Saarland University, Saarbrücken, Germany;
E-mail: mfouz@cs.uni-saarland.de

V. Gurvich
MSIS Department and RUTCOR, Rutgers University, 100 Rockafellar Road, Livingston Campus, Piscataway, NJ 08854, USA; Dep. of Computer Sciences, National Research University Higher School of Economics (HSE), Moscow
E-mail: Vladimir.Gurvich@rutgers.edu

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

B. Manthey
Department of Applied Mathematics, University of Twente, Enschede, the Netherlands;
E-mail: b.manthey@utwente.nl

bilities, admits a polynomial time approximation scheme, both in the relative and absolute sense.

1 Introduction

The rise of the Internet has led to an explosion in research in game theory, the mathematical modeling of competing agents in strategic situations. The central concept in such models is that of a *Nash equilibrium*, which defines a state where no agent gains an advantage by changing to another strategy. Nash equilibria serve as predictions for the outcome of strategic situations in which selfish agents compete.

A fundamental result in game theory states that if the agents can choose a *mixed strategy* (i.e., probability distributions of deterministic strategies), a Nash equilibrium is guaranteed to exist in finite games [24,25]. Often, however, already *pure* (i.e., deterministic) strategies lead to a Nash equilibrium. Still, the existence of Nash equilibria might be irrelevant in practice since their computation would take too long. (Finding mixed Nash equilibria in two-player games is PPAD-complete in general [11].) Thus, algorithmic aspects of game theory have gained a lot of interest. Following the dogma that only polynomial time algorithms are feasible algorithms, it is desirable to show polynomial time complexity for the computation of Nash equilibria.

We consider two-player zero-sum stochastic mean payoff games with perfect information. In this case the concept of Nash equilibria coincides with *saddle points* or *mini-max/maxi-min strategies*. The decision problem associated with computing such strategies and the values of these games is in the intersection of NP and co-NP, but it is unknown whether it can be solved in polynomial time. In cases where efficient algorithms are not known to exist, an approximate notion of a saddle point has been suggested. In an approximate saddle point, no agent can gain a *substantial* advantage by changing to another strategy. In this paper, we design approximation schemes for saddle points for such games when the number of random positions is fixed (see Section 1.2 for a definition).

In the remainder of this section, we introduce the concepts used in this paper. Our results are summarized in Section 1.4. After that, we present our approximation schemes (Section 2). We conclude with a list of open problems (Section 3), where we address in particular the question of polynomial smoothed complexity of mean payoff games. In the conference version of this paper [2], we wrongly claimed that stochastic mean payoff games can be solved in smoothed polynomial time.

1.1 Stochastic Mean Payoff Games

1.1.1 Definition and Notation

The model that we consider is described is a *stochastic mean payoff game with perfect information*, or equivalently a *BWR-game* $\mathcal{G} = (G, P, r)$:

- $G = (V, E)$ is a directed graph that may have loops and multiple edges, but no terminal positions, i.e., no positions of out-degree 0. The vertex set V of G is partitioned into three disjoint subsets $V = V_B \cup V_W \cup V_R$ that correspond to black, white, and random *positions*, respectively. The edges stand for *moves*. The black and white positions are owned by two players: BLACK – the *minimizer* – owns the black positions in V_B, and WHITE – the *maximizer* – owns the white positions in V_W. The positions in V_R are owned by nature.
- P is the vector of probability distributions for all positions $v \in V_R$ owned by nature. We assume that $\sum_{u:(v,u)\in E} p_{vu} = 1$ for all $v \in V_R$ and $p_{vu} > 0$ for all $v \in V_R$ and $(v, u) \in E$.
- r is the vector of rewards; each edge e has a local reward r_e.

Starting from some vertex $v_0 \in V$, a token is moved along one edge e in every round of the game. If the token is on a black vertex, BLACK selects an outgoing edge e and moves the token along e. If the token is on a white vertex, then WHITE selects an outgoing edge 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 edges of v. In all cases, BLACK pays WHITE the reward r_e on the selected edge e.

Starting from a given initial position $v_0 \in V$, the game yields an infinite walk $(v_0, v_1, v_2, \ldots)$, called a *play*. Let b_i denote the reward $r_{(v_{i-1}, v_i)}$ received by WHITE in step i. The *undiscounted limit average effective payoff* is defined as the *Cesàro average* $c = \liminf_{n\to\infty} \frac{\sum_{i=1}^{n} \mathbb{E}[b_i]}{n}$. WHITE's objective is to maximize c, while the objective of BLACK is to minimize it.

In this paper, we will restrict our attention to the sets of *pure* (that is, non-randomized) and *stationary* (that is, history-independent) strategies of players WHITE and BLACK, denoted by S_W and S_B, respectively; such strategies are called *positional strategies*. Formally, a positional strategy $s_W \in S_W$ for WHITE is a mapping that assigns a move $(v, u) \in E$ to each position in V_W. We sometimes abbreviate $s_W(v) = (v, u)$ by $s_W(v) = u$. Strategies $s_B \in S_B$ for BLACK are analogously defined. A pair of strategies $s = (s_W, s_B)$ is called a *situation*. By abusing notation, let $s(v) = u$ if $v \in V_W$ and $s_W(v) = u$ or $v \in V_B$ and $s_B(v) = u$.

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:

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

Here, $E(s) = \{e = (v, u) \in E \mid p_{vu}(s) > 0\}$ 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 $(v, u) \in E$, we have $\rho_{vu}(s) = \pi_v(s)p_{vu}(s)$, 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 $\rho(s, v_0)$ if 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 $R = R(\mathcal{G}) = r^+ - r^-$ be the size of the range.

1.1.2 Strategies and Saddle Points

If we consider $c_{v_0}(s)$ for all possible situations s, 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 [19,29]. Such a saddle point defines an equilibrium state in which no player has an incentive to switch to another strategy. The value at that state coincides with the limiting payoff in the corresponding BWR-game [19,29].

We call a pair of strategies *optimal* if they correspond to a saddle point. It is well-known that there exist optimal strategies (s_W^*, s_B^*) that do not depend on the starting position v_0. Such strategies are called *uniformly* optimal. Of course there might be several optimal strategies, but 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 an arbitrary situation s, $\mu_{v_0}(\mathcal{G}(s))$ denotes the effective payoff $c_{v_0}(s)$ in the Markov chain $\mathcal{G}(s)$.

An algorithm is said to solve the game if it computes an optimal pair of strategies.

1.2 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

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

An alternative concept of an approximate equilibrium are *ε-relative equilibria*. They are determined by a situation (s_W^*, s_B^*) such that

$$\max_{s_W} \mu_v(\mathcal{G}(s_W, s_B^*)) \leq (1 + \varepsilon)\mu_v(\mathcal{G}(s_W^*, s_B^*))$$

$$\text{and} \quad \min_{s_B} \mu_v(\mathcal{G}(s_W^*, s_B)) \geq (1 - \varepsilon)\mu_v(\mathcal{G}(s_W^*, s_B^*)). \quad (2)$$

Note that, for sufficiently small ε, an ε-relative approximation implies a $\Theta(\varepsilon)$-relative equilibrium, and vice versa. Thus, in what follows, we will use these notions interchangeably. When considering relative approximations and relative equilibria, we assume that the rewards are non-negative integers.

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

$$\max_{s_W} \mu_v(\mathcal{G}(s_W, s_B^*)) \leq \mu_v(\mathcal{G}) + \varepsilon \quad \text{and} \quad \min_{s_B} \mu_v(\mathcal{G}(s_W^*, s_B)) \geq \mu_v(\mathcal{G}) - \varepsilon. \quad (3)$$

Similarly, for an ε-absolute equilibrium, we have the following condition:

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

Again, an ε-absolute approximation implies a 2ε-absolute equilibrium, and vice versa. When considering absolute equilibria and absolute approximations, we assume that the rewards come from the interval $[-1, 1]$.

A situation (s_W^*, s_B^*) is called relatively ε-optimal, if satisfies (1), and it is called absolutely ε-optimal if it satisfies (3). In the following, we will drop the specification of absolute and relative if it is clear from the context. If the pair (s_W^*, s_B^*) is (absolutely or relatively) ε-optimal for all starting positions, it is called *uniformly* (absolutely or relatively) ε-optimal (also called *subgame perfect*).

We note that, under the above assumptions, the notion of relative approximation is stronger. Indeed, consider a BWR-game $\mathcal{G}$ with rewards in $[-1, 1]$. A relatively ε-optimal situation (s_W^*, s_B^*) of the game $\hat{\mathcal{G}}$ with local rewards given by $\hat{r} = r + \mathbf{1} \geq 0$ (where $\mathbf{1}$ is the vector of all ones, and the addition and comparison is meant component-wise) satisfies

$$\max_{s_W} \mu_v(\mathcal{G}(s_W, s_B^*)) = \max_{s_W} \mu_v(\hat{\mathcal{G}}(s_W, s_B^*)) - 1 \leq (1 + \varepsilon)\mu_v(\hat{\mathcal{G}}) - 1$$

$$= \mu_v(\mathcal{G}) + \varepsilon\mu_v(\mathcal{G}) + \varepsilon \leq \mu_v(\mathcal{G}) + 2\varepsilon$$

and

$$\min_{s_B} \mu_v(\mathcal{G}(s_W^*, s_B)) = \min_{s_B} \mu_v(\hat{\mathcal{G}}(s_W^*, s_B)) - 1 \geq (1 - \varepsilon)\mu_v(\hat{\mathcal{G}}) - 1$$

$$= \mu_v(\mathcal{G}) - \varepsilon\mu_v(\mathcal{G}) - \varepsilon \geq \mu_v(\mathcal{G}) - 2\varepsilon.$$

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

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

1.3 Previous Results

BWR-games are an equivalent formulation [21] of the stochastic games with perfect information and mean payoff that were introduced in 1957 by Gillette [19]. As it was noticed already in [21], the BWR model generalizes a variety of games

and problems: BWR-games without random positions ($V_R = \emptyset$) are called
cyclic or *mean payoff* games [21,33,34,16,17]; we call these *BW-games*. If one
of the sets V_B or V_W is empty, we obtain a *Markov decision process* for which
polynomial-time algorithms are known [32]. If both are empty ($V_B = V_W = \emptyset$),
we get a *weighted Markov chain*. If $V = V_W$ or $V = V_B$, we obtain the *minimum
mean-weight cycle problem*, which can be solved in polynomial time [27].

