The complexity of interior point methods for solving discounted turn-based stochastic games

Thomas Dueholm Hansen[*†] Rasmus Ibsen-Jensen[*]

Abstract

We study the problem of solving discounted, two player, turn based, stochastic games (2TBSGs). Jurdziński and Savani showed that in the case of deterministic games the problem can be reduced to solving P-matrix linear complementarity problems (LCPs). We show that the same reduction also works for general 2TBSGs. This implies that a number of interior point methods can be used to solve 2TBSGs. We consider two such algorithms: the unified interior point method of Kojima, Megiddo, Noma, and Yoshise, and the interior point potential reduction algorithm of Kojima, Megiddo, and Ye. The algorithms run in time $O((1+\kappa)n^{3.5}L)$ and $O(\frac{-\delta}{\theta}n^4 \log \epsilon^{-1})$, respectively, when applied to an LCP defined by an $n \times n$ matrix M that can be described with L bits, and where the potential reduction algorithm returns an ϵ-optimal solution. The parameters κ, δ, and θ depend on the matrix M. We show that for 2TBSGs with n states and discount factor γ we get $\kappa = \Theta(\frac{n}{(1-\gamma)^2})$, $-\delta = \Theta(\frac{\sqrt{n}}{1-\gamma})$, and $1/\theta = \Theta(\frac{n}{(1-\gamma)^2})$ in the worst case. The lower bounds for κ, $-\delta$, and $1/\theta$ are all obtained using the same family of deterministic games.

1 Introduction

Two-player turn-based stochastic games (2TBSGs). A *two-player turn-based stochastic game* (2TBSG) is a game played by two players (Player 1 and Player 2) on a finite state graph for an infinite number of rounds. The graph is partitioned into two sets of states S^1 (belonging to Player 1) and S^2 (belonging to Player 2). Whenever the current state i is from S^k, Player k chooses an action a emanating from state i, and the next state is then given by a probability distribution, depending on a. In each round there is a probability of $1-\gamma > 0$ of ending the game, where γ is the *discount factor* of the game. Every action has an associated cost. The objective of Player 1 is to *minimize* the expected sum of costs, and the objective of Player 2 is to *maximize* the expected sum of costs, i.e., the game is a zero-sum game. Our results will be for the case when all states have 2 actions.

The class of (turn-based) stochastic games was introduced by Shapley [19] in 1953, and it has received much attention over the following decades. For books on the subject see, e.g., Neyman

[*]Department of Computer Science, Aarhus University, Denmark. E-mail: {tdh, rij}@cs.au.dk. Supported by the Sino-Danish Center for the Theory of Interactive Computation, funded by the Danish National Research Foundation and the National Science Foundation of China (under the grant 61061130540); and by the Center for research in the Foundations of Electronic Markets (CFEM), supported by the Danish Strategic Research Council.

[†]Thomas Dueholm Hansen is a recipient of the Google Europe Fellowship in Game Theory, and this research is supported in part by this Google Fellowship. Part of the work was performed at the School of Computer Science, Tel Aviv University, Israel, with the support of The Danish Council for Independent Research | Natural Sciences (grant no. 12-126512).

and Sorin [16] and Filar and Vrieze [5]. Shapley showed that states in such games have a *value* that can be enforced by both players (*determinacy*). We will in this paper consider the problem of *solving* such games, that is, for each state i finding the value of that state.

Classical algorithms for solving 2TBSGs. 2TBSGs form an intriguing class of games whose status in many ways resembles that of linear programming 40 years ago. They can be solved efficiently with *strategy iteration* algorithms, resembling the simplex method for linear programming, but no polynomial time algorithm is known. Strategy iteration algorithms were first described by Rao *et al.* [17]. Hansen, Miltersen, and Zwick [8] recently showed that the standard strategy iteration algorithm solves 2TBSGs with a fixed discount, γ, in *strongly* polynomial time. Prior to this result a polynomial bound by Littman [15] was known for the case when γ is fixed. Littman showed that Shapley's [19] *value iteration* algorithm can be used to solve discounted 2TBSGs in time $O(\frac{nmL}{1-\gamma}\log\frac{1}{1-\gamma})$, where n is the number of states, m is the number of actions, and L is the number of bits needed to represent the game. For a more thorough introduction to the background of the problem we refer to Hansen *et al.* [8] and the references therein.

Interior point methods. One may hope that a polynomial time algorithm for solving 2TBSGs in the general case, when the discount factor γ is not fixed (i.e., when it is given as part of the input), can be obtained through the use of *interior point methods*. This was also suggested by Jurdziński and Savani [10] and Hansen *et al.* [8]. The first interior point method was introduced by Karmarkar [11] in 1984 to solve linear programs in polynomial time. Since then the technique has been studied extensively and applied in other contexts. See, e.g., Ye [20]. In particular, interior point methods can be used to solve *P-matrix linear complementarity problems*, which, in turn, can be used to solve 2TBSGs. This will be the focus of the present paper.

P-matrix linear complementarity problems (LCPs). A linear complementarity problem (LCP) is defined as follows: Given an $(n \times n)$-matrix M and a vector $\mathbf{q} \in \mathbb{R}^n$, find two vectors $\mathbf{w}, \mathbf{z} \in \mathbb{R}^n$, such that $\mathbf{w} = \mathbf{q} + M\mathbf{z}$ and $\mathbf{w}^\mathsf{T}\mathbf{z} = 0$ and $\mathbf{w}, \mathbf{z} \geq \mathbf{0}$. LCPs have also received much attention. For books on the subject see, e.g., Cottle *et al.* [4] and Ye [20].

Jurdziński and Savani [10] showed that solving a *deterministic* 2TBSG G, i.e., every action leads to a single state with probability 1, can be reduced to solving an LCP $(M, \mathbf{q})$. Gärtner and Rüst [6] gave a similar reduction from simple stochastic games; a class of games that is polynomially equivalent to 2TBSGs (see [1]). Moreover, Jurdziński and Savani [10], and Gärtner and Rüst [6], showed that the resulting matrix M is a *P-matrix* (i.e., all principal sub-matrices have a positive determinant). We show that the reduction of Jurdziński and Savani also works for general 2TBSGs, and that the resulting matrix M is again a P-matrix.

Krishnamurthy *et al.* [14] recently gave a survey on various stochastic games and LCP formulations of those.

The unified interior point method. There exist various interior point methods for solving P-matrix LCPs. One algorithm that we consider in this paper is the unified interior point method of Kojima, Megiddo, Noma, and Yoshise [12]. The unified interior point method solves an LCP whose matrix $M \in \mathbb{R}^{n\times n}$ is a *$P_*(\kappa)$-matrix* in time $O((1+\kappa)n^{3.5}L)$, where L is the number of bits needed to describe M. A matrix M is a $P_*(\kappa)$-matrix, for $\kappa \geq 0$, if and only if for all vectors $\mathbf{x} \in \mathbb{R}^n$, we have that $\mathbf{x}^\mathsf{T}(M\mathbf{x}) + 4\kappa\sum_{i\in\delta_+(M)}\mathbf{x}_i(M\mathbf{x})_i \geq 0$, where $\delta_+(M) = \{i \in [n] \mid \mathbf{x}_i(M\mathbf{x})_i > 0\}$. If M is a P-matrix then it is also a $P_*(\kappa)$-matrix for some $\kappa \geq 0$. Hence, the algorithm can be used to solve 2TBSGs.

Following the work of Kojima *et al.* [12], many algorithms with complexity polynomial in κ, L, and n have been introduced. For recent examples see, e.g., [3, 2, 9].

An interior point potential reduction algorithm. The second interior point method that we consider in this paper is the potential reduction algorithm of Kojima, Megiddo, and Ye [13]. See also Ye [20]. The potential reduction algorithm is an interior point method that takes as input a P-matrix LCP and a parameter $\epsilon > 0$, and produces an approximate solution $\mathbf{w}, \mathbf{z}$, such that $\mathbf{w}^\mathsf{T}\mathbf{z} < \epsilon$, $\mathbf{w} = \mathbf{q} + M\mathbf{z}$, and $\mathbf{w}, \mathbf{z} \geq \mathbf{0}$. The running time of the algorithm is $O(\frac{-\delta}{\theta} n^4 \log \epsilon^{-1})$, where δ is the least eigenvalue of $\frac{M+M^\mathsf{T}}{2}$, and θ is the positive P-matrix number of M, that is, $\theta = \min_{\|\mathbf{x}\|_2=1} \max_{i\in\{1,\dots,n\}} \mathbf{x}_i(M\mathbf{x})_i$. We refer to $\frac{-\delta}{\theta}$ as the condition number of M. The analysis involving the condition number appears in Ye [20].

In his ph.d. thesis, Rüst [18] shows that there exists a simple stochastic game for which the P-matrix LCPs resulting from the reduction of Gärtner and Rüst [6] has a large condition number. The example of Rüst contains a parameter that can essentially be viewed as the discount factor γ for 2TBSGs, and he shows that the condition number can depend linearly on $\frac{1}{1-\gamma}$. To be more precise, Rüst [18] shows that the matrix M resulting from the reduction of Gärtner and Rüst [6] has positive P-matrix number smaller than 1, and that the smallest eigenvalue of the matrix $\frac{M+M^\mathsf{T}}{2}$ is $-\Omega\big(\frac{1}{1-\gamma}\big)$. This bound can be viewed as a precursor for some of our results.

1.1 Our contributions

Our contributions are as follows. We show that the reduction by Jurdziński and Savani [10] from deterministic 2TBSGs to P-matrix LCPs generalizes to 2TBSGs without modification. Although the reduction is the same we provide an alternative proof that the resulting matrix is a P-matrix. Furthermore, let G be any 2TBSG with n states and let M_G be the matrix obtained from the reduction of Jurdziński and Savani [10].

