

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 (2TB-SGs). Jurdziński and Savani showed that 2TB-SGs with deterministic transitions can be reduced to solving P -matrix linear complementarity problems (LCPs). We show that the same reduction works for general 2TB-SGs. This implies that a number of interior point methods for solving P -matrix LCPs can be used to solve 2TB-SGs. We consider two such algorithms. First, we consider the unified interior point method of Kojima, Megiddo, Noma, and Yoshise, which runs in time $O((1 + \kappa)n^{3.5}L)$, where κ is a parameter that depends on the $n \times n$ matrix M defining the LCP, and L is the number of bits in the representation of M . Second, we consider the interior point potential reduction algorithm of Kojima, Megiddo, and Ye, which runs in time $O(\frac{\delta}{\theta}n^4 \log \epsilon^{-1})$, where δ and θ are parameters that depend on M , and ϵ describes the quality of the solution. For 2TB-SGs with n states and discount factor γ we prove that in the worst case $\kappa = \Theta(n/(1 - \gamma)^2)$, $-\delta = \Theta(\sqrt{n}/(1 - \gamma))$, and $1/\theta = \Theta(n/(1 - \gamma)^2)$. The lower bounds for κ , $-\delta$, and $1/\theta$ are obtained using the same family of deterministic games.

1 Introduction

Two-player turn-based stochastic games (2TB-SGs). A *two-player turn-based stochastic game* (2TB-SG) consists of a finite set of states and for each state a finite set of actions. The game is played by two players (Player 1 and Player 2) for an infinite number of rounds. The states are partitioned into two sets S^1 and S^2 , belonging to Player 1 and Player 2, respectively. In each round the game is in some state, and the player controlling the current state i chooses an action a available from state i . Every action is associated with a probability distribution over states, and the next state is picked at random according to the probability distribution for a . After every transition the game ends with probability $1 - \gamma > 0$, where γ is the *discount factor* of the game. Every action has an associated cost, and the objective of Player 1 is to *minimize* the expected sum of costs, whereas the objective of Player 2 is to *maximize* the expected sum of costs, i.e., the game is a zero-sum game. Our main focus is the case where all states have two actions.

^{*}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).

The class of (turn-based) stochastic games was introduced by Shapley [20] in 1953 and has received much attention since then. For books on the subject see, e.g., Neyman and Sorin [16] and Filar and Vrieze [5]. 2TBSGs with a single player, i.e., one player controls all the states, are known as *Markov decision processes*; a problem that is important in its own right (see, e.g., Puterman [17]). Shapley showed that every state in a 2TBSG has a *value* that can be enforced by both players (a property known as *determinacy*). *Solving* a 2TBSG means finding the values of all the states, and the problem of solving 2TBSGs is the topic of this paper.

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 at the same time no polynomial time algorithm is known. Strategy iteration algorithms were first described for discounted 2TBSGs by Rao *et al.* [18]. Building on a result by Ye [22], Hansen *et al.* [8] 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 where γ is fixed. Littman showed that Shapley’s *value iteration* algorithm [20] solves 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. Several polynomial time algorithms have been discovered for solving linear programs. Perhaps the most notable family of such algorithms is *interior point methods*. The first interior point method was introduced by Karmarkar [11] in 1984, and the technique has since been studied extensively and applied in other contexts. See, e.g., Ye [21]. In particular, interior point methods can be used to solve *P-matrix linear complementarity problems* (LCPs). One may hope that a polynomial time algorithm for solving 2TBSGs whose discount factor γ is not fixed (i.e., the discount factor is part of the input) can be obtained through the use of interior point methods. Indeed, this approach was suggested by Jurdziński and Savani [10] and Hansen *et al.* [8]. In this paper we study a reduction from 2TBSGs to *P-matrix* LCPs, and we study the complexity of two known interior point methods when applied to the resulting *P-matrix* LCPs.

P-matrix linear complementarity problems. 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 $\mathbf{w} = \mathbf{q} + M\mathbf{z}$, $\mathbf{w}^\top \mathbf{z} = 0$, 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 [21].

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] 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 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 $\mathbf{x}^\top(M\mathbf{x}) + 4\kappa \sum_{i \in \delta_+(M)} \mathbf{x}_i(M\mathbf{x})_i \geq 0$, 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 additional examples see [3, 2, 9].

An interior point potential reduction algorithm. The second interior point method we consider is the potential reduction algorithm of Kojima, Megiddo, and Ye [13]. (See also Ye [21].) 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 $\mathbf{w}^\top \mathbf{z} < \epsilon$, $\mathbf{w} = \mathbf{q} + M\mathbf{z}$, 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 $\frac{1}{2}(M + M^\top)$, 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 [21].

