RUTCOR
RESEARCH
REPORT

Nested Family of Cyclic Games with k-total Effective Rewards

Endre Boros[a] Khaled Elbassioni[b]
Vladimir Gurvich[b] Kazuhisa Makino[c]

RRR 06-2014, October 2014

RUTCOR
Rutgers Center for
Operations Research
Rutgers University
640 Bartholomew Road
Piscataway, New Jersey
08854-8003
Telephone: 732-445-3804
Telefax: 732-445-5472
Email: rrr@rutcor.rutgers.edu
http://rutcor.rutgers.edu/∼rrr

[a]MSIS Department and RUTCOR, Rutgers University, 100 Rockafellar Road, Livingston Campus Piscataway, NJ 08854, USA; ({boros,gurvich}@rutcor.rutgers.edu)
[b]Masdar Institute of Science and Technology, P.O.Box 54224, Abu Dhabi, UAE; (kelbassioni@masdar.ac.ae)
[c]Research Institute for Mathematical Sciences (RIMS) Kyoto University, Kyoto 606-8502, Japan; (makino@kurims.kyoto-u.ac.jp)

Rutcor Research Report
RRR 06-2014, October 2014

Nested Family of Cyclic Games with k-total Effective Rewards

Abstract. We consider Gillette's two-person zero-sum stochastic games with perfect information. For each $k \in \mathbb{Z}_+$ we introduce a payoff function, called the k-total reward. For $k = 0$ and 1 these payoffs are known as mean payoff and total reward, respectively. We restrict our attention to the deterministic case, the so called cyclic games. For all k, we prove the existence of a saddle point which can be realized by pure stationary strategies. We also demonstrate that k-total reward games can be embedded into $(k+1)$-total reward games. In particular, all of these classes contain mean payoff cyclic games.
Keywords: stochastic game with perfect information, cyclic games, two-person, zero-sum, mean payoff, total reward

1 Introduction

We consider two-person zero-sum stochastic games with perfect information and, for each positive integer k we define an effective payoff function, called the k-total reward, generalizing the classical *mean payoffs* [4] ($k = 0$), as well as the *total rewards* [18, 19] ($k = 1$).

In this paper, we restrict ourselves by two-person zero-sum games with deterministic states, and the solution concept is Nash equilibrium, which is just a saddle point in the considered case. We call the considered family of games *k-total reward BW-games*, where B and W stand for the two players, BLACK, the minimizer and WHITE, the maximizer.

We denote by $\mathbb{R}$ the set of reals, by $\mathbb{Z}$ the set of integers, and by $\mathbb{Z}_+$ the set of nonnegative integers. For a subset $S \subseteq \mathbb{Z}_+$, let $\mathbb{R}^S$ denote the set of vectors indexed by the elements of S. In particular, $\mathcal{S} = \mathcal{S}(R) \subseteq \mathbb{R}^{\mathbb{Z}_+ \setminus \{0\}}$ denotes the set of infinite integral sequences with elements not larger in absolute value than R, and for $\mathbf{a} \in \mathcal{S}$ we write $\mathbf{a} = (a_1, a_2, \ldots)$. Furthermore, for $n \in \mathbb{Z}_+ \setminus \{0\}$ we define $[n] = \{1, 2, \ldots, n\}$ and write simply $\mathbb{R}^n$ instead of $\mathbb{R}^{[n]}$.

To describe BW-games, let us consider a directed graph (digraph) $G = (V, E)$, whose vertices (positions or states) are partitioned into two sets $V = B \cup W$, a fixed initial state $v_0 \in V$, a real-valued function $r : E \to \mathbb{Z}$ assigning integer weights to the arcs (moves), and a mapping $\pi : \mathcal{S} \to \mathbb{R}$. We call the tuple (G, B, W) a *BW-game form* and r its *local rewards*, while the tuple (G, B, W, r, π) is called a *BW-game* and π is its *effective reward.* Two players, BLACK (the minimizer) and WHITE (the maximizer) control the positions of B and W, respectively. The game is played by starting at time $t = 0$ in the initial node (or position) $s_0 = v_0$. In a general step, in time t, we are at node $s_t \in V$. The player who controls s_t chooses an outgoing arc $e_{t+1} = (s_t, v) \in E$, and the game moves to node $s_{t+1} = v$. We assume, in fact without any loss of generality, that every vertex in G has an outgoing arc. (Indeed, if not, one can add loops to the corresponding vertices.) We assume that an initial vertex v_0 is fixed. However, when we talk about solving a BW-game, we consider (separately) all possible initial vertices.

In the course of this game players generate an infinite sequence of edges $\mathbf{p} = (e_1, e_2, \ldots)$ (a *play*) and the corresponding real sequence $r(\mathbf{p}) = (r(e_1), r(e_2), \ldots) \in \mathcal{S}$ of local rewards. At the end (after infinitely many steps) BLACK pays WHITE $\pi(r(\mathbf{p}))$ amount. Naturally, WHITE's aim is to create a play which maximizes this payoff, while BLACK tries to minimize it. (Let us note that the local reward function $r : E \to \mathbb{R}$ may have negative values, and $\pi(r(\mathbf{p}))$ maybe negative too, in which case WHITE has to pay BLACK.) As usual, a pair of (not necessarily stationary) strategies is a saddle point if neither of the players can improve individually by changing her/his strategy. The corresponding $\pi(r(\mathbf{p}))$ is the value of the game.

As we shall see later, it will be enough to restrict ourselves, and the players, to their stationary strategies in these games. This means that each player chooses, in advance, a move in every position that he controls and makes this move whenever the play comes to this position. Then, the play is uniquely determined by the one time selection of arcs and by the initial position. Such a play always looks like a "lasso": it consists of an initial path entering a directed cycle, which is then repeated infinitely many times.

Mean payoff (undiscounted) stochastic games, introduced in [4], are BW-games (see also [15, 14, 3, 7]) with payoff function $\pi = \phi$:

$$\phi(\mathbf{a}) \ = \ \liminf_{T\to\infty} \frac{1}{T}\sum_{j=1}^{T} a_j. \tag{1}$$

Such a game is known to have a saddle point in pure stationary strategies [4, 11] and these strategies can be computed in pseudo-polynomial time [16, 21]. Let us remark that the problem of deciding if the value of a mean payoff BW-game is below (or above) a given threshold belongs to both NP and co-NP ([7, 10, 21]). The exact complexity of this problem is however still not known; the best known algorithms are either pseudo-polynomial [16, 21] or randomized subexponential [1, 8, 20].

Discounted mean payoff stochastic games were in fact introduced earlier in [17] and have payoff function $\pi = \phi_\beta$:

$$\phi_\beta(\mathbf{a}) \ = \ (1-\beta)\sum_{j=1}^{\infty} \beta^{j-1} a_j. \tag{2}$$

As a consequence of the classical Hardy-Littlewood Tauberian theorems [9] we have the equality

$$\phi(\mathbf{a}) \ = \ \lim_{\beta\to 1} \phi_\beta(\mathbf{a}). \tag{3}$$

Discounted games, in general, are easier to solve, due to the fact that a standard value iteration is in fact a converging contraction. Hence, they are widely used in the literature of stochastic games together with the above limit equality. In fact, for mean payoff BW-games it is known [21] that for two sequences $\mathbf{a}, \mathbf{b} \in \mathcal{S}(R)$ we have $\phi_\beta(\mathbf{a}) < \phi_\beta(\mathbf{b})$ if and only if $\phi(\mathbf{a}) < \phi(\mathbf{b})$ whenever $1-\beta \le \frac{1}{4n^3R}$.

Total reward, introduced in [18] and considered in more detail in [19], is defined by

$$\psi(\mathbf{a}) \ = \ \liminf_{T\to\infty} \frac{1}{T}\sum_{i=1}^{T}\sum_{j=1}^{i} a_j. \tag{4}$$

It was shown in [19] that a total reward game is equivalent with a mean payoff game having countably many states. The authors derive from this that every total reward game has a value. Furthermore, ϵ-optimal Markovian strategies can be constructed by solving a discounted mean payoff game on the same graph with the same rewards using a discount factor β close enough to 1. The proof of the latter is analogous to the proof in [12].

It is worth noting that the 1-total reward games with nonnegative local rewards are polynomially solvable [13]. This contrasts the fact that mean payoff games with nonnegative rewards are as hard as general mean payoff games, and that the fastest known algorithms for mean payoff BW-games are either pseudo-polynomial [7, 16, 21] or randomized subexponential [1, 8, 20].

Main Results

In this paper we extend and generalize the above results. For every $k \in \mathbb{Z}_+$, we define the k-total effective reward, which coincide with the mean payoff when $k = 0$ and with the total reward when $k = 1$.

In general, given a sequence of local rewards $\mathbf{a} = (a_0, a_1, a_2, \ldots)$ let us associate to it another sequence $M(\mathbf{a}) = (a_0, a_0 + a_1, a_0 + a_1 + a_2, \ldots)$. Then the k-total payoff is defined as the mean payoff of the sequence $M^k(\mathbf{a})$.

First, we show that for every k a k-total reward game, when restricted to stationary strategies, has the same optimal strategy as the discounted mean payoff game on the same directed graph and with the same local rewards, if the discount factor is close enough to 1.

Second, we show that for every k in a k-total reward game, to a fixed stationary strategy of one player, there exists a best response of the other that is pure and stationary. This with the previous claim implies the existence of a saddle point in uniformly optimal stationary strategies in any k-total reward BW-game.

Next, we prove that the k-total reward games can be embedded into the family of $(k+1)$-total reward games, for each $k \in \mathbb{Z}_+$. In particular, mean payoff games can be embedded into k-total reward games for all $k \in \mathbb{Z}_+$. This containment and the example in [6] prove that for each $k \in \mathbb{Z}_+$, there is a non-zero sum k-total reward game without Nash equilibria.

Finally, it seems possible (and important) to generalize the above results to general stochastic games with (or even without) perfect information, or in other words, to replace the BW-model considered in this paper by the BWR-model, where R stands for positions of chance. For $k = 1$, it was shown in [2] that there always exists a saddle point, which can be realized by pure stationary uniformly optimal strategies. The similar question for $k > 1$ remains open.