(*i*) We show that M_G is a $P_*(\kappa)$-matrix for $\kappa = \frac{n}{(1-\gamma)^2}$. This implies that the running time of the unified interior point method of Kojima *et al.* [12] for 2TBSGs is at most $O(\frac{n^{4.5}L}{(1-\gamma)^2})$. We also show that there exists a family of 2TBSGs, G_n, such that the corresponding matrices, M_{G_n}, are not $P_*(\kappa)$-matrices for $\kappa = \Omega\big(\frac{n}{(1-\gamma)^2}\big)$.

(*ii*) We show that the matrix $\frac{M_G+M_G^\mathsf{T}}{2}$ has smallest eigenvalue at least $-O\big(\frac{\sqrt{n}}{1-\gamma}\big)$. We also show that there exists a family of 2TBSGs, G_n, such that the corresponding matrices $\frac{M_{G_n}+M_{G_n}^\mathsf{T}}{2}$ have smallest eigenvalue less than $-\Omega\big(\frac{\sqrt{n}}{1-\gamma}\big)$.

(*iii*) Finally, we show that the positive P-matrix number $\theta(M_G)$ is at least $\Omega\big(\frac{(1-\gamma)^2}{n}\big)$. We also show that there exists a familiy of 2TBSGs, G_n, such that the corresponding matrices M_{G_n} have positive P-matrix number, $\theta(M_{G_n})$, at most $O\big(\frac{(1-\gamma)^2}{n}\big)$.

Notice that (*ii*) and (*iii*) together imply that the running time of the potential reduction algorithm of Kojima *et al.* [13] for 2TBSGs is at most $O(\frac{n^{5.5}\log\epsilon^{-1}}{(1-\gamma)^3})$. The familiy of 2TBSGs G_n mentioned in (*i*), (*ii*), and (*iii*) is, in fact, the same. Hence, we get matching upper and lower bounds for the parameters of both the unified interior point method and the potential reduction algorithm. Also, the games G_n are deterministic, so the same lower bounds hold for deterministic 2TBSGs.

It should be noted that although our results for existing interior point methods for solving 2TBSGs are negative, it is still possible that other (possibly new) interior point methods can solve 2TBSGs efficiently. In fact, we believe that this remains an important question for future research.

1.2 Overview

In Section 2 we formally introduce the various classes of problems under consideration. More precisely, in Subsection 2.1 we define LCPs, and in Subsection 2.2 we define 2TBSGs. In Subsection 2.3, we show that the reduction by Jurdziński and Savani [10] from *deterministic* 2TBSGs to P-matrix LCPs generalizes to general 2TBSGs. In Section 3 we estimate the κ for which the matrices of 2TBSGs are $P_*(\kappa)$-matrices, thus, giving a bound on the running time of the unified interior point method of Kojima *et al.* [12]. In Section 4 we bound the smallest eigenvalue and the positive P-matrix number, thus giving a bound on the running time of the potential reduction algorithm of Kojima *et al.* [13].

2 Preliminaries

2.1 Linear complementarity problems

Definition 1 (Linear complementarity problems) *A* linear complementarity problem *(LCP) is a pair* $(M, \mathbf{q})$*, where* M *is an* $(n \times n)$*-matrix and* $\mathbf{q}$ *is an* n*-vector. A* solution *to the LCP* $(M, \mathbf{q})$ *is a pair of vectors* $(\mathbf{w}, \mathbf{z}) \in \mathbb{R}^n$ *such that:*

$$\begin{aligned} \mathbf{w} &= \mathbf{q} + M\mathbf{z} \\ \mathbf{w}^\mathsf{T}\mathbf{z} &= 0 \\ \mathbf{w}, \mathbf{z} &\geq \mathbf{0} \ . \end{aligned}$$

We will now define various types of matrices for which interior point methods are known to solve the corresponding LCPs.

Definition 2 (P-matrix) *A matrix* $M \in \mathbb{R}^{n\times n}$ *is a* P-matrix *if and only if all principal submatrices have a positive determinant.*

The following lemma gives an alternative definition of P-matrices (see, e.g., [4, Theorem 3.3.4]).

Lemma 3 *A matrix* $M \in \mathbb{R}^{n\times n}$ *is a* P-matrix *if and only if for all* n*-vectors* $\mathbf{x} \neq \mathbf{0}$ *there is an* $i \in [n]$ *such that* $\mathbf{x}_i(M\mathbf{x})_i > 0$*.*

Definition 4 (Positive P-matrix number) *For a matrix* $M \in \mathbb{R}^{n\times n}$*, the positive* P*-matrix number is*

$$\theta(M) \;=\; \min_{\|\mathbf{x}\|_2=1} \; \max_{i\in[n]} \; \mathbf{x}_i(M\mathbf{x})_i \ .$$

Note that, according to Lemma 3, $\theta(M) > 0$ if and only if M is a P-matrix.

Definition 5 ($P_*(\kappa)$-matrix) *A matrix $M \in \mathbb{R}^{n\times n}$ is a* $P_*(\kappa)$-matrix*, for $\kappa \geq 0$, if and only if for all vectors $\mathbf{x} \in \mathbb{R}^n$:*

$$\sum_{i\in\delta_-(M)} \mathbf{x}_i(M\mathbf{x})_i \ + \ (1+4\kappa)\sum_{i\in\delta_+(M)} \mathbf{x}_i(M\mathbf{x})_i \ \geq \ 0 \ ,$$

where $\delta_-(M) = \{i \in [n] \mid \mathbf{x}_i(M\mathbf{x})_i < 0\}$ and $\delta_+(M) = \{i \in [n] \mid \mathbf{x}_i(M\mathbf{x})_i > 0\}$. We say that M is a P_*-matrix *if and only if it is a $P_*(\kappa)$-matrix for some $\kappa \geq 0$.*

Kojima *et al.* [12] showed that every P-matrix is also a P_*-matrix. By definition, a matrix M is a $P_*(0)$-matrix if and only if it is *positive semi-definite*. The set of symmetric P-matrices is exactly the set of positive semi-definite matrices.

2.2 Two-player turn-based stochastic games

Definition 6 (Two-player turn-based stochastic games) *A* two-player turn-based stochastic game *(2TBSG) is a tuple, $G = (S^1, S^2, (A_i)_{i\in S^1\cup S^2}, p, c, \gamma)$, where*

- *S^k, for $k \in \{1,2\}$, is the set of* states *belonging to Player k. We let $S = S^1 \cup S^2$ be the set of all states, and we assume that S^1 and S^2 are disjoint.*
- *A_i, for $i \in S$, is the set of* actions *applicable from state i. We let $A = \bigcup_{i\in S} A_i$ be the set of all actions. We assume that A_i and A_j are disjoint for $i \neq j$, and that $A_i \neq \emptyset$ for all $i \in S$.*
- *$p : A \to \Delta(S)$ is a map from actions to probability distributions over states.*
- *$c : A \to \mathbb{R}$ is a function that assigns a* cost *to every action.*
- *$\gamma < 1$ is a (positive)* discount factor.

We let $n = |S|$ and $m = |A|$. Furthermore, we let $A^k = \bigcup_{i\in S^k} A_i$, for $k \in \{1,2\}$. Figure 1 shows an example of a simple 2TBSG. The large circles and squares represent the states controlled by Player 1 and 2, respectively. The edges leaving the states represent actions. The cost of an action is shown inside the corresponding diamond shaped square, and the probability distribution associated with the action is shown by labels on the edges leaving the diamond shaped square.

We say that an action a is *deterministic* if it moves to a single state with probability 1, i.e., if $p(a)_j = 1$ for some $j \in S$. If all the actions of a 2TBSG G are deterministic we say that G is deterministic.

Plays and outcomes. A 2TBSG is played as follows. At the beginning of a *play* a pebble is placed on some state $i_0 \in S$. Whenever the pebble is moved to a state $i \in S^k$, Player k chooses an action $a \in A_i$ and the pebble is moved at random according to the probability distribution $p(a)$ to a new state j. Let a^t be the t'th chosen action for every $t \geq 0$. Then the *outcome* of the play, paid by Player 1 to Player 2 is $\sum_{t\geq 0} \gamma^t \cdot c(a^t)$.

We will now give a way to explicitly represent a 2TBSG using vectors and matrices. It will later simplify the notation in our constructions and proofs. Figure 1 also shows such a representation of a 2TBSG.

Definition 7 (Matrix representation) *Let $G = (S^1, S^2, (A_i)_{i\in S^1\cup S^2}, p, c, \gamma)$ be a 2TBSG. Assume WLOG that $S = [n] = \{1, \ldots, n\}$ and $A = [m] = \{1, \ldots, m\}$.*

- *We define the* probability matrix $P \in \mathbb{R}^{m \times n}$ *by* $P_{a,i} = (p(a))_i$, *for all* $a \in A$ *and* $i \in S$.
- *We define the* cost vector $\mathbf{c} \in \mathbb{R}^m$ *by* $\mathbf{c}_a = c(a)$, *for all* $a \in A$.
- *We define the* source matrix $J \in \{0,1\}^{m \times n}$ *by* $J_{a,i} = 1$ *if and only if* $a \in A_i$, *for all* $a \in A$ *and* $i \in S$.
- *We define the* ownership matrix $\mathcal{I} \in \{-1,0,1\}^{n \times n}$ *by* $\mathcal{I}_{i,j} = 0$ *if* $i \neq j$, $\mathcal{I}_{i,i} = -1$ *if* $i \in S^1$, *and* $\mathcal{I}_{i,i} = 1$ *if* $i \in S^2$.

Note that $P_{a,i}$ is the probability of moving to state i when using action a. For a matrix $M \in \mathbb{R}^{m \times n}$ and a subset of indices $B \subseteq [m]$, we let M_B be the sub-matrix of M consisting of rows with indices in B. Also, for any $i \in [m]$, we let $M_i \in \mathbb{R}^{1 \times n}$ be the i-th row of M. We use similar notation for vectors.