Rüst [19] showed 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 [19] showed 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 $\frac{1}{2}(M + M^\top)$ is $-\Omega(\frac{1}{1-\gamma})$. This bound can be viewed as a precursor for 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. Let M_G be the matrix obtained from the reduction for a given 2TBSG G . We also show that for any 2TBSG G with n states and discount factor γ we have:

- (i) The matrix M_G is a $P_*(\kappa)$ -matrix for $\kappa = O(\frac{n}{(1-\gamma)^2})$.
- (ii) The matrix $\frac{M_G + M_G^\top}{2}$ has smallest eigenvalue δ , where $-\delta = O(\frac{\sqrt{n}}{1-\gamma})$.
- (iii) The reciprocal of the positive P -matrix number of M_G is $\frac{1}{\theta(M_G)} = O(\frac{n}{(1-\gamma)^2})$.

Item (i) implies that the running time of the unified interior point method of Kojima *et al.* [12] for 2TBSGs is at most $O((1 + \kappa)n^{3.5}L) = O(\frac{n^{4.5}L}{(1-\gamma)^2})$. Items (ii) and (iii) together imply that the running time of the potential reduction algorithm of Kojima *et al.* [13] for 2TBSGs is at most $O(\frac{\delta}{\theta} n^4 \log \epsilon^{-1}) = O(\frac{n^{5.5} \log \epsilon^{-1}}{(1-\gamma)^3})$.

Finally, we define a family of *deterministic* 2TBSGs G_n for which we prove matching lower bounds: (i) $\kappa = \Omega(\frac{n}{(1-\gamma)^2})$, (ii) $-\delta = \Omega(\frac{\sqrt{n}}{1-\gamma})$, and (iii) $\frac{1}{\theta(M_{G_n})} = \Omega(\frac{n}{(1-\gamma)^2})$.

Note that the upper bounds we prove for the two algorithms are worse than the $O(\frac{n^2 L}{1-\gamma} \log \frac{1}{1-\gamma})$ bound for the value iteration algorithm for the case where every state has two actions. Although our results for existing interior point methods for solving 2TBSGs are therefore negative, other

interior point methods may solve 2TBSGs efficiently. It is also possible that the known bounds for the algorithms studied in this paper do not reflect their true running time. In fact, we believe that interior point methods are the key to solving 2TBSGs efficiently, and that the study of interior point methods in this context 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, thereby bounding the running time of the unified interior point method of Kojima *et al.* [12]. In Section 4 we similarly bound the smallest eigenvalue and the positive P -matrix number, proving a bound for 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:

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

Definition 2 (P -matrix) A matrix $M \in \mathbb{R}^{n \times n}$ is a P -matrix if and only if all its 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 every non-zero vector $\mathbf{x} \in \mathbb{R}^n$ there is an $i \in [n] = \{1, \dots, n\}$ such that $\mathbf{x}_i(M\mathbf{x})_i > 0$.

Definition 4 (Positive P -matrix number) The positive P -matrix number of a matrix $M \in \mathbb{R}^{n \times n}$ 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 every vector $\mathbf{x} \in \mathbb{R}^n$:

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

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 symmetric matrix M is a $P_*(0)$ -matrix if and only if it is *positive semi-definite*.

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 $i \in S$.
- $p : A \rightarrow \Delta(S)$ is a map from actions to probability distributions over states.
- $c : A \rightarrow \mathbb{R}$ 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 $k \in \{1, 2\}$. We refer to Player 1 as the minimizer and to Player 2 as the maximizer. 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.

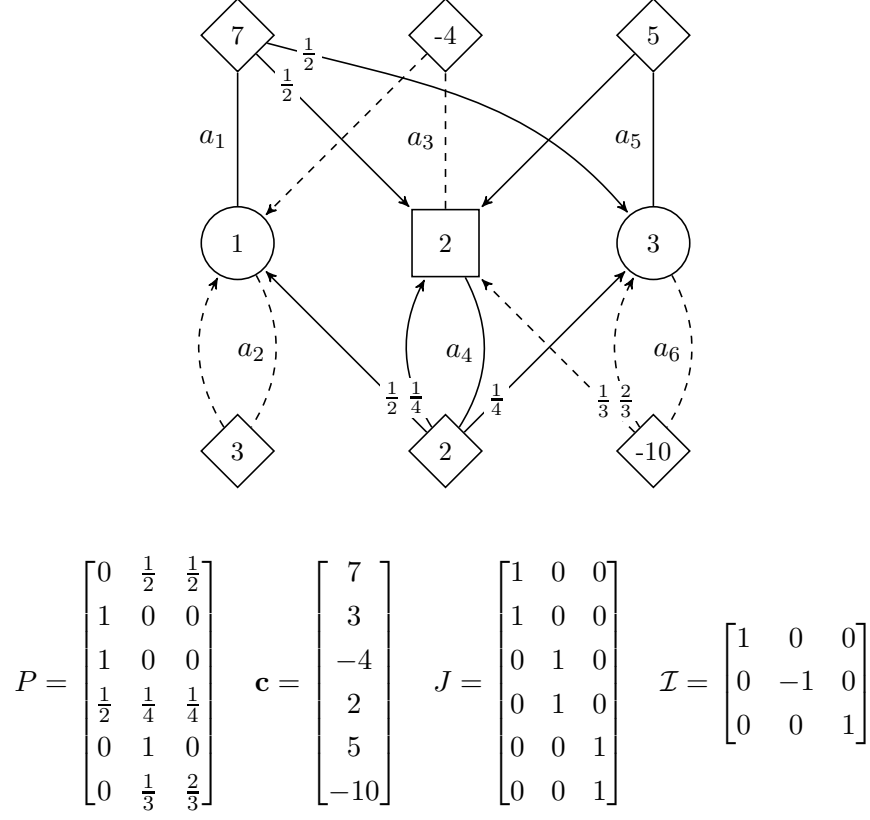
We say that an action a is *deterministic* if it moves to a single state with probability 1, i.e., if $p(a)_j = 1$ for some $j \in S$. If all the actions of a 2TBSG G are deterministic we say that G is deterministic.