Remark 1 *The k-total effective payoff looks somewhat similar to the moment M_k in the theory of probability. The first two cases $k = 0$ and $k = 1$ are easy to interpret and they play a very important role: the first two moments M_0 and M_1 are the expectation and variance, the 0- and 1-total payoffs are the mean and total effective payoffs, respectively. Yet, the higher moments have some important applications too.*

2 Iterated Total Rewards

In this section we introduce a complete hierarchy of payoff functions providing a natural generalization of mean and total payoffs, and their discounted counterparts.

Let us first introduce three operators acting on infinite sequences of reals. The *limiting average* $A : \mathcal{S} \to \mathbb{R}$ and *discounted limiting average* $A_\beta : \mathcal{S} \to \mathbb{R}$ operators map an infinite sequence into the set of reals, while the *moment* $M : \mathcal{S} \to \mathcal{S}$ operator maps it into another infinite sequence.

More precisely, given an infinite sequence $\mathbf{a} = (a_1, a_2, \ldots)$ and a real $0 < \beta \leq 1$ we define

$$A(\mathbf{a}) \;=\; \liminf_{T\to\infty} \frac{1}{T}\sum_{j=1}^{T} a_j \tag{5}$$

and

$$A_\beta(\mathbf{a}) \;=\; (1-\beta)\sum_{j=1}^{\infty} \beta^{j-1} a_j. \tag{6}$$

For convenience, we also extend the definition of the operator A for finite sequences $\mathbf{a}$ to denote the average of the elements of $\mathbf{a}$.

Finally, recall that we define $M(\mathbf{a}) = \mathbf{b} = (b_1, b_2, \ldots) \in \mathcal{S}$ by setting

$$b_i \;=\; \sum_{j=1}^{i} a_j \qquad \text{for all} \;\; i = 1, 2, \ldots \tag{7}$$

For $k = 0, 1, \ldots,$ we call $M^k(\mathbf{a})$ the *kth moment sequence* of $\mathbf{a}$, and define $M^0(\mathbf{a}) = \mathbf{a}$.

We also introduce the following families of functions $\phi^{(k)} : \mathcal{S} \to \mathbb{R}$ and $\phi^{(k)}_\beta : \mathcal{S} \to \mathbb{R}$ for $k = 0, 1, \ldots$, defined by

$$\phi^{(k)}(\mathbf{a}) \;=\; A\left(M^k(\mathbf{a})\right) \qquad \text{and} \qquad \phi^{(k)}_\beta(\mathbf{a}) \;=\; A_\beta\left(M^k(\mathbf{a})\right). \tag{8}$$

Let us note that $\phi^0 = \phi$ is the mean payoff function, $\phi^0_\beta = \phi_\beta$ is the discounted mean payoff, while $\phi^1 = \psi$ is the total reward. Following this terminology, we call $\phi^{(k)}$ the *k-total reward*, and $\phi^{(k)}_\beta$ the *discounted k-total reward*. Thus the above hierarchy of payoff functions provides a natural generalization of mean payoff and total reward.

We show first that the M operator changes the discounted total rewards by a factor that depends only on β.

Fact 1 *For all $\mathbf{a} \in \mathcal{S}$ and for all $0 < \beta < 1$ we have*

$$A_\beta(M(\mathbf{a})) \;=\; \frac{1}{1-\beta} A_\beta(\mathbf{a}).$$

Proof Using definition (6) and (7) we can write

$$\begin{aligned}
A_\beta(M(\mathbf{a})) &= (1-\beta)\sum_{i=1}^{\infty}\beta^{i-1}\left(\sum_{j=1}^{i}a_j\right)\\
&= (1-\beta)\sum_{j=1}^{\infty}a_j\left(\sum_{i=j}^{\infty}\beta^{i-1}\right)\\
&= (1-\beta)\sum_{j=1}^{\infty}\beta^{j-1}a_j\left(\sum_{i=j}^{\infty}\beta^{i-j}\right)\\
&= (1-\beta)\sum_{j=1}^{\infty}\beta^{j-1}a_j\left(\sum_{\ell=0}^{\infty}\beta^{\ell}\right)\\
&= (1-\beta)\sum_{j=1}^{\infty}\beta^{j-1}a_j\left(\frac{1}{1-\beta}\right)\\
&= \frac{1}{1-\beta}A_\beta(\mathbf{a}).
\end{aligned}$$

□

Given two sequences, $\mathbf{x} \in \mathbb{Z}^p$ and $\mathbf{y} \in \mathbb{Z}^q$, let us denote by $\mathbf{a} = (\mathbf{x}(\mathbf{y}))$ the infinite sequence obtained by listing first the elements of $\mathbf{x}$ and then repeating $\mathbf{y}$ cyclically, infinitely many times. Let us call such an $\mathbf{a} \in \mathcal{S}$ a *lasso sequence*, and let us denote the set of lasso sequences that can arise from a graph on n vertices by

$$\mathcal{S}_n(R) \;=\; \left\{ \mathbf{a} = (\mathbf{x}(\mathbf{y})) \;\middle|\; \begin{array}{cc} p,q \in \mathbb{Z}_+, & p+q \leq n \\ \mathbf{x} \in [-R,R]^p, & \mathbf{y} \in [-R,R]^q \end{array} \right\},$$

where $[-R,R]$ denotes the set of integers of absolute value not exceeding R. Note that a BW-game with n states in stationary strategies always produces a play $\mathbf{p}$ such that the corresponding rewards sequence $r(\mathbf{p})$ belongs to $\mathcal{S}_n(R)$, where R is an upper bound on the absolute values of the integral arc rewards. We shall simply write $\mathcal{S}_n$ when R is not specified.

To be able to state and prove our main results about iterated total rewards, we need next to analyze the above operations on lasso sequences.

Fact 2 *For $\mathbf{a} = (\mathbf{x}(\mathbf{y})) \in \mathcal{S}_n$, we have*

$$a_i = \begin{cases} x_i & \text{if } i \leq p, \\ y_r & \text{if } i = p + \ell q + r \text{ for some integers } \ell \geq 0,\ 0 < r \leq q. \end{cases}$$

Fact 3 *For* $\mathbf{a} = (\mathbf{x}(\mathbf{y})) \in \mathcal{S}_n$*, we have*

$$M(\mathbf{a})_i = \begin{cases} \sum_{j=1}^{i} x_j & \text{if } i \le p \\ \sum_{j=1}^{p} x_j + \ell \sum_{j=1}^{q} y_j + \sum_{j=1}^{r} y_j & \text{if } i = p + \ell q + r \text{ for integers } \ell \ge 0,\ 0 < r \le q. \end{cases}$$

Fact 4 *For* $\mathbf{a} = (\mathbf{x}(\mathbf{y})) \in \mathcal{S}_n$*, we have* $A(M(\mathbf{x}(\mathbf{y}))) = \begin{cases} +\infty & \text{if } A(\mathbf{x}(\mathbf{y})) > 0, \\ -\infty & \text{if } A(\mathbf{x}(\mathbf{y})) < 0. \end{cases}$

Note that there is an example in [2] showing that Fact 4 does not necessarily hold for reward sequences corresponding to non-stationary strategies.

Let us also note that $M(\mathbf{x}(\mathbf{y}))$ is not a lasso sequence, in general. However, the above facts, obtained by simple counting arguments, imply the following claim, the second part of which can be obtained from the first part by induction on k:

Corollary 1 *If* $\mathbf{x}(\mathbf{y}) \in \mathcal{S}_n(R)$ *such that* $A(\mathbf{y}) = 0$*, then* $M(\mathbf{x}(\mathbf{y})) = \widetilde{\mathbf{x}}(\widetilde{\mathbf{y}}) \in \mathcal{S}_n(nR)$*, where*

$$\begin{aligned} \widetilde{\mathbf{x}} &= M(\mathbf{x}), \text{ and} \\ \widetilde{\mathbf{y}} &= pA(\mathbf{x}) + M(\mathbf{y}). \end{aligned}$$

Furthermore, if $\mathbf{x}(\mathbf{y}) \in \mathcal{S}_n(R)$ *such that* $A(\mathbf{x}(\mathbf{y})) = A(M(\mathbf{x}(\mathbf{y}))) = \cdots = A(M^{k-1}(\mathbf{x}(\mathbf{y}))) = 0$*, then* $M^k(\mathbf{x}(\mathbf{y})) \in \mathcal{S}_n(n^k R)$*.*

Recall that by adding a scalar to a vector we mean incrementing all components of the vector by the same scalar value.

The above properties allow us to generalize an inequality between the discounted and undiscounted payoffs shown by [21] for the mean payoff case.

Lemma 1 *If* $\mathbf{x}(\mathbf{y}) \in \mathcal{S}_n(R)$ *such that* $A(\mathbf{x}(\mathbf{y})) = A(M(\mathbf{x}(\mathbf{y}))) = \cdots = A(M^{k-1}(\mathbf{x}(\mathbf{y}))) = 0$*, then*

$$|\phi^{(k)}(\mathbf{x}(\mathbf{y})) - \frac{1}{(1-\beta)^k}\phi_\beta(\mathbf{x}(\mathbf{y}))| \le 2(1-\beta)n^{k+2}R.$$

Proof It was shown in [21] that for a lasso sequence $(\mathbf{x}(\mathbf{y})) \in \mathcal{S}_n(R)$ we have

$$|A(\mathbf{x}(\mathbf{y})) - A_\beta(\mathbf{x}(\mathbf{y}))| \le 2(1-\beta)n^2 R.$$

Applying this for an arbitrary lasso sequence $\widetilde{\mathbf{x}}(\widetilde{\mathbf{y}}) \in \mathcal{S}_n(n^k R)$ we get

$$|A(\widetilde{\mathbf{x}}(\widetilde{\mathbf{y}})) - A_\beta(\widetilde{\mathbf{x}}(\widetilde{\mathbf{y}}))| \le 2(1-\beta)n^2(n^k R) = 2(1-\beta)n^{k+2}R.$$

By Corollary 1 we have $M^k(\mathbf{x}(\mathbf{y})) \in \mathcal{S}_n(n^k R)$, thus applying the above for $\widetilde{\mathbf{x}}(\widetilde{\mathbf{y}}) = M^k(\mathbf{x}(\mathbf{y}))$ we get

$$|A(M^k(\mathbf{x}(\mathbf{y}))) - A_\beta(M^k(\mathbf{x}(\mathbf{y})))| \leq 2(1-\beta)n^{k+2}R.$$

By Fact 1 we have $A_\beta(M^k(\mathbf{a})) = \frac{1}{(1-\beta)^k} A_\beta(\mathbf{a})$, and thus the above implies our claim. □

Let us note that we can prove actually a slightly better inequality for $k = 1$; see Lemma 9 in the appendix. This is because not all sequences in $\mathcal{S}_n(nR)$ are of the form $M(\mathbf{x}(\mathbf{y}))$ for some $\mathbf{x}(\mathbf{y}) \in \mathcal{S}_n(R)$. To prove a similarly stronger inequality for any k we would need to understand better the functional dependence of $M^k(\mathbf{x}(\mathbf{y}))$ on $\mathbf{x}(\mathbf{y})$ under the constraints $A(\mathbf{x}(\mathbf{y})) = A(M(\mathbf{x}(\mathbf{y}))) = \cdots = A(M^{k-1}(\mathbf{x}(\mathbf{y}))) = 0$, which might be a harder task. For our purpose the above weaker inequalities are still enough.

