



HAL
open science

A simple random walk game

Stefan Ankirchner, Hugalf Bernburg, Julian Wendt

► **To cite this version:**

| Stefan Ankirchner, Hugalf Bernburg, Julian Wendt. A simple random walk game. 2022. <hal-03607763>

HAL Id: hal-03607763

<https://hal.science/hal-03607763v1>

Preprint submitted on 14 Mar 2022

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.



HAL Authorization

A simple random walk game

S. Ankirchner H. Bernburg J. Wendt

Friedrich Schiller University Jena
Institute for Mathematics
Ernst-Abbe-Platz 2
07743 Jena, Germany.

March 14, 2022

We consider a dynamic two player game where each player can control the step sizes of her own random walk. At any time, the players can choose a step size equal to 0 or 1. Moreover, whenever a player is ahead, i.e. her state is above the opponent's state, she receives a fixed reward. The player behind does not receive anything. We assume that both players aim at maximizing the expected sum of discounted rewards.

We show that there exists a threshold that the correlation between the two random walks must not exceed in order for the game to possess an equilibrium in pure strategies. We explicitly compute the threshold level and show, provided the correlation is smaller than or equal to the threshold, that an equilibrium is given if the player ahead chooses step size 0 and the player behind step size 1.

2020 MSC : 91A05, 91A10, 91A15, 91A25, 91A50

Keywords : two player game; Nash equilibrium; random walk; pure strategy; correlation, step size control

1 Introduction and main results

We start by describing the random walk game in more detail. Let $\alpha \in (0, 1)$ be a discount factor and $\rho \in [-1, 1]$ be the correlation between the players' random walks. Let $A = \{0, 1\}$ be the action set and $\mathcal{A} = \{\pi : \mathbb{Z} \times \mathbb{Z} \rightarrow A\}$ be the set of all *pure* strategies. For the moment we consider only pure strategies, but allow for mixed strategies at the end of this chapter.

Whenever player i ($i \in \{1, 2\}$) chooses action 0, her random walk stays constant over the next period. If she chooses action 1, then her random walk increases or decreases by 1, both with probability $\frac{1}{2}$. To give a rigorous definition of the states, let $Y^{i,\pi}$ denote the random walk of player i , provided she chooses strategy $\pi \in \mathcal{A}$. Given a pair of strategies $(\pi_1, \pi_2) \in \mathcal{A}^2$, the state processes $(Y^{1,\pi_1}, Y^{2,\pi_2})$ are recursively defined as follows: if $(Y_n^{1,\pi_1}, Y_n^{2,\pi_2}) = (j, k) \in \mathbb{Z} \times \mathbb{Z}$ we set

$$Y_{n+1}^{i,\pi_i} = \begin{cases} Y_n^{i,\pi_i} + \xi_n^i, & \text{if } \pi_i(j, k) = 1, \\ Y_n^{i,\pi_i}, & \text{if } \pi_i(j, k) = 0, \end{cases} \quad n \in \mathbb{N}_0, \quad i \in \{1, 2\},$$

where the families $(\xi_n^1)_{n \geq 0}$ and $(\xi_n^2)_{n \geq 0}$ consist of identically distributed random variables satisfying $P(\xi_0^i = 1) = P(\xi_0^i = -1) = \frac{1}{2}$. We assume that

$$E[\xi_n^1 \xi_m^2] = \begin{cases} \rho, & \text{if } n = m, \\ 0, & \text{else,} \end{cases}$$

for $n, m \geq 0$, i.e. the random variables ξ_n^1 and ξ_m^2 have correlation ρ if $n = m$, or are uncorrelated (and hence independent) if $n \neq m$. Note that the state processes are time-homogeneous (discrete time) Markov chains on the countable state space \mathbb{Z} .

At each time step both players receive a reward based on the value of their state processes' difference. The reward is given by the function $R : \mathbb{Z} \rightarrow \mathbb{R}$ defined by

$$R(i) := \begin{cases} 1, & \text{if } i > 0, \\ \frac{1}{2}, & \text{if } i = 0, \\ 0, & \text{if } i < 0. \end{cases}$$

Note that $R(i) + R(-i) = 1$ for all $i \in \mathbb{Z}$. The gain function of player 1 is defined by

$$J(i, \pi, \sigma) = \sum_{n \geq 0} \alpha^n E [R(Y_n^{1,\pi} - Y_n^{2,\sigma}) \mid Y_0^{1,\pi} - Y_0^{2,\sigma} = i]$$

and the gain function of player 2 by

$$I(i, \pi, \sigma) = \sum_{n \geq 0} \alpha^n E [R(Y_n^{2,\sigma} - Y_n^{1,\pi}) \mid Y_0^{1,\pi} - Y_0^{2,\sigma} = i],$$

for the strategies $\pi, \sigma \in \mathcal{A}$. We suppose that both players aim at maximizing their gain functions. Note that the gain functions depend on the players' states only through their difference. Hence, we can, without loss of generality, restrict our considerations to the set of strategies $\{\pi : \mathbb{Z} \rightarrow A\}$, i.e. to controls only depending on the states' difference. We therefore redefine $\mathcal{A} = \{\pi : \mathbb{Z} \rightarrow A\}$ for the rest of the paper.

Notice that $I(i, \pi, \sigma) = \frac{1}{1-\alpha} - J(i, \pi, \sigma)$. Thus maximizing I is equivalent to minimizing J . It is, therefore, enough to consider only J and we can resort to the upper and lower value function for analyzing the game.

The upper value of the game is given by

$$V^+(i) = \inf_{\sigma \in \mathcal{A}} \sup_{\pi \in \mathcal{A}} J(i, \pi, \sigma), \quad i \in \mathbb{Z}.$$

Note that V^+ is the expected total gain of player 1 if at each time player 2 has to choose first and player 1 can choose her strategy afterwards. If conversely player 1 has to choose her strategy before player 2, her total gain is described by the lower value, defined as

$$V^-(i) = \sup_{\pi \in \mathcal{A}} \inf_{\sigma \in \mathcal{A}} J(i, \pi, \sigma), \quad i \in \mathbb{Z}.$$

Notice that it always holds true that $V^- \leq V^+$. In general the value functions are not equal. We say that the game has a value if $V^+ = V^-$. Recall that if there exists a Nash equilibrium, i.e. a pair of strategies $(\pi^*, \sigma^*) \in \mathcal{A} \times \mathcal{A}$ such that

$$J(i, \pi^*, \sigma^*) = \sup_{\pi \in \mathcal{A}} J(i, \pi, \sigma^*) = \inf_{\sigma \in \mathcal{A}} J(i, \pi^*, \sigma), \quad i \in \mathbb{Z},$$

then the game has a value.