Plays and outcomes. A 2TBSG is played as follows. A pebble is moved from state to state starting from some initial state $i_0 \in S$. When the pebble is at state $i \in S^k$, Player k chooses an action $a \in A_i$, and the pebble is moved to a random new state according to the probability distribution $p(a)$. Let a^t be the t -th chosen action for every $t \geq 0$. The sequence of chosen actions is called a *play*, and the *outcome* of the play, paid by Player 1 to Player 2, is $\sum_{t \geq 0} \gamma^t \cdot c(a^t)$.

To simplify notation, we next introduce a way to represent a 2TBSG with vectors and matrices. Figure 1 shows an example of this representation.

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 without loss of generality that $S = [n] = \{1, \dots, n\}$ and $A = [m] = \{1, \dots, m\}$.

- 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 $i \in S$.



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

t	$\mathbf{e}_1^\top P_\sigma^t$
0	$\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$
1	$\begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}$
2	$\begin{bmatrix} \frac{2}{8} & \frac{5}{8} & \frac{1}{8} \end{bmatrix}$
3	$\begin{bmatrix} \frac{10}{32} & \frac{13}{32} & \frac{9}{32} \end{bmatrix}$
\vdots	\vdots

Figure 1: Example of a simple 2TBSG, as well as a strategy profile $\sigma = (\sigma^1, \sigma^2)$ shown with solid lines.

- 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 $\sigma^k : S^k \rightarrow A^k$ 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 $\sigma : S \rightarrow A$ 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 also 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 without loss of generality 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 solid lines, the corresponding matrix P_σ , and the vector \mathbf{c}_σ .

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 at state $i \in S^k$, Player k uses the action $\sigma(i)$. Let $i \in S$ be some state and t some number. The expected cost of the t -th action used when starting in state i 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:*

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

The i -th component of the value vector \mathbf{v}^σ , 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}^σ 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 $\overline{\mathbf{v}} \in \mathbb{R}^n$ by:*

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

Shapley [20] showed that $\underline{\mathbf{v}} = \overline{\mathbf{v}}$. Hence, we may define the *optimal value vector* uniquely as $\mathbf{v}^* := \underline{\mathbf{v}} = \overline{\mathbf{v}}$.

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

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

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

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

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. Equivalently, σ is a Nash equilibrium.

Note that an optimal strategy for Player 1 (Player 2) minimizes (maximizes) the values of all states simultaneously. Hence, it is not obvious that optimal strategies exist. This was shown, however, by Shapley [20]. 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:

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

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 $i \in S$, and 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.

$$(I - \gamma P_\sigma) \mathbf{y} - \mathcal{I} \mathbf{w} = \mathbf{c}_\sigma \tag{2}$$

$$(I - \gamma P_\tau) \mathbf{y} - \mathcal{I} \mathbf{z} = \mathbf{c}_\tau \tag{3}$$

$$\mathbf{w}^\top \mathbf{z} = 0 \tag{4}$$

$$\mathbf{w}, \mathbf{z} \geq \mathbf{0} \tag{5}$$

Lemma 15 The system (2), (3), (4), and (5) has a unique solution $\mathbf{w}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$, where \mathbf{y} is the optimal value vector for G . Also, a strategy profile π is optimal if and only 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 Observe first that if we let $\mathbf{y} = \mathbf{v}^*$, then (2) and (3) ensure that \mathbf{w} and \mathbf{z} correspond to the reduced costs for an optimal strategy profile π^* . More precisely, we have:

$$\begin{aligned} \forall i \in S^1 : \quad & \bar{\mathbf{c}}_{\sigma(i)}^{\pi^*} = \mathbf{w}_i \text{ and } \bar{\mathbf{c}}_{\tau(i)}^{\pi^*} = \mathbf{z}_i \\ \forall i \in S^2 : \quad & \bar{\mathbf{c}}_{\sigma(i)}^{\pi^*} = -\mathbf{w}_i \text{ and } \bar{\mathbf{c}}_{\tau(i)}^{\pi^*} = -\mathbf{z}_i \end{aligned}$$