3 Uniform Optimality within Stationary Strategies

Let us consider a BW-game form (G, B, W) and an integral local reward $r : E \to \mathbb{Z}$. As before, we denote by R the largest absolute value of an arc reward and use $V = B \cup W$ for the set of nodes. For a subset $F \subseteq E$ of the arcs of the directed graph $G = (V, E)$ we denote by $d_F^+(v)$ the out-degree of vertex $v \in V$ in the subgraph (V, F). A subset $F \subseteq E$ of the arcs is a pure stationary strategy of BLACK (resp., WHITE) if $d_F^+(v) = 1$ for all $v \in B$ (resp., for all $v \in W$). Let us denote by $\mathfrak{S}_B$ and $\mathfrak{S}_W$ the sets of stationary strategies of BLACK and WHITE, respectively.

Given a situation (a pair of stationary strategies), $\mathfrak{b} \in \mathfrak{S}_B$ and $\mathfrak{w} \in \mathfrak{S}_W$, we have a unique walk $e_1, e_2, \ldots, e_p, e_{p+1}, \ldots, e_{p+q}, e_{p+1}, e_{p+2}, \ldots, e_{p+q}, \ldots$ from every initial node $v_0 \in V$, which consists of an initial path $e_1, e_2, \ldots, e_p$ followed by a cycle $e_{p+1}, \ldots, e_{p+q}$, which we traverse infinitely many times. We denote by $\mathbf{x} = \mathbf{x}(\mathfrak{b}, \mathfrak{w})$ and $\mathbf{y} = \mathbf{y}(\mathfrak{b}, \mathfrak{w})$ the corresponding reward sequences $\mathbf{x} = (r(e_j) \mid j = 1, \ldots, p)$ and $\mathbf{y} = (r(e_{p+j}) \mid j = 1, \ldots, q)$, and by $c(v_0; \mathfrak{b}, \mathfrak{w})$ the payoff value corresponding to payoff function π:

$$c(v_0; \mathfrak{b}, \mathfrak{w}) \;=\; \pi(\mathbf{x}(\mathbf{y})).$$

We say that a situation $(\mathfrak{b}^*, \mathfrak{w}^*)$, $\mathfrak{b}^* \in \mathfrak{S}_B$, $\mathfrak{w}^* \in \mathfrak{S}_W$ is a *saddle point* for initial position $v_0 \in V$ if

$$c(v_0; \mathfrak{b}^*, \mathfrak{w}) \;\leq\; c(v_0; \mathfrak{b}^*, \mathfrak{w}^*) \;\leq c(v_0; \mathfrak{b}, \mathfrak{w}^*) \tag{9}$$

hold for all $\mathfrak{b} \in \mathfrak{S}_B$ and $\mathfrak{w} \in \mathfrak{S}_W$. We say that $(\mathfrak{b}^*, \mathfrak{w}^*)$ is a *uniform saddle point* if (9) holds for all initial positions $v_0 \in V$.

The following result follows essentially from [17].

Fact 5 *A BW-game with the discounted mean payoff function $\pi = \phi_\beta$ has a uniform saddle point for all $0 < \beta < 1$,*

We shall show next that the same pair of stationary strategies form a uniform saddle point with respect to the total reward $\pi = \psi$, if β is close enough to 1.

Theorem 1 *Consider a BW-game (G, B, W, r) as above, and choose a discount factor satisfying $0 \leq (1-\beta) < \frac{1}{4n^{k+4}R}$. Let us consider a uniform saddle point $(\mathfrak{b}^*, \mathfrak{w}^*)$ with respect to the discounted mean payoff ϕ_β. Then, $(\mathfrak{b}^*, \mathfrak{w}^*)$ is also a uniform saddle point with respect to the k-total reward $\phi^{(k)}$.*

Proof We introduce two preorders $\prec_\beta$ and $\prec_k$ on the set of lasso sequences $\mathcal{S}_n(R)$ as follows: for two sequences $(\mathbf{x}(\mathbf{y}))$ and $\widetilde{\mathbf{x}}(\widetilde{\mathbf{y}}))$ we say that $(\mathbf{x}(\mathbf{y})) \prec_\beta (\widetilde{\mathbf{x}}(\widetilde{\mathbf{y}}))$ (resp., $(\mathbf{x}(\mathbf{y})) \prec_k (\widetilde{\mathbf{x}}(\widetilde{\mathbf{y}}))$) if $\phi_\beta(\mathbf{x}(\mathbf{y})) \leq \phi_\beta(\widetilde{\mathbf{x}}(\widetilde{\mathbf{y}}))$ (resp., $\phi^{(k)}(\mathbf{x}(\mathbf{y})) \leq \phi^{(k)}(\widetilde{\mathbf{x}}(\widetilde{\mathbf{y}}))$). We shall show that the preorder $\prec_\beta$ induced by ϕ_β on $\mathcal{S}_n(R)$ is a refinement of $\prec_k$ induced by $\phi^{(k)}$. For this end let us first prove some partial claim about $\prec_\beta$. Introduce

$$\mathcal{S}(\ell,+) = \left\{\mathbf{x}(\mathbf{y}) \in \mathcal{S}_n(R) \mid \phi^{(j)}(\mathbf{x}(\mathbf{y})) = 0 \;\text{ for }\; j = 0, 1, \ldots, \ell-1, \text{ and } \phi^{(\ell)}(\mathbf{x}(\mathbf{y})) > 0\right\}$$
$$\mathcal{S}(\ell,-) = \left\{\mathbf{x}(\mathbf{y}) \in \mathcal{S}_n(R) \mid \phi^{(j)}(\mathbf{x}(\mathbf{y})) = 0 \;\text{ for }\; j = 0, 1, \ldots, \ell-1, \text{ and } \phi^{(\ell)}(\mathbf{x}(\mathbf{y})) < 0\right\}$$

for $\ell = 0, 1, \ldots, k-1$, and set

$$\mathcal{S}^* \;=\; \left\{\mathbf{x}(\mathbf{y}) \in \mathcal{S}_n(R) \mid \phi^{(j)}(\mathbf{x}(\mathbf{y})) = 0 \;\text{ for }\; j = 0, 1, \ldots, k-1\right\}.$$

It is immediate to see by these definitions that these sets partition $\mathcal{S}_n(R)$. We claim that the following relations hold:

$$\mathcal{S}(0,-) \prec_\beta \cdots \mathcal{S}(k-1,-) \prec_\beta \mathcal{S}^* \prec_\beta \mathcal{S}(k-1,+) \prec_\beta \cdots \prec_\beta \mathcal{S}(0,+). \tag{10}$$

To see this, let us fix an index $0 \leq \ell < k$, and consider $\mathbf{x}(\mathbf{y}) \in \mathcal{S}(\ell,+)$ and $\widetilde{\mathbf{x}}(\widetilde{\mathbf{y}}) \in \mathcal{S}(\ell+1,+)$ (or this could be $\mathcal{S}^*$ if $\ell = k-1$). Lemma 1 implies that

$$\left|\phi^{(\ell)}(\mathbf{x}(\mathbf{y})) - \frac{\phi_\beta(\mathbf{x}(\mathbf{y}))}{(1-\beta)^\ell}\right| \;\leq\; 2(1-\beta)n^{\ell+2}R.$$

Since $M^\ell(\mathbf{x}(\mathbf{y})) \in \mathcal{S}_n(n^\ell R)$ (by Corollary 1) is an integer sequence, thus by the definition of the A operator we get

$$\phi^{(\ell)}(\mathbf{x}(\mathbf{y})) \geq \frac{1}{q} \geq \frac{1}{n}.$$

Thus, since we have $2(1-\beta)n^{\ell+2}R < \frac{1}{2n}$ the above implies

$$\frac{\phi_\beta(\mathbf{x}(\mathbf{y}))}{(1-\beta)^\ell} \;>\; \frac{1}{2n}. \tag{11}$$

On the other hand we can apply Lemma 1 for $\widetilde{\mathbf{x}}(\widetilde{\mathbf{y}})$, too, yielding

$$\left|\phi^{(\ell)}(\widetilde{\mathbf{x}}(\widetilde{\mathbf{y}})) - \frac{\phi_\beta(\widetilde{\mathbf{x}}(\widetilde{\mathbf{y}}))}{(1-\beta)^\ell}\right| \;\leq\; 2(1-\beta)n^{\ell+2}R.$$

Since we have $\phi^{(\ell)}(\widetilde{\mathbf{x}}(\widetilde{\mathbf{y}})) = 0$, we get

$$\frac{\phi_\beta(\widetilde{\mathbf{x}}(\widetilde{\mathbf{y}}))}{(1-\beta)^\ell} \;<\; \frac{1}{2n}. \tag{12}$$

Inequalities (11) and (12) together imply that $\mathcal{S}(\ell+1,+) \prec_\beta \mathcal{S}(\ell,+)$ (or $\mathcal{S}^* \prec_\beta \mathcal{S}(k-1,+)$). Analogous arguments work for the negative side, too, and hence (10) follows.

Since $\phi^{(k)}$ has value $+\infty$ on $\mathcal{S}(\ell,+)$ for $\ell = 0,1,\dots,k-1$ and value $-\infty$ on the sets $\mathcal{S}(\ell,-)$ for $\ell = 0,1,\dots,k-1$ by Corollary 1, the only thing left to prove that $\prec_\beta$ is a refinement of that of $\prec_k$ is to show that these two preorders do not conflict on the set $\mathcal{S}^*$.