Definition 8 (Strategies and strategy profiles) *A* strategy $\sigma^k : S^k \to A^k$ *for Player* $k \in \{1,2\}$ *maps every state* $i \in S^k$ *to an action* $\sigma^k(i) \in A_i$ *applicable from state* i. *A* strategy profile $\sigma = (\sigma^1, \sigma^2)$ *is a pair of strategies, one for each player. We let* Σ^k *be the set of strategies for Player* k, *and* $\Sigma = \Sigma^1 \times \Sigma^2$ *be the set of strategy profiles.*

We view a strategy profile $\sigma = (\sigma^1, \sigma^2)$ as a map $\sigma : S \to A$ from states to actions, such that $\sigma(i) = \sigma^k(i)$ for all $i \in S^k$ and $k \in \{1,2\}$.

A strategy $\sigma^k \in \Sigma^k$ can be viewed as a subset $\sigma^k \subseteq A^k$ of actions such that $\sigma^k \cap A_i = \{\sigma^k(i)\}$ for all $i \in S^k$. A strategy profile $\sigma = (\sigma^1, \sigma^2) \in \Sigma$ can be viewed similarly as a subset of actions $\sigma = \sigma^1 \cup \sigma^2 \subseteq A$. Note that P_σ is an $n \times n$ matrix for every $\sigma \in \Sigma$. We assume WLOG that actions are ordered such that $J_\sigma = I$, where I is the identity matrix, for all $\sigma \in \Sigma$. Figure 1 shows a strategy profile σ represented by bold gray edges, the corresponding matrix P_σ, and the vector $\mathbf{c}_\sigma$.

The matrix P_σ defines a Markov chain. In particular, the probability of being in the j-th state after t steps when starting in state i is $(P_\sigma^t)_{i,j}$. In Figure 1 such probabilities are shown in the table in the lower right corner, where $\mathbf{e}_1$ is the first unit vector. We say that the players *play according to* σ if whenever the pebble is on state $i \in S^k$, Player k uses action $\sigma(i)$. Let $i \in S$ be some state and t some number. The expected cost of the t-th action used is $(P_\sigma^t)_i \mathbf{c}_\sigma$. In particular, the expected outcome is $\sum_{t=0}^{\infty} \gamma^t (P_\sigma^t)_i \mathbf{c}_\sigma$. The following lemma shows that this infinite series always converges.

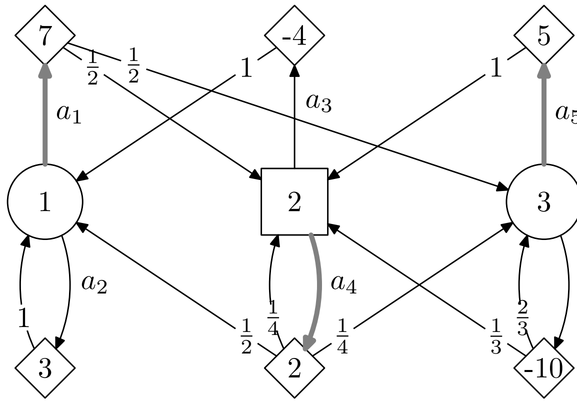
Lemma 9 *For every strategy profile* $\sigma \in \Sigma$ *the matrix* $(I - \gamma P_\sigma)$ *is non-singular, and*

$$(I - \gamma P_\sigma)^{-1} \;=\; \sum_{t=0}^{\infty} \gamma^t P_\sigma^t \; .$$

The simple proof of Lemma 9 has been omitted. For details we refer to, e.g., [7].

Definition 10 (Value vectors) *For every strategy profile* $\sigma \in \Sigma$ *we define the* value vector $\mathbf{v}^\sigma \in \mathbb{R}^n$ *by:*

$$\mathbf{v}^\sigma \;=\; (I - \gamma P_\sigma)^{-1} \mathbf{c}_\sigma \; .$$



$$P = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & 1 & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix} \quad \mathbf{c} = \begin{bmatrix} 7 \\ 3 \\ -4 \\ 2 \\ 5 \\ -10 \end{bmatrix} \quad J = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathcal{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{aligned} S^1 &= \{2\} \\ S^2 &= \{1,3\} \\ \sigma^1 &= \{a_4\} \\ \sigma^2 &= \{a_1, a_5\} \\ \sigma &= \{a_1, a_4, a_5\} \end{aligned} \qquad P_\sigma = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & 1 & 0 \end{bmatrix} \quad \mathbf{c}_\sigma = \begin{bmatrix} 7 \\ 2 \\ 5 \end{bmatrix}$$

t	$\mathbf{e}_1^\mathsf{T} P_\sigma^t$		
0	1	0	0
1	0	$\frac{1}{2}$	$\frac{1}{2}$
2	$\frac{2}{8}$	$\frac{5}{8}$	$\frac{1}{8}$
3	$\frac{10}{32}$	$\frac{13}{32}$	$\frac{9}{32}$
$\vdots$		$\vdots$	

Figure 1: Example of a simple 2TBSG and a strategy profile $\sigma = (\sigma^1, \sigma^2)$.

The i-th component of the value vector $\mathbf{v}^\sigma$, for a given strategy profile σ, is the expected outcome over plays starting in $i \in S$, when the players play according to σ.

It follows from Lemma 9 and Definition 10 that $\mathbf{v}^\sigma$ is the unique solution to:

$$\mathbf{v}^\sigma \;=\; \mathbf{c}_\sigma + \gamma P_\sigma \mathbf{v}^\sigma \;\;. \tag{1}$$

Definition 11 (Lower and upper values) *We define the* lower value vector $\underline{\mathbf{v}} \in \mathbb{R}^n$ *and* upper value vector $\bar{\mathbf{v}} \in \mathbb{R}^n$ *by:*

$$\begin{aligned} \forall i \in S: \quad \underline{\mathbf{v}}_i &= \min_{\sigma^1 \in \Sigma^1} \max_{\sigma^2 \in \Sigma^2} \mathbf{v}_i^{(\sigma^1,\sigma^2)} \\ \forall i \in S: \quad \bar{\mathbf{v}}_i &= \max_{\sigma^2 \in \Sigma^2} \min_{\sigma^1 \in \Sigma^1} \mathbf{v}_i^{(\sigma^1,\sigma^2)} \;\;. \end{aligned}$$

Shapley [19] showed that $\underline{\mathbf{v}} = \bar{\mathbf{v}}$. Hence, we may define the *optimal value vector* as $\mathbf{v}^* := \underline{\mathbf{v}} = \bar{\mathbf{v}}$.

Definition 12 (Optimal strategies) *A strategy $\sigma^1 \in \Sigma^1$ is* optimal *if and only if:*

$$\forall i \in S: \quad \max_{\sigma^2 \in \Sigma^2} \mathbf{v}_i^{(\sigma^1,\sigma^2)} \;=\; \mathbf{v}_i^* \;\;.$$

Similarly, a strategy $\sigma^2 \in \Sigma^2$ is optimal *if and only if:*

$$\forall i \in S: \quad \min_{\sigma^1 \in \Sigma^1} \mathbf{v}_i^{(\sigma^1,\sigma^2)} \;=\; \mathbf{v}_i^* \;\;.$$

We say that a strategy profile $\sigma = (\sigma^1, \sigma^2) \in \Sigma$ is optimal *if and only if σ^1 and σ^2 are optimal.*

Note that an optimal strategy for Player 1 (Player 2) minimizes (maximizes) the values of all states simultaneously. Hence, it is not immediately clear that optimal strategies always exist. This was shown by Shapley [19], however. *Solving* a 2TBSG means finding an optimal strategy profile (or equivalently the optimal value vector).

Definition 13 (Reduced costs) *For every strategy profile $\sigma \in \Sigma$ we define the vector of* reduced costs $\bar{\mathbf{c}}^\sigma \in \mathbb{R}^m$ *by:*

$$\forall i \in S, a \in A_i: \quad \bar{\mathbf{c}}_a^\sigma \;=\; \mathbf{c}_a + \gamma P_a \mathbf{v}^\sigma - \mathbf{v}_i^\sigma \;\;.$$

The following theorem establishes a connection between optimal strategies and reduced costs. For details see, e.g., [8, 7].

Theorem 14 (Optimality condition) *A strategy profile $\sigma \in \Sigma$ is optimal if and only if $(\bar{\mathbf{c}}^\sigma)_{A^1} \geq 0$ and $(\bar{\mathbf{c}}^\sigma)_{A^2} \leq 0$.*

2.3 LCPs for solving 2TBSGs

Jurdziński and Savani [10] showed how the problem of solving *deterministic* 2TBSGs can be reduced to the problem of solving P-matrix LCPs. We next show that the same reduction works for general 2TBSGs.

Throughout this section we let $G = (S^1, S^2, (A_i)_{i \in S}, p, c, \gamma)$ be some 2TBSG and $(P, \mathbf{c}, J, \mathcal{I}, \gamma)$ be the corresponding matrix representation. We assume that there are exactly two actions available from every state, i.e., $|A_i| = 2$ for all $i \in S$. We partition A into two disjoint strategy profiles σ and τ.

An LCP for solving G can be derived as follows. Consider the following system of linear equations and inequalities, where $\mathbf{w}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ are variables.

$$\begin{aligned}
(I - \gamma P_\sigma)\mathbf{y} - \mathcal{I}\mathbf{w} &= \mathbf{c}_\sigma && (2)\\
(I - \gamma P_\tau)\mathbf{y} - \mathcal{I}\mathbf{z} &= \mathbf{c}_\tau && (3)\\
\mathbf{w}^\mathsf{T}\mathbf{z} &= 0 && (4)\\
\mathbf{w}, \mathbf{z} &\geq \mathbf{0} && (5)
\end{aligned}$$

Lemma 15 *If $\mathbf{w}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ is a solution to (2), (3), (4), and (5), then $\mathbf{y}$ is the optimal value vector, a strategy profile π is optimal if $\pi \subseteq \{\sigma(i) \mid i \in [n] \wedge \mathbf{w}_i = 0\} \cup \{\tau(i) \mid i \in [n] \wedge \mathbf{z}_i = 0\}$, and such a strategy profile exists.*

Proof Let $B = \{\sigma(i) \mid i \in [n] \wedge \mathbf{w}_i = 0\} \cup \{\tau(i) \mid i \in [n] \wedge \mathbf{z}_i = 0\}$, and let Π be the set of all strategy profiles contained in B. Since $\mathbf{w}$ and $\mathbf{z}$ satisfy (4), we know that $\Pi \neq \emptyset$.