It then follows from definitions 10 and 13 that $\mathbf{w}^\top \mathbf{z} = 0$, and from Theorem 14 that $\mathbf{w}, \mathbf{z} \geq \mathbf{0}$. Hence the optimal value vector and the corresponding reduced costs are a solution to (2), (3), (4), and (5). We next show that for every solution $\mathbf{w}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$, \mathbf{y} is the optimal value vector. Since the optimal value vector is unique it follows that the system has a unique solution.

Let $\mathbf{w}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ be a solution to (2), (3), (4), and (5), and let $B = \{\sigma(i) \mid i \in [n] \wedge \mathbf{w}_i = 0\} \cup \{\tau(i) \mid i \in [n] \wedge \mathbf{z}_i = 0\}$. Furthermore, 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:

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

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:

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

Equation (6) can now 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:

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

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 .

Jurđziński 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 $\mathbf{x}^\top(I - \gamma P_\sigma)(I - \gamma P_\tau)^{-1}\mathbf{x} = \mathbf{x}^\top(I - \gamma P_\sigma)\mathbf{v}$.

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:*

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

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

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

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}_τ have been replaced by \mathbf{x} . If $\mathbf{v} = \mathbf{0}$ 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 \mathbf{v}_j and \mathbf{x}_j have the same sign we 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 $j \in \operatorname{argmax}_i |((I - \gamma P_\tau)^{-1}\mathcal{I}\mathbf{x})_i|$. Hence, $M_{G,\sigma,\tau}$ is a P -matrix.

The following theorem summarizes the main result of this section.

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 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 cannot be improved further 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 G be a 2TBSG with n states and discount factor γ , where $0 < \gamma < 1$. Furthermore, let σ and τ be two strategy profiles that partition the set of actions of G . Then the matrix $M_{G,\sigma,\tau}$ 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 $M = (I - \gamma P_\sigma)(I - \gamma P_\tau)^{-1}$. We thus 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 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:

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

Similarly, 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:

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

We conclude that:

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

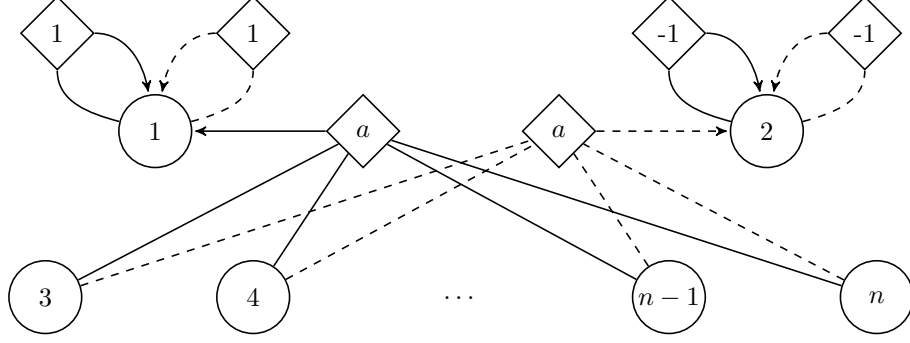


Figure 2: The game G_n and two strategy profiles σ_n (solid) and τ_n (dashed).

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

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

□

We next present a lower bound that matches the upper bound given in Theorem 19. We establish the lower bound using the family of games $\{G_n \mid n > 2\}$ shown in Figure 2. The game G_n consists of n states that all belong to Player 2. Every state has two actions, and all actions are deterministic. The actions from states 1 and 2 form self-loops. The actions from state i , for $i > 2$, lead to state 1 and 2, respectively. Both actions from state 1 have cost 1, both actions from state 2 have cost -1 , and all the remaining actions have cost $a \in \mathbb{R}$, where a will be specified in the analysis. The discount factor $\gamma < 1$ is given along with n .

We also define two strategy profiles σ_n and τ_n that partition the set of actions. Since all states belong to Player 2, σ_n and τ_n are simply strategies for Player 2. The strategies σ_n and τ_n are represented in Figure 2 by solid and dashed arrows, respectively. The strategies σ_n and τ_n both contain self-loops at states 1 and 2. Furthermore, for every state $i > 2$, the strategy σ_n contains the action leading to state 1, and the strategy τ_n contains the action leading to state 2. Observe that since all states belong to Player 2 we have $\mathcal{I} = I$ such that:

$$M_{G_n, \sigma_n, \tau_n} = (I - \gamma P_{\sigma_n})(I - \gamma P_{\tau_n})^{-1} .$$

Note also that $M_{G_n, \sigma_n, \tau_n}$ is independent of the parameter a .

