



**HAL**  
open science

# The Regularity of the Value Function of Repeated Games with Switching Costs

Yevgeny Tsodikovich, Xavier Venel, Anna Zseleva

► **To cite this version:**

Yevgeny Tsodikovich, Xavier Venel, Anna Zseleva. The Regularity of the Value Function of Repeated Games with Switching Costs. *Mathematics of Operations Research*, In press, 10.1287/moor.2022.1325 . hal-03868995v1

**HAL Id: hal-03868995**

**<https://amu.hal.science/hal-03868995v1>**

Submitted on 17 May 2023 (v1), last revised 23 May 2023 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Repeated Games with Switching Costs: Stationary vs History-Independent Strategies

Yevgeny Tsodikovich  
Xavier Venel  
Anna Zseleva

WP 2021 - Nr 29

# Repeated Games with Switching Costs: Stationary vs History-Independent Strategies \*

Yevgeny Tsodikovich<sup>†</sup>, Xavier Venel<sup>‡</sup>, Anna Zseleva<sup>§</sup>

May 10, 2021

## Abstract

We study zero-sum repeated games where the minimizing player has to pay a certain cost each time he changes his action. Our contribution is twofold. First, we show that the value of the game exists in stationary strategies, depending solely on the previous action of the minimizing player, not the entire history. We provide a full characterization of the value and the optimal strategies. The strategies exhibit a robustness property and typically do not change with a small perturbation of the switching costs. Second, we consider a case where the minimizing player is limited to playing simpler strategies that are completely history-independent. Here too, we provide a full characterization of the (minimax) value and the strategies for obtaining it. Moreover, we present several bounds on the loss due to this limitation.

*Journal* classification numbers: C72, C73.

Keywords: Switching Costs, Repeated Games, Stochastic Games, Zero-sum games.

---

\*The authors wish to thank E. Solan, D. Lagziel, J. Flesch, and G. Ashkenazi-Golan for their highly valuable comments. YT acknowledges the support of the French National Research Agency Grant ANR-17-EURE-0020, and by the Excellence Initiative of Aix-Marseille University - A\*MIDEX. XV acknowledges the financial support by the National Agency for Research, Project CIGNE (ANR-15-CE38-0007-01).

<sup>†</sup>Aix Marseille Univ, CNRS, AMSE, Marseille, France. e-mail: [yevgets@gmail.com](mailto:yevgets@gmail.com)

<sup>‡</sup>Dipartimento di Economia e Finanza, LUISS, Viale Romania 32, 00197 Rome, Italy, [xvenel@luiss.it](mailto:xvenel@luiss.it), <https://orcid.org/0000-0003-1150-9139>

<sup>§</sup>Maastricht University, School of Business and Economics, Dept. of Quantitative Economics, P.O. Box 616, 6200 MD Maastricht, The Netherlands, e-mail: [anna.zseleva@maastrichtuniversity.nl](mailto:anna.zseleva@maastrichtuniversity.nl)

# 1 Introduction

No time instance is an island. Past decisions typically affect the present, either by changing the set of available actions, changing the payoff associated with them or simply through the inertia that makes changing actions costly. For example, consider a guard patrolling among several remote warehouses. Whenever he reaches a warehouse, staying there is costless, while continuing the patrol and moving to another location incurs costs (e.g. fuel). Hence, his optimal route depends on the fuel costs and if these are high enough, he might significantly reduce his movements and choose to stay in a single location most of the time.

Switching costs (or “menu costs” [Akerlof and Yellen, 1985a,b]) appear naturally in different economic settings, such as markets [Beggs and Klemperer, 1992] and bargaining [Caruana et al., 2007]. They can reflect the real cost of changing one’s action (costs of setting up machinery to increase production, shut-down costs of reducing production and adapting machinery from one production line to another) or be a method of modeling bounded rationality when changing actions might be “hard” [Lipman and Wang, 2000, 2009]. In a management quality context, switching costs can be related to changing a company policy or marketing strategy. In these examples, the underlying game is a non-zero-sum game and the addition of switching costs creates new opportunities for the players to cooperate to increase their payoffs. This occurs when the switching costs either eliminate existing equilibria (without switching costs) with low payoffs or turn a Pareto-optimal strategy profile into an equilibrium, such as the cooperative strategy profile in the repeated Prisoner’s Dilemma (even the finitely repeated one, see Lipman and Wang [2000, 2009]). On the other hand, switching costs might impede punishments by making switching to a punishment strategy costly.

To the best of our knowledge, apart from studies focused on specific games (such as the Blotto game in Shy [2019]), switching costs have never been thoroughly addressed theoretically in the setting of repeated normal-form zero-sum games. Schoenmakers et al. [2008] studied an equivalent zero-sum model, in which a constant bonus is granted each time the previous action is not changed. In this paper, we generalize their work<sup>1</sup> to non-constant switching costs, discuss how optimal strategies can be approximated using simpler strategies, and study limit cases of small and large switching costs (compared to the payoffs of the game) and action-independent switching costs. In addition, we extend the comparison of stationary and simpler strategies to a more general class of stochastic games than that derived from adding switching costs to a normal-form zero-sum game.

---

<sup>1</sup>Receiving a bonus for repeating an action is equivalent to paying a fine for not repeating the action. Thus, up to an additive constant, their model is equivalent to a private case of our model studied in Section 3.4.

**Our model.** We consider a repeated normal-form zero-sum game. At each time step, the minimizing player pays the maximizing player both the “standard” outcome of the game and an additional fine if he switched his previous action (the maximizing player incurs no switching costs). This fits situations where the minimizing player performs a worst-case analysis and assumes that the other player is an adversary benefiting from any loss of his utility. For example, the security guard can assume that the opponent is an adversary that benefiting from all his losses, including fuel costs. A similar assumption is made when both players have switching costs that are paid to the “bank” rather than to each other. In this case, which is no longer zero-sum, when the minimax value of each player is calculated, the switching costs of the other players play no role. We leave for future research the analysis of situations where both players have switching costs that are paid to each other, or the study of equilibria in models where the switching costs are paid to a “bank”.

Unlike the models most closely resembling ours, Schoenmakers et al. [2008] and Lipman and Wang [2009], we allow the switching costs to depend on the action. For the patrolling guard example, this captures the idea that there may be different costs involved in reaching different locations, based on distance and topography. We assume that there is a multiplicative factor  $c$  applied to the switching costs. When we increase this factor, the switching costs increase, while the ratio between the costs of two different switches remains the same (an increase in fuel costs increases the cost of movement between warehouses without changing the distances between them). By changing factor  $c$ , therefore, we can study how the value of the game and the optimal strategies change when the switching costs change uniformly, without affecting the internal structure of the costs.

**Applications.** Our model fits different real-life situations. In the security guard example (which is in fact an analogy to a cyber-security problem, see Liuzzi et al. [2020]), the attacker benefits both from attacking a warehouse and from causing the defender disutility due to his defensive efforts. In Schoenmakers et al. [2008], the switching costs represent short-term learning, so changing from one action to another is costly in that it disregards the learning from the previous stage. For a non-zero-sum game, our model can be applied to calculate the reservation payoff of each player, i.e., the minimum payoff he can guarantee for himself when his switching costs are taken into account. The reader can find more examples of zero-sum and non-zero-sum games in papers such as Lipman and Wang [2000, 2009].

The additive nature of the payoffs naturally becomes a factor when dealing with multi-objective minimization. Rass and Rainer [2014] proved that any equilibrium in a game with multiple objectives is a minimax strategy in a game with a single objective, where the payoff

function is some weighted average of all the objective functions. Hence, if the minimizing player wishes to minimize both his “regular” payoff in the repeated game and his switching costs, an alternative approach is to consider a game where the two are weighted and summed, as in our model.

Finally, in our model, the players have undiscounted payoffs. This is typical of cybersecurity, where the number of stages played over a short time (hours) is very large. The cybersecurity literature has a variety of reasons for looking at games with switching costs. For example, Liuzzi et al. [2020] mention how changing passwords is costly. They also point out that the patrolling game example can be generalized to other moving target defenses: for example, changing the configuration of a computer system where costs could include downtime of servers and hourly rates of staff. Zhu and Başar [2013] mention, amongst many other examples, the deceptive routing game where a defense strategy misleads a jammer into attacking legitimate routes by deploying fake routes in wireless communication networks.

**Optimal strategies.** A game with switching costs is equivalent to a stochastic game [Shapley, 1953] where the states correspond to the previous action taken by the minimizing player. Since this game is a single-controller stochastic game (only he controls the states), there should be an optimal strategy that is *stationary* (depends solely on the state and not the time) and can be computed using one of several standard tools (see for example the technique in Appendix A, as well as Filar [1980], Stern [1975], Raghavan [2003], Raghavan and Syed [2002] and others). In practice, however, such an approach suffers from two disadvantages. First, if in the original game the player has  $n$  pure actions, a stationary strategy in the stochastic game assigns a mixed action to each state, which yields  $n^2 - n$  decision variables. This poses a computational challenge that leads to the introduction of new simpler strategies, at the expense of some payoff. Second, in some applications such as cybersecurity, there is an additional requirement that the mixed actions of the minimizing player (defender) should be the same for every state [Rass et al., 2017, Wachter et al., 2018]. We therefore also consider strategies that satisfy this additional constraint, and thus do not depend on the past. We call strategies that utilize the same mixed action at each time step “*static strategies*”. Although the problem of finding an optimal static strategy<sup>2</sup> is NP-hard, there exist algorithms that efficiently find an approximate solution [Liuzzi et al., 2020]. In this paper, we compare the optimal payoff in static strategies to the value of the game (achievable through stationary strategies) in different scenarios and provide bounds on the loss incurred from applying this

---

<sup>2</sup>Typically the game has no value in static strategies. Hence, hereafter, when discussing the value in such strategies, we are referring to the minimax value, and when discussing optimal strategies, the minimax strategies.

constraint on strategies.

**Our contribution.** The contribution of this paper is twofold. First, we study the problems arising from zero-sum games with switching costs and provide a complete characterization of the payoff function and the optimal stationary strategies. We show that these strategies belong to a particular finite set that depends only on the underlying one-shot game, not the switching costs. This can be used as a basis for an algorithm to compute them. Second, we study the value that can be obtained by using simpler static strategies and provide bounds on the loss incurred by using them instead of stationary ones. We show that in some cases, the value is obtained by static strategies, so there is no loss at all. In special cases (such as switching costs that are independent of the actions, as in Lipman and Wang [2000, 2009], Schoenmakers et al. [2008]), we can also provide a complete characterization of the payoff function in static strategies and show that in this case too, the optimal strategies belong to a particular finite set which depends solely on the one-shot game.

In particular, we find that the value function in stationary strategies is piece-wise linear in the weight of the switching costs,  $c$ . This means that the optimal strategy depends only on the segment in which  $c$  is situated, not on the exact value of  $c$ . Moreover, our finding has two important implications. First, it means that the player can identify the optimal strategy without knowing the exact weight of the switching costs relative to the game (which is true in the setting of Rass and Rainer [2014]). Second, it allows the player to assess the improvement in value that can be obtained if the switching costs are reduced, as the optimal strategy does not change for small changes of  $c$ . For example, if the guard can switch suppliers and reduce fuel costs by a small amount (say 1%), our result ensures him that he does not need to adapt his strategy to the new cost (it is still optimal) and that the improvement in payoff is linear in this 1%.

We show that the properties of the value function under stationary strategies also hold for the value under static strategies, assuming that switching costs are independent of the actions being switched. Typically this is not the case, and the optimal strategies might strongly depend on  $c$ . We provide additional results on the optimal strategies in specific games by adding stronger assumptions on the structure of the game or the switching costs. For example, if in the game without switching costs the minimizing player has a unique optimal strategy, this strategy remains optimal (both static and stationary) in the presence of small switching costs, as suggested by Lipman and Wang [2009]. Similarly, we show that when switching costs are symmetric (the amount paid when switching from pure action  $i$  to  $j$  is the same as when switching from  $j$  to  $i$ ), the bound on the difference between the value

in stationary and static strategies is smaller than when switching costs are not symmetric. We then generalize our work to wider classes of games and obtain general bounds on the loss arising from using history-independent strategies relative to history-dependent ones. In particular, we show that at least a quarter of the optimal payoff can be obtained using very simple history-independent strategies.

**Structure of the paper.** This section is followed by a short review of the related literature. In Section 2 we formally present the model. In Section 3 we provide our results. We start with a general characterization of the value functions (Section 3.1), followed by a series of theorems concerning the bound on the difference between the value in stationary and in static strategies. Then, in Section 4, we consider static strategies for stochastic games and deduce appropriate bounds. A short discussion follows in Section 5. In Appendix A we present the average cost of optimality (ACOE) algorithm which is used to compute optimal stationary strategies and the value of the game. The solved example in the Appendix also provides some intuition regarding our results and in particular Proposition 1.

## 1.1 Related Literature

Switching costs have been studied in different economic settings, mainly in non-zero-sum games, where the main questions concerned the existence of a Folk Theorem, players' ability to credibly punish each other, and the commitment that switching costs add. A common simplifying assumption is that all switching costs are identical and independent of the actions being switched [Lipman and Wang, 2000, 2009]. This requires only one additional parameter to cover switching costs and allows concise analysis of the effect of the magnitude of switching costs on different aspects of the game. We allow switching costs to depend on actions, but take the same approach, adding one variable and studying its effect on the outcome of the game. In our model, this variable is a multiplicative factor affecting all switching costs in the same manner. This allows us to capture more real-life situations and consider wider-ranging possibilities in different models.

In addition to this generalization, our paper fills two gaps in the switching costs literature. First, except for Schoenmakers et al. [2008], zero-sum games are generally overlooked in explorations of switching costs. Although they can be viewed as merely a sub-case of general games, the considerations addressed for zero-sum games are substantially different from those for non-zero-sum games and therefore require a specific focus. This paper, by considering zero-sum games, both addresses zero-sum situations and provides a reference point for non-zero-sum games. For example, when determining the worst-case payoff (the individually



rational payoff), the auxiliary game to consider is a zero-sum game.