Let $a \in B \cap A_i$ for some $i \in S$. It follows from (2) and (3) that $\mathbf{y}_i - \gamma P_a \mathbf{y} = \mathbf{c}_a$. Hence, we get from Equation (1) that $\mathbf{y} = \mathbf{v}^\pi$, for every $\pi \in \Pi$. Combining this with (2), (3), and (5) we get that:

$$\begin{aligned}
\forall i \in S^1, a \in A_i: \quad \mathbf{v}_i^\pi - \gamma P_a \mathbf{v}^\pi &\leq \mathbf{c}_a\\
\forall i \in S^2, a \in A_i: \quad \mathbf{v}_i^\pi - \gamma P_a \mathbf{v}^\pi &\geq \mathbf{c}_a
\end{aligned}$$

It follows from Definition 13 and Theorem 14 that π is an optimal strategy profile. $\square$

We know from Lemma 9 that $(I - \gamma P_\tau)$ is non-singular. Hence, (3) can be equivalently expressed as:

$$\mathbf{y} = (I - \gamma P_\tau)^{-1}(\mathbf{c}_\tau + \mathcal{I}\mathbf{z}) .$$

Eliminating $\mathbf{y}$ in (2) we then get the following equivalent equation:

$$\begin{aligned}
(I - \gamma P_\sigma)(I - \gamma P_\tau)^{-1}(\mathbf{c}_\tau + \mathcal{I}\mathbf{z}) - \mathcal{I}\mathbf{w} &= \mathbf{c}_\sigma \quad \iff\\
\mathcal{I}(I - \gamma P_\sigma)(I - \gamma P_\tau)^{-1}\mathcal{I}\mathbf{z} - \mathbf{w} &= \mathcal{I}\mathbf{c}_\sigma - \mathcal{I}(I - \gamma P_\sigma)(I - \gamma P_\tau)^{-1}\mathbf{c}_\tau \quad \iff\\
\mathbf{w} &= (\mathcal{I}(I - \gamma P_\sigma)(I - \gamma P_\tau)^{-1}\mathbf{c}_\tau - \mathcal{I}\mathbf{c}_\sigma) + \mathcal{I}(I - \gamma P_\sigma)(I - \gamma P_\tau)^{-1}\mathcal{I}\mathbf{z} \qquad (6)
\end{aligned}$$

To simplify equation (6) we make the following definition.

Definition 16 ($M_{G,\sigma,\tau}$ and $\mathbf{q}_{G,\sigma,\tau}$) *Let G be a 2TBSG with matrix representation $(P, \mathbf{c}, J, \mathcal{I}, \gamma)$, and let the set of actions of G be partitioned into two disjoint strategy profiles σ and τ. We define $M_{G,\sigma,\tau} \in \mathbb{R}^{n\times n}$ and $\mathbf{q}_{G,\sigma,\tau} \in \mathbb{R}^n$ by:*

$$\begin{aligned}
M_{G,\sigma,\tau} &= \mathcal{I}(I - \gamma P_\sigma)(I - \gamma P_\tau)^{-1}\mathcal{I}\\
\mathbf{q}_{G,\sigma,\tau} &= \mathcal{I}(I - \gamma P_\sigma)(I - \gamma P_\tau)^{-1}\mathbf{c}_\tau - \mathcal{I}\mathbf{c}_\sigma .
\end{aligned}$$

I.e., equation (6) can be stated as $\mathbf{w} = \mathbf{q}_{G,\sigma,\tau} + M_{G,\sigma,\tau}\mathbf{z}$. It follows that (2), (3), (4), and (5) can be equivalently stated as $\mathbf{y} = (I - \gamma P_\tau)^{-1}(\mathbf{c}_\tau + \mathcal{I}\mathbf{z})$ and:

$$\begin{aligned}
\mathbf{w} &= \mathbf{q}_{G,\sigma,\tau} + M_{G,\sigma,\tau}\mathbf{z}\\
\mathbf{w}^\mathsf{T}\mathbf{z} &= 0\\
\mathbf{w}, \mathbf{z} &\geq \mathbf{0} .
\end{aligned}$$

Hence, a solution to the LCP $(M_{G,\sigma,\tau}, \mathbf{q}_{G,\sigma,\tau})$ gives a solution to (2), (3), (4), and (5), which, using Lemma 15, solves the 2TBSG G. We say that $(M_{G,\sigma,\tau}, \mathbf{q}_{G,\sigma,\tau})$ *solves* G.

Jurdziński and Savani [10] showed that $M_{G,\sigma,\tau}$ is a P-matrix when G is deterministic. To prove the same for general 2TBSGs we introduce the following lemma. The lemma is also used in the later parts of the paper. To understand the use of $\mathbf{v}$ in the lemma observe that $\mathbf{x}^{\mathsf{T}}(I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1}\mathbf{x} = \mathbf{x}^{\mathsf{T}}(I-\gamma P_\sigma)\mathbf{v}$.

Lemma 17 *Let* $\mathbf{x}$ *be a non-zero vector,* $\mathbf{v} = (I-\gamma P_\tau)^{-1}\mathbf{x}$*, and* $j \in \operatorname{argmax}_i |\mathbf{v}_i|$*. Then we have that:*

$$|\mathbf{x}_j| \;\geq\; (1-\gamma)\,|\mathbf{v}_j| \;\;. \tag{7}$$

$$\forall i : |\mathbf{x}_i| \;\leq\; (1+\gamma)\,|\mathbf{v}_j| \;\;. \tag{8}$$

$$\mathbf{x}_j((I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1}\mathbf{x})_j \;\geq\; (1-\gamma)\,|\mathbf{x}_j\mathbf{v}_j| \;>\; 0 \;\;. \tag{9}$$

Proof Observe first that $\mathbf{v}$ is the unique solution to $\mathbf{v} = \mathbf{x} + \gamma P_\tau \mathbf{v}$. In fact, we can interpret $\mathbf{v}$ as the value vector for τ when the costs $\mathbf{c}_\tau$ have been replaced by $\mathbf{x}$. If $v = \mathbf{0}$ then this implies that $\mathbf{0} = \mathbf{x} + \mathbf{0} \neq \mathbf{0}$ which is a contradiction. Thus, $\mathbf{v} \neq \mathbf{0}$ and in particular $\mathbf{v}_j \neq 0$. Since, for every i, the entries of $(P_\tau)_i$ are non-negative and sum to one we have that $|\gamma(P_\tau)_i\mathbf{v}| \leq \gamma\,|\mathbf{v}_j|$. The equations $\mathbf{v}_i = \mathbf{x}_i + \gamma(P_\tau)_i\mathbf{v}$, for all i, then imply that:

$$\begin{aligned} |\mathbf{x}_j| \;&=\; |\mathbf{v}_j - \gamma(P_\tau)_j\mathbf{v}| \;\geq\; |\mathbf{v}_j| - |\gamma(P_\tau)_j\mathbf{v}| \;\geq\; |\mathbf{v}_j| - |\gamma\mathbf{v}_j| \;=\; (1-\gamma)\,|\mathbf{v}_j| \;\;, \text{ and} \\ \forall i: \;|\mathbf{x}_i| \;&=\; |\mathbf{v}_i - \gamma(P_\tau)_i\mathbf{v}| \;\leq\; |\mathbf{v}_i| + |\gamma(P_\tau)_i\mathbf{v}| \;\leq\; |\mathbf{v}_j| + |\gamma\mathbf{v}_j| \;=\; (1+\gamma)\,|\mathbf{v}_j| \;\;. \end{aligned}$$

This proves (7) and (8).

We next observe that $\mathbf{v}_j$ and $\mathbf{x}_j$ have the same sign. This again follows from $|\gamma(P_\tau)_j\mathbf{v}| \leq \gamma\,|\mathbf{v}_j|$ and $\mathbf{v}_j = \mathbf{x}_j + \gamma(P_\tau)_j\mathbf{v}$. Using that $\operatorname{sgn}(\mathbf{v}_j) = \operatorname{sgn}(\mathbf{x}_j)$ we can now see that:

$$\begin{aligned} \mathbf{x}_j((I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1}\mathbf{x})_j \;&=\; \mathbf{x}_j((I-\gamma P_\sigma)\mathbf{v})_j \;=\; \mathbf{x}_j\mathbf{v}_j - \gamma\mathbf{x}_j(P_\sigma)_j\mathbf{v} \\ &\geq\; \mathbf{x}_j\mathbf{v}_j - \gamma\mathbf{x}_j\mathbf{v}_j \;=\; (1-\gamma)\mathbf{x}_j\mathbf{v}_j \;>\; 0 \;\;. \end{aligned}$$

This proves (9). □

We know from Lemma 3 that the matrix $M_{G,\sigma,\tau}$ is a P-matrix if and only if for every $\mathbf{x} \neq \mathbf{0}$ there exists a $j \in [n]$ such that $\mathbf{x}_j(M_{G,\sigma,\tau}\mathbf{x})_j > 0$. Since $\mathcal{I}\mathbf{x} \neq \mathbf{0}$, inequality (9) in Lemma 17 shows that $\mathbf{x}_j(M_{G,\sigma,\tau}\mathbf{x})_j > 0$ for $j \in \operatorname{argmax}_i \left|((I-\gamma P_\tau)^{-1}\mathcal{I}\mathbf{x})_i\right|$. Hence, $M_{G,\sigma,\tau}$ is a P-matrix.

We summarize the results of this section in the following theorem.