If all rewards are 0 except for m terminal loops, we obtain the so-called
Backgammon-like or *stochastic terminal payoff* games [7]. The special case
$m = 1$, in which every random node has only two outgoing arcs with probability
$1/2$ each, defines the so-called *simple stochastic games* (SSGs), introduced by
Condon [13,14]. 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 has been shown that Gillette games (and hence BWR-
games [3]) are equivalent to SSGs under polynomial-time reductions [1]. Thus,
by recent results of Halman [22], all these games can be solved in randomized
strongly subexponential time $2^{O(\sqrt{n_d \log n_d})} \operatorname{poly}(|V|)$, where $n_d = |V_B| + |V_W|$ is
the number of *deterministic* positions.

Besides their many applications [30,26], all these games are of interest to
complexity theory: The decision problem "whether the value of a BW-game
is positive" is in the intersection of NP and co-NP [28,40]; yet, no polynomial
algorithm is known even in this special case. We refer to Vorobyov [39] for a
survey. A similar complexity claim holds for SSGs and BWR-games [1,3]. On
the other hand, there exist algorithms that solve BW-games in practice very
fast [21]. The situation for these games is thus comparable to linear program-
ming before the discovery of the ellipsoid method: linear programming was
known to lie in the intersection of NP and co-NP, and the simplex method
proved to be fast in practice. In fact, a polynomial algorithm for linear pro-
gramming in the unit cost model would already imply a polynomial algorithm
for BW-games [37]; see also [4] for an extension to BWR-games.

While there are numerous pseudo-polynomial algorithms known for BW-
games [21,40,35], pseudo-polynomiality for BWR-games (with no restriction
on the number of random positions) is in fact equivalent to polynomiality [1].
Gimbert and Horn [20] have shown that a generalization of simple stochastic
games on k random positions having arbitrary transition probabilities (not
necessarily $(1/2, 1/2)$) can be solved in time $O(k!(|V||E| + L))$, where L is the
maximum bit length of a transition probability. There are various improve-
ments with smaller dependence on k [9,15,20,23]. (Note that even though
BWR-games are polynomially reducible to simple stochastic games, under
this reduction the number of random positions does not stay constant, but is
only polynomially bounded in n, even if the original BWR-game had a con-
stant number of random positions.) Recently, a pseudo-polynomial algorithm
was given for BWR-games with a *constant* number of random positions and
polynomial common denominator of transition probabilities, but under the
assumption that the game is ergodic (that is, the value does not depend on
the initial position) [5]. Then, this result was extended for the non-ergodic
case [6]; see also [4].

As for approximation schemes, the only result we are aware [36] of is the observation 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 [21, 35, 40].

On the negative side, it was observed recently [18] that obtaining an ε-absolute FPTAS without the assumption that all rewards are in $[-1, 1]$, or an ε-relative FPTAS without the assumption that all rewards are non-negative, for BW-games, would imply their polynomial time solvability. In that sense, our results below are the best possible unless there is a polynomial algorithm for solving BW-games.

1.4 Our Results

In this paper, we extend the absolute FPTAS for BW-games [36] in two directions. First, we allow a constant number of random positions, and, second, we derive an FPTAS with a relative approximation error. Throughout the paper, we assume the availability of a pseudo-polynomial algorithm $\mathbb{A}$ that solves any BWR-game $\mathcal{G}$ with integral rewards and rational transition probabilities in time polynomial in n, D, and R, where $n = n(\mathcal{G})$ is the total number of positions, $R = R(\mathcal{G}) := r^+(\mathcal{G}) - r^-(\mathcal{G})$ is the size of the range of the rewards, $r^+(\mathcal{G}) = \max_e r_e$ and $r^-(\mathcal{G}) = \min_e r_e$, and $D = D(\mathcal{G})$ is the common denominator of the transition probabilities. Note that the dependence on D is inherent in all known pseudo-polynomial algorithms for BWR-games. Note also that the affine scaling of the rewards does not change the game.

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

The following theorem says that a pseudo-polynomial algorithm can be turned into an *absolute* approximation scheme.

Theorem 1 *Given a pseudo-polynomial algorithm for solving any BWR-game with $k = O(1)$ (in uniformly optimal strategies), there is an algorithm that returns, for any given BWR-game with rewards in $[-1, 1]$, $k = O(1)$, and for any $\varepsilon > 0$, a pair of strategies that (uniformly) approximates the value within an* absolute *error of ε. The running-time of the algorithm is bounded by $\mathrm{poly}(n, 1/p_{\min}, 1/\varepsilon)$ (assuming $k = O(1)$).*

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

Theorem 2 *Given a pseudo-polynomial algorithm for solving any BWR-game with $k = O(1)$, there is an algorithm that returns, for any given BWR-game with non-negative integral rewards, $k = O(1)$, and for any $\varepsilon > 0$, a pair of strategies that approximates the value within a* relative *error of ε. The running-time of the algorithm is bounded by $\mathrm{poly}(n, 1/p_{\min}, \log R, 1/\varepsilon)$ (assuming $k = O(1)$).*

We remark that Theorem 1 (apart from the dependence of the running time on $\log R$) can be obtained from Theorem 2 (see Section 2). However, our reduction in Theorem 1, unlike Theorem 2, has the property that if the pseudo-polynomial algorithm returns uniformly optimal strategies, then the approximation scheme also returns uniformly ε-optimal strategies. For BW-games, i.e., the special case without random positions, we can also strengthen the result of Theorem 2 to return a pair of strategies that is uniformly ε-optimal.

Theorem 3 *Assume that there is a pseudo-polynomial algorithm for solving any BW-game in uniformly optimal strategies. Then for any $\varepsilon > 0$, there is an algorithm that returns, for any given BW-game with non-negative integral rewards, a pair of uniformly relatively ε-optimal strategies. The running-time of the algorithm is bounded by* $\mathrm{poly}(n, \log R, 1/\varepsilon)$.

In deriving these approximation schemes from a pseudo-polynomial algorithm, we face two main technical challenges that distinguish the computation of ε-equilibria of BWR-games from similar standard techniques used in combinatorial optimization. First, 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, in order to obtain a *fully polynomial-time approximation scheme* (FPTAS) with an absolute guarantee whose running-time is independent of D, we have to truncate the probabilities and bound the change in the game value, which is a *non-linear* function of D. Second, in order to obtain an FPTAS with a relative guarantee, one needs (as often in optimization) a (trivial) lower/upper bound on the optimum value. In the case of BWR-games, it is not clear what bound we can use, since the game value can be arbitrarily small. The situation becomes even more complicated if we look for uniformly ε-optimal strategies. This is because we have to output just a single pair of strategies that 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 $O(\delta n^2 / p_{\min}^{2k})$ (see Section 2.1). It is worth mentioning that a somewhat related result was obtained recently for the class of so-called *almost-sure ergodic games* (not necessarily with perfect information) [10]. More precisely, it was shown that for this class of games there is an ε-optimal strategy with rational representation with denominator $D = O(\frac{n^3}{\varepsilon p_{\min}^k})$ [10]. The second issue is resolved through repeated applications of the pseudo-polynomial algorithm on a truncated game. After each such application we have one of the following situations: 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 2.3 and 2.4).

Since BWR-games with a constant number of random positions admit a pseudo-polynomial algorithm, as was recently shown [5,6], we obtain the following results.

Corollary 1 *(i) There is an FPTAS that solves, within an absolute error guarantee, in uniformly ε-optimal strategies, any BWR-game with a constant number of random positions, $1/p_{\min} = \mathrm{poly}(n)$, and rewards in $[-1, 1]$.*

(ii) There is an FPTAS that solves, within a relative error guarantee, in ε-optimal strategies, any BWR-game with a constant number of random positions, $1/p_{\min} = \mathrm{poly}(n)$, and non-negative rational rewards.

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

The proofs of Theorems 1, 2, and 3 will be given in Sections 2.2, 2.3, and 2.4, respectively.

2 Approximation Schemes

2.1 The Effect of Perturbation

Our approximation schemes are based on the following three lemmas. The first one (which is known) says that a linear change in the rewards corresponds 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 need the lemma to bound the error in the game value resulting from the truncation.

Lemma 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 $\theta_1 r_e + \gamma_1 \mathbf{1} \leq \hat{r}_e \leq \theta_2 r_e + \gamma_2 \mathbf{1}$, for all $e \in E$. Then for any $v \in V$, we have $\theta_1 \mu_v(\mathcal{G}) + \gamma_1 \leq \mu_v(\hat{\mathcal{G}}) \leq \theta_2 \mu_v(\mathcal{G}) + \gamma_2$. Moreover, if $(\hat{s}_W, \hat{s}_B)$ is an absolutely ε-optimal situation in $(\hat{\mathcal{G}}, v)$, then*

$$\max_{s_W} \mu_v(\mathcal{G}(s_W, \hat{s}_B)) \leq \frac{\theta_2 \mu_v(\mathcal{G}) + \gamma_2 - \gamma_1 + \varepsilon}{\theta_1} \quad and$$

$$\min_{s_B} \mu_v(\mathcal{G}(\hat{s}_W, s_B)) \geq \frac{\theta_1 \mu_v(\mathcal{G}) + \gamma_1 - \gamma_2 - \varepsilon}{\theta_2}. \tag{5}$$

Proof This uses only standard techniques, and we give the proof only 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_0 and corresponding to pairs (s_W^*, s_B^*), $(\hat{s}_W, \hat{s}_B)$, $(s_W^*, \hat{s}_B)$, and $(\hat{s}_W, s_B^*)$, respectively. By the definition of optimal strategies and the facts that $\|\rho'\|_1 = \|\rho''\|_1 = 1$ (because they are probability distributions), we have the following series of inequalities:

$$\mu_v(\hat{\mathcal{G}}) = (\hat{\rho})^T \hat{r} \geq (\rho')^T \hat{r} \geq \theta_1 (\rho')^T r + \gamma_1 \geq \theta_1 (\rho^*)^T r + \gamma_1 = \theta_1 \mu_v(\mathcal{G}) + \gamma_1 \quad and$$

$$\mu_v(\hat{\mathcal{G}}) = (\hat{\rho})^T \hat{r} \leq (\rho'')^T \hat{r} \leq \theta_2 (\rho'')^T r + \gamma_2 \leq \theta_2 (\rho^*)^T r + \gamma_2 = \theta_2 \mu_v(\mathcal{G}) + \gamma_2.$$