It should be noted that even Schoenmakers et al. [2008] did not consider games with switching costs, but rather games with bonuses, where one player receives a bonus from repeating the same action, due to a “learning by doing” effect. They argued that if the opponent is a computer, then it does not “learn” and should not receive a bonus for playing the same action. However, using equivalent transformations, the model of Schoenmakers et al. [2008] can be rewritten to fit our model, with the additional restriction that switching costs are independent of actions. Thus, our paper generalizes their work in terms of optimal value in stationary and static strategies (which they call “simple strategies”) and their dependence on switching costs relative to the payoffs of the game. Since their model is more specific than ours, they are able to provide additional results. For example, by assuming that the one-shot payoff matrix is regular, they provide a formula for the value of the game.

Second, although static strategies are considered by many (Schoenmakers et al. [2008] and Bomze et al. [2020] are two prominent examples), the main focus is on the computational aspects of the optimal static strategies or approximations of them [Liuzzi et al., 2020]. We provide theoretical background for the use of static strategies and can explain several phenomena reported by these empirical works, such as the fact that optimal strategies only change slightly (if at all) in response to a small change in the ratio between switching costs and stage payoffs. We believe that our work contributes both by examining the quality of approximations of optimal strategies and by suggesting some ways to improve the computational aspects of the algorithms.

In particular, we provide a bound on the difference between the optimal stationary value and the optimal static value. Using our theoretical bound, the “price of using static strategies” can be calculated and compared to other factors, such as computational complexity. Such a comparison has never been made before, since solving the game and finding the value in stationary strategies is generally very challenging computationally.

## 2 The Switching Costs Model

A *zero-sum game with switching costs* is a tuple  $\Gamma = (A, S, c)$  where  $A = (a_{ij})$  is an  $m \times n$  matrix,  $S = (s_{ij})$  is an  $n \times n$  matrix and  $c \geq 0$ . At each time step  $t$ , Player 1 (the row player, the maximizer, she) chooses an integer<sup>3</sup>  $i(t)$  in the set  $\{1, \dots, m\}$  and Player 2 (the column player, the minimizer, he) chooses an integer  $j(t)$  in the set  $\{1, \dots, n\}$ . The stage payoff that

---

<sup>3</sup>We assume action sets are finite. Without loss of generality, we identify them as integers representing the corresponding row or column in the payoff matrix.

Player 2 pays Player 1 is  $a_{i(t)j(t)} + cs_{j(t-1)j(t)}$ , so Player 2 is penalized for switching from previous action  $j(t-1)$  to action  $j(t)$  by  $cs_{j(t-1)j(t)}$ . We assume that  $s_{kj} \geq 0$  and  $s_{jj} = 0$ , for all  $k, j$ . Moreover, we normalize  $S$  by assuming  $\min_{s_{kj} \neq 0} s_{kj} = 1$ . Due to the multiplicative factor  $c$ , this normalization is without loss of generality. In some of the examples below we avoid it for ease of exposition.

At each time period, the players are also allowed to play mixed actions. A mixed action is a probability distribution over the player's action set. Note that at time  $t$ ,  $j(t-1)$  is already known, even if Player 2 played a mixed action at time  $t-1$ . The process repeats indefinitely, and the payoff is the undiscounted average  $(\liminf)^4$ . More precisely, let  $(\sigma, \tau)$  be a pair of strategies in the repeated game and  $(\sigma(t), \tau(t))$  the mixed actions played at stage  $t$  (given the history). We define

$$\gamma(\sigma, \tau) = \mathbb{E}_{\sigma, \tau} \left( \liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T u(\sigma(t), \tau(t)) \right) \quad (1)$$

where  $u(i(t), j(t)) = a_{i(t)j(t)} + cs_{j(t-1)j(t)}$  is the stage payoff defined above (no switching cost is paid at  $t = 1$ ).

When  $c = 0$ , this simply involves indefinitely repeating the one-shot matrix game  $A$ . Without loss of generality, we assume in the theorems that the minimal and the maximal entries in  $A$  are 0 and 1 (but not in the examples, for clarity). We define  $\mathcal{A}$  as the set of optimal strategies for Player 2 in the one-shot game  $A$ ,  $v$  the value of the game and  $\bar{v}$  the minmax value in pure strategies. In the rest of the paper, we will assume that  $A$  and  $S$  are fixed and examine how the optimal strategies of the players and their payoffs change with  $c$ , the relative weight of the switching costs compared to the stage payoff of game  $A$ .

This repeated game is equivalent to a stochastic game [Shapley, 1953] where each state represents the previous action of Player 2. The set of actions in all states is the same as in  $A$  and the payoff in state  $k$  when Player 1 plays  $i$  and Player 2 plays  $j$  is  $a_{ij} + cs_{kj}$ . In this stochastic game, only Player 2 controls the state transitions.<sup>5</sup> This class of games was studied by Filar [1980], who showed that the value exists and obtained in strategies that depend solely on the current state, not on the entire history of the game, namely *stationary strategies*. It follows that the value of the repeated game exists as well, and it is obtained in strategies that depend solely on the previous action of Player 2. We name these strategies

---

<sup>4</sup>Other methods of defining the undiscounted payoff can be considered, all of which coincide in this model according to Filar [1980].

<sup>5</sup>The initial state is irrelevant, as in each time step all states are reachable and the payoff from any single stage is negligible compared to the long-term average. For every state  $s$ , Player 2 can ensure that the game starting at  $t = 2$  will have initial state  $s$ .

too *stationary strategies* and denote the value for each  $c$  by  $v(c)$ .

**Definition 1.** A stationary strategy is a strategy that depends in each  $t$  on the previous action of Player 2,  $j(t-1)$ , but not on the rest of the history or  $t$  itself. Hence, a stationary strategy is a vector of  $n$  mixed actions, one to follow each possible pure action of Player 2.

**Example 1.** A variation of matching pennies with uniform switching costs.

Suppose  $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$  and  $S = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ . Without switching costs ( $c = 0$ ), the optimal strategy for Player 2 is  $\frac{1}{11}(6, 3, 2)$  and the value is  $v = \frac{6}{11}$ . For clarity, we denote the actions of Player 2 as  $L$  (left),  $M$  (middle) and  $R$  (right).

In the equivalent stochastic game, there are three states representing the situation after each of the actions is played. We denote by  $s_i$  the state where Player 2 played action  $i \in \{L, M, R\}$  in the previous time period. The last row represents the next state after each column was chosen:

$$s_L : \begin{pmatrix} 1 & c & c \\ 0 & 2+c & c \\ 0 & c & 3+c \\ s_L & s_M & s_R \end{pmatrix} \quad s_M : \begin{pmatrix} 1+c & 0 & c \\ c & 2 & c \\ c & 0 & 3+c \\ s_L & s_M & s_R \end{pmatrix} \quad s_R : \begin{pmatrix} 1+c & c & 0 \\ c & 2+c & 0 \\ c & c & 3 \\ s_L & s_M & s_R \end{pmatrix}$$

In stationary strategies, each player chooses three mixed strategies, one for each state.  $\diamond$

We also consider a simpler class of strategies, which are history-independent.

**Definition 2.** A static strategy is a strategy that plays the same mixed action in each stage, regardless of the history. Hence, a static strategy is one mixed action played repeatedly.

Typically, when we limit the players to static strategies, the value may not exist. This can be understood if we compare the maximin and the minimax of the repeated game in static strategies. In terms of the maximin, the best response of Player 2 can be pure and no switching is needed. On the other hand, the minimizing strategy of Player 2 is typically mixed and the payoff depends on  $c$ . Therefore, when considering static strategies, the figure of merit we study is the minimax – the maximal cost that Player 2 can guarantee to pay.

If Player 1 uses static strategy  $x$  and Player 2 plays static strategy  $y$ , the payoff can be written in matrix notation as

$$g(c)(x, y) = x^T A y + c y^T S y. \quad (2)$$

If Player 2 is limited to static strategies, Player 1 can best respond with a static strategy (actually, a pure one) and the minimax can be defined as

$$\tilde{v}(c) = \min_y \max_x g(c)(x, y). \quad (3)$$

For simplicity, we sometimes refer to this as the value in static strategies, but it should be understood as only the *minimax value*.

Our results are presented in two parts. First, we characterize both value functions. Next, we bound the difference between them, which represents Player 2’s possible loss from playing a static strategy instead of a stationary strategy (clearly,  $\tilde{v}(c) \geq v(c)$ ).

**The weighted average model.** Another natural way to include switching costs in the payoff is by considering some weighted average between the stage payoff and the switching costs. This was done by Rass and Rainer [2014] and others [Rass et al., 2017, Wachter et al., 2018, Liuzzi et al., 2020], who studied a model where players use static strategies and the payoff is a weighted average of the switching costs and the costs of the game itself:  $\alpha x^T A y + (1 - \alpha) y^T S y$ , with  $\alpha \in [0, 1]$ . Their model is slightly different from our model. For example, in our work, it is straightforward that as  $c$  increases, the situation worsens for Player 2. However, in their case, when the weight on the switching costs tends to 1, the value of the game (with switching costs) tends to 0 and the situation “improves” if the value without switching costs was positive.

Nevertheless, the models are equivalent. Setting  $c = \frac{1-\alpha}{\alpha}$ , we get the same ratio between the two payoffs and, as a result, the games are strategically equivalent. Hence, our work sheds some theoretical light on their results as well as on general economic situations in which the switching costs are added to the real costs and not averaged with them.

### 3 Results

In Subsection 3.1, we characterize the value function  $v(c)$  and the minimax value function in static strategies  $\tilde{v}(c)$ . We benefit from these characterizations in Subsection 3.2, where we bound the difference between the two for any game with switching costs. In Subsections 3.3 and 3.4, we find tighter bounds by considering a more restrictive model. With the appropriate restrictions on  $c$ , the bound can even reduce to zero, which implies that in some cases there is a static strategy that is optimal. We also make restrictions on the structure of the switching-cost matrix  $S$  by making the off-diagonal elements equal.

Clearly, if Player 2 has an optimal *pure* action in  $A$  without switching costs, then this action is also optimal with switching costs and  $v(c) = \tilde{v}(c) = v(0)$  for all  $c \geq 0$ . In the rest of the paper, we therefore assume that the value of the game  $A$  without switching costs cannot be obtained in pure actions.

### 3.1 The Properties of the Value Functions

In this subsection we characterize  $v(c)$  where  $c \geq 0$ , namely the value function in stationary strategies (Proposition 1). We show that there are finitely many segments of the nonnegative real line such that for each segment there exists an optimal stationary strategy of Player 2 that is independent of the exact  $c$  (Corollary 1). We also characterize  $\tilde{v}(c)$  where  $c \geq 0$ , namely the minimax value function in static strategies (Proposition 2). We continue Example 1 to demonstrate these value functions and the corresponding optimal strategies. We then provide two counterexamples to two natural conjectures. In Example 2, we show that Corollary 1 does not hold for  $\tilde{v}(c)$  in general. In Example 3, we show that it is possible for stationary strategies to strictly outperform static strategies for all  $c > 0$ .

**Proposition 1.** *For every  $c \geq 0$ , the game has a value in stationary strategies denoted by  $v(c)$ . This function is continuous, increasing, concave, and piece-wise linear.*

**Proof** This game is equivalent to a stochastic game where the states correspond to the previous pure actions chosen by Player 2. In this stochastic game, only one player controls the transitions, so existence follows from Filar [1980] and Filar and Raghavan [1984].

Moreover, according to Filar and Raghavan [1984], the value of the game is the same as the value of a one-shot matrix game, whose pure actions are the pure stationary actions in the stochastic game (essentially, each pure strategy in the one-shot game is a vector of size  $n$  dictating which pure actions to choose at each state). The payoffs include the corresponding payoffs from matrix  $A$  as well as  $c$ -dependent payoffs from the transitions. Since only Player 2 controls the states, the  $c$ -dependent part is the same for all entries in the same column and is linear in  $c$ . For example, if the pure action of Player 2 is to alternate between actions 2 and 1, the  $c$ -dependent part of the column is  $\frac{s_{12}+s_{21}}{2}c$ . We denote the entry in the  $i, j$  place of this one-shot game by  $b_{ij} + \beta_j c$ .

In order to characterize  $v(c)$ , we look at the value function of this one-shot matrix game. The value of this game is determined by the optimal strategy profiles, which are the same as the Nash-equilibrium strategy profiles. We show that given the support of the Nash-equilibrium strategy profiles, the expected payoff from Player 1's best response is linear in  $c$ . We also show that the value function is pieced together from these finitely many linear functions, so it is piece-wise linear in  $c$ .

Fix  $c' \geq 0$  and let  $I_1$  ( $I_2$ ) be some subset of the rows (columns) of the one-shot game. We check whether Player 2 can make Player 1 indifferent among all the actions in  $I_1$ , using a completely mixed action over  $I_2$ , and can make Player 1 prefer them over actions not in

$I_1$ . Let  $k$  be a pure action in  $I_1$ . Player 1 is indifferent among the actions in  $I_1$  if

$$\forall l \in I_1 \setminus \{k\} : \sum_{j \in I_2} y_j (b_{kj} + \beta_j c') = \sum_{j \in I_2} y_j (b_{lj} + \beta_j c'), \quad (4)$$

and prefers them over other actions if

$$\forall l \notin I_1 : \sum_{j \in I_2} y_j (b_{kj} + \beta_j c') \geq \sum_{j \in I_2} y_j (b_{lj} + \beta_j c'). \quad (5)$$

The  $c'$ -dependent part cancels out. If we add  $\sum_{j \in I_2} y_j = 1$  and  $\forall j \in I_2 : y_j > 0$ , we get a system of linear equations and inequalities in  $|I_2|$  variables which is independent of  $c'$ . If this system has a solution, then there exists at least one solution that is independent of  $c'$ . Let us denote it by  $y(I_1, I_2)$ .