Our main result is that there exists a Nash equilibrium if the correlation does not exceed the following bound

$$B(\alpha) := \frac{1 - \sqrt{1 - \alpha^2}}{2\alpha + 1 - \sqrt{1 - \alpha^2}}.$$

Theorem 1.1. *The strategy $(\pi^*, 1 - \pi^*) \in \mathcal{A} \times \mathcal{A}$, defined by*

$$\pi^*(i) := \begin{cases} 0, & \text{if } i \geq 0, \\ 1, & \text{if } i < 0, \end{cases}$$

is a Nash equilibrium of the two player game if and only if

$$\rho \leq B(\alpha). \tag{1}$$

In this case, the game has a value and the value function is given by

$$V(i) = \begin{cases} \frac{1}{1-\alpha} \left(1 - \frac{1}{2} \left(\frac{1}{\alpha} - \sqrt{\frac{1}{\alpha^2} - 1} \right)^i \right), & \text{if } i \geq 0, \\ \frac{1}{2} \frac{1}{1-\alpha} \left(\frac{1}{\alpha} + \sqrt{\frac{1}{\alpha^2} - 1} \right)^i, & \text{if } i < 0. \end{cases}$$

The proof of Theorem 1.1 is given in Section 2. We proceed here by giving an intuitive explanation of why $(\pi^*, 1 - \pi^*)$ is an equilibrium only if the correlation is bounded from above by the threshold $B(\alpha)$.

Let $X = Y^1 - Y^2$ denote the difference process of the two players' states. If X is positive, then player 1 (player 2) aims at maximizing (minimizing) the return time to her losing region (winning region), the region where X is negative. The higher the step size variance, the faster the process X returns back to zero respectively crosses zero

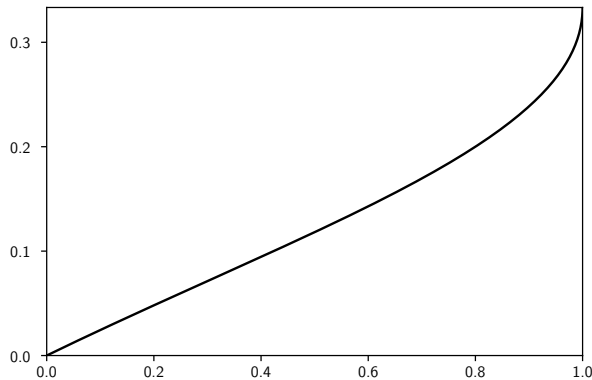


Figure 1: The bound $B(\alpha)$ in dependence on α

(notice that the expectation of the first return time is infinite, however the probability of returning before a fixed time point increases with increasing step size variance). Therefore, the player ahead aims at minimizing the variance of the steps of X , and the player behind aims at maximizing it.

Now for $\rho \leq \frac{1}{2}$ it holds true that $a = 0$ minimizes the step size variance of X and $a = 1$ maximizes the step size variance. Thus, one can expect that $(\pi^*, 1 - \pi^*)$ are mutual best responses if $\rho \leq \frac{1}{2}$. As shown in Section 2, this is indeed true for state pairs such the difference X is not equal to -1 or 1 .

A special situation arises if X lies in $\{-1, 1\}$. To explain this suppose that $X_n = 1$, i.e. the player 1 is ahead by one unit at time n . Suppose that player 2 chooses action $1 - \pi^*(1) = 1$. Then the expected reward of player 1 received at time $n + 1$ is given by

$$\begin{aligned} & \frac{3}{4}, \text{ if she chooses } a = 0, \\ & \frac{3+\rho}{4}, \text{ if she chooses } a = 1. \end{aligned}$$

Thus, if the correlation is positive, player 1 has an incentive to deviate from π^* . The deviation comes at the price that the return time to the losing region is increased. Indeed, if player 1 chooses $a = 0$, then X_{n+1} will always be non-negative, whereas under action $a = 1$ with a positive probability X_{n+1} will be equal -1 . It turns out that π^* is an optimal response to $1 - \pi^*$ if and only if condition (1) is satisfied.

Notice that the discount factor α determines by how much rewards in the present are preferred over rewards in the future. The smaller α , the higher player 1 evaluates the additional benefit from choosing action 1 in $X_n = 1$, and the lower the benefit from choosing $a = 0$. This explains why $B(\alpha)$ is increasing in α (see Figure 1).

Theorem 1.1 states that the pair $(\pi^*, 1 - \pi^*)$ is an equilibrium if (1) is satisfied, but does not make a statement about the existence of an equilibrium if (1) is *not* satisfied. There is a simple argument showing that the correlation must not be too large for an

equilibrium in \mathcal{A} to exist. To this end suppose first that $\rho = 1$. Then the player ahead can keep the distance to the other player constant by choosing the same action. In other words, the best response of the player ahead is to choose the same action as the player behind. However, the best response of the player behind is to choose the other action than the one chosen of the player ahead. Thus, there does not exist an equilibrium in \mathcal{A} if $\rho = 1$. The argument applies also for correlations close to 1, and hence indicates that the correlation must not be too large for an equilibrium in pure strategies to exist.

It seems plausible to conjecture that for all $\rho > B(\alpha)$ the game does not have a value, i.e. $V^- < V^+$ holds true, and that it matters if one player can observe the other. In order to confirm this, we use that the upper and the lower value function solve the associated upper and lower Isaacs equation, respectively. The upper Isaacs equation of our game is given by

$$w^+(i) = R(i) + \alpha \min_{b \in A} \max_{a \in A} \sum_{j \in \mathbb{Z}} p_{ij}(a, b) w^+(j), \quad (2)$$

and the lower Isaacs equation by

$$w^-(i) = R(i) + \alpha \max_{a \in A} \min_{b \in A} \sum_{j \in \mathbb{Z}} p_{ij}(a, b) w^-(j), \quad (3)$$

where the transition probabilities $p_{ij}(a, b)$ are defined in Section 2 below (see Equation (6)). We show in Proposition 3.1 below that V^+ and V^- indeed solve (2) and (3), respectively.