Since all our lower bounds use G_n and many of the calculations are similar, we first establish the following lemma concerning various properties of the game.

Lemma 20 *Let n be given and consider G_n . Let $\mathbf{v} := \mathbf{v}^{\tau_n} = (I - \gamma P_{\tau_n})^{-1} \mathbf{c}_{\tau_n}$ be the value vector*

for τ_n in the game G_n . Define $\mathbf{r} := M_{G_n, \sigma_n, \tau_n} \mathbf{c}_{\tau_n}$, and let $\mathbf{r}' \in \mathbb{R}^n$ satisfy $\mathbf{r}'_i = (\mathbf{c}_{\tau_n})_i \mathbf{r}_i$. Then,

$$\begin{aligned} (\mathbf{c}_{\tau_n})_i &= \begin{cases} 1 & \text{if } i = 1 \\ -1 & \text{if } i = 2 \\ a & \text{if } i > 2 \end{cases} & \mathbf{v}_i &= \begin{cases} \frac{1}{1-\gamma} & \text{if } i = 1 \\ \frac{-1}{1-\gamma} & \text{if } i = 2 \\ a - \frac{\gamma}{1-\gamma} & \text{if } i > 2 \end{cases} \\ \mathbf{r}_i &= \begin{cases} 1 & \text{if } i = 1 \\ -1 & \text{if } i = 2 \\ a - \frac{2\gamma}{1-\gamma} & \text{if } i > 2 \end{cases} & \mathbf{r}'_i &= \begin{cases} 1 & \text{if } i = 1 \\ 1 & \text{if } i = 2 \\ a^2 - \frac{2\gamma a}{1-\gamma} & \text{if } i > 2 \end{cases}. \end{aligned}$$

Proof The calculation of \mathbf{v}^τ follows straightforwardly from (1), i.e., $\mathbf{v} = \mathbf{c}_{\tau_n} + \gamma P_{\tau_n} \mathbf{v}$, which can be stated as:

$$\begin{aligned} \mathbf{v}_1 &= 1 + \gamma \mathbf{v}_1 \\ \mathbf{v}_2 &= -1 + \gamma \mathbf{v}_2 \\ \forall i > 2: \quad \mathbf{v}_i &= a + \gamma \mathbf{v}_2 \end{aligned}$$

Let $\mathbf{r}_i = (M_{G_n, \sigma_n, \tau_n} \mathbf{c}_{\tau_n})_i = ((I - \gamma P_{\sigma_n}) \mathbf{v})_i = \mathbf{v}_i - \gamma (P_{\sigma_n} \mathbf{v})_i$ and observe that:

$$\begin{aligned} \mathbf{r}_1 &= \frac{1}{1-\gamma} - \gamma \frac{1}{1-\gamma} = 1 \\ \mathbf{r}_2 &= \frac{-1}{1-\gamma} + \gamma \frac{1}{1-\gamma} = -1 \\ \forall i > 2: \quad \mathbf{r}_i &= \left(a - \frac{\gamma}{1-\gamma} \right) - \gamma \frac{1}{1-\gamma} = a - \frac{2\gamma}{1-\gamma} \end{aligned}$$

The value of $\mathbf{r}'_i = (\mathbf{c}_{\tau_n})_i \mathbf{r}_i$ is computed straightforwardly as well. \square

Theorem 21 Let $n > 2$ and $0 < \gamma < 1$ be given. The matrix $M_{G_n, \sigma_n, \tau_n}$ is not a $P_*(\kappa)$ -matrix for $\kappa < \frac{\gamma^2(n-2)}{8(1-\gamma)^2} - \frac{1}{4} = \Omega\left(\frac{\gamma^2 n}{(1-\gamma)^2}\right)$.

Proof To simplify notation we denote $M_{G_n, \sigma_n, \tau_n}$ by M . Recall that M is a $P_*(\kappa)$ -matrix if and only if

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

where $\delta_-(M) = \{i \in [n] \mid \mathbf{x}_i(M\mathbf{x})_i < 0\}$ and $\delta_+(M) = \{i \in [n] \mid \mathbf{x}_i(M\mathbf{x})_i > 0\}$. We show that the inequality is violated when $\kappa < \frac{\gamma^2(n-2)}{8(1-\gamma)^2} - \frac{1}{4}$ and $\mathbf{x} = \mathbf{c}_{\tau_n}$ for an appropriate choice of a .