When Player 2 plays according to  $y(I_1, I_2)$ , the expected payoff of each row is a linear function of  $c$  with the same slope (as shown above), and a higher constant for the  $I_1$  rows relative to the others. Hence, when Player 1 best responds, he chooses an action in  $\Delta(I_1)$  and the payoff will be a linear function of  $c$ , denoted by  $l(I_1, I_2)(c)$ . Note that we did not establish that  $y(I_1, I_2)$  is optimal for Player 2. It is optimal only if it is a best response to Player 1's best response to  $y(I_1, I_2)$ , or, equivalently, if Player 1 can make Player 2 indifferent among all actions in  $I_2$  by mixing the actions in  $I_1$  (which typically depends on  $c$ ).

For a given  $c^*$ , take an optimal strategy profile. This profile has a corresponding support pair  $(I_1^*, I_2^*)$ . Then  $y(I_1^*, I_2^*)$  is an optimal strategy for Player 2 and  $v(c^*) = l(I_1^*, I_2^*)(c^*)$ .

There are finitely many support pairs  $(I_1, I_2)$ , and at every  $c$  the value function equals one of their corresponding lines  $l(I_1, I_2)$ . Since the value function  $v(\cdot)$  is continuous in  $c$  (this is a polynomial game, see Shapley and Snow [1950]), it is also piece-wise linear in  $c$ .

In addition, since all the entries in  $S$  are non-negative, all the coefficients of  $c$  (the slopes of the lines) are non-negative, so  $v(c)$  is an increasing function of  $c$ .

Finally, the function  $v(c)$  is concave. Let  $[c_k, c_{k+1}]$  and  $[c_{k+1}, c_{k+2}]$  be two segments where  $v(c)$  is linear. For  $c^* \in (c_k, c_{k+1})$ , let  $y(c^*)$  be the optimal action of Player 2 chosen using the above method. If Player 2 plays  $y(c^*)$  regardless of  $c$ , the payoff is a linear function of  $c$  that coincides with the value on  $[c_k, c_{k+1}]$ . In the region  $[c_{k+1}, c_{k+2}]$ , the value function must be below this line, so the slope of  $c$  must decrease. Hence,  $v(c)$  is a piece-wise linear function with decreasing slope and thus it is concave. ■

If on an interval the value function  $v(c)$  is linear, then there is a Player 2 strategy that is optimal for this entire interval, irrespective of the precise value of  $c$ . A direct result of the piece-wise linearity of  $v(c)$  is that the exact value of  $c$  does not need to be known to

find an optimal strategy, only the segment that contains it matters. This means that the optimal strategies are robust to small changes of  $c$ : knowing the exact  $c$  is not necessary to play optimally, and it is almost universally unnecessary to adjust the optimal strategy as  $c$  changes. Moreover, this shows that when the model arises from a weighted average of different goal functions [Rass and Rainer, 2014], the exact weight given to the switching costs compared to the stage payoff is of smaller significance, if any (as the authors indeed report for static strategies).

Even if there is no information on  $c$ , an optimal strategy for Player 2 can still be found in a finite set that depends only on  $A$ . This observation can serve as a basis for an algorithm to compute an optimal stationary strategy in the game with switching costs. Interestingly, this is not true for Player 1. In general, the optimal strategy for Player 1 depends on the exact  $c$  and changes continuously with  $c$ .

**Corollary 1.** *There exists  $K \in \mathbb{N}$  and  $0 = c_0 < c_1 < \dots < c_K < c_{K+1} = \infty$  such that for each segment  $[c_i, c_{i+1}]$  there exists an optimal stationary strategy of Player 2 that is independent of the exact  $c$ .*

**Proof** Since the function  $v(c)$  is piece-wise linear, the segments are chosen such that in each one,  $v(c)$  is linear. Fix  $c^* \in (c_i, c_{i+1})$  and fix  $\sigma_{c^*}$  an optimal strategy for Player 2 in the stochastic game. Suppose Player 2 plays  $\sigma_{c^*}$  regardless of  $c$  and Player 1 best responds.

When Player 2 plays  $\sigma_{c^*}$ , he fixes the transition probabilities between states, and these are now independent of  $c$ . Hence, the percentage of time spent in each state is constant and independent of  $c$ . In each state, all entries in a particular column have the same slope (in state  $i$ , column  $j$  has slope  $s_{ij}$ ), so their mix according to  $\sigma_{c^*}$  results in parallel lines (parallel in each state, not necessarily parallel between states). Player 1 best responds in each state, but since the lines are parallel, his choice does not depend on  $c$ . Thus, the payoff in each state is a linear function of  $c$ , and the total payoff, which is a weighted average of these functions with the percentage of time spent in each state as weights, is a linear function of  $c$ . Denote this function by  $f_{c^*}(c)$ .

Since  $v(c)$  is the value,  $v(c) \leq f_{c^*}(c)$  for all  $c$  in the segment with equality at  $c^*$ . Thus, these two lines that intersect once must coincide and  $\sigma_{c^*}$  obtains the value in the entire segment. ■

The concavity of  $v(c)$  can be intuitively understood as follows. For a large  $c$ , it is more costly to switch so Player 2 chooses strategies that incur smaller switching costs despite some loss in the stage game  $A$ . Hence, as  $c$  increases, the coefficient of  $c$  decreases and the rate at which the value increases with  $c$  is reduced. This observation can serve in practice

to eliminate from the choice of strategies those with too-high expected per-stage switching costs.

The minimax value in static strategies exhibits the same properties, except piece-wise linearity. Here, it can only be established that the value function is semi-algebraic, so there is no version of Corollary 1 for this case. It is quite possible that the optimal strategy depends on  $c$ , as shown in Example 2.

**Proposition 2.** *The minimax value  $\tilde{v}(c)$  in static strategies is a continuous, increasing, and concave semi-algebraic function.*

**Proof** Let  $g(c)(x, y) = x^T Ay + cy^T Sy$  be a function from  $\mathbb{R}_+ \times \Delta(\{1, \dots, m\}) \times \Delta(\{1, \dots, n\})$  to  $\mathbb{R}$ . By definition,

$$\tilde{v}(c) = \min_y \max_x g(c)(x, y) = \min_{y \in \Delta(\{1, \dots, n\})} \left\{ \max_{x \in \Delta(\{1, \dots, m\})} \{x^T Ay\} + cy^T Sy \right\}. \quad (6)$$

Note that  $\arg \max_x g(c)(x, y)$  depends solely on  $y$  and not on  $c$  (and can always be one of  $\{1, \dots, m\}$ ). Since  $g(c)(x, y)$  is continuous and since the maximum and the minimum of a continuous function is also continuous, so is  $\tilde{v}(c)$ . Moreover,  $\tilde{v}(c)$  is semi-algebraic since  $g(c)(x, y)$  is a polynomial in each variable.

*The function  $\tilde{v}(c)$  is increasing:* Let  $0 \leq c_1 < c_2$ . We show that  $\tilde{v}(c_1) \leq \tilde{v}(c_2)$ . Let  $y_{c_2}$  be the arg min of the minimization in Eq. (6).

$$\tilde{v}(c_2) = \max_{x \in \Delta(\{1, \dots, m\})} \{g(c_2)(x, y_{c_2})\} \geq \max_{x' \in \Delta(\{1, \dots, m\})} \{g(c_1)(x', y_{c_2})\} \geq \tilde{v}(c_1).$$

The equality follows from the definition of  $\tilde{v}(c_2)$ . The first inequality follows from  $c_1 < c_2$  and  $S \geq 0$  (the maximizing  $x$  is the same as it depends solely on  $y_{c_2}$ ). The last inequality follows from the definition of  $\tilde{v}(c_1)$ .

*The function  $\tilde{v}(c)$  is concave:* Let  $0 \leq c_1 < c_2$  and let  $\beta \in (0, 1)$ . Then

$$\begin{aligned} \tilde{v}(\beta c_1 + (1 - \beta)c_2) &= \min_y \left\{ \max_x \{x^T Ay\} + (\beta c_1 + (1 - \beta)c_2)y^T Sy \right\} \\ &= \min_y \left\{ \beta \max_x g(c_1)(x, y) + (1 - \beta) \max_x g(c_2)(x, y) \right\} \\ &\geq \min_y \left\{ \beta \max_x g(c_1)(x, y) \right\} + \min_y \left\{ (1 - \beta) \max_x g(c_2)(x, y) \right\} \\ &= \beta \tilde{v}(c_1) + (1 - \beta) \tilde{v}(c_2). \end{aligned}$$

The first and last equalities are the definition of  $\tilde{v}$ . The second equality is based on the fact that  $\arg \max_x g(c)(x, y)$  does not depend on  $c$ . The inequality is obtained since minimizing



each term individually and then summing results yields a smaller result than first summing and then minimizing the entire expression. ■

In the following example, we demonstrate how the two values behave as a function of  $c$ . Only the results are presented here; the calculations are in Appendix A.

**Example 1 (Continued).** *A variation of matching pennies with uniform switching costs.*

**The Value of the Game** –  $v(c)$

*Case 1:*  $c < \frac{22}{31}$ : The optimal strategy for Player 2 in game  $A$  is still optimal for this case, and the value is  $v(c) = \frac{6}{11} + \frac{72}{121}c$ .

*Case 2:*  $\frac{22}{31} \leq c \leq \frac{121}{156}$ : The switch from  $L$  to  $R$  is too costly, so action  $R$  is removed from state  $s_L$  and the optimal strategy becomes  $\frac{2}{3}(L), \frac{1}{3}(M)$ . The optimal strategies in the other states are the same as in the game without switching costs. The value is  $v(c) = \frac{156c+198}{319}$ .

*Case 3:*  $c \geq \frac{121}{156}$ : The switch from  $L$  to  $M$  is too costly, so  $s_L$  becomes an absorbing state with payoff 1.

**The Minimax Value in Static Strategies** –  $\tilde{v}(c)$

In this simple case there is no intermediate strategy for Player 2. He either plays the optimal strategy of game  $A$  when switching costs are low enough or he plays the pure minimax strategy of  $A$  and never switches when they are high enough. A direct computation reveals that the cut-off  $c$  is  $\frac{55}{72}$ .

**Comparison:** Figure 1 shows the static solution  $\tilde{v}(c)$  (dashed) and the stochastic game solution  $v(c)$  (solid) in the region where they differ. The maximal difference is found at  $c = \frac{55}{71}$  and is equal to  $\frac{37}{2088}$ . ◇

With additional assumptions, the result of Proposition 2 can be strengthened. In particular, in the common case where all switches are equally costly (uniform switching costs), the result can be strengthened to piece-wise linearity (Proposition 7). In the latter case, the set of optimal static strategies can be fully characterized and it is finite. This is not true in the general case, where the only conclusion that can be made (see the remarks after Proposition 7) is that when playing an optimal static strategy, Player 2 keeps Player 1 indifferent in game  $A$  between at least two of his pure actions.

**Example 2.** *A game where  $\tilde{v}(c)$  is not piece-wise linear*

Suppose  $A = \begin{pmatrix} -200 & 1 & 200 \\ 200 & 1 & -200 \end{pmatrix}$  and  $S = \begin{pmatrix} 0 & 1 & 100 \\ 1 & 0 & 1 \\ 100 & 1 & 0 \end{pmatrix}$ . It is easy to verify that for  $c \leq \frac{1}{98}$  the optimal strategy (static and stationary) is the optimal strategy of game  $A$  (choosing the first

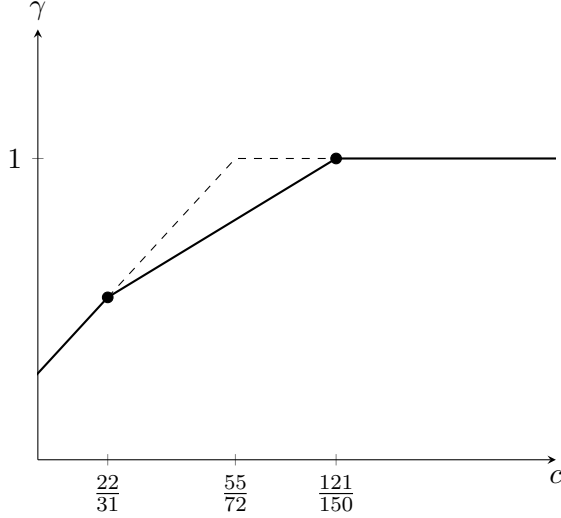


Figure 1: The value in optimal (stationary) strategies  $v(c)$  (solid) and the minimax value in static strategies  $\tilde{v}(c)$  (dashed) of the game from Example 1.

and third columns with equal probability) that leads to a payoff of  $\tilde{v}(c) = 50c$ , and that for  $c \geq \frac{1}{2}$  the optimal strategy (static and stationary) is the pure minimax with a payoff of  $\tilde{v}(c) = 1$ . For  $c \in [\frac{1}{98}, \frac{1}{2}]$ , the optimal static strategy of Player 2 is  $[p(L), 1 - 2p(M), p(R)]$  where  $p = \frac{1-2c}{192c}$ , keeping Player 1 indifferent between his two actions. In this case,  $\tilde{v}(c) = 1 - \frac{(1-2c)^2}{192c}$ , which is not linear in  $c$ .  $\diamond$

Corollary 1 suggests that the optimal stationary strategy in the first segment (which includes  $c = 0$ ) is optimal even without switching costs, so it must be comprised of optimal strategies in game  $A$ . If there is only one optimal strategy in the game (as in Example 1), then the optimal stationary strategy is actually static and in that segment,  $v(c) = \tilde{v}(c)$ . However, in general, this may not be the case. The following example shows that it is possible that stationary strategies will always outperform static strategies (except, of course, at  $c = 0$ ).