To this end let us consider two lasso sequences $\mathbf{x}(\mathbf{y}), \widetilde{\mathbf{x}}(\widetilde{\mathbf{y}}) \in \mathcal{S}^*$ such that $\phi^{(k)}$ has different values on these sequences. Since $\phi^{(k)}(\mathbf{x}(\mathbf{y})) = A(M^k(\mathbf{x}(\mathbf{y})))$ and for $\mathbf{x}(\mathbf{y}) \in \mathcal{S}^*$ we have $M^k(\mathbf{x}(\mathbf{y})) \in \mathcal{S}_n(n^k R)$, we again can conclude that the difference between two nonequal $\phi^{(k)}$ values is always at least $1/n^2$, say:

$$\phi^{(k)}(\mathbf{x}(\mathbf{y})) - \phi^{(k)}(\widetilde{\mathbf{x}}(\widetilde{\mathbf{y}})) \;\geq\; \frac{1}{n^2}.$$

We can apply Lemma 1 for these sequences yielding

$$\begin{aligned}\left|\phi^{(k)}(\mathbf{x}(\mathbf{y})) - \tfrac{\phi_\beta(\mathbf{x}(\mathbf{y}))}{(1-\beta)^k}\right| &\leq 2(1-\beta)n^{k+2}R, \text{ and}\\ \left|\phi^{(k)}(\widetilde{\mathbf{x}}(\widetilde{\mathbf{y}})) - \tfrac{\phi_\beta(\widetilde{\mathbf{x}}(\widetilde{\mathbf{y}}))}{(1-\beta)^k}\right| &\leq 2(1-\beta)n^{k+2}R.\end{aligned}$$

Since we have $2(1-\beta)n^{k+2}R < \frac{1}{2n^2}$, the above three inequalities imply $\phi_\beta(\mathbf{x}(\mathbf{y})) > \phi_\beta(\widetilde{\mathbf{x}}(\widetilde{\mathbf{y}}))$, proving our claim, and hence completing the proof of the theorem. □

Let us remark that the proof of the above theorem in fact provides a more complete picture about these iterated total reward functions than the statement of the theorem alone. Namely, the payoff functions $\phi^{(k)}$, $k = 0,1,\dots$ form a nested sequence, in the sense that the lasso sequences on which $\phi^{(k+1)}$ vanishes form a subset of those where $\phi^{(k)}$ vanishes, and $\phi^{(k+1)}$ has a finite value only on sequences on which $\phi^{(k)}$ vanishes. Furthermore, (10) can be claimed for an arbitrary integer k, providing a complete hierarchy in the limit, and we conjecture that $\mathcal{S}^*$ in the limit contains only zero sequences. It is also interesting to see that all these payoffs rank lasso sequences in agreement with the discounted payoff, and all of them has "essentially" the same discounted version.

4 Best Response to Stationary Strategies

In this section, we will use two operators on finite sequences: for a sequence $\mathbf{a} = (a_1, a_2, \dots, a_n)$ we have $S(\mathbf{a}) = \sum_{i=1}^n a_i$ and $M(\mathbf{a}) = (a_1, a_1 + a_2, \dots, S(\mathbf{a}))$. Note that both M and S are linear operators.

Lemma 2 *For sequences $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ and real $\lambda \in \mathbb{R}$ we have*

$$S(\mathbf{a}+\mathbf{b}) = S(\mathbf{a}) + S(\mathbf{b}) \tag{13a}$$

$$S(\lambda\mathbf{a}) = \lambda S(\mathbf{a}) \tag{13b}$$

$$M(\mathbf{a}+\mathbf{b}) = M(\mathbf{a}) + M(\mathbf{b}) \tag{13c}$$

$$M(\lambda\mathbf{a}) = \lambda M(\mathbf{a}) \tag{13d}$$

Proof By definitions of these operators. □

For a sequence I of integers (indices), we denote by $\mathbf{a}_I$ the sequence of the corresponding $\mathbf{a}$ components. E.g., if $\mathbf{a} = (a_1, a_2, a_3, \ldots)$ and $I = (1, 2, 3, 2)$, then $\mathbf{a}_I = (a_1, a_2, a_3, a_2)$. We denote by $[1, n]$ the sequence of integers from 1 to n. If $\mathbf{a}$ and $\mathbf{b}$ are two finite sequences, then we denote by $[\mathbf{a}; \mathbf{b}]$ their concatenation. For a subset I of indices we denote by $\mathbf{e}_I$ the sequence of 1-s of length $|I|$ indexed by $i \in I$.

We define $M^0(\mathbf{a}) = \mathbf{a}$, and for $k > 1$ we define $M^k(\mathbf{a}) = M(M^{k-1}(\mathbf{a}))$. With these notation, the k-total value of an infinite sequence $\mathbf{a}$ of local rewards (corresponding to a play) can be rewritten as

$$\phi^{(k)}(\mathbf{a}) \;=\; \liminf_{T\to\infty} \frac{1}{T} S(M^k(\mathbf{a}_{[1,T]})).$$

Assume in the sequel that X, Y and Z are subsets of the indices, and $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are finite sequences of appropriate lengths.

Lemma 3

$$M(\mathbf{a}_{[X;Y]}) = \;[M(\mathbf{a}_X);\; M(\mathbf{a}_Y) + S(\mathbf{a}_X)\mathbf{e}_Y] \tag{14a}$$

$$M^k(\mathbf{a}_{[X;Y]}) = \;\left[M^k(\mathbf{a}_X); M^k(\mathbf{a}_Y) + \sum_{\ell=1}^{k} S(M^{k-\ell}(\mathbf{a}_X))M^{\ell-1}(\mathbf{e}_Y)\right]. \tag{14b}$$

Proof Equality (14a) follows by the definitions of M and S. For (14b) we use (14a), (13c), (13d) and induction on k. For $k = 0$ we get $M^0(\mathbf{a}_{[X;Y]}) = [\mathbf{a}_X; \mathbf{a}_Y]$, and for $k = 1$ we get

$$M(\mathbf{a}_{[X;Y]}) \;=\; M([\mathbf{a}_X; \mathbf{a}_Y]) \;=\; [M(\mathbf{a}_X); M(\mathbf{a}_Y) + S(\mathbf{a}_X)\mathbf{e}_Y]$$

by the definition of M, as in (14a). Then, by induction on k we get

$$\begin{aligned}
M^{k+1}(\mathbf{a}_{[X;Y]}) &= M\left(\left[M^k(\mathbf{a}_X); M^k(\mathbf{a}_Y) + \sum_{\ell=1}^{k} S(M^{k-\ell}(\mathbf{a}_X))M^{\ell-1}(\mathbf{e}_Y)\right]\right) \\
&= \left[M^{k+1}(\mathbf{a}_X); M^{k+1}(\mathbf{a}_Y) + S(M^k(\mathbf{a}_X))\mathbf{e}_Y + \sum_{\ell=1}^{k} S(M^{k-\ell}(\mathbf{a}_X))M^{\ell}(\mathbf{e}_Y)\right] \\
&= \left[M^{k+1}(\mathbf{a}_X); M^{k+1}(\mathbf{a}_Y) + \sum_{\ell=1}^{k+1} S(M^{k+1-\ell}(\mathbf{a}_X))M^{\ell-1}(\mathbf{e}_Y)\right].
\end{aligned}$$

□

Corollary 2

$$\begin{aligned}
S(M^k(\mathbf{a}_{[X;Y]})) \;&=\; S(M^k(\mathbf{a}_X)) + S(M^k(\mathbf{a}_Y)) + \sum_{\ell=1}^{k} S(M^{k-\ell}(\mathbf{a}_X))S(M^{\ell-1}(\mathbf{e}_Y)) \\
&=\; S(M^k(\mathbf{a}_X)) + S(M^k(\mathbf{a}_Y)) + \sum_{\ell=1}^{k} S(M^{k-\ell}(\mathbf{a}_X))\binom{|Y| + \ell - 1}{\ell}.
\end{aligned}$$

Proof By Lemmas 2 and 3, and by the equality

$$S(M^{\ell-1}(\mathbf{e}_Y)) \;=\; \binom{|Y|+\ell-1}{\ell}. \tag{15}$$

□

Corollary 3

$$\begin{aligned} M^k(\mathbf{a}_{[X;Y;Z]}) \;=\; \Bigg[\;\; & M^k(\mathbf{a}_X); M^k(\mathbf{a}_Y) + \sum_{\ell=1}^{k} S(M^{k-\ell}(\mathbf{a}_X))M^{\ell-1}(\mathbf{e}_Y); \\ & M^k(\mathbf{a}_Z) + \sum_{\ell=1}^{k} \left(S(M^{k-\ell}(\mathbf{a}_X)) + S(M^{k-\ell}(\mathbf{a}_Y))\right) M^{\ell-1}(\mathbf{e}_Z) \\ & + \sum_{\ell=1}^{k} S(M^{k-\ell}(\mathbf{a}_X)) \sum_{m=1}^{\ell-1} \binom{|Y|+\ell-1-m}{\ell-m} M^{m-1}(\mathbf{e}_Z) \;\Bigg]. \end{aligned}$$

Proof Apply Lemma 3 with $[Y;Z]$ instead of Y, and then again with Y,Z instead of X,Y, and use $S(M^{\ell-1-m}(\mathbf{e}_Y)) = \binom{|Y|+\ell-1-m}{\ell-m}$. □

Corollary 4

$$S(M^k(\mathbf{a}_{[X;Y;Z]})) = S(M^k(\mathbf{a}_X)) + S(M^k(\mathbf{a}_Y)) + S(M^k(\mathbf{a}_Z))$$

$$+ \sum_{\ell=1}^{k} \left(S(M^{k-\ell}(\mathbf{a}_X)) \binom{|Y|+|Z|+\ell-1}{\ell} + S(M^{k-\ell}(\mathbf{a}_Y)) \binom{|Z|+\ell-1}{\ell} \right).$$

Proof Apply (13a) and (13b) of Lemma 2 and Corollary 3 above together with the equality (15) applied for both $\mathbf{e}_Y$ and $\mathbf{e}_Z$, and finally use the binomial identity

$$\sum_{m=0}^{\ell} \binom{|Y|+\ell-1-m}{\ell-m} \binom{|Z|+m-1}{m} \;=\; \binom{|Y|+|Z|+\ell-1}{\ell}. \tag{16}$$