Let $\mathbf{r}'_i = \mathbf{x}_i(M\mathbf{x})_i$. From Lemma 20 we have that

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

We choose $a = \frac{\gamma}{1-\gamma}$ in order to minimize \mathbf{r}'_i for $i > 2$, which gives $\mathbf{r}'_i = -(\frac{\gamma}{1-\gamma})^2$. It follows that:

$$\sum_{i \in \delta_-(M)} \mathbf{x}_i(M\mathbf{x})_i = -(n-2) \left(\frac{\gamma}{1-\gamma} \right)^2 \quad \text{and} \quad \sum_{i \in \delta_+(M)} \mathbf{x}_i(M\mathbf{x})_i = 2 .$$

Hence, we get:

$$\begin{aligned} \sum_{i \in \delta_-(M)} \mathbf{x}_i(M\mathbf{x})_i + (1+4\kappa) \sum_{i \in \delta_+(M)} \mathbf{x}_i(M\mathbf{x})_i &\geq 0 \quad \Longleftrightarrow \\ 2(1+4\kappa) &\geq (n-2) \left(\frac{\gamma}{1-\gamma} \right)^2 \quad \Longleftrightarrow \\ \kappa &\geq \frac{n-2}{8} \left(\frac{\gamma}{1-\gamma} \right)^2 - \frac{1}{4} . \end{aligned}$$

□

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 $\mathbf{w}^\top \mathbf{z} < \epsilon$. Following Ye [21], the running time of the potential reduction algorithm is upper bounded by $O(\frac{-\delta}{\theta} n^4 \log \epsilon^{-1})$, where δ is the smallest eigenvalue of $\frac{M+M^\top}{2}$, and $\theta = \theta(M)$ is the positive P -matrix number of M (Definition 4). In this section we bound the running time of the potential reduction algorithm when applied to 2TBSGs by studying the two quantities δ and 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 $M := M_{G, \sigma, \tau}$. We study the smallest eigenvalue δ of $\frac{M+M^\top}{2}$ 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.

Let δ_n be the smallest eigenvalue of the matrix $\frac{1}{2}(M_{G_n, \sigma_n, \tau_n} + M_{G_n, \sigma_n, \tau_n}^\top)$, where $M_{G_n, \sigma_n, \tau_n}$ is derived from the game in Figure 2. To be precise, we prove the two bounds for δ and δ_n :

$$\begin{aligned} \delta &> -\frac{(1+\gamma)\sqrt{n}}{1-\gamma} \\ \delta_n &\leq 1 - \frac{\gamma\sqrt{(n-2)}}{\sqrt{2}(1-\gamma)} \end{aligned}$$

For the positive P -matrix number we prove the two bounds:

$$\begin{aligned} \theta(M) &\geq \frac{(1-\gamma)^2}{(1+\gamma)^2 n} \\ \theta(M_{G_n, \sigma_n, \tau_n}) &< \frac{(1-\gamma)^2}{(2\gamma)^2 (n-2)} . \end{aligned}$$

Note that the upper bounds for the smallest eigenvalue δ_n and the positive P -matrix number $\theta(M_{G_n, \sigma_n, \tau_n})$ are obtained using the same matrix $M_{G_n, \sigma_n, \tau_n}$, which was also used in the proof of Theorem 21. Hence, for the game G_n 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 $\frac{M+M^\top}{2}$, where $M = \mathcal{I}(I - \gamma P_\sigma)(I - \gamma P_\tau)^{-1}\mathcal{I}$. We let $\mathbb{R}_{\|\cdot\|_\ell=1}^n$ be the set of vectors in \mathbb{R}^n such that each vector $\mathbf{v} \in \mathbb{R}_{\|\cdot\|_\ell=1}^n$ has $\|\mathbf{v}\|_\ell = 1$.

Theorem 22 *The matrix $\frac{M+M^\top}{2}$ has smallest eigenvalue greater than $-\frac{(1+\gamma)\sqrt{n}}{1-\gamma} = -O(\frac{\sqrt{n}}{1-\gamma})$*

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

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

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

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

We can without loss of generality assume that \mathbf{y} has L^2 -norm equal to one, and by the triangle inequality we therefore have:

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

We bound $\|(I - \gamma P_\sigma)(I - \gamma P_\tau)^{-1} \mathbf{y}\|_2$ and $\|(I - \gamma P_\tau^\top)^{-1}(I - \gamma P_\sigma^\top) \mathbf{y}\|_2$ separately. We first observe that:

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

Here, the first inequality follows from the fact that if $\mathbf{v} \in \mathbb{R}^n$ has L^2 -norm equal to 1, then it has L^∞ -norm equal to at most 1. The first equality follows from Lemma 9. To prove the second inequality we use that $\|P_\tau^t \mathbf{v}\|_\infty \leq \|\mathbf{v}\|_\infty$, 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 follows from the fact that if a vector $\mathbf{v} \in \mathbb{R}^n$ has L^∞ -norm equal to 1 then it has L^2 -norm at most \sqrt{n} .