**Example 3. A game with  $\tilde{v}(c) > v(c)$  for all  $c > 0$ .**

Suppose  $A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$  and  $S = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$ . The value of  $A$  is 0.5 and is achieved by any mixed action of Player 2 choosing the set of even columns w.p. 0.5. Switching-cost matrix  $S$  is such that, based on the action of Player 2 in the previous time period, the only costless actions are the same action and the action to the right of the previous action. Due to this cyclical nature of  $S$ , Player 2 can achieve the value of game  $A$  with a stationary strategy, by playing the previous action w.p. 0.5 and the action to the right w.p. 0.5. Clearly, any

static non-pure strategy will contain a  $c$  component, and any pure static strategy will result in  $\tilde{v}(c) = 1$ . To conclude,  $\tilde{v}(c) > v(c) = 0.5$  for all  $c > 0$ .  $\diamond$

### 3.2 A Uniform Bound on the Value Differences in Static and Stationary Strategies

Our goal is to compare  $v(c)$  and  $\tilde{v}(c)$  and evaluate any loss  $\tilde{v}(c) - v(c)$  due to using a static strategy instead of a stationary strategy. Our first bound is very general and does not rely heavily on game-specific assumptions (Proposition 4). Then we improve the bound by taking into account additional properties of the game, such as the multiplicity of equilibria in  $A$  and the structure of  $S$  (Theorem 1). We start by showing that in the simplest set of games,  $2 \times 2$  games, the optimal stationary strategy is always a static strategy (Proposition 3). This is a generalization of Theorems 4.1, 4.2 and Corollary 4.1 in Schoenmakers et al. [2008].

**Proposition 3.** *Let  $A$  be a  $2 \times 2$  zero-sum game. For every  $S$  and every  $c$ , there exists an optimal stationary strategy in switching-costs game  $(A, S, c)$  that is a static strategy. Thus,  $\tilde{v}(c) = v(c)$ .*

**Proof** Let  $A = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$  be a zero-sum  $2 \times 2$  game and  $S = \begin{pmatrix} 0 & s_{12} \\ s_{21} & 0 \end{pmatrix}$  the switching-costs matrix for Player 2. We assume that the value in  $A$  without switching costs cannot be obtained using pure strategies (otherwise the discussion is trivial), so w.l.o.g.,  $\alpha > \beta$ ,  $\alpha > \gamma$ ,  $\delta > \beta$ , and  $\delta > \gamma$ . Let  $p \in (0, 1)$  be the probability of playing the left column (action  $L$ ) in  $A$  that achieves the value (a direct computation shows that only one such  $p$  exists). As in the examples, we denote by  $s_i$  the new state when the previous action of Player 2 was  $i \in \{L, R\}$ .

Based on Proposition 1 and the ACOE method (Appendix A), the optimal stationary strategy for Player 2 is such that in each state Player 1 is either indifferent between his two actions or plays purely one of them. Since the same switching costs and continuation payoffs are added to all entries of the same column of each  $s_i$  (see Eq. (7)), they do not affect Player 1's choice of rows. Thus, in each state either Player 2 plays purely too or he plays the mixed action  $p$ . All the possible combinations of these actions are summarized in Table 1. For example, in case I he plays  $p$  in both states, whereas in case IV he plays  $p$  only at state  $s_R$ , while after playing  $L$  (and reaching  $s_L$ ) he always plays purely action  $R$ . In cases II and III he is absorbed in states  $s_L$  or  $s_R$  (resp.), so the action in the other state is irrelevant.

Note that cases I, II, and III represent a static strategy, so to show that  $v(c) = \tilde{v}(c)$  we must show that the strategies presented in cases IV and V cannot be optimal. Assume by contradiction that the strategy presented in case IV is indeed optimal. Let  $\kappa$  be the

$\mathcal{N}^e$	$s_L$	$s_R$
I	$p$	$p$
II	$L$	$?$
III	$?$	$R$
IV	$R$	$p$
V	$p$	$L$

Table 1: All the possible optimal stationary strategies.

continuation payoff after playing  $R$  (the continuation payoff after playing  $L$  is normalized to 0). When we add the switching costs and the continuation payoffs to the games, we obtain two one-shot games:

$$s_L : \begin{pmatrix} \alpha & \beta + \kappa + cs_{12} \\ \gamma & \delta + \kappa + cs_{12} \end{pmatrix} \quad s_R : \begin{pmatrix} \alpha + cs_{21} & \beta + \kappa \\ \gamma + cs_{21} & \delta + \kappa \end{pmatrix}$$

Let us consider state  $s_L$ . By assumption on the coefficients,  $\alpha > \gamma$  and  $\delta + \kappa + cs > \beta + \kappa + cs$ . Since it is optimal to play purely  $R$ , we necessarily have

$$\alpha \geq \beta + \kappa + cs_{12} \text{ or } \gamma \geq \delta + \kappa + cs_{12}.$$

If only the first equation is true, the equilibrium is mixed and  $R$  is not optimal. If only the second equation is true, then there is a contradiction:

$$\alpha > \gamma \geq \delta + \kappa + cs > \beta + \kappa + cs > \alpha.$$

It follows that both equations are true, i.e. the right column dominates the left including the continuation payoffs (this is only true in  $2 \times 2$  games) and in particular  $\gamma \geq \delta + cs_{12} + \kappa$ .

In state  $s_R$ , however, the optimal strategy is mixed and therefore such dominance is not possible. If the direction of the original inequalities remains correct, then  $\delta + \kappa \geq \gamma + cs_{21}$ . It follows that  $s_{12} + s_{21} \leq 0$ , which is a contradiction. The other option is that, since in  $s_R$  both inequalities changed direction, a contradiction can be constructed using the first row.

A similar approach can be used for case V and the proof is complete. ■

Thus, in the  $2 \times 2$  case, there are two candidates for optimal stationary strategy: either playing the optimal strategy of  $A$  regardless of the previous action, or playing the pure minimax action of  $A$ . Both are also static strategies. It is straightforward to calculate the value in both cases and to determine the cut-off value of  $c$  above which playing the pure minimax becomes profitable.

When either of the players has more than 2 actions in  $A$ , the value may differ under static and stationary strategies. The following Proposition bounds this difference. Naturally,

since this Proposition is general, the resulting bound is not easy to calculate or interpret. However, in the Theorems that follow, we add more assumptions to obtain a bound in a simple analytic manner.

**Proposition 4.** *Player 2's loss from limiting himself to static strategies,  $\tilde{v}(c) - v(c)$ , is bounded by*

$$\Delta(c) = \frac{1}{2}(1 - v(c)) + \frac{c}{4}\Xi + \frac{c}{4}M, \quad (7)$$

where

- $M = \max_{y \in \Delta(\{1, \dots, n\})} y^T S y$  is the maximal switching cost that Player 2 can pay,
- $\Xi = \max_{j, k} |S_{jk} - S_{kj}|$  is a measure of the asymmetry of the switching-cost matrix.

**Proof** By definition of value  $v(c)$ , there exists a state  $j \in S$  and a mixed action  $y \in \Delta(J)$  such that for all  $i \in I$  we have that the cost at  $j$  by playing  $y$  is less than  $v(c)$  (otherwise, the payoff of Player 1 in each state can exceed  $v(c)$  and thus can be the total value).

Consider the static strategy  $\tilde{y}$  of Player 2, that with probability 0.5 chooses action  $j$  and otherwise chooses an action according to mixed action  $y$ . We know that if Player 2 plays a static strategy then Player 1 has a static strategy best response. Let  $x$  be this best response. Then, replacing  $\tilde{y}$  by its definition, we have

$$\begin{aligned} g(c)(x, \tilde{y}) &= x^T A \tilde{y} + c \tilde{y}^T S \tilde{y} \\ &= \frac{1}{2} x^T A j + \frac{1}{2} x^T A y + c \frac{1}{4} j^T S j + c \frac{1}{4} j^T S y + c \frac{1}{4} y^T S j + c \frac{1}{4} y^T S y. \end{aligned}$$

Bounding  $y^T S j$  from above with  $j S y + \Xi$  and reorganizing leads to

$$g(c)(x, \tilde{y}) \leq \frac{1}{2} (x^T A y + c j S y) + c \frac{1}{4} \Xi + \frac{1}{2} x^T A j + c \frac{1}{4} j^T S j + c \frac{1}{4} y^T S y.$$

We can now replace the first part of the bound by  $v(c)$ . Recall that  $A$  is normalized so that  $\max_{i, j} a_{ij} = 1$  and  $S$  defined such that  $s_{ii} = 0$  for every  $i \in \{1, \dots, n\}$  so:

$$g(c)(x, \tilde{y}) \leq \frac{1}{2} v(c) + \frac{c}{4} \Xi + \frac{1}{2} + \frac{c}{4} M.$$

It follows from the definition of  $\tilde{v}(c)$  that

$$\tilde{v}(c) \leq \frac{1}{2} v(c) + \frac{c}{4} \Xi + \frac{1}{2} + \frac{c}{4} M.$$

The bound is obtained by subtracting  $v(c)$  from both sides:

$$\tilde{v}(c) - v(c) \leq \frac{1}{2} (1 - v(c)) + \frac{c}{4} \Xi + \frac{c}{4} M = \Delta(c).$$

■

In Proposition 4, the bound depends on  $\Xi$ , which measures the symmetry of the switching-cost matrix. Naturally, the symmetry of  $S$  has no effect on the minimax value in static strategies. Recall that  $g(c)(x, y)$  from Eq. (2) expressed the expected payoff of Player 2 when the players play a static strategy pair  $(x, y)$ . When the expression for  $g(c)(x, y)$  is expanded, each product  $y_i y_j$  is multiplied by  $(s_{ij} + s_{ji})$ . Hence, when solving for  $\tilde{v}(c)$ , only the sum of symmetric off-diagonal elements of  $S$  matters, not their exact values. This, for example, is the reason why Liuzzi et al. [2020] assumed that  $S$  is a symmetric matrix.

A stationary strategy, on the other hand, can take any asymmetry into account, yielding different results. This can easily be seen in Example 3. If we change  $S$  into a symmetric matrix while keeping the sums  $s_{ij} + s_{ji}$  at their current values, the result will be a matrix where every mixed stationary strategy has a  $c$ -component, whereas in the non-symmetric case, the optimal stationary strategy kept  $v(c) = 0.5$  regardless of  $c$ .

**Theorem 1.** *If matrix  $S$  is symmetric, Player 2's loss from limiting himself to static strategies,  $\tilde{v}(c) - v(c)$ , is bounded by  $\Delta(c) = \frac{1}{2}(1 - v(c)) + \frac{1}{4}M$  (where  $M$  is defined in Proposition 4). This formula implies the following bound*

$$\frac{1+cM-\tilde{v}(c)}{1+cM-v(c)} \geq \frac{1}{2}. \quad (8)$$

**Proof** The first bound is a direct corollary from Proposition 4, setting  $\Xi = 0$ . The second part is obtained by relaxing the bound and reorganizing. By Proposition 4, we obtain

$$\tilde{v}(c) - v(c) \leq \frac{1}{2}(1 - v(c)) + \frac{c}{4}M \leq \frac{1}{2}(1 + cM - v(c)).$$

We can deduce that

$$\begin{aligned} 1 + cM - \tilde{v}(c) &= 1 + cM - v(c) + (v(c) - \tilde{v}(c)), \\ &\geq (1 + cM - v(c)) - \frac{1}{2}(1 + cM - v(c)), \\ &= \frac{1}{2}(1 + cM - v(c)). \end{aligned}$$

■

The reorganization of the bound to the fraction form presented in Eq. (8) has an intuitive interpretation: by using static strategies, Player 2 loses at most half of what he could have obtained with a stationary strategy. Had we done the analysis from the point of view of the maximizing player, this Equation and the intuition would be further simplified into: “with static strategies the player guarantees at least half of the value from stationary ones”. This

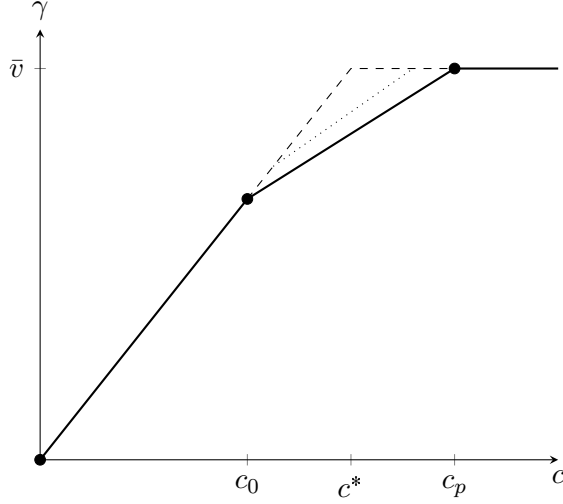


Figure 2: Different value functions for non-symmetric RPS. Solid:  $\Delta = 0$ . Dashed:  $\Delta = 1$ , static strategy. Dotted:  $\Delta = \frac{1}{3}$ .

sets the “price of being static ” at  $\frac{1}{2}$  when  $S$  is symmetric. Such a calculation is possible in other settings too, such as the ones from Theorem 5, where the “price of being static ” is  $\frac{1}{4}$ .

It is important to note, however, that this approach is not always possible. In the following example, when  $S$  is asymmetric the stationary strategy yields a worse result than when  $S$  is symmetric, and in an extreme case – the optimal stationary strategy is actually static.

**Example 4. Rock-Paper-Scissors (RPS) with non-symmetric  $S$ .**

Consider the RPS game with  $A = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}$  and  $S = \begin{pmatrix} 0 & 1+\Delta & 1+\Delta \\ 1-\Delta & 0 & 1+\Delta \\ 1-\Delta & 1-\Delta & 0 \end{pmatrix}$ , where  $-1 \leq \Delta \leq 1$ . The static solution remains the same regardless of  $\Delta$ , and dictates choosing all actions with equal probability if  $c \leq 1.5$  and purely playing one of the actions for larger  $c$ s.