□

Corollary 5

$$S(M^k(\mathbf{a}_{[X;Y;Z]})) = S(M^k(\mathbf{a}_{[X;Z]})) + \sum_{\ell=0}^{k} \binom{|Z|+k-1-\ell}{k-\ell} S(M^{\ell}(\mathbf{a}_{[X;Y]})).$$

Proof Elementary calculations by Corollaries 2, 4, and by the binomial identity (16). □

Corollary 6 *If* $\mathbf{a} = \mathbf{x}(\mathbf{y})$ *is a lasso sequence, where* $\mathbf{x} = \mathbf{a}_X$ *and* $\mathbf{y} = \mathbf{a}_Y$*, and*

$$S(M^\ell(\mathbf{a}_{[X;Y]})) = 0 \text{ for } \ell = 0, 1, \ldots, k-1, \tag{17a}$$

then we have

$$\phi^{(k)}(\mathbf{a}) = \frac{1}{|Y|} S(M^k(\mathbf{a}_{[X;Y]})). \tag{17b}$$

Furthermore, if condition (17a) *does not hold, and* $0 \leq m < k$ *is the smallest index with* $S(M^{(m)}(\mathbf{a}_{[X;Y]})) \neq 0$*, then we have*

$$\phi^{(k)}(\mathbf{a}) = \begin{cases} -\infty & \text{if} \quad S(M^{(m)}(\mathbf{a}_{[X;Y]})) < 0, \\ +\infty & \text{if} \quad S(M^{(m)}(\mathbf{a}_{[X;Y]})) > 0. \end{cases} \tag{17c}$$

Proof For a positive integer T let us define $r(T) = (T - |X|) \mod |Y|$ and $\alpha(T) = \frac{T-|X|-r_T}{|Y|}$. Furthermore, let us denote by R_r the first r elements of Y, for $0 \leq r < |Y|$, and by $\alpha Y = [Y, Y, \ldots, Y]$ the concatenation of Y α times. Then, for the lasso sequence $\mathbf{a} = \mathbf{a}_X(\mathbf{a}_Y)$ we have

$$\phi^{(k)}(\mathbf{a}) = \liminf_{T\to\infty} \frac{1}{T} S(M^k(\mathbf{a}_{[X;\alpha(T)Y;R_{r(T)}]}). \tag{18}$$

Let us now apply Corollary 5 with $Z = [(\beta - 1)Y; R_{r(T)}]$ for an arbitrary positive integer β to get

$$S(M^k(\mathbf{a}_{[X;\beta Y;R_{r(T)}]}) = S(M^k(\mathbf{a}_{[X;(\beta-1)Y;R_{r(T)}]}) + \sum_{\ell=0}^{k} \binom{r(T) + (\beta-1)|Y| + k - 1 - \ell}{k - \ell} S(M^\ell(\mathbf{a}_{[X;Y]})). \tag{19}$$

Assume first that condition (17a) holds. Then the above equality simplifies to

$$S(M^k(\mathbf{a}_{[X;\beta Y;R_{r(T)}]}) = S(M^k(\mathbf{a}_{[X;(\beta-1)Y;R_{r(T)}]}) + S(M^k(\mathbf{a}_{[X;Y]})).$$

Summing up the above equalities for $\beta = 1, \ldots, \alpha(T)$, we get

$$S(M^k(\mathbf{a}_{[X;\alpha(T)Y;R_{r(T)}]}) = S(M^k(\mathbf{a}_{[X;R_{r(T)}]}) + \alpha(T) S(M^k(\mathbf{a}_{[X;Y]})). \tag{20}$$

Since $S(M^k(\mathbf{a}_{[X;R_{r(T)}]})$ and $|X| + r(T)$ are both bounded independently of T, and since

$$\lim_{T\to\infty} \frac{\alpha(T)}{T} = \lim_{T\to\infty} \frac{\alpha(T)}{|X| + \alpha(T)|Y| + r(T)} = \frac{1}{|Y|},$$

we get from (18) and (20) that

$$\phi^{(k)}(\mathbf{a}) = \frac{1}{|Y|} S(M^k(\mathbf{a}_{[X;Y]}))$$

as claimed in (17b).

Let us assume next that condition (17a) does not hold, and let m be the smallest index such that $S(M^m(\mathbf{a}_{[X;Y]})) \neq 0$. Consider equality (19), and note that the summation on the right hand side can be viewed as a polynomial $p(\beta)$, which is of degree $k - m$ and in which the sign of the leading term is the same as the sign of $S(M^m(\mathbf{a}_{[X;Y]}))$. Then,

$$\sum_{\beta=1}^{\alpha(T)} p(\beta) = q(\alpha(T))$$

is a polynomial of $\alpha(T)$ of degree $k - m + 1$, and the sign of its leading term is the same as the sign of $S(M^m(\mathbf{a}_{[X;Y]}))$. Thus, by summing up (19) for $\beta = 1, \ldots, \alpha(T)$ we get

$$S(M^k(\mathbf{a}_{[X;\alpha(T)Y;R_{r(T)}]}) \;=\; S(M^k(\mathbf{a}_{[X;R_{r(T)}]}) + q(\alpha(T)). \tag{21}$$

Since $k-m+1 > 1$ and since $S(M^k(\mathbf{a}_{[X;R_{r(T)}]})$ and $|X|+r(T)$ are both bounded independently of T, we obtain (17c) from (18) and (21) by dividing by T and taking limits. □

Given a BW-game Γ, we denote by $\mathfrak{P}_B(\Gamma)$ and $\mathfrak{P}_W(\Gamma)$ the sets of (not necessarily pure and/or stationary) strategies of BLACK and WHITE, respectively. Let us note that $\mathfrak{S}_B$ and $\mathfrak{S}_W$ are proper subsets of $\mathfrak{P}_B$ and $\mathfrak{P}_W$, respectively.

Now we are ready to prove the main result of this section.

Theorem 2 *There is a pure stationary best response to a pure stationary strategy in a k-total reward BW-game, for any $k \in \mathbb{Z}_+$.*

Proof Let us consider a positive integer k, and a k-total reward game $\Gamma = (G, B, W, v_0, r, \phi^{(k)})$. Assume that BLACK fixes an arbitrary pure stationary strategy $\mathfrak{b}^* \in \mathfrak{S}_B = \mathfrak{S}_B(\Gamma)$, and we will show that WHITE has a pure stationary best response. Though there is a mild asymmetry due to the definition of the value of the game, the case when WHITE fixes a pure stationary strategy can be handled similarly.

Let us note that at this moment we may not (yet) assume that a best response exists, at all. We however know that for every $1 > \epsilon > 0$ there exists a strategy $\mathfrak{w}^* \in \mathfrak{P}_W = \mathfrak{P}_W(\Gamma)$ of WHITE (not necessarily pure and/or stationary) for which

$$\phi^{(k)}(\mathfrak{b}^*, \mathfrak{w}^*) \;\geq\; \sup_{\mathfrak{w} \in \mathfrak{P}_W} \phi^{(k)}(\mathfrak{b}^*, \mathfrak{w}) \quad - \epsilon \tag{22}$$

simply by the definition of sup. Let us also note that we can assume

$$\phi^{(k)}(\mathfrak{b}^*, \mathfrak{w}^*) > -\infty$$

since otherwise any pure and stationary strategy of White is a best response.

To prove our theorem, let us first use the fact that $\mathfrak{b}^*$ is pure and stationary, delete all the arcs going out from vertices controlled by Black which are not chosen in $\mathfrak{b}^*$, and denote by $G^* = (V, E^*)$ the resulting directed graph.

Let us note that any walk in G^* starting from v_0 arises as the play corresponding to strategies $(\mathfrak{b}^*, \mathfrak{w})$ for some strategy $\mathfrak{w}$ of White. Furthermore, if $\mathfrak{w}$ is a pure and stationary strategy then this walk is a lasso, and every lasso sequence starting at v_0 corresponds in this way to some pure and stationary strategy of White.

Let us next consider the pair $(\mathfrak{b}^*, \mathfrak{w}^*)$ satisfying (22), which defines an infinite walk in the directed graph G^* along the vertices $W = (v_0, v_1, \ldots)$. As before, let us denote by $e_t = (v_{t-1}, v_t)$ the edges of this walk, and by $a_t = r(e_t)$ the local rewards, for $t = 1, 2, \ldots$, and set $F = \{e_1, e_2, \ldots\}$ to denote the sequence of arcs in this infinite walk in G^*.

Let us denote for integers $i < j$ by $[i, j]$ the set of integers $\{i, i+1, \ldots, j\}$, and set $J_0 = \{1, 2, \ldots\}$ to denote the set of positive integers. To an increasing subset $I = \{i_1, i_2, \ldots\} \subseteq J_0$ of the indices, $i_1 < i_2 < \cdots$, we associate the edge set $F(I) = \{e_i \mid i \in I\}$ and we say that $F(I)$ is a *walk* in G^* if the endpoint of e_{i_s} is the beginning of $e_{i_{s+1}}$, that is if $v_{i_s} = v_{i_{s+1}-1}$ for all $s = 1, 2, \ldots$. Note that if $s \in I$, then $F(I \cap [1, s])$ is also a walk in G^*.

Let us consider an arbitrary infinite increasing sequence I of integers such that $F(I)$ is a walk in G^* and let $q = q(I)$ be the smallest integer $q \in I$ such that vertex v_q is repeated in the walk $F(I \cap [1, q])$. Let us denote then by $p = p(I)$ the unique index $p < q$, $p \in I$ for which $v_p = v_q$. Let us introduce $C(I) = I \cap [p+1, q]$, set $P(I) = I \cap [1, p]$ and define $J(I) = I \setminus [p+1, q]$. Observe that $F([P(I); C(I)])$ is a lasso sequence with $F(C(I))$ as its cycle, and that $F(J(I))$ is again a walk in G^*.