We also have that:

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

Here, the first inequality follows from the fact that if a vector $\mathbf{v} \in \mathbb{R}^n$ has L^2 -norm equal to 1 then it has L^1 -norm equal to at most \sqrt{n} . The second inequality follows from the fact that the columns of P_σ^\top sum to 1 such that $\|(I - \gamma P_\sigma^\top) \mathbf{v}\|_1 \leq (1 + \gamma) \|\mathbf{v}\|_1$. The first equality follows from Lemma 9. For the third inequality we again use that the columns of P_σ^\top sum to 1, which implies that $\mathbf{e}^\top P_\sigma^\top \mathbf{v} = \mathbf{e}^\top \mathbf{v}$ such that $\|P_\sigma^\top \mathbf{v}\|_1 \leq \|\mathbf{v}\|_1$. The last equality follows from the fact that if a vector $\mathbf{v} \in \mathbb{R}^n$ has L^1 -norm equal to 1 then it has L^2 -norm at most 1.

Hence,

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

which completes the proof. \square

We next 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 are defined in Section 3 (Figure 2).

Theorem 23 *Let $n > 2$ 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 Recall that

$$(\mathbf{c}_\tau)_i = \begin{cases} 1 & \text{if } i = 1 \\ -1 & \text{if } i = 2 \\ a & \text{if } i > 2 . \end{cases}$$

We show that, for an appropriate choice of a , $\mathbf{x} = \mathbf{c}_\tau$ is an eigenvector for the matrix $\frac{M+M^\top}{2}$ with eigenvalue $\lambda = 1 - \frac{\gamma\sqrt{n-2}}{\sqrt{2}(1-\gamma)}$. Hence, we prove that the following equation is satisfied:

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

Recall that $\mathcal{I} = I$. We evaluate the two terms on the right-hand-side separately and from right to left.

We first evaluate $\mathbf{r} := (I - \gamma P_\sigma)(I - \gamma P_\tau)^{-1} \mathbf{x}$. From Lemma 20 we get that

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

We next evaluate $\mathbf{r}'' := (I - \gamma P_\tau^\top)^{-1}(I - \gamma P_\sigma^\top) \mathbf{x}$. Let $\mathbf{v}' = (I - \gamma P_\sigma^\top) \mathbf{x}$. Observe that no actions move to state i for $i > 2$, which means that $(P_\sigma^\top)_i = \mathbf{0}$ and $(P_\tau^\top)_i = \mathbf{0}$ for $i > 2$. By similarly considering the incoming actions for states 1 and 2, it is easy to check that

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

Since $\mathbf{r}'' = (I - \gamma P_\tau^\top)^{-1} \mathbf{v}'$ we see that $\mathbf{r}''_i = \mathbf{v}'_i + \gamma(P_\tau^\top)_i \mathbf{r}''$. For $i > 2$, it follows that $\mathbf{r}''_i = \mathbf{v}'_i + 0 = a$. For $i = 1$, only the self-loop moves to state 1 in τ , and we get that

$$\mathbf{r}''_1 = \mathbf{v}'_1 + \gamma \mathbf{r}''_1 = 1 - \gamma - \gamma a(n-2) + \gamma \mathbf{r}''_1 = 1 - \frac{\gamma a(n-2)}{1-\gamma} .$$

For $i = 2$, all actions in τ , except the one from state 1, move to state 2. We therefore see that

$$\mathbf{r}''_2 = \mathbf{v}'_2 + \gamma \sum_{j=2}^n \mathbf{r}''_j = -1 + \gamma + \gamma a(n-2) + \gamma \mathbf{r}''_2 = -1 + \frac{\gamma a(n-2)}{1-\gamma} .$$

Hence, we have that

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

We conclude that

$$\left(\frac{M + M^\top}{2} \mathbf{x} \right)_i = \frac{\mathbf{r}_i + \mathbf{r}''_i}{2} = \begin{cases} 1 - \frac{\gamma a(n-2)}{2(1-\gamma)} & \text{if } i = 1 \\ -1 + \frac{\gamma a(n-2)}{2(1-\gamma)} & \text{if } i = 2 \\ a - \frac{\gamma}{1-\gamma} & \text{if } i > 2 . \end{cases}$$

It remains to show that $\lambda \mathbf{x} = \frac{\mathbf{r} + \mathbf{r}''}{2}$, which corresponds to the system of equations:

$$\begin{aligned}\lambda &= 1 - \frac{\gamma a(n-2)}{2(1-\gamma)} \\ \lambda a &= a - \frac{\gamma}{1-\gamma}.\end{aligned}$$

By eliminating λ , using the first equation, we get:

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

Thus for $a = \sqrt{\frac{2}{n-2}}$ we get that

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

is an eigenvalue for $\frac{M+M^T}{2}$ with eigenvector $\mathbf{x} = \mathbf{c}_\tau$, and in particular the smallest eigenvalue for $\frac{M+M^T}{2}$ is at most $1 - \frac{\gamma \sqrt{n-2}}{\sqrt{2}(1-\gamma)}$. \square