The stationary strategy for  $\Delta = 0$  is to choose all actions with equal probability for  $c \leq 1$  and to play purely for  $c \geq 2$ . For  $c \in [1, 2]$ , the strategy involves cyclic behavior: after  $L$  never play  $R$ , after  $M$  never play  $L$ , and after  $R$  never play  $M$ . In all cases, the previous action is replayed w.p.  $\frac{2}{3}$ .

For  $\Delta = \frac{1}{3}$ , the optimal strategy is  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$  for  $c \leq \frac{9}{8} = 1.125$ . For a larger  $c$ , one action is never played in each state, as discussed in the  $\Delta = 0$  case. This is true until  $c = \frac{9}{5}$ . From then on, the optimal strategy is to play purely.

For  $\Delta = 1$ , the optimal stationary strategy is in fact the static one, and the two cases coincide. All the cases are shown in Figure 2.  $\diamond$

The bound from Proposition 4 is a general bound without making any assumptions on game  $A$  or matrix  $S$ . As shown in Theorem 1, tighter bounds can be tailored to special cases.

In Subsection 3.3 we show that in certain cases, when  $c$  is either small or large, the optimal strategies are static. This implies that the two value functions coincide on these segments and the bound is 0.

In Subsection 3.4, we make further assumptions on the matrix  $S$ . While the diagonal elements in  $S$  are all zeros, the off-diagonal elements are ones. This means that all switches are equally costly, as suggested by Lipman and Wang [2000, 2009]. In this case, not only  $v(c)$  but also  $\tilde{v}(c)$  is piece-wise linear, which strengthens the robustness result and simplifies the process of finding the optimal static strategy. The search domain narrows down to a finite (although large) predetermined set of mixed actions.

### 3.3 Optimality of a static strategy for Small and Large Values of $c$

In this section we tackle two possible assumptions affecting the two extremes of  $c$ . We first show how to identify the case where, for small values of  $c$ , switching costs do not play a role and the optimal stationary strategy is a static one. We next assume that there are no free switches, i.e., that all non-diagonal elements of  $S$  are non-zero. In this case, again for large enough  $c$ , the optimal stationary strategy is a static one, and moreover, pure.

We start with the small  $c$  case. In most models, switching costs are either ignored or considered irrelevant, so comparing a case with no switching costs ( $c = 0$ ) to one with small switching costs is important as a robustness check on this claim. Lipman and Wang [2009] provide one such comparison in discussing the continuity of equilibrium payoffs and equilibrium strategies when switching costs converge to zero. Here, we provide another comparison: between the optimal stationary and static strategies when  $c$  is very small. We have already established that in some cases (Example 1) small switching costs do not matter and the optimal stationary strategy is also a static one, while in other cases (Example 3) there is a difference between the two for every  $c > 0$ . Hence, we are interested in finding conditions enabling us to distinguish between the two types of cases.

Clearly, for  $c = 0$ , value  $v(0)$  and the minimax value in static strategies  $\tilde{v}(0)$  coincide. Let  $\underline{c}$  be the maximal  $c$  such that for every  $c \in [0, \underline{c}]$  the value is attained by a static strategy, so  $v$  and  $\tilde{v}$  still coincide. As Proposition 5 shows, the exact value of  $\underline{c}$  can be found in polynomial time  $O(n^4)$ .<sup>6</sup> If  $\underline{c} > 0$ , it immediately reveals the optimal strategy in the entire

---

<sup>6</sup>With additional assumptions, this can be theoretically studied for particular classes of games. For example, Schoenmakers et al. [2008] showed in their Corollary 6.1 that if  $A$  is a square matrix, the unique optimal strategy for both players without switching costs is completely mixed and all off-diagonal entries in  $S$  are



$[0, \underline{c}]$  region, both static and stationary. This optimal strategy in the game with switching costs is repeatedly playing an optimal mixed action in game  $A$  that incurs the lowest switching costs according to  $S$ .

**Proposition 5.** *Let  $\underline{c}$  be the maximal  $c$  such that  $v(c) = \tilde{v}(c)$  for every  $c \in [0, \underline{c}]$ . If  $\underline{c} > 0$ , then for every  $c \in [0, \underline{c}]$  the value is attained by the static strategy  $y^* = \arg \min_{y \in \mathcal{A}} y^T S y$ . This  $\underline{c}$  can be found in polynomial time  $O(n^4)$ .*

**Proof** First, note that  $y^* = \arg \min_{y \in \mathcal{A}} y^T S y$ , the optimal strategy for Player 2 in game  $A$  with the lowest average switching costs is well defined:  $\mathcal{A}$  is a compact set,  $y^T S y$  is continuous, so it has a minimum and  $y^*$  can be taken as any of them. We assume that  $y^*$  is not pure, otherwise the entire discussion is trivial.

We still need to find  $\underline{c}$ . Let  $I_1$  be the largest support of optimal actions of Player 1 in  $A$  and  $x_0 \in I_1$  (arbitrarily chosen). Use ACOE (see Appendix A) with undiscounted payoff  $\gamma(c) = v + c y^{*T} S y^*$  and actions  $x_0, y^*$  in each state. This results in an  $n \times n$  linear equation, where the unknowns are the continuation payoffs, their coefficients are the elements of  $y^*$ , and the free elements are linear functions of  $c$ . Solve it to find the continuations payoffs (they are linear functions of  $c$ ).

Then, plug in the continuation payoffs to each state and consider each state as a one-shot game. Since both the switching costs and the continuation payoffs are the same in each column, as in the proof of Proposition 1, they cancel out. Thus when Player 2 plays  $y^*$ , he makes Player 1 indifferent among all the actions in  $I_1$  and prefer them over all the other actions.

Next, check whether in each such game,  $y^*$  can actually be played by Player 2 in equilibrium, i.e. check whether Player 1 has a mixed action over  $I_1$  that keeps Player 2 indifferent among all the columns for which  $y_i^* > 0$  and still preferring them over the ones where  $y_i^* = 0$  (this time, including the switching costs and the continuation payoffs!). Each such test consists of solving another (order of)  $n \times n$  linear equation (and verifying some inequalities) and results in an upper bound on  $c$  for which the solution is indeed a distribution over the rows. The minimum of all these  $n$  upper bounds is the required  $\underline{c}$ . ■

Recall that we assume that the diagonal of  $S$  is zero, while the off-diagonal elements are non-negative and for normalization, the minimal non-zero element is 1. Now we also assume that there are no free switches, i.e., all off-diagonal elements of  $S$  are strictly positive. In this case, increasing  $c$  increases the cost of playing mixed strategies and, at some point, it will equal (as in our §3.4), then  $\underline{c} > 0$ .

be optimal to play purely. The cutoff after which the optimal stationary strategy is pure is denoted by  $\bar{c}$ . The following proposition proves its existence and generalizes Theorem 6.3 in Schoenmakers et al. [2008].

**Proposition 6.** *Let  $S$  be a matrix where all off-diagonal elements are strictly positive. Then there exists  $\bar{c}$  s.t. Player 2's optimal strategy for every  $c \geq \bar{c}$  is the pure minimax strategy in the matrix game  $A$ , but not for  $c < \bar{c}$ . Hence,  $v(c) = \tilde{v}(c) = \bar{v}$  for  $c \geq \bar{c}$ , where  $\bar{v}$  is the minimax of game  $A$  in pure actions.*

**Proof** In each state, when  $c \rightarrow \infty$ , switching a pure action is strictly dominated by replaying it (alternatively, when  $c \rightarrow \infty$ , in the matrix from Filar and Raghavan [1984], all pure actions that represent switches become dominated since  $c$  is multiplied by positive constants). Hence, for  $c$  large enough, Player 2 never changes his action and plays the same pure action for all  $t$ . Under this restriction, the optimal strategy is to play the pure minimax action of matrix game  $A$ . Clearly, this is also a static strategy. ■

The case of large  $c$  is interesting, as it represents a situation where the switching cost is higher than the possible one-stage gain. For example, the profit of a firm that adjusts its price to dominate a market can be much smaller than the adjustment cost, provided that the other firms in the market respond quickly enough. Chakrabarti [1990] studied this scenario for supergames and showed how such high switching costs affect the set of possible equilibria and the optimal strategies, which must include only finitely many action switches.

Although we have no efficient algorithm to calculate  $\bar{c}$ , an upper bound for it can be found in the following manner. A sufficient condition for the optimal stationary strategy to be pure is that one pure action of the one-shot game presented in Filar and Raghavan [1984] strictly dominates all others (see the proof of Proposition 1). Hence, every mixed entry in the matrix should suffice  $\bar{v} \leq b_{ij} + \beta_j c$ . Assuming  $\beta_j > 0$ , we get  $c \geq \frac{\bar{v} - b_{ij}}{\beta_j}$ . When this condition is met for all  $i, j$ , we get  $\tilde{v}(c) = \bar{v}$ , hence  $\bar{c} \leq \max_{i,j} \left( \frac{\bar{v} - b_{ij}}{\beta_j} \right)$ .

Ideally, we would like to have the exact value of  $\bar{c}$  or the smallest possible upper bound (this will more accurately provide the difference between  $v(c)$  and  $\tilde{v}(c)$ ). For example, checking all the values in the matrix and finding the maximal for the above expression would be a better algorithm than what follows. However, given the large matrix (order of  $n^m \times n^n$ ), computation would be hard and time-consuming. Instead, we can maximize the expression element-by-element – requiring that  $\beta_j$  and  $b_{ij}$  be minimal but not necessarily for the same  $i, j$ .

By normalization, the lowest off-diagonal switching cost is 1. Since  $\beta_j$  is the average switching cost per stage, we take it to be the minimum, which is 1. Similarly, for  $b_{ij}$  we can

choose the lowest entry in  $A$ . An upper bound on  $\bar{c}$  can be written as

$$\hat{c} = \bar{v} - \min_{i,j} a_{ij} = \bar{v}, \quad (9)$$

since  $A$  was normalized s.t.  $\min_{i,j} a_{ij} = 0$ . In all the following calculations and formulas, if the exact value of  $\bar{c}$  is unknown, it is possible to use  $\hat{c}$  remembering that  $\hat{c} \geq \bar{c}$  and hence the bound will be less precise (larger).

Subsequently, the only region where stationary strategies might out-perform static ones is  $(\underline{c}, \bar{c})$ . As the following example shows, however, a static strategy could be optimal for some  $c$ s in this region.

**Example 5. A game where  $v(c)$  and  $\tilde{v}(c)$  intersect in  $(\underline{c}, \bar{c})$**

Suppose  $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$  and  $S = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 2 & 3 & 0 \end{pmatrix}$ . For  $c < \frac{1}{6}$ , the optimal stationary strategy is the static strategy that chooses each action w.p.  $\frac{1}{3}$ , regardless of the history. Starting from  $c = \frac{1}{6}$ , it is too costly to play  $R$  after  $M$ , so the optimal stationary strategy is to play  $[0.5(L), 0.5(M)]$  in  $s_M$  (and each action w.p.  $\frac{1}{3}$  in the two other states). In this region, the stationary strategy outperforms static strategies.

Next, at  $c = \frac{3}{13}$ , it is also too costly to play  $R$  in  $s_L$  and this action is never played. The optimal stationary strategy is the static strategy that plays  $[0.5(L), 0.5(M)]$  at every stage. This remains true until  $c$  is high enough to make switching too costly and the optimal strategy is playing a pure action which leads to  $v(c) = \tilde{v}(c) = 1$ .  $\diamond$

In the above example,  $v(c) = \tilde{v}(c)$  before  $\bar{c}$ , where the value functions flatten. As outside  $(\underline{c}, \bar{c})$  the value functions are equal, the only region of interest is  $(\underline{c}, \bar{c})$ , and the following theorem bounds the difference in payoff functions there.

**Theorem 2.** *Player 2's loss from limiting himself to static strategies,  $\tilde{v}(c) - v(c)$ , is bounded by*

$$\Delta(c) = \begin{cases} (c - \underline{c}) \frac{\bar{c} - c_0}{\bar{c} - \underline{c}} s & \text{if } c \in (\underline{c}, c_0), \\ (\bar{c} - c) \frac{c_0 - \underline{c}}{\bar{c} - \underline{c}} s & \text{if } c \in (c_0, \bar{c}), \\ 0 & \text{otherwise,} \end{cases} \quad (10)$$

where  $s = y^{*T} S y^*$ ,  $c_0 = \frac{\bar{v} - v}{s}$  and  $\underline{c}$  and  $y^*$  are defined in Proposition 5, and  $\bar{c}$  in Proposition 6.