Theorem 18 *Let* G *be a 2TBSG, and let* σ *and* τ *be two disjoint strategy profiles that form a partition of the set of actions of* G*. Then the optimal value vector for* G *is* $\mathbf{v}^* = (I-\gamma P_\tau)^{-1}(\mathbf{c}_\tau + \mathcal{I}\mathbf{z})$*, where* $(\mathbf{w}, \mathbf{z})$ *is a solution to the LCP* $(M_{G,\sigma,\tau}, \mathbf{q}_{G,\sigma,\tau})$*. Furthermore,* $M_{G,\sigma,\tau}$ *is a* P*-matrix.*

Recall that Kojima *et al.* [12] showed that every P-matrix is a P_*-matrix. Hence, we have shown that $M_{G,\sigma,\tau}$ is a P_*-matrix.

3 The $P_*(\kappa)$ property for 2TBSGs

Let G be a 2TBSG with matrix representation $(P, \mathbf{c}, J, \mathcal{I}, \gamma)$, and let σ and τ be two disjoint strategy profiles that form a partition of the set of actions of G. Recall that G can be solved by solving the LCP $(M_{G,\sigma,\tau}, \mathbf{q}_{G,\sigma,\tau})$. In this section we provide essentially tight upper and lower bounds on the smallest number κ for which the matrix $M_{G,\sigma,\tau}$ is guaranteed to be a $P_*(\kappa)$-matrix. More precisely, we first show that for $\kappa = \frac{n}{(1-\gamma)^2}$, the matrix $M_{G,\sigma,\tau}$ is always a $P_*(\kappa)$-matrix. We then also show that for every $n > 2$ and $\gamma < 1$ there exists a game G_n, and two strategy profiles σ_n and τ_n, such that M_{G_n,σ_n,τ_n} is not a $P_*(\kappa)$-matrix for any $\kappa < \frac{\gamma^2(n-2)}{8(1-\gamma)^2} - \frac{1}{4}$. It follows that the unified interior point method of Kojima *et al.* [12] solves the 2TBSG G in time $O(\frac{n^{4.5}L}{(1-\gamma)^2})$, where L is the number of bits required to describe G, and that this bound can not be improved further only by bounding κ.

Recall that $M_{G,\sigma,\tau} = \mathcal{I}(I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1}\mathcal{I}$, and define $M := \mathcal{I} M_{G,\sigma,\tau} \mathcal{I} = (I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1}$. It is easy to see that $M_{G,\sigma,\tau}$ is a $P_*(\kappa)$-matrix for some $\kappa \geq 0$ if and only if M is. Indeed, the inequality of Definition 5 must hold for all $\mathbf{x} \in \mathbb{R}^n$, and we can therefore substitute $\mathbf{x}$ by $\mathcal{I}\mathbf{x}$. Hence, for the remainder of this section we will bound the κ for which M is a $P_*(\kappa)$-matrix.

Theorem 19 *Let n and $0 < \gamma < 1$ be given. For any γ-discounted 2TBSG G with n states, the matrix $M_{G,\sigma,\tau}$, where σ and τ partition the actions of G, is a $P_*(\kappa)$-matrix for $\kappa = \frac{n}{(1-\gamma)^2}$.*

Proof As discussed above we may prove the theorem by bounding κ for $M = (I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1}$ instead of for $M_{G,\sigma,\tau}$. Thus, we need to find a number κ, such that

$$\forall \mathbf{x} \in \mathbb{R}^n : \sum_{i \in \delta_-(M)} \mathbf{x}_i((I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1}\mathbf{x})_i \;+\; (1+4\kappa) \sum_{i \in \delta_+(M)} \mathbf{x}_i((I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1}\mathbf{x})_i \;\geq\; 0 \;,$$

where $\delta_-(M) = \{i \in [n] \mid \mathbf{x}_i(M\mathbf{x})_i < 0\}$ and $\delta_+(M) = \{i \in [n] \mid \mathbf{x}_i(M\mathbf{x})_i > 0\}$.

Let $\mathbf{x}$ be any non-zero vector (the expression is trivially satisfied for $\mathbf{x} = 0$), $\mathbf{v} = (I-\gamma P_\tau)^{-1}\mathbf{x}$, and $j \in \operatorname{argmax}_i |\mathbf{v}_i|$. To prove the lemma we will estimate $\sum_{i \in \delta_-(M)} \mathbf{x}_i((I-\gamma P_\sigma)\mathbf{v})_i$ and $\sum_{i \in \delta_+(M)} \mathbf{x}_i((I-\gamma P_\sigma)\mathbf{v})_i$ separately.

Using Lemma 17 we see that:

$$\forall i : \; |\mathbf{x}_i((I-\gamma P_\sigma)\mathbf{v})_i| \;\leq\; |\mathbf{x}_i|\,(|\mathbf{v}_j| + \gamma\,|\mathbf{v}_j|) \;\leq\; (1+\gamma)^2|\mathbf{v}_j|^2 \;<\; 4|\mathbf{v}_j|^2 \;,$$

which implies that:

$$\left| \sum_{i \in \delta_-(M)} \mathbf{x}_i((I-\gamma P_\sigma)\mathbf{v})_i \right| \;<\; 4n|\mathbf{v}_j|^2 \;.$$

Similiarly, from Lemma 17 we have that:

$$\mathbf{x}_j(M\mathbf{x})_j \;\geq\; (1-\gamma)\mathbf{x}_j\mathbf{v}_j \;=\; (1-\gamma)\,|\mathbf{x}_j|\,|\mathbf{v}_j| \;\geq\; (1-\gamma)^2|\mathbf{v}_j|^2 \;,$$

which implies that:

$$\sum_{i \in \delta_+(M)} \mathbf{x}_i((I-\gamma P_\sigma)\mathbf{v})_i \;\geq\; (1-\gamma)^2|\mathbf{v}_j|^2 \;.$$

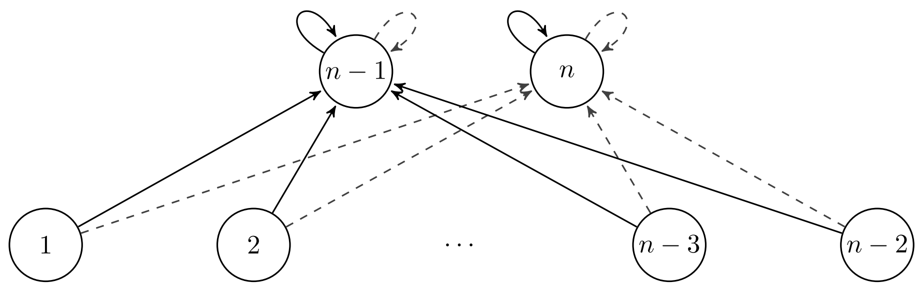


Figure 2: An example game G_n and two strategy profiles σ_n (solid) and τ_n (dashed), where M_{G_n,σ_n,τ_n} essentially matches the upper bound given in Theorem 19.

We conclude that:

$$\sum_{i\in\delta_-(M)} \mathbf{x}_i((I-\gamma P_\sigma)\mathbf{v})_i \;+\; (1+4\kappa)\sum_{i\in\delta_+(M)} \mathbf{x}_i((I-\gamma P_\sigma)\mathbf{v})_i \;>\; -4n|\mathbf{v}_j|^2 + (1+4\kappa)(1-\gamma)^2|\mathbf{v}_j|^2 \;.$$

It follows that M is a $P_*(\kappa)$-matrix when:

$$\begin{aligned} -4n|\mathbf{v}_j|^2 + (1+4\kappa)(1-\gamma)^2|\mathbf{v}_j|^2 \;&\geq\; 0 \qquad \iff \\ 4\kappa(1-\gamma)^2|\mathbf{v}_j|^2 \;&\geq\; (4n-(1-\gamma)^2)|\mathbf{v}_j|^2 \qquad \iff \\ \kappa \;&\geq\; \frac{n}{(1-\gamma)^2} - \frac{1}{4} \;. \end{aligned}$$

□

We next present a lower bound that essentially matches the upper bound given in Theorem 19. The gap between the upper and lower bounds is close to a factor of 8 for γ going to 1. Note that we are mostly interested in the case when γ is very close to 1, since it is known that the problem can be solved in strongly polynomial time when γ is a fixed constant [8]. We establish the lower bound using the family of games $\{G_n \mid n > 2\}$ shown in Figure 2. Figure 2 also shows two strategies σ_n and τ_n shown as solid and dashed arrows, respectively. Formally, the games are defined as follows.

The game G_n. For a given n, let G_n be the following game containing n states, all belonging to Player 2. For $i \leq n-2$, state i has two actions: one leading to state $n-1$ and one leading to state n. State $n-1$ and n have two self-loops each. I.e., the game is deterministic. The cost vector $\mathbf{c}$ can be arbitrary, and the discount factor γ will be specified by the analysis. We also define two disjoint strategy profiles σ_n and τ_n that partition the set of actions. The strategy profile σ_n contains for all states $i \leq n-2$ the action leading to state $n-1$, and the strategy profile τ_n contains for all states $i \leq n-2$ the action leading to state n. Furthermore, at states $n-1$ and n, the strategy profiles σ_n and τ_n contain a self-loop each.