For $\rho = 1$ one can solve both equations explicitly.

Proposition 1.2. *Let $\rho = 1$. Then $V^- < V^+$ and the upper and lower value function are given by*

$$V^+(i) = \begin{cases} \frac{1}{\alpha(1-\alpha)} \left(\frac{1}{\alpha} + \sqrt{\frac{1}{\alpha^2} - 1} \right)^{i-1}, & \text{if } i < 0, \\ \frac{1}{2} + \frac{\alpha}{2(1-\alpha)} \left(1 + \frac{1}{\alpha} \left(\frac{1}{\alpha} - \sqrt{\frac{1}{\alpha^2} - 1} \right)^2 \right), & \text{if } i = 0, \\ \frac{1}{1-\alpha}, & \text{if } i > 0, \end{cases} \quad (4)$$

$$V^-(i) = \begin{cases} 0, & \text{if } i < 0, \\ \frac{1}{2} + \frac{\alpha}{2(1-\alpha)} \left(1 - \frac{1}{\alpha} \left(\frac{1}{\alpha} - \sqrt{\frac{1}{\alpha^2} - 1} \right)^2 \right), & \text{if } i = 0, \\ \frac{1}{1-\alpha} \left(1 - \frac{1}{\alpha} \left(\frac{1}{\alpha} - \sqrt{\frac{1}{\alpha^2} - 1} \right)^{i+1} \right), & \text{if } i > 0. \end{cases} \quad (5)$$

For $\rho \in (B(\alpha), 1)$ it is more difficult to obtain an explicit formula for the upper and lower value function. A probabilistic explanation is that if $\rho \leq B(\alpha)$, then the equilibrium difference process X is a simple random walk (i.e. in each step X moves up or down by one unit, with probability $\frac{1}{2}$, respectively). In the case $\rho \in (B(\alpha), 1)$, however, the transitions of the difference process X in the minmax and maxmin game are not as simple.

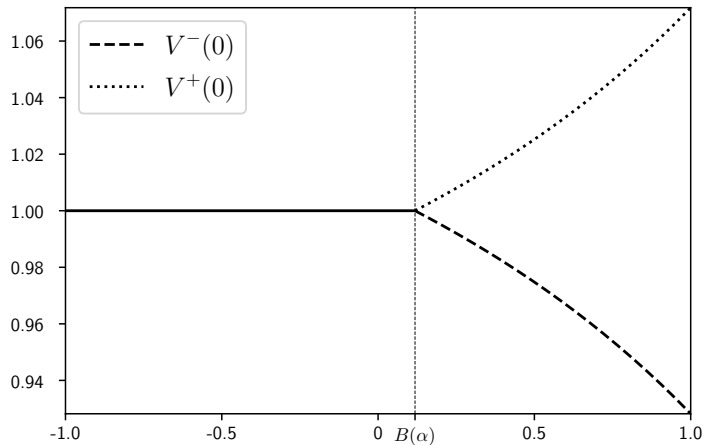


Figure 2: The picture shows the dependence of $V^+(0)$ and $V^-(0)$ on the correlation ρ for $\alpha = \frac{1}{2}$. The vertical line marks the value $\rho = B(\alpha) \approx 0.12$. For $\rho \leq B(\alpha)$ the upper and lower value function coincide and assume the value $\frac{1}{2} \frac{1}{1-\alpha} = 1$. At the point $\rho = B(\alpha)$ the graphs separate into two branches. Hence, for $\rho > B(\alpha)$ the upper and lower value function in pure strategies differ and the game does not have a value in pure strategies.

To show that the upper and lower value function do not coincide for $\rho \in (B(\alpha), 1)$, one can fall back on numerical approximations of the Isaacs equations. A standard approximation method interprets the right-hand side of (2) and (3) as operators applied to w^+ and w^- , respectively. The operators can be shown to be contracting on $l^\infty(\mathbb{Z})$, the space of bounded functions on \mathbb{Z} . Applying this operator iteratively leads to a numerical scheme that can be shown to converge as the number of iterations and the cutoff in the state space tends to infinity. One can even obtain an explicit bound for the approximation error depending on these quantities. By using the numerical approximation one can thus show that the game does *not* have a value for $\rho > B(\alpha)$ in pure strategies. For $\alpha = \frac{1}{2}$, we illustrate this in Figure 2.

So far we have only considered pure strategies. The game can be extended, in a straightforward manner, so that players are allowed to choose mixed strategies. Mixed strategies can be represented as functions in $\mathcal{M} = \{\pi : \mathbb{Z} \times \mathbb{Z} \rightarrow [0, 1]\}$. Irrespective of the size of the correlation, the game always possesses an equilibrium in mixed strategies.

Proposition 1.3. *For all $\rho \in [-1, 1]$ there exists a Nash equilibrium of the game in mixed strategies.*

Proposition 1.3 follows from more general results, provided, e.g., in Parthasarathy [13], Theorem 5.1. The proof in the latter reference relies on Kakutani's fixed point theorem. One can also prove Proposition 1.3 by utilizing von Neumann's minimax theorem or the more general minimax theorem of Sion (see e.g. [22] and [18]).

Theorem 1.1 and Proposition 1.3 reveal that if the correlation exceeds $B(\alpha)$, then at every time n each player aims at hiding her action so that the opponent has to choose an action *before* seeing the other player's one.

Comparison with the literature

Von Neumann [22] was one of the first to study two player zero-sum matrix games. He showed the existence of an equilibrium in mixed strategies by presenting his minimax theorem. Later von Neumann and Morgenstern [23] proposed a general theory of games that was groundbreaking at this time. The theory of non-cooperative game theory started in the 1950s. Nash introduced in [11] and [12] the notion of equilibria in n -player games and showed their existence for finite games. Nowadays this equilibrium is known as *Nash equilibrium*.

Discounted dynamic games on an infinite time horizon were first described by Shapley [17] in 1953. He studied a two player zero-sum game with finite action and state space and positive stop probability. His work marks the beginning of the theory of stochastic games. After that numerous generalizations of discounted dynamic games have been studied in the literature: For example Takahashi [20], Maitra and Parthasarathy [9], Parthasarathy [13] contributed to generalizations to more general action and state spaces, Fink [5], Takahashi [21], Rogers [15], Sobel [19] and Federgruen [3] considered n -player analogues. We particularly highlight the article [13] by Parthasarathy. To the best of our knowledge it was the first that covers our setting. While also considering discounted two player games with quite general state and action spaces, Parthasarathy showed that there exists an equilibrium in mixed strategies for discounted two player games with countable state space and finite action sets (see Theorem 5.1 in [13]). We refer the reader to [4], [8], [2], [6] for further bibliographical and game theoretical details.