**Proof** Define  $s = y^{*T} S y^*$  ( $y^*$  is defined in Proposition 5). From concavity,  $v(c)$  lies above the line that connects  $(\underline{c}, v + \underline{c}s)$  and  $(\bar{c}, \bar{v})$  (denoted by  $l$ ). Function  $\tilde{v}(c)$  must be below lines  $\bar{v}$  (corresponding to a pure static strategy) and  $v + cs$  (corresponding to static strategy  $y^*$ ). These lines meet at  $c_0 = \frac{\bar{v} - v}{s}$ . Hence, for  $c < c_0$  the difference between the

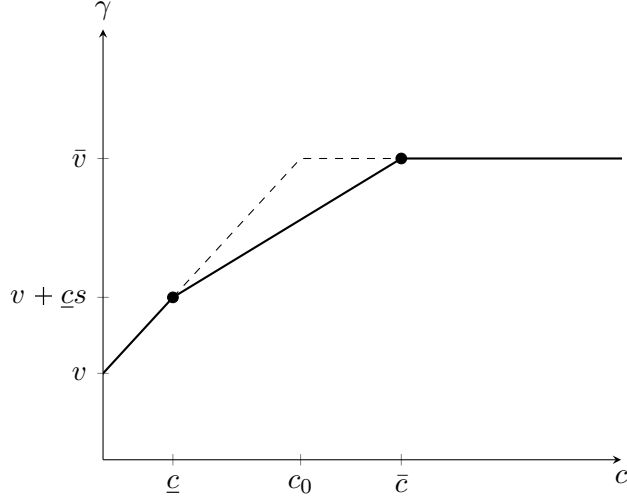


Figure 3: The intuition for the bound in Eq. (10). The maximal difference between  $v(c)$  and  $\tilde{v}(c)$  is smaller than the maximal difference between the dashed and the solid lines, which is attained at  $c_0$ .

value in static and in stationary strategies is at most the difference between line  $v + cs$  and  $l$ , and for  $c > c_0$  it is the difference between  $\underline{v}$  and  $l$  (see Fig. 3). A direct computation reveals that these differences are as in Eq. (10).  $\blacksquare$

In Theorem 2, we construct a bound by comparing a lower bound on  $v(c)$  with an upper bound on  $\tilde{v}(c)$ . This upper bound is constructed by using two feasible and simple static strategies: the minimax pure action of game  $A$  and an optimal mixed action of game  $A$ , the one with the lowest switching cost. Any convex combination of these two static strategies is also a static strategy and can also be considered to improve the bound.

In particular, assume  $y^*$  is, for Player 2 in game  $A$ , the optimal mixed action with the lowest switching costs ( $y^* = \arg \min_{y \in \mathcal{A}} y^T S y$ ) and  $\bar{y}$  is the pure minimax action of Player 2 (so  $\bar{y}^T S \bar{y} = 0$ ). For every  $\alpha \in [0, 1]$ ,  $y_\alpha = \alpha y^* + (1 - \alpha) \bar{y}$  is also a feasible static strategy for Player 2 and its payoff bounds  $\tilde{v}(c)$  from above.

**Theorem 3.** *Let  $\bar{y}$  be a pure minimax action in game  $A$ , and denote by  $\hat{s} = y^{*T} S \bar{y} + \bar{y}^T S y^*$ .<sup>7</sup> If  $s > \hat{s}$ , then for every  $c \in [c_1, c_2] \cap (\underline{c}, \bar{c})$ , the bound from Theorem 2 can be improved by using*

$$\Delta(c) = (\bar{v} - v - \underline{c}s) \frac{\bar{c} - c}{\bar{c} - \underline{c}} - \frac{(\bar{v} - v - c\hat{s})^2}{4c(s - \hat{s})}, \quad (11)$$

<sup>7</sup>Different bounds can be attained by using different pure minimax actions (if there are several). In that case, all of them can be constructed using the method presented, and the lowest chosen.

where  $c_1 = \frac{s}{2s-\hat{s}}c_0$  and  $c_2 = \frac{s}{\hat{s}}c_0$ .

**Proof** For every  $x$ , the payoff of the strategy pair  $(x, y_\alpha)$  is

$$\begin{aligned} g(c)(x, y_\alpha) &= xAy_\alpha + cy_\alpha^T Sy_\alpha \\ &= \alpha xAy^* + (1-\alpha)xA\bar{y} + c\alpha^2 y^{*T} Sy^* + c\alpha(1-\alpha) \left[ y^{*T} S\bar{y} + \bar{y}^T Sy^* \right]. \end{aligned}$$

Since  $y^*$  is optimal in  $A$ , it guarantees the value and  $xAy^* \leq v$ . Similarly,  $\bar{y}$  guarantees the minimax value so  $xA\bar{y} \leq \bar{v}$ . As previously, we let  $s = y^{*T} Sy^*$  and to simplify the expressions also  $\hat{s} = y^{*T} S\bar{y} + \bar{y}^T Sy^*$  (note that if  $S$  is symmetric,  $\hat{s} = 2y^{*T} S\bar{y}$ ). Therefore  $g(c)(x, y_\alpha)$  can be bounded by the following quadratic function of  $\alpha$ :

$$g(c)(x, y_\alpha) \leq \alpha v + (1-\alpha)\bar{v} + \alpha^2 cs + \alpha(1-\alpha)c\hat{s} = \alpha^2 c(s - \hat{s}) + \alpha(c\hat{s} + v - \bar{v}) + \bar{v} := f(\alpha).$$

Since  $\tilde{v}(c) \leq g(c)(x, y_\alpha)$ , it follows that  $\tilde{v}(c) \leq \min_{\alpha \in [0,1]} f(\alpha)$ . Hence, by finding this minimum, we can obtain a tighter bound on  $\tilde{v}(c) - v(c)$  relative to Theorem 2.

If  $f$  is concave, the minimum is obtained on one of the extreme points (depending on  $c$ ), which takes us back to the case of Theorem 2. In fact, we depart from the case presented in Theorem 2 only if  $f$  is convex and the minimum is inside  $(0, 1)$ , i.e. only if  $s > \hat{s}$  and  $0 \leq \frac{\bar{v}-v-c\hat{s}}{2c(s-\hat{s})} \leq 1$ .

For the rest of the proof we assume therefore that  $s > \hat{s}$ . The condition  $0 \leq \frac{\bar{v}-v-c\hat{s}}{2c(s-\hat{s})} \leq 1$  can be rearranged as a condition on  $c$ . The “ $0 \leq$ ” part becomes  $c \leq c_2 = \frac{\bar{v}-v}{\hat{s}}$  and the “ $\leq 1$ ” becomes  $c \geq c_1 = \frac{\bar{v}-v}{2s-\hat{s}}$ .

Hence, for  $c \in [c_1, c_2] \cap (\underline{c}, \bar{c})$ , the bound from Theorem 2 can be improved by using

$$f\left(\frac{\bar{v}-v-c\hat{s}}{2c(s-\hat{s})}\right) = \bar{v} - \frac{(c\hat{s}+v-\bar{v})^2}{4c(s-\hat{s})}$$

instead of  $\min\{\bar{v}, v + cs\}$ . Thus, the bound for this domain is

$$\Delta(c) = (\bar{v} - v - \underline{c}s) \frac{\bar{c}-c}{\bar{c}-\underline{c}} - \frac{(\bar{v}-v-c\hat{s})^2}{4c(s-\hat{s})}.$$

■

An illustration of this bound can be found in Figure 4, applied to Example 2. Interestingly, in this case the optimal static strategy is actually a mixture of  $y^*$  and  $\bar{y}$ , so the bound is tight and achieves the minimax value in static strategies. We demonstrate this bound through Example 2 due to the strong variance in the switching-costs matrix. Without this variance, this bound might not be applicable, as shown in Observation 1 for equal switching costs ( $s_{ij} = 1$  for all  $i \neq j$ ).

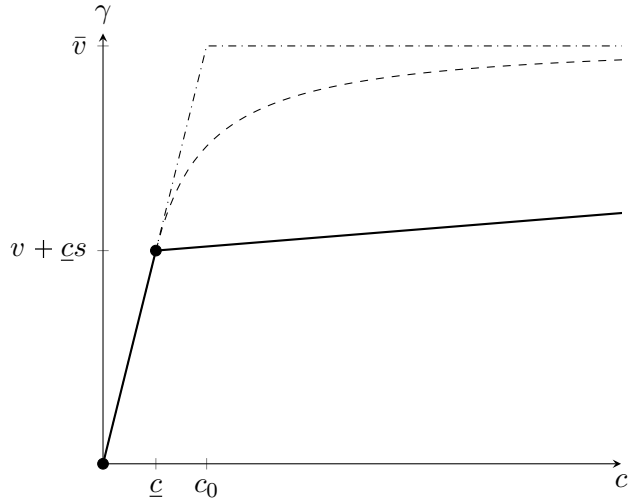


Figure 4: The bound of Theorem 3, applied to Example 2. The solid line is  $v(c)$ , the dotted line is the bound (in this example, it is also  $\tilde{v}(c)$ ) and the dot-dashed line is the bound from Theorem 2. For clarity, the graph was zoomed on the region near  $c_0$  where the difference is the largest. All three meet at  $\bar{c}$ .

### 3.4 Action-Independent Switching Costs

In this section, we consider the special case of switching costs which are independent of the actions being switched (uniform switching costs), i.e.  $s_{ij} = 1$  for all  $i \neq j$ . This case naturally arises when the switching stem from the act of “switching” itself and do not depend on the actions being switched to or from. For example, consider a person trying to multi-task. Each time he switches from one task to another, it takes him a short time to regain focus on the task he is about to perform. This time is generally fixed and does not depend on the nature of the task (whereas the time spent actually performing different tasks can differ significantly). This case was explored in previous papers, such as Lipman and Wang [2000, 2009] and Schoenmakers et al. [2008]. In fact, the model Schoenmakers et al. [2008] studied is exactly the one presented here, up to an additive constant (as they pay a positive bonus for not switching actions and nothing when actions are switched). We refer the interested reader to Schoenmakers et al. [2008], who provide additional results using this model. For example, they analyze in depth the case where  $A$  is an  $m \times 2$  matrix and provide a formula to compute the optimal strategy when  $A$  is a square matrix.

A significant benefit from the uniform switching-costs assumption is that finding the optimal static strategy becomes simpler. Indeed, note first that under this assumption,  $g(c)(x, y)$  can be significantly simplified. Let  $\mathbb{1}$  be the ones matrix (for all  $i, j : (\mathbb{1})_{ij} = 1$ ).

Then  $S = \mathbb{1} - I$  and

$$y^t S y = y^t (\mathbb{1} - I) y = y^t \mathbb{1} y - \|y\|^2 = 1 - \|y\|^2. \quad (12)$$

So  $g(c)(x, y) = x^T A y + c(1 - \|y\|^2)$  and  $\tilde{v}(c) = c + \min_{y \in \Delta(\{1, \dots, n\})} \left\{ \max_{x \in \Delta(\{1, \dots, m\})} \{x^T A y\} - c\|y\|^2 \right\}$ .

Using this expression we can strengthen Proposition 2, as it turns out that now  $\tilde{v}(c)$  is piece-wise linear. In a similar fashion to Corollary 1, this implies the existence of a finite set of strategies to consider when looking for the optimal static strategy.

**Proposition 7.** *Whenever switching costs are uniform ( $s_{ij} = 1$  for  $i \neq j$  and 0 otherwise), the minimax value  $\tilde{v}(c)$  in static strategies is a continuous, increasing, and concave piece-wise linear function.*

**Proof** It follows from Proposition 2 that  $\tilde{v}(c)$  is continuous, increasing, and concave. Now we prove that it is also piece-wise linear. For each  $i \in \{1, \dots, m\}$ , let  $I_i$  be the subset of  $\Delta(\{1, \dots, n\})$  such that  $i$  is a best response for Player 1 in the one-shot game  $A$  to all  $y \in I_i$ . It is well known that  $I_i$  is a compact, convex polyhedron in  $\mathbb{R}^n$ .

Now, suppose Player 2 is restricted to playing only in set  $I_i$ . The best response for Player 1 is  $i$  and the payoff (based on the above calculation) is  $c + \min_{y \in I_i} \{A_i y - c\|y\|^2\}$ , where  $A_i$  is the  $i$ th row of  $A$ . The expression inside the minimization is a concave function in  $y$  and the minimization domain is a polyhedron, so the minimum is attained at one of its vertices.

There are  $m$  polyhedrons and each one has finitely many vertices, so the set of candidate optimal strategies is finite. Using a similar argument as in Proposition 1, it follows that  $\tilde{v}(c)$  is piece-wise linear. ■

The reasoning behind this proof can be used to slightly strengthen Proposition 2 (even without the additional assumption on  $S$ ). If we restrict Player 2 to  $I_i$  and write  $g(c)(x, y)$  in the general case explicitly (including the constraint  $\sum y_i = 1$ ), we obtain a function which is concave in each of the variables  $y_i$ . Thus, the arg min cannot be an interior point of  $I_i$  and must lie on the boundary. However, since the function is not concave, the arg min is not guaranteed to be on a vertex, and could be on the boundary, where it moves as  $c$  changes, as in Example 2. Still, the fact that the arg min is on the boundary of one of the  $I_i$ s means that when playing an optimal static strategy, Player 2 keeps Player 1 indifferent between at least two of her actions, and this can be used to reduce the dimension of the search.

As mentioned, a direct result of Proposition 7 is that the optimal strategy comes from a finite set and is robust to the exact value of  $c$ . Unlike Corollary 1 however, this time it is also true for the best responses of Player 1 (since she best-responds purely). This provides a

theoretical explanation for the numerical results of Rass and Rainer [2014] and Liuzzi et al. [2020], which indicate that slightly changing  $c$  (or  $\alpha$ , in their model) has a small effect, if any, on optimal strategies.

**Corollary 2.** *When switching costs are uniform ( $s_{ij} = 1$  for  $i \neq j$  and 0 otherwise), there exists  $K \in \mathbb{N}$  and  $0 = c_0 < c_1 < c_2 < \dots < c_K < c_{K+1} = \infty$  such that for each segment  $[c_i, c_{i+1}]$  there exists a Player 2 strategy to achieve the minimax value  $\tilde{v}$  in static strategies that is suited to all  $c$ s in the segment.*

The proof is identical to the proof of Corollary 1. Note that in general the number of segments differs between  $v(c)$  and  $\tilde{v}(c)$ .