4.2 Bounds for the positive P -matrix number

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

Theorem 24 *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 actions 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 $\mathbf{x} \in \mathbb{R}_{\|\cdot\|_2=1}^n$ 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 $\mathbf{v}_j^2 \geq \frac{\mathbf{x}_j^2}{(1+\gamma)^2}$ for all $i \in [n]$. Hence, we see that $\mathbf{x}_j(M\mathbf{x})_j \geq \frac{(1-\gamma)^2 \mathbf{x}_j^2}{(1+\gamma)^2}$ 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 $\mathbf{x} \in \mathbb{R}_{\|\cdot\|_2=1}^n$ we see that $\theta(M) \geq \frac{(1-\gamma)^2}{(1+\gamma)^2 n}$. \square

We next upper bound the positive P -matrix number of M_{G_n,σ_n,τ_n} , where once again G_n , σ_n , and τ_n are from the construction shown in Figure 2.

Theorem 25 Let $n > 2$ 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:

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

Recall also that

$$(\mathbf{c}_\tau)_i = \begin{cases} 1 & \text{if } i = 1 \\ -1 & \text{if } i = 2 \\ a & \text{if } i > 2 . \end{cases}$$

Define $\mathbf{x} := \mathbf{c}_\tau$ and $\mathbf{x}' := \frac{\mathbf{x}}{\|\mathbf{x}\|_2}$. We show that, for an appropriate choice of a , \mathbf{x}' satisfies

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

This implies that $\theta(M) < \frac{(1-\gamma)^2}{(2\gamma)^2(n-2)}$, completing the proof.

Let $\mathbf{r}' \in \mathbb{R}^n$ satisfy $\mathbf{r}'_i = \mathbf{x}_i(M\mathbf{x})_i$ for all $i \in [n]$. From Lemma 20, we then get that

$$(\mathbf{r}')_i = \mathbf{x}_i(M\mathbf{x})_i = \begin{cases} 1 & \text{if } i = 1 \\ 1 & \text{if } i = 2 \\ a^2 - \frac{2\gamma a}{1-\gamma} & \text{if } i > 2 . \end{cases}$$

Since $\|\mathbf{x}\|_2 = \sqrt{2 + (n-2)a^2}$ it follows that

$$\mathbf{x}'_i(M\mathbf{x}')_i = \frac{\mathbf{x}_i(M\mathbf{x})_i}{\|\mathbf{x}\|_2^2} = \begin{cases} \frac{1}{2+(n-2)a^2} & \text{if } i = 1 \\ \frac{1}{2+(n-2)a^2} & \text{if } i = 2 \\ \frac{a^2 - \frac{2\gamma a}{1-\gamma}}{2+(n-2)a^2} & \text{if } i > 2 , \end{cases}$$

and for $a = \frac{2\gamma}{1-\gamma}$ we see that:

$$\mathbf{x}'_i(M\mathbf{x}')_i = \begin{cases} \frac{1}{2+(n-2)\frac{(2\gamma)^2}{(1-\gamma)^2}} & \text{if } i = 1 \\ \frac{1}{2+(n-2)\frac{(2\gamma)^2}{(1-\gamma)^2}} & \text{if } i = 2 \\ 0 & \text{if } i > 2 . \end{cases}$$

Since

$$\frac{1}{2 + (n-2)\frac{(2\gamma)^2}{(1-\gamma)^2}} < \frac{1}{(n-2)\frac{(2\gamma)^2}{(1-\gamma)^2}} = \frac{(1-\gamma)^2}{(2\gamma)^2(n-2)} ,$$

there is a vector \mathbf{x}' , with $\|\mathbf{x}'\|_2 = 1$, such that

$$\max_{i \in [n]} \mathbf{x}'_i(M\mathbf{x}')_i < \frac{(1-\gamma)^2}{(2\gamma)^2(n-2)} ,$$

implying that

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

□

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. *J. ACM*, 60(1):1, 2013.
- [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] M. Puterman. *Markov decision processes*. Wiley, 1994.
- [18] 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.
- [19] L. Y. Rüst. *The P-matrix linear complementarity problem*. PhD thesis, ETH Zürich, 2007.
- [20] L. S. Shapley. Stochastic games. *Proc. Nat. Acad. Sci. U.S.A.*, 39(10):1095–1100, 1953.
- [21] Y. Ye. *Interior Point Algorithms: Theory and Analysis*. Wiley-Interscience series in discrete mathematics and optimization. Wiley, 1998.
- [22] Y. Ye. The simplex and policy-iteration methods are strongly polynomial for the markov decision problem with a fixed discount rate. *Mathematics of Operations Research*, 36(4):593–603, 2011.