One way to obtain explicit results for more general assumptions, e.g. on the action set, is to use diffusions as state processes, i.e. to consider stochastic differential games. Controlling the step size of a random walk then corresponds to controlling the diffusion rate of a Brownian motion. E.g. McNamara [10] studies a two player zero-sum stochastic differential game where both players control the diffusion coefficient of a single state process. The objective of one player is to maximize the probability that this state process is positive at some final time T while the other players' objective is to minimize it. McNamara derives an explicit equilibrium for the game. The article [1] addresses an n -player diffusion control game on a finite time horizon where the rewards depend on the empirical quantile of all players. For the control any value in a bounded interval $[\sigma_1, \sigma_2] \subseteq (0, \infty)$ can be chosen. The authors derive an approximate Nash equilibrium and also present an explicit Nash equilibrium in the two player case. In contrast to the game considered in the present article, [1] does not allow for correlation. We believe, however, that the results for the two player case can be generalized to the game version where the driving Brownian motions are correlated.

The rest of the paper is organized as follows: In Section 2 we prove our main results Theorem 1.1 and Proposition 1.2, and finally in Section 3 we prove that the upper and

lower value function solve the upper and lower Isaacs equation, respectively, to make the paper self-contained.

2 Proof of the main results

2.1 Proof of Theorem 1.1

The goal of this section is to prove our main result Theorem 1.1. We start by formulating the game in an equivalent way: The reward function R and hence the target function J only depend on the difference of the state processes $Y^{1,\pi}$ and $Y^{2,\sigma}$. Therefore, we define the process $X^{\pi,\sigma} := Y^{1,\pi} - Y^{2,\sigma}$ for $\pi, \sigma \in \mathcal{A}$ and restrict our considerations to the game where both players control the common state process $X^{\pi,\sigma}$. The process $X^{\pi,\sigma}$ is again a time-homogeneous Markov chain with transition probabilities given by

$$\begin{aligned} p_{ij}(a, b) &:= abP(\xi_n^1 - \xi_n^2 = j - i) \\ &\quad + (1 - a)bP(-\xi_n^2 = j - i) \\ &\quad + a(1 - b)P(\xi_n^1 = j - i) \\ &\quad + (1 - a)(1 - b)\mathbf{1}_{\{0\}}(j - i) \\ &= \begin{cases} (1 - a)(1 - b) + \frac{1+\rho}{2}ab, & \text{if } j = i, \\ \frac{1}{2}(1 - a)b + \frac{1}{2}a(1 - b), & \text{if } j = i - 1, i + 1, \\ \frac{1-\rho}{4}ab, & \text{if } j = i - 2, i + 2, \\ 0, & \text{else,} \end{cases} \end{aligned} \quad (6)$$

for $i, j \in \mathbb{Z}$ and actions $a, b \in A$. For strategies $\pi, \sigma \in \mathcal{A}$ we set

$$p_{ij}(\pi, \sigma) = p_{ij}(\pi, \sigma(i)) = p_{ij}(\pi(i), \sigma) = p_{ij}(\pi(i), \sigma(i)).$$

Moreover, we write $X^{i,\pi,\sigma}$ for the Markov chain $X^{\pi,\sigma}$ with initial state $i \in \mathbb{Z}$.

We aim at proving that $(\pi^*, 1 - \pi^*)$ is a Nash equilibrium. To this end, we show that

$$J(\cdot, \pi^*, 1 - \pi^*) = \sup_{\pi \in \mathcal{A}} J(\cdot, \pi, 1 - \pi^*) = \inf_{\sigma \in \mathcal{A}} J(\cdot, \pi^*, \sigma),$$

by proving that $J(\cdot, \pi^*, 1 - \pi^*)$ solves the two corresponding optimality equations

$$v(i) = R(i) + \max_{a \in A} \sum_{j \in \mathbb{Z}} p_{ij}(a, 1 - \pi^*)v(j), \quad (7)$$

$$w(i) = R(i) + \min_{b \in A} \sum_{j \in \mathbb{Z}} p_{ij}(\pi^*, b)w(j), \quad i \in \mathbb{Z}. \quad (8)$$

Note that for the strategy tuple $(\pi^*, 1 - \pi^*)$ the transition probabilities are given by

$$p_{ij}(\pi^*, 1 - \pi^*) = \begin{cases} \frac{1}{2}, & \text{if } j = i - 1, i + 1, \\ 0, & \text{else,} \end{cases} \quad i, j \in \mathbb{Z}.$$

This means $X^{i,\pi^*,1-\pi^*}$ is a simple random walk on \mathbb{Z} with steps uniformly distributed in $\{-1,1\}$. We abbreviate the value of the game for this strategy tuple by $W := J(\cdot, \pi^*, 1 - \pi^*)$. W has the following properties:

Lemma 2.1. *The function W satisfies the following:*

- (i) $W(i) + W(-i) = \frac{1}{1-\alpha}$ for $i \in \mathbb{Z}$,
- (ii) $\lim_{i \rightarrow \infty} W(i) = \frac{1}{1-\alpha}$ and $\lim_{i \rightarrow -\infty} W(i) = 0$,
- (iii) W solves the difference equation

$$W(i) = R(i) + \frac{\alpha}{2} (W(i+1) + W(i-1)), \quad i \in \mathbb{Z}, \quad (9)$$

and is given by

$$W(i) = \begin{cases} \frac{1}{1-\alpha} \left(1 - \frac{1}{2} \left(\frac{1}{\alpha} - \sqrt{\frac{1}{\alpha^2} - 1} \right)^i \right), & \text{if } i \geq 0, \\ \frac{1}{2} \frac{1}{1-\alpha} \left(\frac{1}{\alpha} + \sqrt{\frac{1}{\alpha^2} - 1} \right)^i, & \text{if } i < 0. \end{cases} \quad (10)$$

- (iv) W is convex for $i \leq 0$ and concave for $i \geq 0$.