Theorem 20 *Let $n > 2$ and $0 < \gamma < 1$ be given. For the 2TBSG G_n, the matrix M_{G_n,σ_n,τ_n} is not a $P_*(\kappa)$-matrix, for $\kappa < \frac{\gamma^2(n-2)}{8(1-\gamma)^2} - \frac{1}{4} = \Omega\big(\frac{\gamma^2 n}{(1-\gamma)^2}\big)$.*

Proof Notice first that since all states belong to the Player 2, $\mathcal{I}$ is the identity matrix. Thus $M = M_{G_n,\sigma_n,\tau_n}$. We then need to find a number $\kappa \geq 0$, such that:

$$\forall \mathbf{x}: \sum_{i\in\delta_-(M)} \mathbf{x}_i((I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1}\mathbf{x})_i + (1+4\kappa)\sum_{i\in\delta_+(M)} \mathbf{x}_i((I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1}\mathbf{x})_i \;\geq 0 \;.$$

Let $\mathbf{x} \in \mathbb{R}^n$ be defined by, for all $i \in [n]$:

$$\mathbf{x}_i = \begin{cases} 1 - \frac{1}{1-\gamma} & \text{if } i \leq n-2 \\ 1 & \text{if } i = n-1 \\ -1 & \text{if } i = n \ . \end{cases}$$

Let $\mathbf{v} = (I - \gamma P_\tau)^{-1}\mathbf{x}$ be the value vector for τ when using costs $\mathbf{x}$. By straightforward calculation, using Equation (1), we see that

$$\mathbf{v}_i = \begin{cases} 0 & \text{if } i \leq n-2 \\ \frac{1}{1-\gamma} & \text{if } i = n-1 \\ \frac{-1}{1-\gamma} & \text{if } i = n \ . \end{cases}$$

Also, let $\mathbf{r}_i = \mathbf{x}_i((I - \gamma P_\sigma)\mathbf{v})_i = \mathbf{x}_i\mathbf{v}_i - \gamma\mathbf{x}_i(P_\sigma\mathbf{v})_i$. Again, by straightforward calculations we see that:

$$\begin{aligned} \forall i \leq n-2: \ \mathbf{r}_i &= -\gamma\left(1 - \frac{1}{1-\gamma}\right)\frac{1}{1-\gamma} &= -\left(\frac{\gamma}{1-\gamma}\right)^2 \\ \mathbf{r}_{n-1} &= \frac{1}{1-\gamma} - \gamma\frac{1}{1-\gamma} &= 1 \\ \mathbf{r}_n &= \frac{1}{1-\gamma} - \gamma\frac{1}{1-\gamma} &= 1 \end{aligned}$$

Therefore we have that:

$$\sum_{i\in\delta_-(M)} \mathbf{x}_i((I - \gamma P_\sigma)\mathbf{v})_i = -(n-2)(\frac{\gamma}{1-\gamma})^2 \ ,$$

and that:

$$\sum_{i\in\delta_+(M)} \mathbf{x}_i((I - \gamma P_\sigma)\mathbf{v})_i = 2 \ .$$

Hence, for $\sum_{i\in\delta_-(M)} \mathbf{x}_i((I - \gamma P_\sigma)\mathbf{v})_i + (1+4\kappa)\sum_{i\in\delta_+(M)} \mathbf{x}_i((I - \gamma P_\sigma)\mathbf{v})_i \geq 0$ to be true we need that:

$$\begin{aligned} (n-2)\left(\frac{\gamma}{1-\gamma}\right)^2 &\leq (1+4\kappa)2 \quad \Longleftrightarrow \\ \kappa &\geq \frac{n-2}{8}\left(\frac{\gamma}{1-\gamma}\right)^2 - \frac{1}{4} \ . \end{aligned}$$

□

4 Bounds for the potential reduction algorithm

The interior point potential reduction algorithm of Kojima *et al.* [13] for solving a P-matrix LCP $(M, \mathbf{q})$ takes as input a parameter $\epsilon > 0$ and produces a feasible solution $(\mathbf{w}, \mathbf{z})$ for which $\mathbf{w}^\mathsf{T}\mathbf{z} < \epsilon$. Following Ye [20], the running time of the potential reduction algorithm is upper bounded by $O(\frac{-\delta}{\theta}n^4 \log \epsilon^{-1})$, where (i) δ is the smallest eigenvalue of $\frac{M+M^\mathsf{T}}{2}$; and (ii) θ is the positive P-matrix number of M (Definition 4). In this section we are interested in bounding the running time of the potential reduction algorithm when applied to 2TBSGs by studying the two quantities (i) δ and (ii) θ.

Throughout the section we let G be a 2TBSG with matrix representation $(P, \mathbf{c}, J, \mathcal{I}, \gamma)$, and σ and τ be two disjoint strategy profiles that form a partition of the set of actions of G. To simplify notation we let $M := M_{G,\sigma,\tau}$. To bound the running time of the potential reduction algorithm we need to bound the smallest eigenvalue δ of $\frac{M+M^\mathsf{T}}{2}$ and the positive P-matrix number $\theta(M)$ of M. We study the smallest eigenvalue of $\frac{M+M^\mathsf{T}}{2}$ in Section 4.1, and the positive P-matrix number $\theta(M)$ in Section 4.2. For both quantities we provide upper and lower bounds that are essentially tight.

More precisely, we show that we always have smallest eigenvalue $\delta > -\frac{(1+\gamma)\sqrt{n}}{1-\gamma}$, and that there exists a family of 2TBSGs G_n, with corresponding strategy profiles σ_n and τ_n, for which the smallest eigenvalue is at most $1 - \frac{\gamma\sqrt{(n-2)}}{\sqrt{2}(1-\gamma)}$. I.e., the gap is only a factor of $2\sqrt{2}$ for γ going to 1. We also show that we always have $\theta(M) \geq \frac{(1-\gamma)^2}{(1+\gamma)^2 n}$, and that there exists a family of 2TBSGs G_n, with corresponding strategy profiles σ_n and τ_n, for which the positive P-matrix number satisfies $\theta(M_{G_n,\sigma_n,\tau_n}) < \frac{(1-\gamma)^2}{(2\gamma)^2(n-2)}$. I.e., the bound is tight when γ goes to 1.

It is important to note that the upper bound for the smallest eigenvalue δ and the upper bound for the positive P-matrix number $\theta(M_{G_n,\sigma_n,\tau_n})$ are obtained using the same game G_n and the same strategy profiles σ_n and τ_n. In fact, for both bounds we use the same game and strategy profile as were used in the proof of Theorem 20. I.e., G_n, σ_n, and τ_n are shown in Figure 2. Hence, for this particular game we achieve the worst-case ratio of $\frac{-\delta}{\theta} = \Omega(\frac{\gamma n^{3/2}}{(1-\gamma)^3})$.

4.1 Bounds for the smallest eigenvalue

We first lower bound the smallest eigenvalue of $\frac{M+M^\mathsf{T}}{2}$, where $M = \mathcal{I}(I - \gamma P_\sigma)(I - \gamma P_\tau)^{-1}\mathcal{I}$. We let $\mathbb{R}^n_{\|\cdot\|_i = k}$ be the set of vectors in $\mathbb{R}^n$ such that each vector $\mathbf{v} \in \mathbb{R}^n_{\|\cdot\|_i = k}$ has $\|\mathbf{v}\|_i = k$.

Theorem 21 *The matrix* $\frac{M+M^\mathsf{T}}{2}$ *has smallest eigenvalue greater than* $-\frac{(1+\gamma)\sqrt{n}}{1-\gamma} = -O(\frac{\sqrt{n}}{1-\gamma})$

Proof Look at the equation $\lambda\mathbf{x} = \frac{M+M^\mathsf{T}}{2}\mathbf{x}$, where λ is the smallest eigenvalue. We have that:

$$\begin{aligned}
\lambda\mathbf{x} &= \frac{M+M^\mathsf{T}}{2}\mathbf{x} \\
&= \frac{\mathcal{I}(I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1}\mathcal{I} + \mathcal{I}(I-\gamma P_\tau^\mathsf{T})^{-1}(I-\gamma P_\sigma^\mathsf{T})\mathcal{I}}{2}\mathbf{x} \\
&= \mathcal{I}\frac{(I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1} + (I-\gamma P_\tau^\mathsf{T})^{-1}(I-\gamma P_\sigma^\mathsf{T})}{2}\mathcal{I}\mathbf{x} \ .
\end{aligned}$$

By letting $\mathbf{y} = \mathcal{I}\mathbf{x}$ we obtain the equation:

$$\lambda\mathbf{y} = \frac{(I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1} + (I-\gamma P_\tau^\mathsf{T})^{-1}(I-\gamma P_\sigma^\mathsf{T})}{2}\mathbf{y} \ .$$

We can without loss of generality assume that $\mathbf{y}$ has two-norm equal to one, and by the triangle inequality we therefore have:

$$\begin{aligned}|\lambda| = \|\lambda\mathbf{y}\|_2 &= \left\|\frac{(I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1} + (I-\gamma P_\tau^\mathsf{T})^{-1}(I-\gamma P_\sigma^\mathsf{T})}{2}\mathbf{y}\right\|_2 \\ &\leq \frac{1}{2}\left\|(I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1}\mathbf{y}\right\|_2 + \frac{1}{2}\left\|(I-\gamma P_\tau^\mathsf{T})^{-1}(I-\gamma P_\sigma^\mathsf{T})\mathbf{y}\right\|_2 \ .\end{aligned}$$

We will bound $\frac{1}{2}\left\|(I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1}\mathbf{y}\right\|_2$ and $\frac{1}{2}\left\|(I-\gamma P_\tau^\mathsf{T})^{-1}(I-\gamma P_\sigma^\mathsf{T})\mathbf{y}\right\|_2$ separately. We first observe that:

$$\begin{aligned}\left\|(I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1}\mathbf{y}\right\|_2 &\leq \max_{\mathbf{v}\in\mathbb{R}^n_{\|\cdot\|_\infty=1}} \left\|(I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1}\mathbf{v}\right\|_2 \\ &= \max_{\mathbf{v}\in\mathbb{R}^n_{\|\cdot\|_\infty=1}} \left\|(I-\gamma P_\sigma)\sum_{t=0}^{\infty}\gamma^t P_\tau^t\mathbf{v}\right\|_2 \\ &\leq \max_{\mathbf{v}\in\mathbb{R}^n_{\|\cdot\|_\infty=1}} \left\|(I-\gamma P_\sigma)\sum_{t=0}^{\infty}\gamma^t\mathbf{v}\right\|_2 \\ &= \max_{\mathbf{v}\in\mathbb{R}^n_{\|\cdot\|_\infty=1}} \left\|(I-\gamma P_\sigma)\frac{\mathbf{v}}{1-\gamma}\right\|_2 \\ &\leq \max_{\mathbf{v}\in\mathbb{R}^n_{\|\cdot\|_\infty=1+\gamma}} \left\|\frac{\mathbf{v}}{1-\gamma}\right\|_2 \\ &= \frac{1}{1-\gamma}\max_{\mathbf{v}\in\mathbb{R}^n_{\|\cdot\|_\infty=1+\gamma}} \|\mathbf{v}\|_2 \\ &= \frac{(1+\gamma)\sqrt{n}}{1-\gamma} \ .\end{aligned}$$

Here, the first inequality comes from the fact that if a vector $\mathbf{v}\in\mathbb{R}^n$ has two-norm equal to 1, then it has infinity-norm equal to at most 1. The first equality follows from Lemma 9. To prove the second inequality we use that $\left\|P_\tau^t\mathbf{v}\right\|_\infty \leq \|\mathbf{v}\|_\infty$, for all $t \geq 0$, since the entries of P_τ are in $[0,1]$. The third inequality follows from the fact that $\|(I-\gamma P_\sigma)\mathbf{v}\|_\infty \leq (1+\gamma)\|\mathbf{v}\|_\infty$. The last equality comes from the fact that if a vector $\mathbf{v}\in\mathbb{R}^n$ has infinity-norm equal to $1+\gamma$ then it has two-norm at most $(1+\gamma)\sqrt{n}$.