According to Proposition 7 and Corollary 2, there is a finite set of static strategies from which an optimal static strategy can be selected for any  $c$ . This set can be large (exponential in  $n$ ), and the above results can be used to search more efficiently for the optimal strategy within the set. For example, by calculating  $y^T S y$  for each  $y$  in the set, the strategies in the set can be ordered from largest to smallest. Since  $y^T S y$  are the possible slopes of  $c$  in  $\tilde{v}(c)$  and since  $\tilde{v}(c)$  is concave, if a particular  $y$  is found to be optimal for a specific  $c$ , all  $y$ s with a higher slope can be excluded from the search domain for all larger  $c$ s. A similar approach can be used to find the optimal stationary strategy as a consequence of Proposition 1, keeping in mind that calculating the slope of  $c$  for a given stationary strategy is slightly more complicated.

We can improve the bound from Proposition 4 by plugging in the explicit expression of  $S$ .

**Theorem 4.** *When switching costs are uniform ( $s_{ij} = 1$  for  $i \neq j$  and 0 otherwise), Player 2's loss from limiting himself to static strategies,  $\tilde{v}(c) - v(c)$ , is bounded by*

$$\Delta(c) = \frac{1}{2} \left( 1 + c \left( 1 - \frac{1}{n} \right) - v(c) \right). \quad (13)$$

**Proof** The matrix  $S$  is symmetric, so  $\Xi = 0$ . Moreover,  $y^T S y = 1 - \|y\|^2$  so  $M = \max_{y \in \Delta(\{1, \dots, m\})} (1 - \|y\|^2)$ . Inside the simplex, the maximum is achieved at  $y = (\frac{1}{n}, \dots, \frac{1}{n})$  and is equal to  $1 - \frac{1}{n}$ . Plugging these into Proposition 4, the bound becomes  $\Delta(c) = \frac{1}{2} \left( 1 + c \left( 1 - \frac{1}{n} \right) - v(c) \right)$ .  $\blacksquare$

Interestingly, this bound cannot be improved by harnessing Theorem 3 to this case. The reason is that the condition  $s > \hat{s}$  can never hold in this kind of switching-costs matrix, and requires inequality of some kind in the costs (see, for example, Example 2 and Figure 4).



**Observation 1.** *When switching costs are uniform ( $s_{ij} = 1$  for  $i \neq j$  and 0 otherwise), the condition  $s > \hat{s}$  from Theorem 3 never holds.*

**Proof** In this case, we can compute  $s$  and  $\hat{s}$  directly. Following Eq. (12),  $s = 1 - \|y^*\|^2$  and by definition,  $\hat{s} = 2(1 - \bar{y} \cdot y^*)$ .<sup>8</sup> The condition  $s > \hat{s}$  from Theorem 3 becomes  $1 - \|y^*\|^2 > 2(1 - \bar{y} \cdot y^*)$  or equivalently  $(2\bar{y} - y^*) \cdot y^* > 1$ .

However, the maximum of the function  $(2\bar{y} - x) \cdot x$  (over all  $\mathbb{R}^n$ ) is at  $x = \bar{y}$ , and is equal to  $\|\bar{y}\|^2 = 1$ . Thus,  $(2\bar{y} - y^*) \cdot y^* \leq 1$  and the desired inequality never holds. ■

## 4 Static Strategies in Stochastic Games

So far, we have compared the values guaranteed by stationary and static strategies in a particular class of games resulting from adding switching costs to normal-form repeated games. In this section, we consider general stochastic games and study which game structure is needed for our results (or similar) to hold. Naturally, we cannot consider the most general stochastic games and must at least assume that the same set of actions is available to the players in each state; otherwise, static strategies are not well defined.

Let  $\Gamma$  be a stochastic game. We denote by  $S$  the finite set of states when  $I$  and  $J$  are finite action sets of Player 1 and Player 2 (resp.). The transition function is denoted by  $q : S \times I \times J \rightarrow \Delta(S)$  and the normalized payoff function for Player 1 (zero-sum) by  $r : S \times I \times J \rightarrow [0, 1]$ . We assume that the normalization is tight:  $\min_{s,i,j} r(s, i, j) = 0$  and  $\max_{s,i,j} r(s, i, j) = 1$ .

As before, a stationary strategy is a strategy that depends at each  $t$  on the state, but not on the rest of the history or on  $t$  itself. We denote by  $\Sigma$  the set of stationary strategies for Player 1, and by  $\mathcal{T}$  the set of stationary strategies for Player 2. A static strategy is a strategy that dictates the same mixed action in each stage, independent of  $t$  or the history. We denote by  $\Sigma_f$  the set of static strategies for Player 1, and by  $\mathcal{T}_f$  the set of static strategies for Player 2.

Without additional requirements, limiting Player 2 to static strategies can be arbitrarily bad for him, as the following example shows:

**Example 6.** *An arbitrarily bad payoff in static strategies relative to stationary strategies.*

Consider the following three-state stochastic game. For simplicity, Player 1 has only one action and has no role in the game. Arrows represent the deterministic transitions, so  $s_R$  is an absorbing state:

---

<sup>8</sup>Since  $\bar{y}$  is pure, this dot product is simply the probability of choosing pure action  $\bar{y}$  when following mixed  $y^*$ .

$$s_L : (1 \text{ } \cup, 0 \rightarrow) \quad s_M : (0 \leftarrow, 0 \rightarrow) \quad s_R : (1 \text{ } \cup, 1 \text{ } \cup)$$

It is clear that the optimal strategy for Player 2 is the stationary strategy that plays  $R$  at  $S_L$  and  $L$  at  $S_M$ , since it guarantees him the minimal payoff in this game. By contrast, let us consider a static strategy that plays  $L$  with probability  $x$  (and  $R$  otherwise). If  $x < 1$ , then every trajectory eventually reaches state  $s_R$  where the payoff is 1, whereas if  $x = 1$ , the game remains in  $s_L$  and the payoff is 1. It follows that the difference between the two values is equal to the maximum difference in payoff.  $\diamond$

We therefore consider only stochastic games where only Player 2 controls the states and the states correspond to the previously played action. Hence,  $S = J$  and every transition satisfies

$$\forall (s, i, j) \in S \times I \times J, q(s, i, j) = \delta_j.$$

Informally, the state at stage  $t + 1$  is the action played by Player 2 at stage  $t$ . This class of games is denoted by  $\mathcal{G}$ . Moreover, we denote by  $\mathcal{G}_S$  the set of games in  $\mathcal{G}$  such that there exist two matrices  $A$  of size  $|I| \times |J|$ , and  $S$  of size  $|J| \times |J|$  such that

$$\forall (j', i, j) \in S \times I \times J, r(j', i, j) = A_{i,j} + S_{j',j}.$$

Thus, Sections 2 and 3 dealt only with games from  $\mathcal{G}_S$  (with the additional constraints  $s_{ij} \geq 0$  and  $s_{ii} = 0$ ). In Example 7 we show that indeed  $\mathcal{G} \neq \mathcal{G}_S$ . The following characterization of  $\mathcal{G}_S$  can be used to see whether a game in  $\mathcal{G}$  is indeed in  $\mathcal{G}_S$  and how to construct the appropriate matrices  $A$  and  $S$ . The intuition behind this result is that there must be no interaction between the actions of the maximizing player and the states.

**Proposition 8.** *Let  $\Gamma$  be in  $\mathcal{G}$ , then  $\Gamma$  is in  $\mathcal{G}_S$  if and only if*

$$\forall i, i' \in I, \forall j, j', j'' \in J, r(j', i, j) - r(j'', i, j) = r(j', i', j) - r(j'', i', j). \quad (14)$$

**Proof** Let us first check that if  $\Gamma$  is in  $\mathcal{G}_S$ , then it satisfies the equation. Let  $i \in I$  and  $j, j', j'' \in J$ , then

$$r(j', i, j) - r(j'', i, j) = (A_{ij} + S_{j'j}) - (A_{ij} + S_{j''j}) = S_{j'j} - S_{j''j}.$$

Hence, it is independent of  $i$  and therefore for two actions  $i$  and  $i'$  of Player 1, the two differences are equal.

Let us now prove the converse. Let  $r$  be a payoff function of a stochastic game in  $\mathcal{G}$  that satisfies

$$\forall i, i' \in I, \forall j, j', j'' \in J, r(j', i, j) - r(j'', i, j) = r(j', i', j) - r(j'', i', j).$$

Fix  $i_0 \in I, j_0 \in J$ . Define for every  $j \in J$  and for every  $j' \in J$ ,

$$S_{jj'} = r(j', i_0, j) - r(j_0, i_0, j).$$

Notice that  $S_{jj_0} = 0$  for every  $j \in J$ . Moreover for every  $(j', i, j) \in S \times I \times J$ , let

$$B_{ijj'} = r(j', i, j) - S_{jj'} = r(j', i, j) - r(j', i_0, j) + r(j_0, i_0, j).$$

To conclude, we need to check that  $B_{ijj'}$  does not depend on  $j'$  and is therefore equal to some  $A_{ij}$ . To do so, letting  $j'_0$  and  $j'_1$ , we have

$$\begin{aligned} & B_{ijj'_0} - B_{ijj'_1} \\ &= [r(j'_0, i, j) - r(j'_0, i_0, j) + r(j_0, i_0, j)] - [r(j'_1, i, j) - r(j'_1, i_0, j) + r(j_0, i_0, j)] \\ &= [r(j'_0, i, j) - r(j'_1, i, j)] - [r(j'_0, i_0, j) - r(j'_1, i_0, j)] + [r(j_0, i_0, j) - r(j_0, i_0, j)] \\ &= 0. \end{aligned}$$

Hence the result. ■

Given a game  $\Gamma$  in  $\mathcal{G}$ , a pair of stationary strategies  $(\sigma, \tau) \in \Sigma \times \mathcal{T}$ , and an initial state  $s \in S$ , we define

$$\gamma(s, \sigma, \tau) = \mathbb{E}_{s, \sigma, \tau} \left( \liminf_{T \rightarrow +\infty} \frac{1}{T} \sum_{t=1}^T r(s_t, i_t, j_t) \right),$$

and we can define  $\forall s \in S$ ,

$$v_\Gamma(s) = \inf_{\tau \in \mathcal{T}} \sup_{\sigma \in \Sigma} \gamma(s, \sigma, \tau).$$

which is the minimax value of Player 2 when using stationary strategies. Note that the value is independent of the original state due to the structure of the transitions, so it will be dropped hereafter.

We also consider

$$\tilde{v}_\Gamma = \inf_{\tau_f \in \mathcal{T}_f} \sup_{\sigma \in \Sigma} \gamma(\sigma_f, \tau),$$

which is the minimax value of Player 2 when Player 2 is restricted to static strategies. Note that in this case, even when Player 2 uses a static strategy, the best response of Player 1 might be stationary and not static.

**Example 7.** *A stochastic game where the best response of Player 1 to a static strategy of Player 2 is not static*

Consider the following two-state stochastic game

$$s_L : \begin{pmatrix} 1 \oslash & 0 \rightarrow \\ 1 \oslash & 1 \rightarrow \end{pmatrix} \quad s_R : \begin{pmatrix} 1 \leftarrow & 1 \oslash \\ 0 \leftarrow & 1 \oslash \end{pmatrix}$$

and suppose that Player 2 plays the static non-pure action  $[x(L), 1 - x(L)]$  for  $x \in (0, 1)$ . The expected payoff of  $T$  in state  $s_L$  is  $x < 1$  and the expected payoff of  $B$  in state  $s_R$  is  $1 - x < 1$ . Hence, each of these actions is dominated by the other action in the state and the best response for Player 1 is to play  $B$  in state  $s_L$  and  $T$  in state  $s_R$ , with an overall payoff of 1. Any static strategy by Player 1 would result in a payoff strictly smaller than 1 (which is worse for him, as the maximizer).

In the model presented in Section 2, the best response to a static strategy is static. Since that is not true here, this example shows that  $\mathcal{G} \neq \mathcal{G}_S$  and that the payoff function of this game cannot be written as two additive payoffs, one corresponding to a normal-form game and another to the switching costs (see also Proposition 8).  $\diamond$

Our main result generalizes Proposition 4 for the games in  $\mathcal{G}$ , rather than just in  $\mathcal{G}_S$ .

**Theorem 5.** *Fix  $\Gamma \in \mathcal{G}$ . The minimax value  $\tilde{v}_\Gamma$  Player 2 can guarantee in static strategies is at least a quarter of the value  $v_\Gamma$  he can guarantee in stationary ones:*

$$\tilde{v}_\Gamma - v_\Gamma \leq \Delta_\Gamma = \frac{3}{4}(1 - v_\Gamma). \quad (15)$$

*This equation is equivalent to the following*

$$\frac{1 - \tilde{v}_\Gamma}{1 - v_\Gamma} \geq \frac{1}{4}. \quad (16)$$

**Proof** Let  $v_\Gamma$  be the value of the game (in stationary strategies). There exists a state  $j^* \in S$  and a mixed action  $y^* \in \Delta(J)$  such that for all  $i \in I$  we have  $r(j^*, i, y^*) \leq v_\Gamma$  (otherwise, the payoff of Player 1 in each state can exceed  $v_\Gamma$ , as can the total value). Consider the static strategy  $\tau_f$  of Player 2, who with probability 0.5 chooses action  $j^*$  and otherwise chooses an action according to mixed action  $y^*$ . The payoff of this strategy against the stationary strategy  $\sigma(s)$  of Player 1 is

$$\begin{aligned} & \frac{1}{4}r(j^*, \sigma(j^*), y^*) + \frac{1}{4}r(j^*, \sigma(j^*), j^*) + \frac{1}{4} \sum_{s \in S} y^*(s)r(s, \sigma(s), j^*) + \frac{1}{4} \sum_{s \in S} y^*(s)r(s, \sigma(s), y^*) \\ & \leq \frac{1}{4}v_\Gamma + \frac{3}{4}, \end{aligned}$$

where we bound the payoff by the maximal payoff 1. It follows that

$$\tilde{v}_\Gamma \leq \frac{1}{4}v_\Gamma + \frac{3}{4}.$$

Therefore, we obtain by subtracting  $v_\Gamma$  on both sides that

$$\tilde{v}_\Gamma - v_\Gamma \leq \frac{3}{4}(1 - v_\Gamma),$$

or by subtracting 1 on both sides and then dividing by  $v_\Gamma - 1$  that

$$\frac{1-\tilde{v}_\Gamma}{1-v_\Gamma} \geq \frac{1}{4}.$$

The gain obtained by a static strategy (compared to the maximal payoff 1) is at least  $\frac{1}{4}$  of the gain obtained by a stationary strategy.