Proof. (i): The process $X^{\pi^*,1-\pi^*}$ is a standard random walk. Note that for any $i \in \mathbb{Z}$ the distributions of $X^{i,\pi^*,1-\pi^*}$ and $-X^{-i,\pi^*,1-\pi^*}$ coincide. Hence, we have

$$\begin{aligned} W(i) + W(-i) &= \sum_{n=0}^{\infty} \alpha^n E \left(R(X_n^{i,\pi^*,1-\pi^*}) + R(X_n^{-i,\pi^*,1-\pi^*}) \right) \\ &= \sum_{n=0}^{\infty} \alpha^n E \left(R(X_n^{i,\pi^*,1-\pi^*}) + R(-X_n^{i,\pi^*,1-\pi^*}) \right) = \sum_{n=0}^{\infty} \alpha^n = \frac{1}{1-\alpha}. \end{aligned}$$

(ii): By definition $W(i) = \sum_{n=0}^{\infty} \alpha^n E [R(X_n^{i,\pi^*,1-\pi^*})]$ and hence the results follow by dominated convergence.

(iii): Standard results for Markov decision processes imply that W solves the difference equation

$$\begin{aligned} W(i) &= R(i) + \sum_j p_{ij}(\pi^*(i), 1 - \pi^*(i)) W(j) \\ &= R(i) + \frac{\alpha}{2} (W(i+1) + W(i-1)), \quad i \in \mathbb{Z}, \end{aligned}$$

see, e.g., Proposition II.2.4 in [16] or Theorem 6.1.1 in [14]. Hence, W solves (9). Moreover, by definition W is bounded and satisfies (ii) and $W_0 = \frac{1}{2} \frac{1}{1-\alpha}$. These properties uniquely characterize W as solution of (9). The formula (10) follows from solving (9) explicitly respecting the boundedness of W .

(iv): Follows from the explicit formula (10). □

To show that π^* and $1 - \pi^*$ are mutually best responses we consider the identities

$$\sum_j p_{ij}(\pi^*, 1 - \pi^*)W(j) = \max_{a \in A} \sum_j p_{ij}(a, 1 - \pi^*)W(j), \quad (11)$$

$$\sum_j p_{ij}(\pi^*, 1 - \pi^*)W(j) = \min_{a \in A} \sum_j p_{ij}(\pi^*, a)W(j), \quad (12)$$

that appear on the right-hand side of the optimality equations (7) and (8). We have the following:

Lemma 2.2. *Equation (11) holds true for*

(i) $i \leq 0$,

(ii) $i \geq 2$ if and only if $\rho \leq \frac{1}{1+\alpha}$,

(iii) $i = 1$ if and only if condition (1) is satisfied.

Moreover, equation (12) holds true for

(iv) $i \geq 0$,

(v) $i \leq -2$ if and only if $\rho \leq \frac{1}{1+\alpha}$,

(vi) $i = -1$ if and only if condition (1) is satisfied.

Proof. First of all, note that for any $a, b \in A$

$$\sum_j p_{0j}(a, b)W(j) = \frac{1}{2(1-\alpha)}.$$

Hence, (11) and (12) hold true for $i = 0$ independently of the choice of the control.

(i): For $i < 0$ we have due to convexity (Lemma 2.1(iv))

$$W(i) \leq \frac{1}{2}(W(i+1) + W(i-1)).$$

Hence, (11) holds true for $i < 0$.

(ii): For $i \geq 2$ the equation (11) is equivalent to

$$0 \leq \frac{1}{2}(W(i+1) + W(i-1)) - \frac{1-\rho}{4}(W(i+2) + W(i-2)) - \frac{1+\rho}{2}W(i) =: Z. \quad (13)$$

Hence, we study the right-hand side of (13) and determine under which conditions Z is non-negative. By Lemma 2.1 we can write Z as

$$Z = \frac{1}{4(1-\alpha)}z^{i-2} \left(\frac{1-\rho}{2}z^4 - z^3 + (1+\rho)z^2 - z + \frac{1-\rho}{2} \right),$$

with $z := \frac{1}{\alpha} - \sqrt{\frac{1}{\alpha^2} - 1}$. Note that $z \in (0, 1)$ since $\alpha \in (0, 1)$. Firstly, we consider the case $\rho = 1$: Then Z can be written as

$$Z = \frac{1}{4(1-\alpha)} z^{i-1} (-z^2 + 2z - 1) = -\frac{1}{4(1-\alpha)} z^{i-1} (z-1)^2 < 0.$$

This means that (13) is not fulfilled for $\rho = 1$. Secondly, we study the case $\rho < 1$: Then Z is given by

$$\begin{aligned} Z &= \frac{1}{4(1-\alpha)} \frac{1-\rho}{2} z^{i-2} \left(z^4 - \frac{2}{1-\rho} z^3 + \frac{2(1+\rho)}{1-\rho} z^2 - \frac{2}{1-\rho} z + 1 \right) \\ &= \frac{1}{4(1-\alpha)} \frac{1-\rho}{2} z^{i-2} (z-1)^2 \left(z^2 - \frac{2\rho}{1-\rho} z + 1 \right) \end{aligned}$$

This means that it suffices to consider the polynomial $f(x) := x^2 - \frac{2\rho}{1-\rho}x + 1$ to determine the sign of Z . If $\rho \in (\frac{1}{2}, 1)$, then the roots of f are real-valued and given by

$$x_1 := \frac{\rho - \sqrt{2\rho - 1}}{1-\rho} \in (0, 1) \quad \text{and} \quad x_2 := \frac{\rho + \sqrt{2\rho - 1}}{1-\rho} > 1.$$

We see that $f(x) < 0$ for $x \in (x_1, x_2)$, and $f(x) \geq 0$ for $x \leq x_1$. This implies that Z is non-negative if and only if $z \leq x_1$, which is equivalent to $\rho \leq \frac{1}{1+\alpha}$. If $\rho = \frac{1}{2}$, then

$$Z = \frac{1}{4(1-\alpha)} \frac{1-\rho}{2} z^{i-2} (z-1)^4 \geq 0.$$

If $\rho < \frac{1}{2}$ we have that $f(x) \geq \frac{1-2\rho}{(1-\rho)^2} > 0, x \in \mathbb{R}$, and thus $Z \geq 0$. Summing up, (13) is satisfied for $i \geq 2$ if and only if $\rho \leq \frac{1}{1+\alpha}$.