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

The second lemma, which is new as far as we are aware, states that if we truncate the transition probabilities within a small error ε, then the change in the game value is bounded by $O(\varepsilon^2 n^3 / p_{\min}^{2k})$. More precisely, for a BWR-game $\mathcal{G}$ and a constant $\varepsilon > 0$, define

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

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

Lemma 2 *Let $\mathcal{G} = (G = (V, E), P, r)$ be a BWR-game with $r \in [-1, 1]^E$, and let $\varepsilon \leq p_{\min}/2 = p_{\min}(\mathcal{G})/2$ be a positive constant. Let $\hat{\mathcal{G}}$ be a game $(G = (V, E), \hat{P}, r)$ with $\|P - \hat{P}\|_\infty \leq \varepsilon$ (and $\hat{p}_{uv} = 0$ if $p_{uv} = 0$ for all arcs (u, v)). Then we have $|\mu_v(\mathcal{G}) - \mu_v(\hat{\mathcal{G}})| \leq \delta(\mathcal{G}, \varepsilon)$ for any $v \in V$. Moreover, if the pair $(\tilde{s}_W, \tilde{s}_B)$ is absolutely ε'-optimal in $(\hat{\mathcal{G}}, v)$, then it is absolutely $(\varepsilon' + 2\delta(\mathcal{G}, \varepsilon))$-optimal in $(\mathcal{G}, v)$.*

Proof We apply Lemma 10. 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 10 imply the following two series of inequalities:

$$\mu_v(\hat{\mathcal{G}}) = \mu_v(\hat{\mathcal{G}}(\hat{s}_W, \hat{s}_B)) \geq \mu_v(\hat{\mathcal{G}}(s_W^*, \hat{s}_B))$$
$$\geq \mu_v(\mathcal{G}(s_W^*, \hat{s}_B)) - \delta \geq \mu_v(\mathcal{G}(s_W^*, s_B^*)) - \delta = \mu_v(\mathcal{G}) - \delta \quad \text{and}$$
$$\mu_v(\hat{\mathcal{G}}) = \mu_v(\hat{\mathcal{G}}(\hat{s}_W, \hat{s}_B)) \leq \mu_v(\hat{\mathcal{G}}(\hat{s}_W, s_B^*))$$
$$\leq \mu_v(\mathcal{G}(\hat{s}_W, s_B^*)) + \delta \leq \mu_v(\mathcal{G}(s_W^*, s_B^*)) + \delta = \mu_v(\mathcal{G}) + \delta.$$

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

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

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

Since we assume that the running-time of the pseudo-polynomial algorithm for the original game $\mathcal{G}$ depends on the common denominator D of the transition probabilities, we have to truncate the probabilities to remove this dependence on D. By Lemma 2, the value of the game does not change too much after such a truncation.

The third result that 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 lemma says that we can reduce all large rewards without changing the value of the game.

Lemma 3 *Let $\mathcal{G} = (G = (V, E), P, r)$ be a BWR-game with $r \geq 0$, and let v be any vertex with $\mu_v(\mathcal{G}) < t$. Suppose that $r_e \geq t' = ntp_{\min}^{-(2k+1)}$ for some $e \in E$. Let $\hat{\mathcal{G}} = (G = (V, E), P, \hat{r})$, where $\hat{r}_e = \min\{r_e, t''\}$, $t'' \geq (1 + \varepsilon)t'$ 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 relatively ε-optimal situation in $(\hat{\mathcal{G}}, v)$ is also relatively ε-optimal in $(\mathcal{G}, v)$.*

Proof We assume that $\hat{r}_e = t'' \geq (1 + \varepsilon)t'$, since otherwise there is nothing to prove. Let $s^* = (s_W^*, s_B^*)$ be an optimal situation for $(\mathcal{G}, v)$. This means that $\mu_v(\mathcal{G}) = \mu_v(\mathcal{G}(s^*)) = \rho(s^*)^T r < t$. Lemma 8 says that $\rho_e(s^*) > 0$ implies $\rho_e(s^*) \geq p_{\min}^{2k+1}/n$. Hence, $r_e \rho_e(s^*) \leq \rho(s^*)^T r = \mu_v(\mathcal{G}) < t$ implies that $r_e < t'$, if $\rho_e(s^*) > 0$. We conclude that $\rho_e(s^*) = 0$, and hence $\mu_v(\hat{\mathcal{G}}(s^*)) = \mu_v(\mathcal{G})$.

Since $\hat{r} \leq r$, we have $\mu_v(\hat{\mathcal{G}}(s)) \leq \mu_v(\mathcal{G}(s))$ 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 that also $\mu_v(\hat{\mathcal{G}}(s_W^*, s_B)) \geq \mu_v(\hat{\mathcal{G}}(s_W^*, s_B^*))$ for all $s_B \in S_B$. Indeed, if there is a strategy s_B for BLACK 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 as above, we must have $\rho_e(s_W^*, s_B) = 0$ (since $\rho_e(s_W^*, s_B)(1 + \varepsilon)t' \leq \rho_e(s_W^*, s_B)t'' = \rho_e(s_W^*, s_B)\hat{r}_e \leq \rho(s_W^*, s_B)^T \hat{r} = \mu_v(\hat{\mathcal{G}}(s_W^*, s_B)) < t$). This, however, implies that

$$\mu_v(\mathcal{G}(s_W^*, s_B)) = \mu_v(\hat{\mathcal{G}}(s_W^*, s_B)) < \mu_v(\hat{\mathcal{G}}(s_W^*, s_B^*)) = \mu_v(\mathcal{G}(s_W^*, s_B^*)),$$

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

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

$$\leq (1 + \varepsilon)\mu_v(\hat{\mathcal{G}}) = (1 + \varepsilon)\mu_v(\mathcal{G}) < (1 + \varepsilon)t,$$

gives a contradiction with Lemma 8 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 $s_B \in S_B$,

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

2.2 Absolute Approximation

In this section, we assume that $r^- = -1$ and $r^+ = 1$, i.e., all rewards are from the interval $[-1, 1]$. We may assume also that $\varepsilon \in (0, 1)$ and $\frac{1}{\varepsilon} \in \mathbb{Z}_+$. 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 $\tilde{r} := \lfloor r \rfloor_{\frac{\varepsilon}{4}}$) and truncating the vector of probabilities $(p_{(v,u)})_{u \in V}$ for each random node $v \in V_R$, as described in the following lemma.

Lemma 4 *Let $\alpha \in [0,1]^n$ with $\|\alpha\|_1 = 1$. Let $B \in \mathbb{N}$ 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 $c_i \in \mathbb{N}$ is an integer;
(iii) for all $i = 1, \ldots, n$, $\alpha_i' > 0$ if and only $\alpha_i > 0$; and
(vi) $\|\alpha - \alpha'\|_\infty \leq 2^{-B}$.

Proof This is straight-forward, and we include the proof only for completeness. Without loss of generality, we assume $\alpha_i > 0$ for all i (set $\alpha_i' = 0$ for all i such that $\alpha_i = 0$). Initialize $\varepsilon_0 = 0$ and iterate for $i = 1, \ldots, n$: set $\alpha_i' = \lfloor\alpha_i + \varepsilon_{i-1}\rceil_{2^{-B}}$ and $\varepsilon_i = \alpha_i + \varepsilon_{i-1} - \alpha_i'$. The construction implies (4). Note that $|\varepsilon_i| \leq 2^{-(B+1)}$ for all i, and $\varepsilon_n = \sum_i \alpha_i - \sum_i \alpha_i'$, which implies (4). Furthermore, $|\alpha_i - \alpha_i'| = |\varepsilon_i - \varepsilon_{i-1}| \leq 2^{-B}$, which implies (4). Note finally that (4) follows from (4) since $\min_{i:\alpha_i>0}\{\alpha_i\} > 2^{-B}$.

Lemma 5 *Let $\mathbb{A}$ be a pseudo-polynomial algorithm that solves, in (uniformly) optimal strategies, any BWR-game $\mathcal{G} = (G, P, r)$ in time $\tau(n, D, R)$. Then for any $\varepsilon > 0$, there is an algorithm that solves, in (uniformly) absolutely ε-optimal strategies, any given BWR-game $\mathcal{G} = (G, P, r)$ in time bounded by $\tau\big(n, \frac{2^{k+4}n^2(3k+1)}{\varepsilon p_{\min}^k}, \frac{8}{\varepsilon}\big)$.*

Proof We apply $\mathbb{A}$ to the game $\tilde{\mathcal{G}} = (G, \tilde{P}, \tilde{r})$, where $\tilde{r} := \frac{4}{\varepsilon}\lfloor r\rceil_{\frac{\varepsilon}{4}}$. The probabilities $\tilde{P}$ are obtained from P by applying Lemma 4 with $B = \lceil\log_2(1/\varepsilon')\rceil$, where we select ε' such that $\delta(\mathcal{G}, \varepsilon') \leq \frac{\varepsilon}{4}$ (as defined by (6)). It is easy to check that $\delta(\mathcal{G}, \varepsilon') \leq \varepsilon/4$ for $\varepsilon' = \frac{\varepsilon p_{\min}^k}{2^{k+3}n^2(3k+1)}$, as $r^* = 1$. Note that all rewards in $\tilde{\mathcal{G}}$ are integers in the range $[-\frac{4}{\varepsilon}, \frac{4}{\varepsilon}]$. Since $D(\tilde{\mathcal{G}}) = 2^B$ and $R(\tilde{\mathcal{G}}) = 8/\varepsilon$, the statement about the running-time follows.