We also have that:

$$
\begin{aligned}
\left\|(I-\gamma P_\tau^\top)^{-1}(I-\gamma P_\sigma^\top)\mathbf{y}\right\|_2 &\leq \max_{\mathbf{v}\in\mathbb{R}^n_{\|\cdot\|_1=1}} \left\|(I-\gamma P_\tau^\top)^{-1}(I-\gamma P_\sigma^\top)\sqrt{n}\mathbf{v}\right\|_2 \\
&\leq (1+\gamma)\sqrt{n} \max_{\mathbf{v}\in\mathbb{R}^n_{\|\cdot\|_1=1}} \left\|(I-\gamma P_\tau^\top)^{-1}\mathbf{v}\right\|_2 \\
&= (1+\gamma)\sqrt{n} \max_{\mathbf{v}\in\mathbb{R}^n_{\|\cdot\|_1=1}} \left\|\sum_{t=0}^{\infty}\gamma^t (P_\tau^t)^\top \mathbf{v}\right\|_2 \\
&\leq (1+\gamma)\sqrt{n} \max_{\mathbf{v}\in\mathbb{R}^n_{\|\cdot\|_1=1}} \left\|\sum_{t=0}^{\infty}\gamma^t \mathbf{v}\right\|_2 \\
&= (1+\gamma)\sqrt{n} \max_{\mathbf{v}\in\mathbb{R}^n_{\|\cdot\|_1=1}} \left\|\frac{\mathbf{v}}{1-\gamma}\right\|_2 \\
&= \frac{(1+\gamma)\sqrt{n}}{1-\gamma} \max_{\mathbf{v}\in\mathbb{R}^n_{\|\cdot\|_1=1}} \|\mathbf{v}\|_2 \\
&= \frac{(1+\gamma)\sqrt{n}}{1-\gamma} \ .
\end{aligned}
$$

Here, the first inequality comes from the fact that if a vector $\mathbf{v} \in \mathbb{R}^n$ has two-norm equal to 1 then it has one-norm equal to at most $\sqrt{n}$. The second inequality follows from the fact that the columns of $P_\sigma^\top$ sum to 1 such that $\left\|(I-\gamma P_\sigma^\top)\mathbf{v}\right\|_1 \leq (1+\gamma)\|\mathbf{v}\|_1$. The first equality follows from Lemma 9. For the third inequality we again use that the columns of $P_\sigma^\top$ sum to 1, which implies that $\left\|P_\sigma^\top \mathbf{v}\right\|_1 = \|\mathbf{v}\|_1$. The last equality comes from the fact that if a vector $\mathbf{v} \in \mathbb{R}^n$ has one-norm equal to 1 then it has two-norm at most 1.

Hence,

$$
|\lambda| \ \leq \ \frac{1}{2}\frac{(1+\gamma)\sqrt{n}}{1-\gamma} + \frac{1}{2}\frac{(1+\gamma)\sqrt{n}}{1-\gamma} \ = \ \frac{(1+\gamma)\sqrt{n}}{1-\gamma} \ ,
$$

completing the proof. □

We will now upper bound the smallest eigenvalue of $\frac{M_{G_n,\sigma_n,\tau_n}+M_{G_n,\sigma_n,\tau_n}^\top}{2}$, where G_n, σ_n, and τ_n were defined in Section 3 (Figure 2).

Theorem 22 *Let n and $0 < \gamma < 1$ be given, and let $M := M_{G_n,\sigma_n,\tau_n}$. The matrix $\frac{M+M^\top}{2}$ has smallest eigenvalue at most $1 - \frac{\gamma\sqrt{(n-2)}}{\sqrt{2}(1-\gamma)}$.*

Proof Let $\mathbf{x} \in \mathbb{R}^n$ be defined by, for all $i \in [n]$:

$$
\mathbf{x}_i = \begin{cases} a & \text{if } i < n-1 \\ 1 & \text{if } i = n-1 \\ -1 & \text{if } i = n \ , \end{cases}
$$

where a is a parameter that will be specified later.

We will show that there is an a such that $\mathbf{x}$ is an eigenvector with eigenvalue $2 - \frac{\gamma\sqrt{2(n-2)}}{1-\gamma}$. Hence, we look at the equation

$$\begin{aligned}\lambda \mathbf{x} &= \frac{M + M^{\mathsf{T}}}{2}\mathbf{x} \\ &= \frac{\mathcal{I}(I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1}\mathcal{I} + \mathcal{I}(I-\gamma P_\tau^{\mathsf{T}})^{-1}(I-\gamma P_\sigma^{\mathsf{T}})\mathcal{I}}{2}\mathbf{x} \ .\end{aligned}$$

Notice that $\mathcal{I}$ is the identity matrix. We will look at each term separately. In each case we will evaluate the expression from right to left.

We first evaluate $(I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1}\mathbf{x}$. Let $\mathbf{v} = (I-\gamma P_\tau)^{-1}\mathbf{x}$. I.e., $\mathbf{v}$ is the value vector of τ then the costs are replaced by $\mathbf{x}$. Then, by simple calculation using Equation (1), we see that:

$$\mathbf{v}_i = \begin{cases} a + \frac{\gamma}{1-\gamma} & \text{if } i \le n-2 \\ \frac{1}{1-\gamma} & \text{if } i = n-1 \\ \frac{-1}{1-\gamma} & \text{if } i = n \ . \end{cases}$$

Let $\mathbf{r} = (I-\gamma P_\sigma)\mathbf{v} = (I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1}\mathbf{x}$. Then, by simple calculations, we see that:

$$\mathbf{r}_i = \begin{cases} a + \frac{2\gamma}{1-\gamma} & \text{if } i \le n-2 \\ 1 & \text{if } i = n-1 \\ -1 & \text{if } i = n \ . \end{cases}$$

Next, we evaluate $(I-\gamma P_\tau^{\mathsf{T}})^{-1}(I-\gamma P_\sigma^{\mathsf{T}})\mathbf{x}$. Let $\mathbf{x}' = (I-\gamma P_\sigma^{\mathsf{T}})$. Then it is easy to see that for all $i \in [n]$:

$$\mathbf{x}'_i = \begin{cases} a & \text{if } i \le n-2 \\ 1-\gamma-\gamma a(n-2) & \text{if } i = n-1 \\ -1+\gamma & \text{if } i = n \ . \end{cases}$$

Let $\mathbf{r}' = (I-\gamma P_\tau^{\mathsf{T}})^{-1}\mathbf{x}'$. Note that no actions move to state i for $i \le n-2$, meaning that $(P_\tau^{\mathsf{T}})_i = \mathbf{0}$. Then by multiplying from the right with $(I-\gamma P_\tau^{\mathsf{T}})$ we see that:

$$\begin{aligned} \forall i \le n-2 : \ \mathbf{r}'_i \ &= \ \mathbf{x}'_i + \gamma(P_\tau^{\mathsf{T}})_i\mathbf{r}' \ = \ \mathbf{x}'_i \ = \ a \\ \mathbf{r}'_{n-1} \ &= \ \mathbf{x}'_{n-1} + \gamma(P_\tau^{\mathsf{T}})_{n-1}\mathbf{r}' \ = \ \mathbf{x}'_{n-1} + \gamma\mathbf{r}'_{n-1} \ = \ \frac{1-\gamma-\gamma a(n-2)}{1-\gamma} \\ \mathbf{r}'_n \ &= \ \mathbf{x}'_n + \gamma(P_\tau^{\mathsf{T}})_n\mathbf{r}' \ = \ \mathbf{x}'_n + \gamma\mathbf{r}'_n + \gamma\sum_{i=1}^{n-2}\mathbf{r}'_i \ = \ \frac{-1+\gamma+\gamma a(n-2)}{1-\gamma} \end{aligned}$$

Therefore, if we let:

$$\mathbf{v}' = \frac{(I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1} + (I-\gamma P_\tau^{\mathsf{T}})^{-1}(I-\gamma P_\sigma^{\mathsf{T}})}{2}\mathbf{x} \ ,$$

we have that:

$$\mathbf{v}'_i = \begin{cases} a + \frac{\gamma}{1-\gamma} & \text{if } i < n-1 \\ 1 - \frac{\gamma a(n-2)}{2(1-\gamma)} & \text{if } i = n-1 \\ -1 + \frac{\gamma a(n-2)}{2(1-\gamma)} & \text{if } i = n \ . \end{cases}$$

We want $\lambda \mathbf{x} = \mathbf{v}'$, which is then the same as the equation system:

$$\lambda a = a + \frac{\gamma}{1-\gamma}$$
$$\lambda = 1 - \frac{\gamma a(n-2)}{2(1-\gamma)} \ .$$

By eliminating λ we get:

$$\begin{aligned} a + \frac{\gamma}{1-\gamma} &= a + \frac{\gamma a^2(n-2)}{2(1-\gamma)} \quad \Rightarrow \\ \frac{\gamma}{1-\gamma} &= \frac{\gamma a^2(n-2)}{2(1-\gamma)} \quad \Rightarrow \\ 2 &= a^2(n-2) \quad \Rightarrow \\ a &= \pm\sqrt{\frac{2}{n-2}} \ . \end{aligned}$$

Since we want to minimize λ over a we get:

$$\lambda = 1 - \frac{\gamma\sqrt{(n-2)}}{\sqrt{2}(1-\gamma)} \ .$$

Hence, the matrix $\frac{M+M^\top}{2}$ has smallest eigenvalue at most $1 - \frac{\gamma\sqrt{(n-2)}}{\sqrt{2}(1-\gamma)}$. □

4.2 Bounds for the positive P-matrix number

We will now lower bound the positive P-matrix number for any 2TBSG.