Next, starting with $I = J_0$ and $F = F(J_0)$, let us define $p_s = p(J_{s-1})$, $q_s = q(J_{s-1})$, $P_s = P(J_{s-1})$, $C_s = C(J_{s-1})$, $L_s = [C_s; P_s]$ and set $J_s = J(J_{s-1})$, recursively for $s = 1, 2, \ldots$. Note that by the above observation $F(J_k) \subseteq F$ is a walk in G^*, $F(C_s)$ is a cycle, and $F(L_s)$ is a lasso with $F(C_s)$ as its cycle, for every index $s = 1, 2, \ldots$.

Let us finally denote by $\mathfrak{w}_s$, $s = 1, 2, \ldots$ a pure stationary strategy of White for which $F(L_s)$ is the corresponding play when Black chooses $\mathfrak{b}^*$ and White chooses $\mathfrak{w}_s$. (As we observed above, every lasso sequence in G^*, starting from v_0 corresponds to such a pure stationary strategy of White.)

Recall that we have the notation $a_t = r(e_t)$ for all $t \in J_0$, and consider $\mathbf{a} = \mathbf{a}_{J_0} = (a_1, a_2, \ldots)$ the infinite sequence of local rewards of the play $(\mathfrak{b}^*, \mathfrak{w}^*)$.

Let us first observe that if for some index s and $m < k$ we have $S(M^\ell(\mathbf{a}_{P_s}, \mathbf{a}_{C_s})) = 0$ for $\ell < m$ and $S(M^m(\mathbf{a}_{P_s}, \mathbf{a}_{C_s})) > 0$, then by Corollary 6 we have

$$\phi^{(k)}(\mathfrak{b}^*, \mathfrak{w}_s) \;=\; \phi^{(k)}(\mathbf{a}_{P_s}(\mathbf{a}_{C_s})) \;=\; +\infty$$

implying that $\mathfrak{w}_s$ is a pure and stationary best response of White to $\mathfrak{b}^*$, completing the proof of the claim.

Consequently, for the rest of the proof we can assume that

$$S(M^m(\mathbf{a}_{P_s},\mathbf{a}_{C_s})) \;\le\; 0 \text{ for all } 0\le m<k \text{ and } s=1,2,\dots \tag{23}$$

For a positive integer t let us define $s(t)$ as the largest index s for which $q_s \le t$, and set $Q_t = J_{s(t)} \cap [1,t]$. Let us then note that $F(Q_t)$ is a simple path in G^*.

By repeated applications of Corollary 5 and by the inequalities (23) we can write for any integer T that

$$S(M^k(\mathbf{a}_{[1,T]})) \;\le\; S(M^k(\mathbf{a}_{Q_T})) + \sum_{s=1}^{s(T)} S(M^k(\mathbf{a}_{P_s},\mathbf{a}_{C_s})).$$

Since $F(Q_T)$ and $F(P_s \cup C_s)$ for $s=1,\dots,s(T)$ are all subgraphs of at most n arcs, there is a polynomial $p(n,k)$, independent of T, such that all quantities on the right hand side are bounded form above by $p(n,k)R$. Consequently,

$$\frac{1}{T} S(M^k(\mathbf{a}_{[1,T]})) \;\le\; p(n,k)R$$

follows, implying that $\phi^{(k)}(\mathfrak{b}^*,\mathfrak{w}^*) \le p(n,k)R$, in other words, that $\phi^{(k)}(\mathfrak{b}^*,\mathfrak{w}^*)$ is finite.

We shall show next that in fact we have equalities in (23).

To this end, let us note first that if we have $S(M^m(\mathbf{a}_{P_i},\mathbf{a}_{C_i}))=0$ for all $m=0,\dots,k-1$ and $i=1,2,\dots,\ell-1$, then for every integer $T\ge q_\ell$ we can write by Corollaries 5 and 6 that

$$S(M^k(\mathbf{a}_{[1,T]})) \;=\; S(M^k(\mathbf{a}_{[P_1;Z]})) + \sum_{i=1}^{\ell-1} |C_i|\phi^{(k)}(\mathbf{a}_{P_i}(\mathbf{a}_{C_i})),$$

where $Z=[q_{\ell-1}+1,T]=[1,T]\setminus \cup_{i=1}^{\ell-1}(P_i\cup C_i)$. Assume now that s is the smallest index such that $S(M^m(\mathbf{a}_{P_s},\mathbf{a}_{C_s})) \;<\; 0$ for some $0\le m<k$. Then for all $T>q_s$ we can write, analogously to the above by Corollaries 5 and 6 that

$$\begin{aligned} S(M^k(\mathbf{a}_{[1,T]})) - S(M^k(\mathbf{a}_{[P_1;Z]})) \;&=\; \sum_{i=1}^{s-1} |C_i|\phi^{(k)}(\mathbf{a}_{P_i}(\mathbf{a}_{C_i})) \\ &+ \sum_{m=0}^{k} \binom{|Z|+k-1-m}{k-m} S(M^m(\mathbf{a}_{P_s},\mathbf{a}_{C_s})), \end{aligned}$$

where $Z=[q_{s-1}+1,T]=[1,T]\setminus \cup_{i=1}^{s-1}(P_i\cup C_i)$. We can view the right hand side here as a polynomial of $T=|Z|$ (of degree at most k). According to (23) and our choice of s, the leading term in this polynomial have a negative coefficient. Thus, for all values of T, large enough, we have

$$S(M^k(\mathbf{a}_{[1,T]})) - S(M^k(\mathbf{a}_{[P_1;Z]})) < -1,$$

implying that for the strategy $\widetilde{\mathfrak{w}}$ of WHITE corresponding to $F([P_1;Z])$ we have $\phi^{(k)}(\mathfrak{b}^*,\widetilde{\mathfrak{w}}) > \phi^{(k)}(\mathfrak{b}^*,\mathfrak{w}^*)$, contradicting the choice of $\mathfrak{w}^*$. This contradiction proves that we must have equalities in (23) for all $s=1,2,\dots$ and $m=0,\dots,k-1$.

The above then implies that for all T we have the equality

$$S(M^k(\mathbf{a}_{[1,T]})) \; = \; S(M^k(\mathbf{a}_{Q_T})) + \sum_{i=1}^{s(T)} |C_i| \phi^{(k)}(\mathbf{a}_{P_i}(\mathbf{a}_{C_i})).$$

Since there are only a finite number of different pure stationary strategies of White, let us choose $\mathfrak{w}^{**}$ such that $\phi^{(k)}(\mathfrak{b}^*, \mathfrak{w}^{**})$ is the maximum among all pure stationary strategies of White. In particular, we have $\phi^{(k)}(\mathbf{a}_{P_i}(\mathbf{a}_{C_i})) \leq \phi^{(k)}(\mathfrak{b}^*, \mathfrak{w}^{**})$ for all $i = 1, \ldots, s(T)$.

Then, since $|Q_T| \leq n$ and hence $\sum_{i=1}^{s(T)} |C_i| \geq T - n$, from the above equality we get

$$\begin{aligned} \phi^{(k)}(\mathfrak{b}^*, \mathfrak{w}^*) \; &= \; \liminf_{T\to\infty} \frac{1}{T} S(M^k(\mathbf{a}_{[1,T]})) \\ &\leq \; \liminf_{T\to\infty} \left[\frac{p(n,k)R}{T} + \frac{\sum_{i=1}^{s(T)} |C_i|}{T} \phi^{(k)}(\mathfrak{b}^*, \mathfrak{w}^{**}) \right] \\ &\leq \; \liminf_{T\to\infty} \frac{T-n}{T} \phi^{(k)}(\mathfrak{b}^*, \mathfrak{w}^{**}) \; = \; \phi^{(k)}(\mathfrak{b}^*, \mathfrak{w}^{**}) \end{aligned}$$

Since this inequality holds for every choice of $\epsilon > 0$, it follows that $\mathfrak{w}^{**}$ is a pure stationary best response to $\mathfrak{b}^*$, completing the proof of the theorem. □

Remark 2 *Let us observe that the above proof goes through if we replace* lim inf *in the definition of the payoff function by any convex combination of* lim inf *and* lim sup.

5 Hierarchy of k-Total Rewards

In what follows we shall show that the k-total reward is a special case of the $(k+1)$-total reward. In particular, mean payoff games are a special case of 1-total reward games, and in general of k-total reward games.

To state and prove our results we need to obtain a functional relation between $M(\mathbf{x}(\mathbf{y}))$ and the vectors $\mathbf{x}$ and $\mathbf{y}$. To start our analysis, let us consider the all-one vector $\mathbf{e} = (1, 1, \ldots, 1) \in \mathbb{R}^n$, and compute its iterated M images and the corresponding sums.

Lemma 4 *For* $\mathbf{e} = (1, 1, \ldots, 1) \in \mathbb{R}^n$ *and for every* $k \in \mathbb{Z}_+$ *we have*

$$M^k(\mathbf{e}) \; = \; \left(\binom{k-1+j}{k} \;\middle|\; j = 1, \ldots, n \right) \tag{24}$$

and correspondingly

$$S(M^k(\mathbf{e})) \; = \; \binom{n+k}{k+1}. \tag{25}$$

Proof The above expressions are clearly correct for $k = 0$. We can prove them by induction on k using the binomial identity

$$\sum_{j=0}^{k} \binom{a+j}{a} = \binom{a+k+1}{k} \tag{26}$$

for all integers a and k. $\square$

Lemma 5 *For* $\mathbf{z} = (z_1, z_2, \ldots, z_n) \in \mathbb{R}^n$ *and for every* $k \in \mathbb{Z}_+$ *we have*

$$M^k(\mathbf{z}) = \left(\sum_{j=1}^{i} \binom{k-1+i-j}{k-1} z_j \;\middle|\; i = 1, \ldots, n \right) \tag{27}$$

and correspondingly

$$S(M^k(\mathbf{z})) = \sum_{j=1}^{n} \binom{k+n-j}{k} z_j. \tag{28}$$

Proof For $k = 0$, the above formula with an extended definition of the binomial coefficients [5] shows that M^0 is the identity operator as assumed. For $k = 1$ the above expressions coincide with the definitions of the M and S operators. Thus, by induction on k we can write

$$\begin{aligned} M^{k+1}(\mathbf{z}) = M(M^k(\mathbf{z})) &= \left(\sum_{\ell=1}^{i} \sum_{j=1}^{\ell} \binom{k-1+\ell-j}{k-1} z_j \;\middle|\; i = 1, \ldots, n \right) \\ &= \left(\sum_{j=1}^{i} z_j \sum_{\ell=j}^{i} \binom{k-1+\ell-j}{k-1} \;\middle|\; i = 1, \ldots, n \right) \\ &= \left(\sum_{j=1}^{i} \binom{k+i-j}{k} z_j \;\middle|\; i = 1, \ldots, n \right) \end{aligned}$$

where the second equality follows by (27), and the last one by (26). Finally (28) follows by (27) and (26). $\square$

Let us recall a few combinatorial identities from [5], which we shall need in the sequel.