Let $\tilde{s}$ be the pair of (uniformly) optimal strategies returned by $\mathbb{A}$ on input $\tilde{\mathcal{G}}$. Let $\hat{\mathcal{G}}$ be the game $(G, \tilde{P}, r)$. Since $\|\tilde{r} - \frac{4}{\varepsilon}r\|_\infty \leq 1$, we can apply Lemma 1 (with $\hat{r} = \tilde{r}$, $\theta_1 = \theta_2 = \frac{4}{\varepsilon}$ and $\gamma_1 = -\gamma_2 = -1$) to conclude that $\tilde{s}$ is a (uniformly) absolutely $\frac{\varepsilon}{2}$-optimal pair for $\hat{\mathcal{G}}$. Now we apply Lemma 2 and conclude that $\tilde{s}$ is (uniformly) $(\frac{\varepsilon}{2} + 2\delta(\mathcal{G}, \varepsilon'))$-optimal for $\mathcal{G}$.

Note that the above technique yields an approximation algorithm with polynomial running-time only for $k = O(1)$, even if the pseudo-polynomial algorithm $\mathbb{A}$ works for *arbitrary* k.

2.3 Relative Approximation

Let $\mathcal{G} = (G, P, r)$ be a BWR-game on G with non-negative integral rewards, that is, $r^- = 0$ and $\min_{e:r_e>0} r_e \geq 1$. The algorithm is given as Algorithm 1. The main idea is to truncate the rewards, scaled by a certain factor of $1/K$, and use the pseudo-polynomial algorithm on the truncated game $\hat{\mathcal{G}}$. If the value $\mu_w(\hat{\mathcal{G}})$ in the truncated game from the starting node w is large enough (step 4), then we get a good relative approximation of the original value and we

---

Data: a BWR-game $\mathcal{G} = (G = (V, E), P, r)$, a starting vertex $w \in V$, and an accuracy ε

Result: an ε-optimal pair $(\tilde{s}_W, \tilde{s}_B)$ for the game $(\mathcal{G}, w)$

1 **if** $r^+(\mathcal{G}) \leq 2$ **then**

2 $\quad$ $\hat{\mathcal{G}} \leftarrow (G, \tilde{P}, r)$ **return** $\mathbb{A}(\hat{\mathcal{G}}, v)$

3 **end**

4 $K \leftarrow r^+(\mathcal{G})/\theta(\mathcal{G})$ $\hat{r}_e \leftarrow \lfloor r_e/K \rfloor$ for $e \in E$ $\hat{\mathcal{G}} = (G, \tilde{P}, \hat{r})$ $(\tilde{s}_W, \tilde{s}_B) \leftarrow \mathbb{A}(\hat{\mathcal{G}}, w)$ **if** $\mu_w(\hat{\mathcal{G}}) \geq 3/\varepsilon$ **then**

5 $\quad$ **return** $(\tilde{s}_W, \tilde{s}_B)$

6 **end**

7 **else**

8 $\quad$ **for** *all* $e \in E$ **do**

9 $\quad\quad$ $\tilde{r}_e \leftarrow \min\left\{ r_e, \lceil \frac{r^+(\mathcal{G})}{2} \rceil \right\}$

10 $\quad$ **end**

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

12 **end**

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

---

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 9); we invoke Lemma 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 4 with $B = \lceil \log(1/\varepsilon') \rceil$, where we select $\varepsilon' = \frac{p_{\min}^{2k}}{2^{k+1} n^2 (k+2)^2 \theta}$ with $\theta = \theta(\mathcal{G}) := \frac{2(1+\varepsilon)(3+2\varepsilon)n}{\varepsilon p_{\min}^{2k+1}}$. Thus, we have $2\delta(\mathcal{G}, \varepsilon') \leq 2^{k+1} \varepsilon' n^2 (k+2)^2 p_{\min}^{-2k} \leq r^+(\mathcal{G})/\theta(\mathcal{G}) = K(\mathcal{G})$.

Lemma 6 *Let $\mathbb{A}$ be a pseudo-polynomial algorithm that solves any BWR-game $(\mathcal{G} = (G, P, r), w)$, from any given starting position w, in time $\tau(n, D, R)$. Then, for any $\varepsilon \in (0, 1)$, there is an algorithm that solves, in relatively ε-optimal strategies, any BWR-game $(\mathcal{G} = (G, P, r), w)$ from any given starting position w in time*

$$O\left(\left(\tau\left(n, \frac{2^{k+3} n^3 (k+2)^2 (1+\varepsilon)(3+2\varepsilon)}{\varepsilon p_{\min}^{4k+1}}, \frac{2(1+\varepsilon)(3+2\varepsilon)n}{\varepsilon p_{\min}^{2k+1}} \right) + \text{poly}(n) \right) \cdot \log(R) \right).$$

Proof The algorithm FPTAS-BWR$(\mathcal{G}, w, \varepsilon)$ is given as Algorithm 1. The bound on the running-time follows since, by step (9), 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 rewards in the truncated game $\hat{\mathcal{G}}$ are non-negative integers with a maximum value of $r^+(\hat{\mathcal{G}}) \leq \theta$, and the smallest common denominator of the transition probabilities is at most $\tilde{D} := \frac{2}{\varepsilon'}$. Thus the time taken by algorithm $\mathbb{A}$ for each recursive call is at most $\tau(n, \tilde{D}, \theta)$.

What remains to be done is to argue by induction (on $r^+(\mathcal{G})$) that the algorithm returns a pair $\tilde{s} = (\tilde{s}_W, \tilde{s}_B)$ of ε-optimal strategies. For the base case, we have either $r^+(\mathcal{G}) \leq 2$ or the value returned by the pseudo-polynomial $\mathbb{A}$ satisfies $\mu_w(\hat{\mathcal{G}}) \geq 3/\varepsilon$. In the former case, note that since $\|P - \tilde{P}\|_\infty \leq \varepsilon'$

and $r^+(\mathcal{G}) \leq 2$, Lemma 2 implies that the pair $\tilde{s} = (\tilde{s}_W, \tilde{s}_B)$ returned in step 2 is absolutely ε''-optimal, where $\varepsilon'' = 2\delta(\mathcal{G}, \varepsilon') < \frac{\varepsilon p_{\min}^{2k+1}}{n}$. Lemma 8 and the integrality of the non-negative rewards imply that, for any situation s, $\mu_w(\mathcal{G}(s)) \geq p_{\min}^{2k+1}/n$ if $\mu_w(\mathcal{G}(s)) > 0$. Thus, if $\mu_w(\mathcal{G}) > 0$, then $\varepsilon'' \leq \varepsilon\mu_w(\mathcal{G})$, and it follows that $(\tilde{s}_W, \tilde{s}_B)$ is relatively ε-optimal. On the other hand, if $\mu_w(\mathcal{G}) = 0$, then $\mu_w(\mathcal{G}(\tilde{s})) \leq \mu_w(\mathcal{G}) + \varepsilon'' < p_{\min}^{2k+1}/n$, implying that $\mu_w(\mathcal{G}(\tilde{s})) = 0$. Thus, we get a relative ε-approximation in both cases.

Suppose now that $\mathbb{A}$ determines that $\mu_w(\hat{\mathcal{G}}) \geq 3/\varepsilon$ in step 4, and hence the algorithm returns $(\tilde{s}_W, \tilde{s}_B)$. Note that $\frac{1}{K} \cdot r_e - 1 \leq \hat{r}_e \leq \frac{1}{K} \cdot r_e$ for all $e \in E$, and $\|P - \tilde{P}\|_\infty \leq \varepsilon'$. Hence, by Lemmas 1 and 2, we have

$$K\mu_w(\hat{\mathcal{G}}) - \delta(\mathcal{G}, \varepsilon') \leq \mu_w(\mathcal{G}) \leq K\mu_w(\hat{\mathcal{G}}) + K + \delta(\mathcal{G}, \varepsilon'), \tag{7}$$

and the pair $(\tilde{s}_W, \tilde{s}_B)$ returned in step 5 is absolutely $K + 2\delta(\mathcal{G}, \varepsilon') \leq 2K$-optimal for $\mathcal{G}$. (To see (7), let $\tilde{\mathcal{G}} := (G, \tilde{P}, r)$. Then by Lemma 2, we have

$$\mu_w(\tilde{\mathcal{G}}) - \delta(\mathcal{G}, \varepsilon') \leq \mu_w(\mathcal{G}) \leq \mu_w(\tilde{\mathcal{G}}) + \delta(\mathcal{G}, \varepsilon'). \tag{8}$$

Furthermore, as $\hat{\mathcal{G}}$ is obtained from $\tilde{\mathcal{G}}$ by scaling and truncating the local rewards, we have by Lemma 1 (applied with $\theta_1 = \theta_2 = \frac{1}{K}$, $\gamma_1 = -1$ and $\gamma_2 = 0$),

$$\frac{1}{K}\mu_w(\tilde{\mathcal{G}}) - 1 \leq \mu_w(\hat{\mathcal{G}}) \leq \frac{1}{K}\mu_w(\tilde{\mathcal{G}}). \tag{9}$$

Combining (8) and (9), we get (7).)

Then (7) implies that

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

and we are done.

On the other hand, if $\mu_w(\hat{\mathcal{G}}) < 3/\varepsilon$ then, by (7), $\mu_w(\mathcal{G}) < \frac{K(3+2\varepsilon)}{\varepsilon} = \frac{p_{\min}^{2k+1} r^+}{2(1+\varepsilon)n}$. By Lemma 3, applied with $t = K(3 + 2\varepsilon)/\varepsilon$, the game $\tilde{\mathcal{G}}$ defined in step 11 satisfies $\mu_w(\mathcal{G}) = \mu_w(\tilde{\mathcal{G}})$, and any (relatively) ε-optimal strategy in $(\tilde{\mathcal{G}}, w)$ (in particular the one returned by induction in step 11) is also ε-optimal for $(\mathcal{G}, w)$.

Note that the running-time in the above lemma simplifies to $\text{poly}(n, 1/\varepsilon, 1/p_{\min}) \cdot \log R$ for $k = O(1)$.