(iii): For $i = 1$ the equation (11) is equivalent to

$$\begin{aligned} 0 &\leq \frac{1}{2} (W(2) + W(0)) - \frac{1-\rho}{4} (W(3) + W(-1)) - \frac{1+\rho}{2} W(1) \\ &= \frac{1}{8(1-\alpha)} ((1-\rho)z^3 - 2z^2 + (1+3\rho)z - 2\rho) =: Z, \end{aligned}$$

again with $z := \frac{1}{\alpha} - \sqrt{\frac{1}{\alpha^2} - 1} \in (0, 1)$. If $\rho = 1$, then

$$Z = -\frac{1}{4(1-\alpha)} (z-1)^2 < 0.$$

If $\rho < 1$, then

$$Z = \frac{1-\rho}{8(1-\alpha)} (z-1)^2 \left(z - \frac{2\rho}{1-\rho} \right),$$

i.e. Z is non-negative if and only if $z \geq \frac{2\rho}{1-\rho}$. But this is equivalent to condition (1).

(iv)-(vi): Note that (6) implies that the transition probabilities are symmetric in the control variables, i.e. $p_{ij}(a, b) = p_{ij}(b, a)$ for $a, b \in A$. Then Lemma 2.1 implies for $i, j \in \mathbb{Z}$ that

$$\sum_j p_{ij}(a, b)W(j) = \sum_j p_{ij}(b, a) \left(\frac{1}{1-\alpha} - W(-j) \right) = \frac{1}{1-\alpha} - \sum_j p_{(-i)j}(b, a)W(j).$$

Thus, (i)-(iii) imply (iv)-(vi). \square

Finally, we are in the position to prove Theorem 1.1 as a result of Lemma 2.1 and 2.2.

Proof of Theorem 1.1. In the first step we show that under condition (1) the tuple $(\pi^*, 1 - \pi^*)$ is a Nash equilibrium. Then in the second step we prove the converse statement: If $(\pi^*, 1 - \pi^*)$ is a Nash equilibrium then also (1) holds true.

Step 1: Let condition (1) hold true. Then also $\rho \leq \frac{1}{1+\alpha}$, i.e. under the condition (1) the equations (11) and (12) hold true due to Lemma 2.2. Note that the left-hand side of (11) and (12) is equal to $\frac{1}{2}(W(i-1) + W(i+1))$ and that W solves (9). Hence, W also solves the optimality equations (7) and (8) of the two control problems $\sup_{\pi} J(\cdot, \pi, 1 - \pi^*)$ and $\inf_{\sigma} J(\cdot, \pi^*, \sigma)$. Standard results for Markov decision processes (see, e.g., Theorem II.2.1 in [16] or Theorem 6.2.5 in [14]) imply that

$$W = \sup_{\pi \in \mathcal{A}} J(\cdot, \pi, 1 - \pi^*) = \inf_{\sigma \in \mathcal{A}} J(\cdot, \pi^*, \sigma).$$

This means that $(\pi^*, 1 - \pi^*)$ is a Nash equilibrium in pure strategies. Moreover, the game has a value and W coincides with the value function of the game.

Step 2: Conversely, suppose that $(\pi^*, 1 - \pi^*)$ is a pure strategy Nash equilibrium. This means on one hand that $W = \sup_{\pi \in \mathcal{A}} J(\cdot, \pi, 1 - \pi^*)$ and W satisfies the optimality equation (see e.g. Theorem II.2.1 in [16] or Theorem 6.2.5 in [14])

$$W(i) = R(i) + \alpha \max_{a \in A} \sum_{j \in \mathbb{Z}} p_{ij}(a, 1 - \pi^*)W(j), \quad i \in \mathbb{Z}.$$

On the other hand $W = J(i, \pi^*, 1 - \pi^*)$ and hence W satisfies the difference equation

$$\begin{aligned} W(i) &= R(i) + \alpha \sum_{j \in \mathbb{Z}} p_{ij}(\pi^*, 1 - \pi^*)W(j), \\ &= R(i) + \frac{\alpha}{2} (W(i-1) + W(i+1)), \quad i \in \mathbb{Z}. \end{aligned}$$

This means that (11) holds true. Lemma 2.2 (i)-(iii) implies that (1) holds true. \square

2.2 Proof of Proposition 1.2

If the correlation is equal to 1 the transition probabilities given in (6) take the following simple form

$$p_{ij}(0,0) = p_{ij}(1,1) = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{else,} \end{cases}$$

$$p_{ij}(1,0) = p_{ij}(0,1) = \begin{cases} \frac{1}{2}, & \text{if } j = i - 1, i + 1, \\ 0, & \text{else.} \end{cases}$$

We prove Proposition 1.2 by defining candidates for optimal responses to the control chosen by the opposing player. We show that if player 2 chooses the strategy $\sigma \in \mathcal{A}$, then the strategy π^σ defined by

$$\pi^\sigma(i) = \begin{cases} 1 - \sigma(i), & i \leq 0, \\ \sigma(i), & i > 0, \end{cases}$$

is an optimal response to σ for player 1. Similarly, if player 1 chooses a control $\pi \in \mathcal{A}$ the control σ^π defined by

$$\sigma^\pi(i) = \begin{cases} \pi(i), & i < 0, \\ 1 - \pi(i), & i \geq 0, \end{cases}$$

is an optimal response to π for player 2. These results enable us to find an explicit representation of the upper and lower value function. In more detail, we aim at showing that $J(\cdot, \pi^\sigma, \sigma) = V^+$ and $J(\cdot, \pi, \sigma^\pi) = V^-$. Let $W^+, W^- : \mathbb{Z} \rightarrow \mathbb{R}$ be equal to the right-hand sides of (4) and (5), respectively.