Lemma 6

$$\sum_{u=0}^{N} (-1)^u \binom{N}{u} \binom{X-u}{R} = \binom{X-N}{X-R}.$$

Lemma 7

$$\sum_{u=0}^{N} \binom{X+u}{u} \binom{Y+N-u}{N-u} = \binom{X+Y+N+1}{N}.$$

Now we are ready to provide an algebraic description of the M operator over the set of lasso sequences.

Theorem 3 *Let us consider integers $p, q > 0$, $k \geq 0$ and a lasso sequence $\mathbf{x}(\mathbf{y}) \in \mathcal{S}_n(R)$ with $\mathbf{x} \in \mathbb{Z}^p$ and $\mathbf{y} \in \mathbb{Z}^q$, satisfying*

$$\phi^{(\ell)}(\mathbf{x}(\mathbf{y})) = 0 \quad \textit{for} \quad 0 \leq \ell \leq k-1. \tag{29}$$

Then, we have

$$M^k(\mathbf{x}(\mathbf{y})) \;=\; M^k(\mathbf{x}) \left(\sum_{\ell=1}^{k} S\left(M^{k-\ell}(\mathbf{x})\right) M^{\ell-1}(\mathbf{e}) \;+\; M^k(\mathbf{y}) \right) \tag{30}$$

where $\mathbf{e} = (1, 1, \ldots, 1) \in \mathbb{Z}^q$, and

$$\begin{aligned} \phi^{(k)}(\mathbf{x}(\mathbf{y})) \;&=\; A\left(\sum_{\ell=1}^{k} S\left(M^{k-\ell}(\mathbf{x})\right) M^{\ell-1}(\mathbf{e}) \;+\; M^k(\mathbf{y}) \right) \\ &=\; \frac{1}{q} S\left(\sum_{\ell=1}^{k} S\left(M^{k-\ell}(\mathbf{x})\right) M^{\ell-1}(\mathbf{e}) \;+\; M^k(\mathbf{y}) \right) \\ &=\; \frac{1}{q}\sum_{j=1}^{p} x_j \left[\binom{q+p-j+k}{k} - \binom{p-j+k}{k} \right] \;+\; \frac{1}{q}\sum_{i=1}^{q} \binom{k+q-i}{k} y_i. \end{aligned} \tag{31}$$

Proof Let us note first that for $k = 0$ the condition (29) is empty, and the summation in (30) is an empty sum, yielding $M^0(\mathbf{x}(\mathbf{y})) = \mathbf{x}(\mathbf{y})$. Furthermore, for $k = 1$ we have by Corollary 1 that $M(\mathbf{x}(\mathbf{y})) = M(\mathbf{x})\,(S(\mathbf{x})\mathbf{e} + M(\mathbf{y}))$, in agreement with (30), since condition (29) is equivalent with saying $S(\mathbf{y}) = 0$ by definition (8). Thus, (30) follows by induction on k using the linearity of M as in Lemma 2.

Finally, (31) follows from (30) after applying the S operator to both sides of (30) yielding the second line of (31). Then, by using the linearity of S by Lemma 2 and applying Lemmas 4 and 5 we get

$$\begin{aligned} &\frac{1}{q}\sum_{\ell=1}^{k} \binom{q+\ell-1}{\ell} \sum_{j=1}^{p} \binom{k-\ell+p-j}{k-\ell} x_j \;+\; \frac{1}{q}\sum_{i=1}^{q} \binom{k+q-i}{k} y_i \\ =\; &\frac{1}{q}\sum_{j=1}^{p} x_j \left[\sum_{\ell=1}^{k} \binom{q+\ell-1}{\ell} \binom{k-\ell+p-j}{k-\ell} \right] \;+\; \frac{1}{q}\sum_{i=1}^{q} \binom{k+q-i}{k} y_i. \end{aligned}$$

Finally, applying Lemma 7 with $u = \ell$, $N = k$, $X = q-1$ and $Y = p-j$ and subtracting the $u = 0$ term from both sides we get the last line of (31). $\square$

Remark 3 *Formula (31) shows that the value $\phi^{(k)}(\mathbf{x}(\mathbf{y}))$ is a linear combination of the components of $\mathbf{x}$ and $\mathbf{y}$. Furthermore, it can be verified that these linear combinations for $k = 0, 1, \ldots, n-1$, are linearly independent. Consequently, if ϕ^n takes a finite value on a lasso sequence, then all local rewards on this sequence must be equal to 0. On the other hand, there are BW-games such that the $\phi^{(n-1)}$ value, from a certain starting position, is finite and different from zero.*

For the main claim of this section, we introduce a split operation on sequences: to a given sequence $\mathbf{x} = (x_1, x_2, \ldots)$, we associate $\mathbf{x}^{(1)} = (x_1, -x_1, x_2, -x_2, \ldots)$.

We need an additional technical lemma.

Lemma 8 *For every integer k there exist reals α_j, $j = 0, 1, \ldots, k$ such that*

$$\binom{2X+k+1}{k} = \sum_{j=0}^{k} \alpha_j \binom{X+j}{j}$$

holds for all integers X. In particular, we have $\alpha_k = 2^k$.

Proof Let us note that $g_j(X) = \binom{X+j}{j}$ is a polynomial of X of degree exactly j, for all $j = 0, \ldots, k$. Thus, the polynomials g_j, $j = 0, \ldots, k$ form a basis for polynomials in X of degree at most k. Hence, every polynomial in X of degree k can be expressed as a linear combination of the g_j, $j = 0, \ldots, k$ functions, proving the existence of the α_j, $j = 0, \ldots, k$ coefficients.To see that $\alpha_k = 2^k$, we need to compare the coefficients of X^k on both sides. □

Remark 4 *One can also obtain the explicit values of the coefficients α_j in Lemma 8; for completeness, we give these in Lemma 10 in the appendix.*

The main result in this section is the following.

Theorem 4 *For a nonnegative integer k and any lasso sequence $\mathbf{x}(\mathbf{y})$ we have*

$$\phi^{(k+1)}\left(\mathbf{x}^{(1)}\left(\mathbf{y}^{(1)}\right)\right) = 2^{k-1}\phi^{(k)}\left(\mathbf{x}(\mathbf{y})\right). \tag{32}$$

Proof We prove this claim by induction on k. For $k = 0$, let us note that $M(x_1, -x_1, x_2, -x_2, \ldots) = (x_1, 0, x_2, 0, \ldots)$ implying that $\phi^1(\mathbf{x}^{(1)}) = \frac{1}{2}\phi^{(0)}(\mathbf{x})$.

Let us next assume that we have the equalities

$$\phi^{(\ell+1)}\left(\mathbf{x}^{(1)}\left(\mathbf{y}^{(1)}\right)\right) = 2^{\ell-1}\phi^{(\ell)}\left(\mathbf{x}(\mathbf{y})\right)$$

for all $0 \leq \ell < k$. Thus, in particular, the signs of $\phi^{(\ell+1)}\left(\mathbf{x}^{(1)}\left(\mathbf{y}^{(1)}\right)\right)$ and $\phi^{(\ell)}\left(\mathbf{x}(\mathbf{y})\right)$ are the same for all $\ell < k$. Therefore, by Fact 4 and the definition of $\phi^{(k)}$, both sides of (32) are simultaneously equal to $\pm\infty$, whenever $\phi^{(k)}\left(\mathbf{x}^{(1)}\left(\mathbf{y}^{(1)}\right)\right) \neq 0$. Hence it is enough to prove the claim for lasso sequences satisfying

$$\phi^{(\ell+1)}\left(\mathbf{x}^{(1)}\left(\mathbf{y}^{(1)}\right)\right) \;=\; \phi^{(\ell)}\left(\mathbf{x}(\mathbf{y})\right) \;=\; 0 \quad \text{for all} \quad \ell = 0, 1, \ldots, k-1. \tag{33}$$

Under these conditions Theorem 3 can be applied and we get for the split sequence $\mathbf{x}^{(1)}\left(\mathbf{y}^{(1)}\right)$ that $2q\phi^{(k+1)}\left(\mathbf{x}^{(1)}\left(\mathbf{y}^{(1)}\right)\right)$ is equal to

$$\sum_{j=1}^{p} x_j \sum_{r=0}^{1} (-1)^r \left[\binom{2(q+p-j)+1-r+k+1}{k+1} - \binom{2(p-j)+1-r+k+1}{k+1}\right]$$
$$+ \sum_{i=1}^{q} y_i \sum_{r=0}^{1} (-1)^r \binom{2(q-i)+1-r+k+1}{k+1}$$
$$= \sum_{j=1}^{p} x_j \left[\binom{2(q+p-j)+k+1}{k} - \binom{2(p-j)+k+1}{k}\right] + \sum_{i=1}^{q} y_i \binom{2(q-i)+k+1}{k}.$$

Let us also note that under conditions (33) we can apply Theorem 3 to $\phi^{(\ell)}\left(\mathbf{x}(\mathbf{y})\right)$ and express it as in (31) for all $\ell < k$. Let us finally note that using these expressions and using Lemma 8 three times, with $X = q+p-j$, $X = p-j$ and $X = q-i$ we get

$$2\phi^{(k+1)}\left(\mathbf{x}^{(1)}\left(\mathbf{y}^{(1)}\right)\right) \;=\; \sum_{j=0}^{k} \alpha_j \phi^{(j)}(\mathbf{x}(\mathbf{y})).$$

By condition (33) this further simplifies to

$$2\phi^{(k+1)}\left(\mathbf{x}^{(1)}\left(\mathbf{y}^{(1)}\right)\right) \;=\; \alpha_k \phi^{(k)}\left(\mathbf{x}(\mathbf{y})\right) \;=\; 2^k \phi^{(k)}\left(\mathbf{x}(\mathbf{y})\right)$$

from which the statement follows. □

Corollary 7 *k-total reward BW-games are special cases of $(k+1)$-total reward BW-games.*