2.4 Uniformly Relative Approximation for BW-Games

The FPTAS in Theorem 6 does not necessarily return a uniformly ε-optimal situation, even if the given pseudo-polynomial algorithm $\mathbb{A}$ provides a uniformly optimal solution. For BW-games, we can modify this FPTAS to return a uniformly ε-optimal situation. The algorithm is given as Algorithm 2. The main difference is that when we recurse on a game with reduced rewards (step 11), we also have to delete all positions that have large values $\mu_v(\tilde{\mathcal{G}})$ in the

---

Data: a BW-game $\mathcal{G} = (G = (V = V_B \cup V_W, E), r)$, and an accuracy ε
Result: a uniformly ε-optimal pair $(\tilde{s}_W, \tilde{s}_B)$ for $\mathcal{G}$

1 **if** $r^+(\mathcal{G}) \leq 2$ **then**
2 $\quad$ **return** $\mathbb{A}(\mathcal{G})$
3 **end**
4 $K \leftarrow \frac{\varepsilon' r^+(\mathcal{G})}{2(1+\varepsilon')^2 n}$ $\hat{r}_e \leftarrow \lfloor r_e/K \rfloor$ for $e \in E$ $\hat{\mathcal{G}} \leftarrow (G, \hat{r})$ $(\hat{s}_W, \hat{s}_B) \leftarrow \mathbb{A}(\hat{\mathcal{G}})$
$\quad U \leftarrow \{u \in V \mid \mu_u(\hat{\mathcal{G}}) \geq 1/\varepsilon'\}$ **if** $U = V$ **then**
5 $\quad$ **return** $(\tilde{s}_W, \tilde{s}_B) = (\hat{s}_W, \hat{s}_B)$
6 **end**
7 **else**
8 $\quad$ $\tilde{G} \leftarrow G[V \setminus U]$ **for** *all* $e \in E(\tilde{G})$ **do**
9 $\quad\quad$ $\tilde{r}_e \leftarrow \min\left\{r_e, \lceil \frac{r^+(\mathcal{G})}{2} \rceil \right\}$
10 $\quad$ **end**
11 $\quad$ $\tilde{\mathcal{G}} \leftarrow (\tilde{G}, \tilde{r})$ $(\tilde{s}_W, \tilde{s}_B) \leftarrow$ FPTAS-BW$(\tilde{\mathcal{G}}, \varepsilon)$ $\tilde{s}(w) \leftarrow \hat{s}(w)$ for all $w \in U$ **return**
$\quad\quad \tilde{s} = (\tilde{s}_W, \tilde{s}_B)$
12 **end**

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

truncated game. This is similar to the approach used to decompose a BW-game into ergodic classes [21]. However, the main technical difficulty is that, with approximate equilibria, WHITE or BLACK might still have some incentive to move to a lower- or higher-value class, respectively, since the values obtained are just approximations of the optimal values. We show that such a move will not be much profitable for either WHITE nor for BLACK. Recall that we assume that the rewards are non-negative integers.

Lemma 7 *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 relatively ε-optimal strategies, any BW-game $\mathcal{G}$, in time $O\left(\left(\tau\left(n, \frac{2(1+\varepsilon')^2 n}{\varepsilon'}\right) + \text{poly}(n)\right) \cdot h\right)$, where $h = \lceil \log R \rceil + 1$, and $\varepsilon' = \frac{\ln(1+\varepsilon)}{3h} \approx \frac{\varepsilon}{3h}$.*

Proof The algorithm FPTAS-BW$(\mathcal{G}, \varepsilon)$ is given as Algorithm 2. The bound on the running-time is obvious: in step (9), 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 rewards in the truncated game $\hat{\mathcal{G}}$ are integral with a maximum value of $r^+(\hat{\mathcal{G}}) \leq \frac{r^+(\mathcal{G})}{K} \leq \frac{2(1+\varepsilon')^2 n}{\varepsilon'}$. Thus, the time that algorithm $\mathbb{A}$ needs in each recursive call is bounded from above by $\tau\left(n, \frac{2(1+\varepsilon')^2 n}{\varepsilon'}\right)$.

So it remains to argue (by induction) that the algorithm returns a pair $(\tilde{s}_W, \tilde{s}_B)$ of (relatively) uniformly ε-optimal strategies. Let us index the different recursive calls of the algorithm by $i = 1, 2, \ldots, h' \leq h$ and denote by $\mathcal{G}^{(i)} = (G^{(i)} = (V^{(i)}, E^{(i)}), r^{(i)})$ 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, 4, 5, or 11. Similarly, we denote by $V^{(i)} = V_W^{(i)} \cup V_B^{(i)}$, $U^{(i)}$, $r^{(i)}$, $K^{(i)}$ $\hat{r}^{(i)}$, $\hat{\mathcal{G}}^{(i)}$, $\tilde{\mathcal{G}}^{(i)}$ the instantiations of V, V_B, V_W, U, r, $\hat{r}$, $\hat{\mathcal{G}}$, K, $\tilde{\mathcal{G}}$, respectively, 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 of

positions, a game $\mathcal{G}$, and a situation s, we denote by $\mathcal{G}[U] = (G[U], r)$ and $s[U]$, respectively, the game and situation induced on U.

Claim (i) There does not exist an edge $(v, u) \in E$ such that $v \in V_B^{(i)} \cap U^{(i)}$ and $u \in V^{(i)} \setminus U^{(i)}$.

(ii) For all $v \in V_W^{(i)} \cap U^{(i)}$, there exists a $u \in U^{(i)}$ with $(v, u) \in E$.

(i') There does not exist an edge $(v, u) \in E$ such that $v \in V_W^{(i)} \setminus U^{(i)}$ and $u \in U^{(i)}$.

(ii') For all black positions $v \in V_B^{(i)} \setminus U^{(i)}$, there exists a $u \in V^{(i)} \setminus U^{(i)}$ such that $(v, u) \in E$.

(iii) Let $\hat{s}^{(i)} = (\hat{s}_W^{(i)}, \hat{s}_B^{(i)})$ be the situation returned in step 4. Then, for all $v \in U^{(i)}$, we have $\hat{s}^{(i)}(v) \in U^{(i)}$, and, for all $v \in V^{(i)} \setminus U^{(i)}$, we have $\hat{s}^{(i)}(v) \in V^{(i)} \setminus U^{(i)}$.

Proof By the optimality conditions in $\hat{\mathcal{G}}^{(i)}$ (see, e.g., [21]), we have

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

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

(12) and (12), 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 for all $v \in V^{(i)}$, $\mu_v(\hat{\mathcal{G}}^{(i)}) = \mu_{\hat{s}^{(i)}(v)}(\hat{\mathcal{G}}^{(i)})$, which in turn implies (iii).

Note that Claim 12 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 (and similarly for a strategy s_B) and a subset $V' \subseteq V$, we write $S_W(V') = \{s_W(u) \mid u \in V'\}$. The following two claims state respectively that the values of the positions in $U^{(i)}$ are well-approximated by the pseudo-polynomial algorithm and that these values are sufficiently larger than those in the residual set $V^{(i)} \setminus U^{(i)}$.

Claim For $i = 1, \ldots, h'$, let $\hat{s}^{(i)}$ be the situation returned by the pseudo-polynomial algorithm on the game $\hat{\mathcal{G}}^{(i)}$ in step 4. Then, for any $w \in U^{(i)}$, we have

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

$$\min_{s_B \in S_B^{(i)}} \mu_w(\mathcal{G}^{(i)}(\hat{s}_W^{(i)}, s_B)) \geq (1 - \varepsilon') \mu_w(\mathcal{G}^{(i)}).$$

Proof This follows from Lemma 1 by the uniform optimality of $\hat{s}^{(i)}$ in $\hat{\mathcal{G}}^{(i)}$ and the fact that $\mu_w(\hat{\mathcal{G}}^{(i)}) \geq 1/\varepsilon'$ for every $w \in U^{(i)}$.

Claim For all $u \in U^{(i)}$ and $v \in V^{(i)} \setminus U^{(i)}$, we have $(1 + \varepsilon') \mu_u(\mathcal{G}^{(i)}) > \mu_v(\mathcal{G}^{(i)})$.

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

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

We observe that the strategy $\tilde{s}^{(i)}$, returned by the ith call to the algorithm, is determined as follows (c.f. steps 11 and 11): 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 $w \in V^{(i)} \setminus U^{(i)}$, $\tilde{s}^{(i)}(w)$ 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 (relative) approximation of the value in the original game $\mathcal{G}^{(i)}$.

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

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

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 by the pseudo-polynomial algorithm step 4. Then, by Claim 12(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,

$$\mu_u(\mathcal{G}^{(i)}) = \mu_u(\mathcal{G}^{(i)}(s_W^*, s_B^*)) \le \mu_u(\mathcal{G}^{(i)}(s_W^*, \bar{s}_B))$$
$$= \mu_u(\bar{\mathcal{G}}^{(i)}(s_W^*, \bar{s}_B)) \le \mu_u(\bar{\mathcal{G}}^{(i)}(\bar{s}_W, \bar{s}_B)) = \mu_u(\bar{\mathcal{G}}^{(i)}).$$

For similar reasons, $\mu_u(\mathcal{G}^{(i)}) \ge \mu_u(\bar{\mathcal{G}}^{(i)})$, 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 (by Lemma 1) $\mu_u(\mathcal{G}^{(i)}) = \mu_{u'}(\mathcal{G}^{(i)}) \ge K^{(i)}\mu_{u'}(\hat{\mathcal{G}}^{(i)}) \ge \frac{K^{(i)}}{\varepsilon'}$. Moreover, the optimality of $(\hat{s}_W, \hat{s}_B)$ in $\hat{\mathcal{G}}^{(i)}$ and the fact that $\frac{1}{K^{(i)}}r^{(i)} - 1 \le \hat{r}^{(i)} \le \frac{1}{K^{(i)}}r^{(i)}$ imply by Lemma 1 that

$$\forall s_W \in S_W^{(i)}: \ \mu_u(\mathcal{G}^{(i)}(\hat{s}_W, \hat{s}_B)) \ge K^{(i)}\mu_u(\hat{\mathcal{G}}^{(i)}(\hat{s}_W, \hat{s}_B)) \ge K^{(i)}\mu_u(\hat{\mathcal{G}}^{(i)}(s_W, \hat{s}_B))$$
$$\ge \mu_u(\mathcal{G}^{(i)}(s_W, \hat{s}_B)) - K^{(i)}$$
$$\ge \mu_u(\mathcal{G}^{(i)}(s_W, \hat{s}_B)) - \varepsilon'\mu_u(\mathcal{G}^{(i)})$$

and

$$\forall s_B \in S_B^{(i)}: \ \mu_u(\mathcal{G}^{(i)}(\hat{s}_W, \hat{s}_B)) \le K^{(i)}\mu_u(\hat{\mathcal{G}}^{(i)}(\hat{s}_W, \hat{s}_B)) + K^{(i)} \le K^{(i)}\mu_u(\hat{\mathcal{G}}^{(i)}(\hat{s}_W, s_B)) + K^{(i)}$$
$$\le \mu_u(\mathcal{G}^{(i)}(\hat{s}_W, s_B)) + K^{(i)} \le \mu_u(\mathcal{G}^{(i)}(\hat{s}_W, s_B)) + \varepsilon'\mu_u(\mathcal{G}^{(i)}).$$