Lemma 2.3. *For any $\sigma \in \mathcal{A}$ we have $\sup_{\pi \in \mathcal{A}} J(\cdot, \pi, \sigma) = J(\cdot, \pi^\sigma, \sigma) = W^+$, and for any $\pi \in \mathcal{A}$ we have $\inf_{\sigma \in \mathcal{A}} J(\cdot, \pi, \sigma) = J(\cdot, \pi, \sigma^\pi) = W^-$.*

Proof. We only show the statement for W^+ since the other statement follows along the same lines. Let $\sigma \in \mathcal{A}$. First we show that $J(\cdot, \pi^\sigma, \sigma) = W^+$. It holds that

$$J(i, \pi^\sigma, \sigma) = \frac{\alpha}{2} (J(i-1, \pi^\sigma, \sigma) + J(i+1, \pi^\sigma, \sigma)), \quad i < 0,$$

$$J(0, \pi^\sigma, \sigma) = \frac{1}{2} + \frac{\alpha}{2} (J(-1, \pi^\sigma, \sigma) + J(1, \pi^\sigma, \sigma)),$$

$$J(i, \pi^\sigma, \sigma) = 1 + \alpha J(i, \pi^\sigma, \sigma), \quad i > 0.$$

Solving these equations implies that $J(\cdot, \pi^\sigma, \sigma) = W^+$. Now one can verify that W^+ satisfies the optimality equation of the control problem $\sup_{\pi} J(\cdot, \pi, \sigma)$ by using the explicit formula for W^+ . This yields the result $W^+ = \sup_{\pi \in \mathcal{A}} J(\cdot, \pi, \sigma) = J(\cdot, \pi^\sigma, \sigma)$. \square

Proof of Proposition 1.2. Lemma 2.3 implies that for any $\pi, \sigma \in \mathcal{A}$ we have

$$W^+ = \sup_{\pi \in \mathcal{A}} J(\cdot, \pi, \sigma)$$

$$W^- = \inf_{\sigma \in \mathcal{A}} J(\cdot, \pi, \sigma).$$

Hence, we have $W^+ = V^+$ and $W^- = V^-$. \square

3 Isaacs equations and the upper and lower value function

To make this paper self-contained we show in this section that the upper and lower value function solve the upper and lower Isaacs equations, respectively. Note that we are seeking solutions of (2) and (3) in the space $l^\infty(\mathbb{Z})$, the space of bounded integer-indexed sequences, equipped with the supremum norm $\|\cdot\|_\infty$. This is indeed a reasonable choice given that the upper and lower value functions are bounded by definition. The main statement is as follows:

Proposition 3.1. *The upper value function V^+ and the lower value function V^- are the unique solutions of (2) and (3) in $l^\infty(\mathbb{Z})$, respectively.*

We prove Proposition 3.1 with the help of two lemmas.

Lemma 3.2. *The upper Isaacs equation (2) and the lower Isaacs equation (3) are uniquely solvable in $l^\infty(\mathbb{Z})$.*

Proof. We only show that the upper Isaacs equation (2) admits a unique solution since the other result follows analogously. To this end, we apply Banach's fixed point theorem to a suitable mapping Φ .

Let $\Phi : l^\infty(\mathbb{Z}) \rightarrow l^\infty(\mathbb{Z})$ be defined by

$$\Phi(w) = \left(R(i) + \alpha \min_{b \in A} \max_{a \in A} \sum_{j \in \mathbb{Z}} p_{ij}(a, b) w(j) \right)_{i \in \mathbb{Z}}.$$

Note that any fixed point of Φ is a solution of (2) and vice versa. Then we have for $v, w \in l^\infty(\mathbb{Z})$ and any $i \in \mathbb{Z}$

$$\begin{aligned} \Phi(v)(i) - \Phi(w)(i) &= \alpha \left(\min_{b \in A} \max_{a \in A} \sum_{j \in \mathbb{Z}} p_{ij}(a, b) v(j) - \min_{b \in A} \max_{a \in A} \sum_{j \in \mathbb{Z}} p_{ij}(a, b) w(j) \right) \\ &\leq \alpha \max_{b \in A} \left(\max_{a \in A} \sum_{j \in \mathbb{Z}} p_{ij}(a, b) v(j) - \max_{a \in A} \sum_{j \in \mathbb{Z}} p_{ij}(a, b) w(j) \right) \\ &\leq \alpha \max_{b \in A} \max_{a \in A} \left(\sum_{j \in \mathbb{Z}} p_{ij}(a, b) v(j) - \sum_{j \in \mathbb{Z}} p_{ij}(a, b) w(j) \right) \\ &\leq \alpha \|v - w\|_\infty. \end{aligned}$$

By interchanging v and w we obtain $\Phi(w)(i) - \Phi(v)(i) \leq \alpha \|v - w\|_\infty$. Hence, Φ is a contraction and there exists an unique fixed point $v^* \in l^\infty(\mathbb{Z})$. \square

For $\pi, \sigma \in \mathcal{A}$ we define the linear operator $\mathcal{P}_{\pi, \sigma} : l^\infty(\mathbb{Z}) \rightarrow l^\infty(\mathbb{Z})$ by

$$(\mathcal{P}_{\pi, \sigma} v)(i) = \sum_{j \in \mathbb{Z}} p_{ij}(\pi, \sigma) v(j), \quad i \in \mathbb{Z}, \quad v \in l^\infty(\mathbb{Z}).$$

$\mathcal{P}_{\pi, \sigma}$ has the following properties:

Lemma 3.3. *Let $\pi, \sigma \in \mathcal{A}$. Then the linear operator $\mathcal{P}_{\pi, \sigma}$ satisfies*

(i) $\|\mathcal{P}_{\pi, \sigma}\| = 1$, and

(ii) $(\text{Id} - \alpha\mathcal{P}_{\pi, \sigma})$ is invertible with

$$(\text{Id} - \alpha\mathcal{P}_{\pi, \sigma})^{-1} = \sum_{n=0}^{\infty} \alpha^n \mathcal{P}_{\pi, \sigma}^n. \quad (14)$$

Proof. (i) is obvious. Note that (i) implies that $\|\alpha\mathcal{P}_{\pi, \sigma}\| = \alpha < 1$. Standard results from functional analysis (see e.g. [7], Theorem 7.3-1) yield that $(\text{Id} - \alpha\mathcal{P}_{\pi, \sigma})$ is invertible and (14) holds true. \square

Now, we are able to prove Proposition 3.1.

Proof of Proposition 3.1. We only prove that the upper value function V^+ solves the upper Isaacs equation (2). The statement for the lower value function V^- follows analogously. The proof is inspired by [14], Theorem 6.2.2.

Let $w \in l^\infty(\mathbb{Z})$ be a solution of (2). We prove that $w = V^+$.