Theorem 23 *Let n and $0 < \gamma < 1$ be given. For any 2TBSG G with n states, the matrix $M := M_{G,\sigma,\tau}$, where σ and τ partition the states of G, has positive P-matrix number, $\theta(M)$, at least $\frac{(1-\gamma)^2}{(1+\gamma)^2 n} = \Omega(\frac{(1-\gamma)^2}{n})$*

Proof Recall that the positive P-matrix number of $M = M_{G,\sigma,\tau}$ is defined as:

$$\theta(M) \ = \ \min_{\|\mathbf{x}\|_2=1} \ \max_{i\in[n]} \ \mathbf{x}_i(M\mathbf{x})_i \ .$$

Let $\mathbf{x} \in \mathbb{R}^n_{\|\cdot\|_2=1}$ be given. Let $\mathbf{v} = (I-\gamma P_\tau)^{-1}\mathbf{x}$ and $j \in \operatorname{argmax}_i |\mathbf{v}_i|$. From Lemma 17 we know that $\mathbf{x}_j(M\mathbf{x})_j \geq (1-\gamma)\,|\mathbf{x}_j\mathbf{v}_j| \geq (1-\gamma)^2(\mathbf{v}_j)^2$. We also know from Lemma 17 that $\mathbf{v}_j^2 \geq \frac{\mathbf{x}_i^2}{(1+\gamma)^2}$ for all $i \in [n]$. Hence, we see that $\mathbf{x}_j(M\mathbf{x})_j \geq \frac{(1-\gamma)^2\mathbf{x}_i^2}{(1+\gamma)^2}$ for all $i \in [n]$. Since $\|\mathbf{x}\|_2 = 1$ there must exist an index i such that $|\mathbf{x}_i| \geq \frac{1}{\sqrt{n}}$. It follows that $\mathbf{x}_j(M\mathbf{x})_j \geq \frac{(1-\gamma)^2}{(1+\gamma)^2 n}$. Since this inequality holds for all $\mathbf{x} \in \mathbb{R}^n_{\|\cdot\|_2=1}$ we see that $\theta(M) \geq \frac{(1-\gamma)^2}{(1+\gamma)^2 n}$.

□

We will now upper bound the positive P-matrix number of M_{G_n,σ_n,τ_n}. I.e., we once again use the construction from Figure 2.

Theorem 24 *Let n and $0 < \gamma < 1$ be given. The matrix $M := M_{G_n,\sigma_n,\tau_n}$ has positive P-matrix number $\theta(M) < \frac{(1-\gamma)^2}{(2\gamma)^2(n-2)}$.*

Proof Recall that the positive P-matrix number of $M = M_{G_n,\sigma_n,\tau_n}$ is defined as:

$$\theta(M) = \min_{\|\mathbf{x}\|_2=1} \max_{i\in[n]} \mathbf{x}_i(M\mathbf{x})_i \ .$$

In the following we consider a concrete vector $\mathbf{x} \in \mathbb{R}^n$ defined by, for all $i \in [n]$:

$$\mathbf{x}_i = \begin{cases} 1 & \text{if } i \leq n-2 \\ -a & \text{if } i = n-1 \\ a & \text{if } i = n \ , \end{cases}$$

where a is a parameter that we specify later. We later normalize $\mathbf{x}$, such that $\|\mathbf{x}\|_2 = 1$. I.e., we want to evaluate the following expression for all $i \in [n]$:

$$\frac{\mathbf{x}_i(M\mathbf{x})_i}{\|\mathbf{x}\|_2^2} = \frac{\mathbf{x}_i(\mathcal{I}(I-\gamma P_\sigma)(I-\gamma P_\tau)^{-1}\mathcal{I}\mathbf{x})_i}{\|\mathbf{x}\|_2^2} \ .$$

Recall that $\mathcal{I}$ is the identity matrix. We will evaluate the expression from right to left.

Let $\mathbf{v} = (I-\gamma P_\tau)^{-1}\mathbf{x}$. Then, by simple calculations using that $\mathbf{v} = \mathbf{x} + \gamma P_\tau \mathbf{v}$, we see that:

$$\mathbf{v}_i = \begin{cases} 1 - \frac{a\gamma}{1-\gamma} & \text{if } i \leq n-2 \\ \frac{-a}{1-\gamma} & \text{if } i = n-1 \\ \frac{a}{1-\gamma} & \text{if } i = n \ . \end{cases}$$

Let $\mathbf{r} = (I-\gamma P_\sigma)\mathbf{v}$. Then, by simple calculations, we see that:

$$\mathbf{r}_i = \begin{cases} 1 - \frac{2a\gamma}{1-\gamma} & \text{if } i \leq n-2 \\ -a & \text{if } i = n-1 \\ a & \text{if } i = n \ . \end{cases}$$

Hence, $\mathbf{x}_i\mathbf{r}_i$ is:

$$\mathbf{x}_i\mathbf{r}_i = \begin{cases} 1 - \frac{2a\gamma}{1-\gamma} & \text{if } i \leq n-2 \\ a^2 & \text{if } i = n-1 \\ a^2 & \text{if } i = n \end{cases}$$

We now let $a = \frac{1-\gamma}{2\gamma}$, in which case we see that:

$$\mathbf{x}_i\mathbf{r}_i = \begin{cases} 0 & \text{if } i \leq n-2 \\ \frac{(1-\gamma)^2}{(2\gamma)^2} & \text{if } i = n-1 \text{ or } i = n \end{cases}$$

We also see that $\|\mathbf{x}\|_2^2 = n-2+2a^2 > n-2$. It follows that the positive P-matrix number, $\theta(M)$, is at most:

$$\max_{i\in[n]} \frac{\mathbf{x}_i\mathbf{r}_i}{\|\mathbf{x}\|_2^2} < \frac{(1-\gamma)^2}{(2\gamma)^2(n-2)}.$$

∎

References

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

[2] H. Chen, M. Zhang, and Y. Zhao. A class of new large-update primal-dual interior-point algorithms for $P_*(\kappa)$ linear complementarity problems. In *Proc. of 6th ISNN*, pages 77–87, 2009.

[3] G.-M. Cho. A new large-update interior point algorithm for $P_*(\kappa)$ linear complementarity problems. *Journal of Computational and Applied Mathematics*, 216:265–278, 2008.

[4] R. W. Cottle, J.-S. Pang, and R. E. Stone. *The Linear Complementarity Problem*. Computer science and scientific computing. Academic Press, Boston, 1992.

[5] J. Filar and K. Vrieze. *Competitive Markov Decision Processes*. Springer, 1997.

[6] B. Gärtner and L. Rüst. Simple stochastic games and P-matrix generalized linear complementarity problems. In *Proc. of 15th FCT*, pages 209–220, 2005.

[7] T. D. Hansen. *Worst-case Analysis of Strategy Iteration and the Simplex Method*. PhD thesis, Aarhus University, 2012.

[8] T. D. Hansen, P. B. Miltersen, and U. Zwick. Strategy iteration is strongly polynomial for 2-player turn-based stochastic games with a constant discount factor. In *Proc. of 2nd ICS*, pages 253–263, 2011.

[9] T. Illés, M. Nagy, and T. Terlaky. A polynomial path-following interior point algorithm for general linear complementarity problems. *Journal of Global Optimization*, 47:329–342, 2010.

[10] M. Jurdziński and R. Savani. A simple P-matrix linear complementarity problem for discounted games. In *Proc. of 4th CiE*, pages 283–293, 2008.

[11] N. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4(4):373–395, 1984.

[12] M. Kojima, N. Megiddo, T. Noma, and A. Yoshise. A unified approach to interior point algorithms for linear complementarity problems: A summary. *Oper. Res. Lett.*, 10(5):247–254, 1991.

[13] M. Kojima, N. Megiddo, and Y. Ye. An interior point potential reduction algorithm for the linear complementarity problem. *Mathematical Programming*, 54:54–267, 1992.

[14] N. Krishnamurthy, T. Parthasarathy, and G. Ravindran. Solving subclasses of multi-player stochastic games via linear complementarity problem formulations – a survey and some new results. *Optimization and Engineering*, 13:435–457, 2012.

[15] M. L. Littman. *Algorithms for sequential decision making*. PhD thesis, Brown University, 1996.

[16] A. Neyman and S. Sorin. *Stochastic games and applications*. Kluwer Academic Publishers, Dordrecht Boston, 2003.

[17] S. S. Rao, R. Chandrasekaran, and K. P. K. Nair. Algorithms for discounted stochastic games. *Journal of Optimization Theory and Applications*, 11:627–637, 1973.

[18] L. Y. Rüst. *The P-matrix linear complementarity problem.* PhD thesis, ETH Zürich, 2007.

[19] L. S. Shapley. Stochastic games. *Proc. Nat. Acad. Sci. U.S.A.*, 39(10):1095–1100, 1953.

[20] Y. Ye. *Interior Point Algorithms: Theory and Analysis.* Wiley-Interscience series in discrete mathematics and optimization. Wiley, 1998.