In particular,

$$\mu_u(\mathcal{G}^{(i)}) = \mu_u(\mathcal{G}^{(i)}(s_W^*, s_B^*)) \ge \mu_u(\mathcal{G}^{(i)}(\hat{s}_W, s_B^*)) \ge \mu_u(\mathcal{G}^{(i)}(\hat{s}_W, \hat{s}_B)) - \varepsilon'\mu_u(\mathcal{G}^{(i)})$$
$$\ge \mu_u(\mathcal{G}^{(i)}(\bar{s}_W, \hat{s}_B)) - 2\varepsilon'\mu_u(\mathcal{G}^{(i)}) = \mu_u(\bar{\mathcal{G}}^{(i)}(\bar{s}_W, \hat{s}_B)) - 2\varepsilon'\mu_u(\mathcal{G}^{(i)})$$
$$\ge \mu_u(\bar{\mathcal{G}}^{(i)}(\bar{s}_W, \bar{s}_B)) - 2\varepsilon'\mu_u(\mathcal{G}^{(i)}) = \mu_u(\bar{\mathcal{G}}^{(i)}) - 2\varepsilon'\mu_u(\mathcal{G}^{(i)}),$$

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 from Claim 12 (since $(\bar{s}_W, \hat{s}_B)(v) \in V^{(i)} \setminus U^{(i)}$). It follows that $\mu_u(\mathcal{G}^{(i)}) \ge \frac{1}{1+2\varepsilon'}\mu_u(\bar{\mathcal{G}}^{(i)})$.

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

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

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

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

Proof The proof is by induction on $i = h', h'-1, \ldots, 1$. For $i = h'$, the statement follows directly from Claim 12 since $U^{(h')} = V^{(h')}$. So suppose that $i < h'$.

By induction, $\bar{s}^{(i)} = (\bar{s}_W^{(i)}, \bar{s}_B^{(i)}) := (\tilde{s}_W^{(i)}, \tilde{s}_B^{(i)})[V^{(i)} \setminus U^{(i)}]$ 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 step 9. Thus, Lemma 3 yields that $\mu_v(\bar{\mathcal{G}}^{(i)}) = \mu_v(\tilde{\mathcal{G}}^{(i)})$, and hence,

$$\max_{s_W' \in S_W^{(i+1)}} \mu_v(\bar{\mathcal{G}}^{(i)}(s_W', \bar{s}_B^{(i)})) \leq (1 + \varepsilon_{i+1})\mu_v(\bar{\mathcal{G}}^{(i)}) \tag{13}$$

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

Proof of (11)*:* Consider an arbitrary strategy $s_W \in S_W^{(i)}$ for WHITE. Suppose first that $w \in U^{(i)}$. Note that, by Claim 12(iii), $\tilde{s}_B^{(i)}(u) \in U^{(i)}$ for all $u \in V_B \cap U^{(i)}$. If also $s_W(u) \in U^{(i)}$ for all $u \in V_W \cap U^{(i)}$, such that u is reachable from w in the graph $G(s_W, \tilde{s}_B^{(i)})$, then Claim 12 implies $\mu_w(\mathcal{G}^{(i)}(s_W, \tilde{s}_B^{(i)})) \leq (1 + \varepsilon')\mu_w(\mathcal{G}^{(i)}) \leq (1 + \varepsilon_i)\mu_w(\mathcal{G}^{(i)})$.

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

Note that $\tilde{s}_B^{(i)}(v') \in V^{(i)} \setminus U^{(i)}$ for all $v' \in V_B^{(i)} \setminus U^{(i)}$, and by Claim 12($i'$), $S_W^{(i+1)}$ is the restriction of $S_W^{(i)}$ to $V^{(i)} \setminus U^{(i)}$. Thus, we get the following series of inequalities:

$$\mu_w(\mathcal{G}^{(i)}(s_W, \tilde{s}_B^{(i)})) = \mu_v(\mathcal{G}^{(i)}(s_W, \tilde{s}_B^{(i)})) \leq (1 + \varepsilon_{i+1})\mu_v(\mathcal{G}^{(i)}[V^{(i)} \setminus U^{(i)}]) \tag{15}$$

$$\leq (1 + \varepsilon_{i+1})(1 + 2\varepsilon')\mu_v(\mathcal{G}^{(i)}) \tag{16}$$

$$< (1 + \varepsilon_{i+1})(1 + 2\varepsilon')(1 + \varepsilon')\mu_w(\mathcal{G}^{(i)}) \leq (1 + \varepsilon_i)\mu_w(\mathcal{G}^{(i)}).$$

The equality holds since v is reachable from w in the graph $G(s_W, \tilde{s}_B^{(i)})$; the first inequality holds by (13); the second inequality holds because of (10); the third one follows from Claim 12; the fourth inequality holds since $(1 + \varepsilon_{i+1})(1 + 2\varepsilon')(1 + \varepsilon') \leq (1 + \varepsilon_i)$.

If $w \in V^{(i)} \setminus U^{(i)}$, then a similar argument as in (15) and (16) shows that $\mu_w(\mathcal{G}^{(i)}(s_W, \tilde{s}_B^{(i)})) \leq (1 + \varepsilon_{i+1})(1 + 2\varepsilon')\mu_w(\mathcal{G}^{(i)}) \leq (1 + \varepsilon_i)\mu_w(\mathcal{G}^{(i)})$. Thus, (11) follows.

Proof of (12)*:* Consider an arbitrary strategy $s_B \in S_B^{(i)}$ for BLACK. If $w \in U^{(i)}$, then we have $\mu_w(\mathcal{G}^{(i)}(\tilde{s}_W^{(i)}, s_B)) \geq (1 - \varepsilon')\mu_w(\mathcal{G}^{(i)}) \geq (1 - \varepsilon_i)\mu_w(\mathcal{G}^{(i)})$ from Claims 12(i), 12(iii), 12 and $\varepsilon_i \geq \varepsilon'$.

Suppose now that $w \in V^{(i)} \setminus U^{(i)}$. If $s_B(v) \in V^{(i)} \setminus U^{(i)}$ for all $v \in V_B^{(i)} \setminus U^{(i)}$, 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 BLACK has no escape from $U^{(i)}$ in this case (by Claim 12(i)), Claims 12 and 12 yield

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

where the last inequality follows from the fact that, for all $i = 1, \ldots, h' - 1$, $1 + \varepsilon_i \geq (1 + 2\varepsilon')(1 + \varepsilon')^2 \geq (1 + \varepsilon')^3$, and hence, $1 - \varepsilon_i \leq 2 - (1 + \varepsilon')^3 \leq (1 - \varepsilon')^2$.

Finally, to finish the proof of Lemma 7, we set the ε_i's and ε' such that $\varepsilon_1 = \left((1 + 2\varepsilon')(1 + \varepsilon')\right)^{h'-1}(1 + \varepsilon') - 1 \leq \varepsilon$. Note that our choice of $\varepsilon' = \frac{\ln(1+\varepsilon)}{3h}$ satisfies this as

$$\left((1 + 2\varepsilon')(1 + \varepsilon')\right)^{h'-1}(1 + \varepsilon') = \frac{(1 + 2\varepsilon')^{h'}(1 + \varepsilon')^{h'}}{(1 + 2\varepsilon')}$$
$$\leq \frac{e^{3h'\varepsilon'}}{(1 + 2\varepsilon')} \leq \frac{(1 + \varepsilon)}{(1 + 2\varepsilon')} \leq (1 + \varepsilon).$$

3 Concluding Remarks

In this paper, we have shown that computing the game values of classes of stochastic mean payoff games with perfect information and a constant number of random positions admits approximation schemes, provided that the class of games at hand can be solved in pseudo-polynomial time.

To conclude this paper, let us raise a number of open questions:

1. First, in the conference version of this paper [2], we claimed that, up to some technical requirements, a pseudo-polynomial algorithm for a class of stochastic mean payoff games implies that this class has polynomial smoothed complexity. (Smoothed analysis is a paradigm to analyze algorithms with poor worst-case and good practical performance. Since its invention, it has been applied to a variety of algorithms and problems to explain their performance or complexity, respectively [31, 38].)
   However, the proof of this result is flawed. In particular, the proof of a lemma that is not contained in the proceedings version, but only in the accompanying technical report (Oberwolfach Preprints, OWP 2010-22, Lemma 4.3)) is flawed. The reason for this is relatively simple: If we are just looking for an optimal solution, then we can show that the second-best solution is significantly worse than the best solution. For two-player games, where one player maximizes and the other player minimizes, we have an optimization problem for either player, given an optimal strategy of the other player. However, the optimal strategy of the other player depends on the random rewards of the edges. Thus, the two strategies are dependent. As a consequence, we cannot use the full randomness of the rewards to use

an isolation lemma to compare the best and second-best response to the optimal strategy of the other player.

Therefore, we raise the open problem to prove that stochastic mean payoff games have polynomial smoothed complexity.

2. In Section 2.3 we gave an approximation scheme that relatively approximates the value of a BWR-game from any starting position. If we apply this algorithm from different positions, we are likely to get two different relatively ε-optimal strategies. In Section 2.4 we have shown that a modification of the algorithm in section 2.3 yields a *uniformly* relatively ε-optimal strategies when there are no random positions. It remains an interesting question whether this can be extended to BWR-games with a constant number of random positions.