Proof Let us consider an arbitrary BW-game $\Gamma = (G, B, W, r)$ and let us define its split, denoted by $\widetilde{\Gamma} = (\widetilde{G}, \widetilde{B}, \widetilde{W}, \widetilde{r})$, as the game obtained from Γ by subdividing each arc of local reward $r(u,v)$ by a vertex $w = w_{uv}$ and defining $\widetilde{r}(u,w) = r(u,v)$ and $\widetilde{r}(w,v) = -r(u,v)$. Clearly, there is a one-to-one correspondence between the strategies in Γ and $\widetilde{\Gamma}$. Then, it is also clear that the expected reward sequence arising from a play in $\widetilde{\Gamma}$ is the split sequence of the reward sequence arising from the corresponding play in Γ. Theorem 4 implies that the games $(G, B, W, r, \phi^{(k)})$ and $(\widetilde{G}, \widetilde{B}, \widetilde{W}, \widetilde{r}, \phi^{(k+1)})$ are equivalent; more precisely, the effective values of the corresponding plays are equal up to a multiplicative factor of 2^{k-1}. □

Remark 5 *Due to an example of a non-zero sum mean payoff BW-game [6] that does not have a Nash equilibrium, we can conclude that, for any* $k \in \mathbb{Z}_+$*, there are non-zero sum* k*-total reward games that have no Nash equilibria either.*

Let us note that the above reduction creates games in which every cycle has zero total length. In fact, we have no NE-free examples with 1*-total effective payoff and without directed cycles of zero total length.*

It is also interesting to notice that in [19] the 1*-total reward was viewed as a refinement of the* 0*-total one. The above result shows that in fact every* 0*-total game can be viewed as a special* 1*-total game. Yet, a polynomial reduction in the other direction is not known.*

References

[1] H. Björklund and S. Vorobyov. Combinatorial structure and randomized subexponential algorithms for infinite games. *Theoretical Computer Science*, 349(3):347–360, 2005.

[2] E. Boros, K. Elbassioni, V. Gurvich, and K. Makino. Markov decision processes and stochastic games with total effective payoff. RRR-4-2014, RUTCOR, Rutgers University, 2014.

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

[4] D. Gillette. Stochastic games with zero stop probabilities. In *Contributions to the Theory of Games, Vol. III*, volume 39 of *Annals of Mathematics Studies*, pages 179–187. Princeton, N.J., 1957.

[5] Henry W. Gould. Combinatorial indentities, 2012. `http://www.math.wvu.edu/ gould/`.

[6] V. Gurvich. A stochastic game with complete information and without equilibrium situations in pure stationary strategies. *Russian Mathematical Surveys*, 43(2):171–172, 1988.

[7] V.A. Gurvich, A.V. Karzanov, and L.G. Khachiyan. Cyclic games and an algorithm to find minimax cycle means in directed graphs. *USSR Comput. Math. Math. Phys.*, 28:8591, 1988.

[8] 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.

[9] G.H. Hardy and J.E. Littlewood. Notes on the theory of series (xvi): two tauberian theorems. *J. of London Mathematical Society*, 6:281–286, 1931.

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

[11] Thomas M. Liggett and Steven A. Lippman. Stochastic games with perfect information and time average payoff. *SIAM Review*, 11(4):604–607, 1969.

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

[13] Rolf H. Möhring, Martin Skutella, and Frederik Stork. Scheduling with and/or precedence constraints. *SIAM J. Comput.*, 33(2):393–415, 2004.

[14] H. Moulin. Extension of two person zero sum games. *Journal of Mathematical Analysis and Aplication*, 55:2:490–507, 1976.

[15] H. Moulin. Prolongement des jeux à deux joueurs de somme nulle. une théorie abstraite des duels. *Mémoires de la Soc. Math. France*, 45:5–111, 1976.

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

[17] L. Shapley. Stochastic games. *Proc. Nat. Acad. Sci. USA*, 39:1095–1100, 1953.

[18] F. Thuijsman and O. J. Vrieze. The bad match, a total reward stochastic game. *Operations Research Spektrum*, 9:93–99, 1987.

[19] F. Thuijsman and O. J. Vrieze. Total reward stochastic games and sensitive average reward strategies. *Journal of Optimization Theory and Applications*, 98:175–196, 1998.

[20] S. Vorobyov. Cyclic games and linear programming. *Discrete Applied Mathematics*, 156(11):2195–2231, 2008.

[21] Uri Zwick and Mike Paterson. The complexity of mean payoff games on graphs. *Theoretical Computer Science*, 158(1-2):343 – 359, 1996.

A Stronger Version of Lemma 1 for $k = 1$

Lemma 9 *Let ε be a positive constant, $0 < \varepsilon < \min\{\frac{1}{R}, 1 - \frac{1}{2^{1/n}}\}$, and let $\mathbf{x}(\mathbf{y})$ be a reward sequence from the set $\mathcal{S}^*$ such that $A(\mathbf{y}) = 0$. Then, we have*

$$\left|\varepsilon\phi^{(1)}(\mathbf{x}(\mathbf{y})) - \phi_{1-\varepsilon}(\mathbf{x}(\mathbf{y}))\right| \;\leq\; 2\varepsilon^2 n^2 R.$$

Proof Using $\beta = 1 - \varepsilon$ we can write

$$\phi^{(1)}(\mathbf{x}(\mathbf{y})) - \frac{\phi_\beta(\mathbf{x}(\mathbf{y}))}{\varepsilon} = \sum_{i=1}^{p} x_i(1 - \beta^{i-1}) + \sum_{j=1}^{q} y_j \left(\frac{q-j}{q} - \frac{\beta^{p+j-1}}{1-\beta^q}\right).$$

Since $A(\mathbf{y}) = 0$ is assumed, we have $\left(1 - \frac{\beta^{p-1}}{1-\beta^q}\right) \sum_{j=1}^{q} y_j = 0$ and thus we can rewrite the right hand side as

$$\phi^{(1)}(\mathbf{x}(\mathbf{y})) - \frac{\phi_\beta(\mathbf{x}(\mathbf{y}))}{\varepsilon} = \varepsilon \sum_{i=2}^{p} x_i(1 + \beta + \cdots + \beta^{i-2}) + \sum_{j=1}^{q} y_j \left(\frac{\beta^{p-1}(1-\beta^j)}{1-\beta^q} - \frac{j}{q} \right).$$

Here we have $(1 + \beta + \cdots + \beta^{i-2}) \leq i - 1$ and

$$\left(\frac{\beta^{p-1}(1-\beta^j)}{1-\beta^q} - \frac{j}{q} \right) \leq \frac{2\varepsilon n j}{q}$$

from which, together with $|x_i| \leq R$, $i = 1, \ldots, p$ and $|y_j| \leq R$, $j = 1, \ldots, q$, one can derive

$$\left| \phi^{(1)}(\mathbf{x}(\mathbf{y})) - \frac{\phi_\beta(\mathbf{x}(\mathbf{y}))}{\varepsilon} \right| \leq 2\varepsilon n^2 R$$

implying the claim of the lemma. □

B Explicit coefficients in Lemma 8

Lemma 10 *Let $A, k \geq 0$ be integers. Then*

$$\sum_{j=0}^{\lfloor k/2 \rfloor} (-1)^j 2^{k-2j} \binom{k-j}{j} \binom{A+k-j}{k-j} = \binom{2A+k+1}{k}. \tag{34}$$

Proof Let $g(A, k)$ denote the summation on the left hand side. Thus, we want to show that $g(A, k) = \binom{2A+k+1}{k}$. We apply induction on $A \geq 0$. For $A = 0$, we have $g(0, k) = k + 1$ by (2.4) in [5], and hence (34) holds in this case. We assume now that it holds for A, and verify it for $A + 1$.

First we can show the following claim.

Claim 1

$$g(A+1, k) = \frac{1}{A+1} \sum_{j=0}^{\lfloor k/2 \rfloor} (A + k + 1 - 2j) g(A, k - 2j). \tag{35}$$

Proof

$$\begin{aligned} g(A+1,k) &= \sum_{j=0}^{\lfloor k/2 \rfloor} (-1)^j 2^{k-2j} \binom{k-j}{j} \binom{A+1+k-j}{k-j} \\ &= \sum_{j=0}^{\lfloor k/2 \rfloor} (-1)^j 2^{k-2j} \binom{k-j}{j} \binom{A+k-j}{k-j} \cdot \frac{A+k+1-j}{A+1} \\ &= \frac{A+k+1}{A+1} g(A,k) + h(k), \end{aligned}$$

where

$$\begin{aligned} h(k) &:= -\sum_{j=0}^{\lfloor k/2\rfloor}(-1)^j 2^{k-2j}\binom{k-j}{j}\binom{A+k-j}{k-j}\cdot\frac{j}{A+1} \\ &= \sum_{j=0}^{\lfloor (k-2)/2\rfloor}(-1)^j 2^{(k-2)-2j}\binom{(k-2)-j}{j}\binom{A+(k-2)-j}{(k-2)-j}\cdot\frac{A+k-1-j}{A+1} \\ &= \frac{A+k-1}{A+1}g(A,k-2)+h(k-2). \end{aligned} \tag{36}$$

The claim follows by iterative application of (36). □

By induction, we have $g(A,k-2j)=\binom{2A+k-2j+1}{k-2j}$. Thus, it remains to prove the following claim.

Claim 2

$$\frac{1}{A+1}\sum_{j=0}^{\lfloor k/2\rfloor}(A+k+1-2j)\binom{2A+k-2j+1}{k-2j}=\binom{2(A+1)+k+1}{k}. \tag{37}$$

Proof We use induction on k. The base cases $k=0$ and 1 are easily verified. Assume the statement holds for all integers less than k. Denote by $f(k)$ the left hand side of (37). Then

$$\begin{aligned} f(k) &= \frac{1}{A+1}(A+k+1)\binom{2A+k+1}{k}+f(k-2) \\ &= \frac{1}{A+1}(A+k+1)\binom{2A+k+1}{k}+\binom{2(A+1)+(k-2)+1}{k-2} \\ &= \frac{(2A+k+1)!}{k!(2A+3)!}[2(2A+3)(A+k+1)+k(k-1)] \\ &= \binom{2(A+1)+k+1}{k}. \end{aligned}$$

□