Step 1: We show $w \geq V^+$. There exists $\sigma^* \in \mathcal{A}$ such that for $i \in \mathbb{Z}$

$$\begin{aligned} w(i) &= R(i) + \alpha \min_{b \in A} \max_{a \in A} \sum_{j \in \mathbb{Z}} p_{ij}(a, b) w(j) \\ &= R(i) + \alpha \max_{a \in A} \sum_{j \in \mathbb{Z}} p_{ij}(a, \sigma^*) w(j) \\ &\geq R(i) + \alpha \sum_{j \in \mathbb{Z}} p_{ij}(\pi, \sigma^*) w(j), \end{aligned}$$

for some $\pi \in \mathcal{A}$. Therefore, $(\text{Id} - \alpha\mathcal{P}_{\pi, \sigma^*})w \geq R$. Lemma 6.1.2 in [14] implies

$$w(i) \geq (\text{Id} - \alpha\mathcal{P}_{\pi, \sigma^*})^{-1} R(i) = \sum_{n=0}^{\infty} \alpha^n (\mathcal{P}_{\pi, \sigma^*}^n R)(i) = J(i, \pi, \sigma^*), \quad i \in \mathbb{Z}.$$

Since π was chosen arbitrarily we obtain

$$w(i) \geq \sup_{\pi \in \mathcal{A}} J(i, \pi, \sigma^*) \geq V^+(i), \quad i \in \mathbb{Z}.$$

Step 2: We show $w \leq V^+$. Let $\sigma \in \mathcal{A}$. Then there exists $\pi^* \in \mathcal{A}$ such that for all $i \in \mathbb{Z}$

$$\begin{aligned} w(i) &= R(i) + \alpha \min_{b \in A} \max_{a \in A} \sum_{j \in \mathbb{Z}} p_{ij}(a, b) w(j) \\ &\leq R(i, a, \sigma) + \alpha \max_{a \in A} \sum_{j \in \mathbb{Z}} p_{ij}(a, \sigma) w(j) \\ &= R(i) + \alpha \sum_{j \in \mathbb{Z}} p_{ij}(\pi^*, \sigma) w(j) \\ &= R(i) + \alpha (\mathcal{P}_{\pi^*, \sigma} w)(i) \end{aligned}$$

Iterating this inequality n -times we see that for $i \in \mathbb{Z}$

$$w(i) \leq \sum_{k=0}^n \alpha^k (\mathcal{P}_{\pi^*, \sigma}^k R)(i) = J(i, \pi^*, \sigma) - \sum_{k=n+1}^{\infty} \alpha^k (\mathcal{P}_{\pi^*, \sigma}^k R)(i).$$

Because R is bounded by 1 and $\|\mathcal{P}_{\pi^*, \sigma}\| = 1$ we have

$$- \sum_{k=n+1}^{\infty} \alpha^k (\mathcal{P}_{\pi^*, \sigma}^k R)(i) \leq \frac{\alpha^{n+1}}{1 - \alpha}, \quad i \in \mathbb{Z}.$$

Now, let $\varepsilon > 0$. Because $\alpha \in (0, 1)$ we have for n sufficiently large

$$w(i) \leq J(i, \pi^*, \sigma) + \varepsilon \leq \sup_{\pi \in \mathcal{A}} J(i, \pi, \sigma) + \varepsilon.$$

Since $\varepsilon > 0$ and $\sigma \in \mathcal{A}$ were chosen arbitrarily we obtain $w \leq V^+$. □

Acknowledgement

Support from the German Research Foundation through the project AN 1024/5-1 is gratefully acknowledged.

References

- [1] S. Ankirchner, N. Kazi-Tani, J. Wendt, and C. Zhou. Large ranking games with diffusion control. working paper or preprint, Nov. 2021.
- [2] T. Başar and G. J. Olsder. *Dynamic noncooperative game theory*. Academic Press, Ltd., London, second edition, 1995.
- [3] A. Federgruen. On N -person stochastic games with denumerable state space. *Advances in Applied Probability*, 10(2):452–471, 1978.
- [4] J. Filar and K. Vrieze. *Competitive Markov decision processes*. Springer-Verlag, New York, 1997.
- [5] A. M. Fink. Equilibrium in a stochastic n -person game. *Journal of Science of the Hiroshima University. Series A-I. Mathematics*, 28:89–93, 1964.
- [6] D. Fudenberg and J. Tirole. *Game theory*. MIT Press, Cambridge, MA, 1991.
- [7] E. Kreyszig. *Introductory functional analysis with applications*. John Wiley & Sons, New York-London-Sydney, 1978.
- [8] R. Laraki, J. Renault, and S. Sorin. *Mathematical foundations of game theory*. Universitext. Springer, Cham, 2019.

- [9] A. Maitra and T. Parthasarathy. On stochastic games. *Journal of Optimization Theory and Applications*, 5:289–300, 1970.
- [10] J. M. McNamara. A stochastic differential game with safe and risky choices. *Probability in the Engineering and Informational Sciences*, 2(1):31–39, 1988.
- [11] J. F. Nash, Jr. Equilibrium points in n -person games. *Proceedings of the National Academy of Sciences of the United States of America*, 36:48–49, 1950.
- [12] J. F. Nash, Jr. Non-cooperative games. *Annals of Mathematics. Second Series*, 54:286–295, 1951.
- [13] T. Parthasarathy. Discounted, positive, and noncooperative stochastic games. *International Journal of Game Theory*, 2(1):25–37, 1973.
- [14] M. L. Puterman. *Markov decision processes: discrete stochastic dynamic programming*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons, Inc., New York, 1994. A Wiley-Interscience Publication.
- [15] P. D. Rogers. *NONZERO-SUM STOCHASTIC GAMES*. PhD thesis, 1969. University of California, Berkeley.
- [16] S. Ross. *Introduction to stochastic dynamic programming*. Probability and Mathematical Statistics. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York, 1983.
- [17] L. S. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences of the United States of America*, 39:1095–1100, 1953.
- [18] M. Sion. On general minimax theorems. *Pacific Journal of Mathematics*, 8:171–176, 1958.
- [19] M. J. Sobel. Noncooperative stochastic games. *Annals of Mathematical Statistics*, 42:1930–1935, 1971.
- [20] M. Takahashi. Stochastic games with infinitely many strategies. *Journal of Science of the Hiroshima University. Series A-I. Mathematics*, 26:123–134, 1962.
- [21] M. Takahashi. Equilibrium points of stochastic non-cooperative n -person games. *Journal of Science of the Hiroshima University. Series A-I. Mathematics*, 28:95–99, 1964.
- [22] J. von Neumann. Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen*, 100(1):295–320, 1928.
- [23] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, New Jersey, 1944.