3. Is it true that pseudo-polynomial solvability of a class of stochastic mean payoff games implies polynomial smoothed complexity? In particular, do mean payoff games have polynomial smoothed complexity?

4. Related to Question 3: is it possible to prove an isolation lemma for (classes of) stochastic mean payoff games? We believe that this is not possible and that different techniques are required to prove smoothed polynomial complexity of these games.

5. While stochastic mean payoff games include parity games as a special case, the probabilistic model that we used here does not make sense for parity games. However, parity games can be solved in quasi-polynomial time [8]. One wonders if they also have polynomial smoothed complexity under a reasonable probabilistic model.

6. Finally, let us remark that removing the assumption that k is constant in the above results remains a challenging open problem that seems to require totally new ideas. Another interesting question is whether stochastic mean payoff games with perfect information can be solved in *parameterized* pseudo-polynomial time with the number k of stochastic positions as the parameter?

References

1. Daniel Andersson and Peter Bro Miltersen. The complexity of solving stochastic games on graphs. In *20th Int. Symp. on Algorithms and Computation (ISAAC)*, volume 5878 of *Lecture Notes in Computer Science*, pages 112–121. Springer, 2009.
2. Endre Boros, Khaled Elbassioni, Mahmoud Fouz Vladimir Gurvich, Kazuhisa Makino, and Bodo Manthey. Stochastic mean payoff games: Smoothed analysis and approximation schemes. In *Proc. of the 38th Int. Colloquium on Automata, Languages and Programming (ICALP), Part I*, volume 6755 of *Lecture Notes in Computer Science*, pages 147–158. Springer, 2011.
3. Endre Boros, Khaled Elbassioni, Vladimir Gurvich, and Kazuhisa Makino. Every stochastic game with perfect information admits a canonical form. RRR-09-2009, RUT-COR, Rutgers University, 2009.
4. Endre Boros, Khaled Elbassioni, Vladimir Gurvich, and Kazuhisa Makino. A convex programming-based algorithm for mean payoff stochastic games with perfect information. To appear *in Optimization Letters*, 2017.
5. Endre Boros, Khaled M. Elbassioni, Vladimir Gurvich, and Kazuhisa Makino. A pumping algorithm for ergodic stochastic mean payoff games with perfect information. In

Proc. of the 14th Int. Conf. on Integer Programming and Combinatorial Optimization (IPCO), volume 6080 of *Lecture Notes in Computer Science*, pages 341–354. Springer, 2010.

6. Endre Boros, Khaled M. Elbassioni, Vladimir Gurvich, and Kazuhisa Makino. A pseudo-polynomial algorithm for mean payoff stochastic games with perfect information and a few random positions. In Fedor V. Fomin, Rusins Freivalds, Marta Z. Kwiatkowska, and David Peleg, editors, *Proc. 40th Int. Colloquium on Automata, Languages and Programming, Part I*, volume 7965 of *Lecture Notes in Computer Science*, pages 220–231. Springer, 2013.

7. Endre Boros and Vladimir Gurvich. Why chess and backgammon can be solved in pure positional uniformly optimal strategies? RRR-21-2009, RUTCOR, Rutgers University, 2009.

8. Cristian S. Calude, Sanjay Jain, Bakhadyr Khoussainov, Wei Li, and Frank Stephan. Deciding parity games in quasipolynomial time. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017*, pages 252–263, 2017.

9. Krishnendu Chatterjee, Luca de Alfaro, and Thomas A. Henzinger. Termination criteria for solving concurrent safety and reachability games. In *Proc. of the 20th Ann. ACM-SIAM Symp. on Discrete Algorithms (SODA)*, pages 197–206, 2009.

10. Krishnendu Chatterjee and Rasmus Ibsen-Jensen. The complexity of ergodic mean-payoff games. In Javier Esparza, Pierre Fraigniaud, Thore Husfeldt, and Elias Koutsoupias, editors, *Proc. of the 41st Int. Colloquium on Automata, Languages and Programming, Part II*, volume 8572 of *Lecture Notes in Computer Science*, pages 122–133. Springer, 2014.

11. Xi Chen, Xiaotie Deng, and Shang-Hua Teng. Settling the complexity of computing two-player nash equilibria. *Journal of the ACM*, 56(3), 2009.

12. Grace E. Cho and Carl D. Meyer. Markov chain sensitivity measured by mean first passage times. *Linear Algebra and its Applications*, 316(1–3):21–28, 2000.

13. Anne Condon. The complexity of stochastic games. *Information and Computation*, 96(2):203–224, 1992.

14. Anne Condon. On algorithms for simple stochastic games. In Jin-Yi Cai, editor, *Advances in Computational Complexity Theory*, volume 13 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 51–73, 1993.

15. D. Dai and R. Ge. Another sub-exponential algorithm for the simple stochastic game. *Algorithmica*, 61:1092–1104, 2011.

16. Andrzej Ehrenfeucht and Jan Mycielski. Positional games over a graph. *Notices of the American Mathematical Society*, 20:A–334, 1973.

17. Andrzej Ehrenfeucht and Jan Mycielski. Positional strategies for mean payoff games. *International Journal of Game Theory*, 8:109–113, 1979.

18. Raffaella Gentilini. A note on the approximation of mean-payoff games. *Inf. Process. Lett.*, 114(7):382–386, 2014.

19. Dean Gillette. Stochastic games with zero stop probabilities. In M. Dresher, A. W. Tucker, and P. Wolfe, editors, *Contributions to the Theory of Games, Volume 3*, volume 39 of *Annals of Mathematics Studies*, pages 179–187. Princeton University Press, 1957.

20. Hugo Gimbert and Florian Horn. Simple stochastic games with few random vertices are easy to solve. In *Proc. 11th Int. Conf. Foundations of Software Science and Computational Structures (FoSSaCS)*, volume 4962 of *Lecture Notes in Computer Science*, pages 5–19. Springer, 2008.

21. Vladimir Gurvich, Alexander V. Karzanov, and Leonid Khachiyan. Cyclic games and an algorithm to find minimax cycle means in directed graphs. *USSR Computational Mathematics and Mathematical Physics*, 28:85–91, 1988.

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

23. Rasmus Ibsen-Jensen and Peter Bro Miltersen. Solving simple stochastic games with few coin toss positions. In Leah Epstein and Paolo Ferragina, editors, *Proc. 20th Ann. European Symposium on Algorithms (ESA)*, volume 7501 of *Lecture Notes in Computer Science*, pages 636–647. Springer, 2012.

24. Jr. John F. Nash. Equilibrium points in n-person games. In *Proceedings of the National Academy of Sciences*, volume 36, pages 48–49, 1950.
25. Jr. John F. Nash. Non-cooperative games. *Annals of Mathematics*, 54(1):286–295, 1951.
26. Marcin Jurdziński. *Games for Verification: Algorithmic Issues*. PhD thesis, University of Aarhus, BRICS, 2000.
27. Richard M. Karp. A characterization of the minimum cycle mean in a digraph. *Discrete Mathematics*, 23:309–311, 1978.
28. Alexander V. Karzanov and Vasilij N. Lebedev. Cyclical games with prohibition. *Mathematical Programming*, 60:277–293, 1993.
29. Thomas M. Liggett and Steven A. Lippman. Stochastic games with perfect information and time-average payoff. *SIAM Review*, 4:604–607, 1969.
30. Michael L. Littman. *Algorithms for Sequential Decision Making*. PhD thesis, Brown University, Department of Computer Science, 1996.
31. Bodo Manthey and Heiko Röglin. Smoothed analysis: Analysis of algorithms beyond worst case. *it – Information Technology*, 53(6):280–286, 2011.
32. H. Mine and S. Osaki. *Markovian decision process*. Elsevier, 1970.
33. Hervé Moulin. Extension of two person zero sum games. *Journal of Mathematical Analysis and Application*, 5(2):490–507, 1976.
34. Hervé Moulin. Prolongement des jeux à deux joueurs de somme nulle. *Bull. Soc. Math. France, Memoire*, 45, 1976.
35. Nicolai N. Pisaruk. Mean cost cyclical games. *Mathematics of Operations Research*, 24(4):817–828, 1999.
36. Aaron Roth, Maria-Florina Balcan, Adam Kalai, and Yishay Mansour. On the equilibria of alternating move games. In *Proc. of the 21st Ann. ACM-SIAM Symp. on Discrete Algorithms (SODA)*, pages 805–816. SIAM, 2010.
37. Sven Schewe. From parity and payoff games to linear programming. In *Proc. of the 34th Int. Symp. on Mathematical Foundations of Computer Science (MFCS)*, volume 5734 of *Lecture Notes in Computer Science*, pages 675–686. Springer, 2009.
38. Daniel A. Spielman and Shang-Hua Teng. Smoothed analysis: An attempt to explain the behavior of algorithms in practice. *Communications of the ACM*, 52(10):76–84, 2009.
39. Sergei Vorobyov. Cyclic games and linear programming. *Discrete Applied Mathematics*, 156(11):2195–2231, 2008.
40. Uri Zwick and Mike Paterson. The complexity of mean payoff games on graphs. *Theoretical Computer Science*, 158(1-2):343–359, 1996.

A Lemmas About Markov Chains

For a situation s, let $d_{G(s)}(u, v)$ be 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 (i.e., one with $p_e(s) = 1$) to 0 and of every stochastic arc (i.e., one with $p_e(s) \in (0, 1)$) to 1. Let

$$\lambda = \lambda(\mathcal{G}) = \max\{d_{G(s)}(v, u) \mid v, u \in V, \, d_{G(s)}(v, u) \text{ is finite, and } s \text{ is a situation}\}$$

be the *stochastic diameter* of $\mathcal{G}$. Clearly, $\lambda(\mathcal{G}) \leq k(\mathcal{G})$. Some of our bounds will be given in terms of λ instead of k, which implies 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 8 *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 at least $p_{\min}^{2\lambda+1}/n$.*

Proof Let π and ρ denote the limiting vertex- and arc-distribution, respectively. Let C be any absorbing class 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 every $v' \in C$, we have $\pi_{v'} \geq p_{\min}^{2\lambda}/n$. Denote by $\pi_C = \sum_{v \in C} \pi_v$ the total limiting probability of C. Note that π_C is equal to the probability that we reach some vertex $v \in C$ starting from u. Since there is a simple path in G from u to C with at most λ stochastic vertices, this probability is at least $p_{\min}^{\lambda}$. Furthermore, there exists a vertex $v \in C$ with $\pi_v \geq \pi_C/|C| \geq p_{\min}^{\lambda}/n$. Now for any $v' \in C$, there exists again a simple path in G from v to v' with at most λ stochastic positions, so the probability that we reach v' starting from v is at least $p_{\min}^{\lambda}$. It follows that $\pi_{v'} \geq p_{\min}^{2\lambda}/n$.