To finalize the proof, we use an example to show that this bound is indeed tight. Suppose  $J = \{0, \dots, 4\}$ ,  $I = \{1\}$  and the payoff function is 1, except for  $r(i, 1, i + 1) = 0$  and  $r(4, 0, 1) = 0$ . The optimal stationary strategy is to choose in state  $i$  action  $i + 1$  (modulo 5) with  $v_\Gamma = 0$ . Let  $\sigma_f$  be a static strategy that chooses action  $i$  with probability  $x_i$ . Then the payoff is  $1 - x_0x_1 - x_1x_2 - \dots - x_4x_0$  and it is straightforward to verify that the minimum of this function in domain  $x_i \geq 0$  and  $\sum_{i=0}^4 x_i = 1$  is  $\frac{3}{4}$ . ■

The above result gives a tight bound on the static value compared to the stationary value, as well as offering a strategy that attains it in some cases. For specific games, the bound can be tighter and the optimal static strategy can be different. A particular set of games where the bound is different are games where  $|J| \leq 4$ . In this case, the difference between the values is at most  $\Delta_\Gamma = \left(1 - \frac{1}{|J|}\right) (1 - v_\Gamma)$  provided that Player 2 can guarantee to pay at most  $v_\Gamma$  in each state as a one-shot game. This is why the example in the proof of Theorem 5 has 5 actions.

## 5 Conclusions

In this paper, we analyze zero-sum games with switching costs and compare the value under optimal strategies (that are stationary) to an approximate solution with time-independent strategies (“static strategies”). Our main findings are summarized in a series of theorems which provide bounds on the difference between the two types of strategies for different scenarios and under various assumptions. Our bounds can be used to assess whether the additional computational complexity is worth the extra payoff gain or, symmetrically, assess the price of playing sub-optimal but simple strategies when stationary ones are forbidden.

In addition, we study and fully characterize the value in stationary strategies as a function of the ratio between switching costs and stage payoff ( $c$ ). Our key finding is that this function,  $v(c)$ , is piece-wise linear. This implies that the solutions are robust in the sense that an optimal stationary strategy for a particular  $c$  is also optimal for all the  $c$ s in its segment. Hence, knowing the exact value of  $c$  is not essential to play optimally and the strategy should not be changed in response to small changes in  $c$ .

We generalize our work on repeated games to compare static and stationary strategies for general stochastic games. This generalization allows us to consider a wider range of games with more elaborate switching-cost structures. Still, we are able to propose a bound on the price of being history-dependent. This generalization is unique to our paper, and considers the widest possible set of repeated zero-sum games to which the comparison between history-dependent and stationary strategies is applicable.

Our analysis can serve as a basis for future work regarding non-zero-sum games since, when only one player’s payoffs are considered,  $v(c)$  is his minimax value. We therefore portray the behavior of the individually rational payoffs of the players, which serves as a lower bound for the possible equilibrium payoff via a version of a Folk Theorem suitable for this model. A particularly interesting model is that where switching costs paid to a “bank” are added to a zero-sum stage game. The addition of switching costs transforms the game into a non-zero-sum game and facilitates cooperation in an otherwise zero-sum game.

Other possible extensions are games where the switching costs depend on the outcome of the previous round and not only on the action played. For example, Wang et al. [2014] report that when playing rock-paper-scissors, players tend to repeat an action that led to a victory. This can be modeled theoretically as a switching cost added to all the other actions. Although the model is significantly different, some of the techniques and observations of this paper might be useful in its analysis. This and the above non-zero-sum models are left for future research.

## References

- G.A. Akerlof and J.L. Yellen. Can small deviations from rationality make significant differences to economic equilibria? *The American Economic Review*, 75(4):708–720, 1985a.
- G.A. Akerlof and J.L. Yellen. A near-rational model of the business cycle, with wage and price inertia. *The Quarterly Journal of Economics*, 100:823–838, 1985b.
- A. Arapostathis, V.S. Borkar, E. Fernández-Gaucherand, M.K. Ghosh, and S.I. Marcus. Discrete-time controlled markov processes with average cost criterion: a survey. *SIAM Journal on Control and Optimization*, 31(2):282–344, 1993.
- A. Beggs and P. Klemperer. Multi-period competition with switching costs. *Econometrica: Journal of the Econometric Society*, pages 651–666, 1992.

- I. Bomze, W. Schachinger, and J. Weibull. Does moral play equilibrate? *Economic Theory*, pages 1–11, 2020.
- G. Caruana, L. Einav, and D. Quint. Multilateral bargaining with concession costs. *Journal of Economic Theory*, 132(1):147–166, 2007.
- S.K. Chakrabarti. Characterizations of the equilibrium payoffs of inertia supergames. *Journal of Economic Theory*, 51(1):171–183, 1990.
- J.A. Filar. Algorithms for solving some undiscounted stochastic games[ph. d. thesis]. 1980.
- J.A. Filar and T.E.S. Raghavan. A matrix game solution of the single-controller stochastic game. *Mathematics of Operations Research*, 9(3):356–362, 1984.
- B. L. Lipman and R. Wang. Switching costs in infinitely repeated games. *Games and Economic Behavior*, 66(1):292–314, 2009.
- B.L Lipman and R. Wang. Switching costs in frequently repeated games. *Journal of Economic Theory*, 93(2):149–190, 2000.
- G. Liuzzi, M. Locatelli, V. Piccialli, and S. Rass. Computing mixed strategies equilibria in presence of switching costs by the solution of nonconvex qp problems. 2020.
- TES Raghavan. Finite-step algorithms for single-controller and perfect information stochastic games. In *Stochastic games and applications*, pages 227–251. Springer, 2003.
- TES Raghavan and Zamir Syed. Computing stationary nash equilibria of undiscounted single-controller stochastic games. *Mathematics of Operations Research*, 27(2):384–400, 2002.
- S. Rass and B. Rainer. Numerical computation of multi-goal security strategies. In *International Conference on Decision and Game Theory for Security*, pages 118–133. Springer, 2014.
- S. Rass, S. Konig, and S. Schauer. Defending against advanced persistent threats using game-theory. *PloS one*, 12(1):e0168675, 2017.
- G. Schoenmakers, J. Flesch, F. Thuijsman, and O.J. Vrieze. Repeated games with bonuses. *Journal of optimization theory and applications*, 136(3):459–473, 2008.
- L.S. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences*, 39(10):1095–1100, 1953.

- L.S. Shapley and R.N. Snow. Basic solutions of discrete games, *annals of mathematics studies*. 1950.
- O. Shy. Colonel blotto games with switching costs. *Available at SSRN 3306880*, 2019.
- M.A. Stern. *On Stochastic Games with Limiting Average Payoff*. PhD thesis, University of Illinois at Chicago, 1975.
- J. Wachter, S. Rass, and S. Konig. Security from the adversarys inertia—controlling convergence speed when playing mixed strategy equilibria. *Games*, 9(3):59, 2018.
- Z. Wang, B. Xu, and H.J. Zhou. Social cycling and conditional responses in the rock-paper-scissors game. *Scientific reports*, 4:5830, 2014.
- Q. Zhu and T. Başar. Game-theoretic approach to feedback-driven multi-stage moving target defense. In Sajal K. Das, Cristina Nita-Rotaru, and Murat Kantarcioglu, editors, *Decision and Game Theory for Security*, pages 246–263, Cham, 2013. Springer International Publishing. ISBN 978-3-319-02786-9.



## A Average Cost Optimality Equation (ACOE)

In this Appendix we show how the value of the game and the optimal stationary strategies can be computed using the Average Cost Optimality Equation. We achieve this by solving Example 1 in detail. More about these methods can be found in Arapostathis et al. [1993], the references within, and similar papers.

In general terms, the algorithm consists of the following steps:

1. “Guess” a stationary strategy for Player 2. Based on the ideas of Proposition 1, this stationary strategy should in each state play a mixed action that keeps Player 1 indifferent among several of his actions and makes him prefer them over the rest in the one-shot game  $A$ . This well defines the set of actions to consider and ensures that the algorithm eventually terminates.
2. Player 1 does not control transitions, so once the strategy of Player 2 is determined, he can best respond by myopically best responding in each state. This can be achieved even by pure actions. Fix such a best response.
3. Let  $q_{ij}$  be the probability of Player 2 playing action  $j$  in state  $s_i$ , let  $r_i$  be the value of one-shot game  $A$  when the players use their mixed actions chosen for state  $s_i$ , let  $v_i$  be the continuation payoff of the repeated game starting from state  $s_i$  and let  $\gamma$  be the value of the game (Eq.(1)). The ACOE connects these variables in the following manner:

$$\gamma + v_i = r_i + c \sum_j q_{ij} s_{ij} + \sum_j q_{ij} v_j, \quad (17)$$

where  $s_{ij}$  is an element in the matrix  $S$ .

4. The above equation actually represents  $n$  linear equations (one for each state) with  $n + 1$  variables:  $\gamma$  and the  $v_i$ s. W.l.o.g. set  $v_1 = 0$  and solve the remaining equations to obtain the value of the game  $\gamma$  and the continuation payoffs  $v_i$ .
5. Find the real mixed actions of Player 1. To do so, treat each state as a one-shot game by writing the switching costs and the continuation payoffs in the payoff matrix. Check that Player 1 has a best response to Player 2’s action in this game that keeps Player 2 indifferent among the actions he plays and makes him prefer them over all other actions.
6. If it is possible to find such a strategy for Player 1, this pair of strategies is optimal and  $\gamma$  is the value of the game. Otherwise, restart the algorithm with another guess.

We now apply this algorithm to the zero-sum game  $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$  and the switching-costs matrix  $S = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ . Without switching costs ( $c = 0$ ), the optimal strategy for Player 2 is  $\frac{1}{11}(6, 3, 2)$  and the value is  $v = \frac{6}{11}$ . For clarity, we denote the actions of Player 2 as  $L$  (left),  $M$  (middle) and  $R$  (right), and the state after action  $i \in \{L, M, R\}$  was played by  $s_i$ .

Suppose that Player 2 plays  $\frac{1}{11}(6, 3, 2)$  in all states. This makes Player 1 indifferent among all his actions, and we can assume that the best response is simply playing the first row. This results in  $r_i = \frac{6}{11}$  and  $(q_{iL}, q_{iM}, q_{iR}) = \frac{1}{11}(6, 3, 2)$  for all  $i$ . The three ACOE equations are

$$\begin{aligned} \gamma + v_L &= \frac{6}{11} + c\frac{5}{11} + \frac{6}{11}v_L + \frac{3}{11}v_M + \frac{2}{11}v_R, \\ \gamma + v_M &= \frac{6}{11} + c\frac{8}{11} + \frac{6}{11}v_L + \frac{3}{11}v_M + \frac{2}{11}v_R, \\ \gamma + v_R &= \frac{6}{11} + c\frac{9}{11} + \frac{6}{11}v_L + \frac{3}{11}v_M + \frac{2}{11}v_R. \end{aligned}$$

Setting  $v_L = 0$  and solving these equations leads to  $v_M = \frac{3c}{11}$ ,  $v_R = \frac{4c}{11}$  and  $\gamma = \frac{6}{11} + \frac{72}{121}c$ . Our final step is to add these continuation payoffs to the payoff matrix of each state and find optimal strategies for Player 1. The resulting three one-shot games are:

$$s_L : \begin{pmatrix} 1 & \frac{14c}{11} & \frac{15c}{11} \\ 0 & 2 + \frac{14c}{11} & \frac{15c}{11} \\ 0 & \frac{14c}{11} & 3 + \frac{15c}{11} \end{pmatrix} \quad s_M : \begin{pmatrix} 1 + c & \frac{3c}{11} & \frac{15c}{11} \\ c & 2 + \frac{3c}{11} & \frac{15c}{11} \\ c & \frac{3c}{11} & 3 + \frac{15c}{11} \end{pmatrix} \quad s_R : \begin{pmatrix} 1 + c & \frac{14c}{11} & \frac{4c}{11} \\ c & 2 + \frac{14c}{11} & \frac{4c}{11} \\ c & \frac{14c}{11} & 3 + \frac{4c}{11} \end{pmatrix}$$

In game  $s_L$ , solving the standard equation of indifference leads to a mixed strategy in which Player 1 plays his actions (from top to bottom) with probabilities  $\frac{1}{121}(72c + 66, 33 - 41c, 22 - 31c)$  (and similar solutions for the other states). This mixed action is indeed a distribution for  $c \leq \frac{22}{31}$  and for these values of  $c$  we have found the value  $v(c) = \frac{6}{11} + \frac{72c}{121}$  and the optimal strategies.

For  $c > \frac{22}{31}$ , action  $R$  in state  $s_L$  becomes dominated and Player 1 cannot keep Player 2 indifferent between it and the other actions. We therefore assume Player 2 never plays action  $R$  in state  $s_L$ . As a result, Player 1 will never play his bottom action in  $s_L$  and the new guess for a mixed action for Player 2 in state  $s_L$  is  $\frac{1}{3}(2, 1, 0)$ . In the rest of the states, action  $R$  is not dominated so Player 2 continues to play mixed action  $\frac{1}{11}(6, 3, 2)$  in them.

Repeating the same process with this strategy (which is not static this time) leads to the next critical value of  $c = \frac{121}{156}$ . Above it, action  $M$  in state  $s_L$  (including, again, continuation payoffs and switching costs) is dominated by  $L$  and so  $s_L$  becomes absorbing.