Now for ρ, note that $\rho_{(u,v)} \geq \pi_u p_{\min}$, 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} denote 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 relates these values to the sensitivity of the limiting probabilities of a Markov chain.

Lemma 9 (Cho, Meyer [12]) *Let $\varepsilon > 0$. Let $\mathcal{M} = (G = (V,E), P)$ be an irreducible Markov chain. For any transition probabilities $\tilde{P}$ with $\|\tilde{P} - P\|_\infty \leq \varepsilon$ such that the corresponding Markov chain $\tilde{\mathcal{M}}$ is also irreducible, we have $\|\tilde{\pi} - \pi\|_\infty \leq \frac{1}{2}\varepsilon \cdot \max_v \frac{\max_{u \neq v} m_{uv}}{m_{vv}}$, where m_{vu} are the mean values defined with respect to $\mathcal{M}$.*

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 $\pi = (\pi_v : v \in V)$ is the limiting distribution when u is the starting position. We will write μ_u when $\mathcal{M}$ is clear from the context.

Lemma 10 *Let $\mathcal{M} = (G = (V,E), P, r)$ be a weighted Markov chain with arc weights in $[r^-, r^+]$, and let $\varepsilon \leq \frac{1}{2}p_{\min} = \frac{1}{2}p_{\min}(\mathcal{M})$ be a positive constant. Let $\tilde{\mathcal{M}} = (G = (V,E), \tilde{P}, r)$ be a weighted Markov chain with transition probabilities $\tilde{P}$ such that $\|\tilde{P} - P\|_\infty \leq \varepsilon$ and $\tilde{p}_{uv} = 0$ if $p_{uv} = 0$. Then, for any $u \in V$, we have $|\mu_u(\tilde{\mathcal{M}}) - \mu_u(\mathcal{M})| \leq \delta(\mathcal{M}, \varepsilon)$, where δ is defined as in (6):*

$$\delta(\mathcal{M}, \varepsilon) := \left(\frac{\varepsilon n^2}{2} \left(\frac{p_{\min}}{2} \right)^{-k} \left(\varepsilon n k(k+1) \left(\frac{p_{\min}}{2} \right)^{-k} + 3k+1 \right) + \varepsilon n \right) r^*,$$

where $n = |V|$, $p_{\min} = p_{\min}(\mathcal{M})$, $k = k(\mathcal{M})$, and $r^* = r^*(\mathcal{M}) := \max\{|r^+(\mathcal{M})|, |r^-(\mathcal{M})|\}$

.

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 to $\mathcal{M}$ and $\tilde{\mathcal{M}}$, respectively, 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 9, we have $\|\pi^{|i} - \tilde{\pi}^{|i}\|_\infty \leq \frac{1}{2}\varepsilon \cdot \max_{v \in C_i} \max_{\substack{u \in C_i \\ u \neq v}} \frac{m_{uv}}{m_{vv}}$.

Claim For any $u, v \in C_i$, we have $m_{uv} \leq \frac{(\lambda+1)|C_i|}{p_{\min}^{\lambda}}$.

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

$$m_{uv} = \sum_{w \neq v} p_{uw}(1 + m_{wv}) + p_{uv}. \tag{17}$$

Let $h = \max\{d_G(u,v) \mid u \in C_i\}$. For $j = 0, 1, \ldots, h$, let $X_j = \max\{m_{uv} \mid u \in C_i, d_G(u,v) = j\}$. Let ℓ be in $\operatorname{argmax}\{X_j \mid j \in \{1, \ldots, h\}\}$. Then $X_0 \leq |C_i|$ and, for $j = 1, \ldots, h$, (17) implies that

$$X_j \leq |C_i| + p_{\min}X_{j-1} + (1 - p_{\min})X_\ell. \tag{18}$$

Indeed, for a vertex for $u \in V$ such that $d_G(u,v) = j$, there is a path Q from u to v with j stochastic arcs. Let u' be the vertex closest to u on Q such that $d_G(u',v) = j-1$, and let u'' be the vertex on Q preceding u'. Then u'' is stochastic, and hence by (17)

$$m_{u''v} \leq p_{u''u'}(1 + X_{j-1}) + \sum_{w \neq u'} p_{u''w}(1 + X_\ell)$$
$$= p_{u''u'}(1 + X_{j-1}) + (1 - p_{u''u'})(1 + X_\ell)$$
$$\leq p_{\min}(1 + X_{j-1}) + (1 - p_{\min})(1 + X_\ell),$$

using the fact that $X_j \leq X_\ell$ for all j and $p_{u''u'} \geq p_{\min}$. Finally, $m_{uv} \leq |C_i| - 1 + m_{u''v}$ implies (18).

Applying (18) for $j = 1, \ldots, \ell$ yields

$$X_\ell \leq |C_i| \cdot \frac{1 - p_{\min}^{\ell+1}}{1 - p_{\min}} + X_\ell(1 - p_{\min}^\ell).$$

This implies that $X_\ell \leq |C_i|\frac{1-p_{\min}^{\ell+1}}{1-p_{\min}}p_{\min}^{-\ell} \leq |C_i|(\lambda + 1)p_{\min}^{-\lambda}$.

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

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

Proof Without loss of generality we assume that $u_0 \notin C_i$. For a transient vertex v (i.e., one for which $\pi_v = 0$), let y_v and $\tilde{y}_v$ be the absorption probability into class C_i in $\mathcal{M}$ and $\tilde{\mathcal{M}}$, respectively. In particular $y_{u_0} = \pi_{C_i}$. Let $p_{vC_i} = \sum_{u \in C_i} p_{vu}$ and $\tilde{p}_{vC_i} = \sum_{u \in C_i} \tilde{p}_{vu}$. Then we have

$$y_v = \sum_{u \notin C_i} p_{vu}y_u + p_{vC_i}. \tag{19}$$

Similarly,

$$\tilde{y}_v = \sum_{u \notin C_i} \tilde{p}_{vu}\tilde{y}_u + \tilde{p}_{vC_i} = \sum_{u \notin C_i} p_{vu}\tilde{y}_u + \sum_{u \notin C_i} (\tilde{p}_{vu} - p_{vu})\tilde{y}_u + \tilde{p}_{vC_i}. \tag{20}$$

Write $\Delta_v := |\tilde{y}_v - y_v|$. Subtracting (19) from (20) yields

$$\Delta_v \leq \sum_{u \notin C_i} p_{vu}\Delta_u + \sum_{u \notin C_i} |\tilde{p}_{vu} - p_{vu}|\tilde{y}_u + |\tilde{p}_{vC_i} - p_{vC_i}|$$
$$\leq \sum_{u \notin C_i} p_{vu}\Delta_u + (n - |C_i|)\varepsilon + |C_i|\varepsilon = \sum_{u \notin C_i} p_{vu}\Delta_u + \varepsilon n. \tag{21}$$

Let $h = \max\{d_G(u, C_i) \mid u \notin C_i \wedge d_G(u, C_i) < \infty\}$, where $d_G(u, C_i) = \min\{d_G(u,v) \mid v \in C_i\}$ is the stochastic distance in G from u to C_i. For $j = 0, 1, \ldots, h$, let $X_j = \max\{\Delta_u \mid u \notin C_i \wedge d(u, C_i) = j\}$, and let ℓ be in $\operatorname{argmax}\{X_j \mid j \in \{1 \ldots, h\}\}$. Then $X_0 = 0$ (since deterministic vertices in $\mathcal{M}$ remain deterministic in $\tilde{\mathcal{M}}$) and, for $j = 1, \ldots, h$, (21) (for a stochastic vertex and (19) for a deterministic vertex) implies

$$X_j \leq \varepsilon n + p_{\min}X_{j-1} + (1 - p_{\min})X_\ell. \tag{22}$$

Applying this iteratively gives us $X_\ell \leq \varepsilon n\frac{1-p_{\min}^\ell}{1-p_{\min}}p_{\min}^{-\ell} \leq \varepsilon n\lambda p_{\min}^{-\lambda}$.

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

$$\pi_v = \pi_{C_i} \pi_v^{|i} \leq (\tilde{\pi}_{C_i} + \varepsilon n \lambda p_{\min}^{-\lambda})(\tilde{\pi}_v^{|i} + \frac{\varepsilon}{2}(\lambda+1)|C_i|p_{\min}^{-\lambda})$$

$$\leq \tilde{\pi}_v + \frac{\varepsilon}{2} n p_{\min}^{-\lambda}[\varepsilon n \lambda(\lambda+1)p_{\min}^{-\lambda} + 3\lambda + 1] := \tilde{\pi}_v + \delta'(\mathcal{M}, \varepsilon).$$

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

$$|\mu_{u_0}(\mathcal{M}) - \mu_{u_0}(\tilde{\mathcal{M}})| \leq \sum_{(u,v) \in E} |\pi_u p_{uv} - \tilde{\pi}_u \tilde{p}_{uv}||r_{uv}|$$

$$\leq \sum_{(u,v) \in E} (|\pi_u - \tilde{\pi}_u|p_{uv} + \tilde{\pi}_u|\tilde{p}_{uv} - p_{uv}|)r^*$$

$$\leq \sum_{(u,v) \in E} (\max\{\delta'(\mathcal{M}, \varepsilon), \delta'(\tilde{\mathcal{M}}, \varepsilon)\}p_{uv} + \tilde{\pi}_u \varepsilon)r^*$$

$$\leq (\max\{\delta'(\mathcal{M}, \varepsilon), \delta'(\tilde{\mathcal{M}}, \varepsilon)\} + \varepsilon)n r^* \leq \delta(\mathcal{M}, \varepsilon)r^*,$$

which completes the proof.