Counteracting Erratic Behaviour: A Game-Theoretical Analysis of Parametrised Strategies

High-School Thesis
(Maturaarbeit)

Mathematical-Computational
Analysis

Author:
Adrian Hossner

Supervisor:	Dr. Geoffrey Ostrin
Submission Date:	September 7, 2025
School:	Gymnasium Thun
Word Count (Sections 1-6):	8000+

Abstract

This high-school thesis addresses the politically and socially relevant question of how to best counteract erratic behaviour. To this end, game theory is used as a theoretical framework, in particular the Prisoner's Dilemma game. In its extended version of the iterated continuous Prisoner's Dilemma, two persons repeatedly play against each other and win or lose points depending on how they mutually treated each other in the previous round. A discrete and a continuous random strategy are chosen as erratic behaviours, which compete against each other as well as against a rigid strategy and a discrete and a continuous version of an adaptive strategy. All strategies are parametrised (from 0 to 10) in terms of the rigour with which the respective strategy is pursued. The results are presented in the form of individual absolute- and relative-gain plots as well as an overall-gain plot that reflects the benefit achieved jointly by both players. On the one hand, the mathematical-computational analysis reveals surprising details, which are discussed in detail under the headings of shallow randomness, dispersion and risk, threshold values for discrete adaptations, and dynamics of continuous adaptations. On the other hand, when examining the interactions in terms of gains accumulated against randomness, no strong strategy could be identified that successfully counteracts unpredictable behaviour while maximising relative- and overall-gain values simultaneously. In the case of encountering erratic behaviour, the summarising recommendation is that, if one's own advantage is emphasised, cooperation should be completely rejected, while some cooperation should be offered if one is dealing with an at least partially rational interaction partner.

Preface

The present thesis has been written in times of global uncertainty. At least partially, this is ascribed to a more or less erratic behaviour of political leaders—as a prime example the 47$^{\text{th}}$ President of the U.S.A. Donald J. Trump. When discussing this issue among my family and friends, a water-polo team member of mine drew my attention to the fact that this problem of erratic behaviour could be mathematically modelled by means of game theory. Of particular interest in this context would be the so-called 'Prisoner's Dilemma'. This game has been popularised in the 'A Beautiful Mind' film that narrates a story about the famous mathematician John Forbes Nash. The opportunity of combining my politically-social concerns with my subject-related interests in Computer Science and Mathematics captivated me immediately. Hence, I was more than delighted that my school teacher Mr Ostrin accepted my request to serve as a supervisor of a related high-school thesis.

Since then, Mr Ostrin has been always approachable and an extremely valuable supporter of my work. Likewise, I am grateful for all the help I have received from my parents, be it as inspiring discussants, critical proofreaders, or supportive coaches when I was temporarily stuck in problems. Many thanks from the bottom of my heart!

Table of Contents

1 Introduction

On August 9-10, 2025, the 'Süddeutsche Zeitung' (p. 20) reported on massive irritations of the incumbent Federal President of Switzerland Karin Keller-Sutter, stemming from the ongoing tariff dispute with the United States of America. In this context, the U.S. President's unpredictable character is perceived as particularly bothersome. Consequently, the question arises how to successfully counteract such erratic behaviour on the political stage.

Game theory offers a powerful framework for studying human decision making, for the issue sketched above, especially in the form of the Prisoner's Dilemma (PD) game. When translating 'erratic behaviour' as pursuing a random PD strategy, the resulting leading question of the present thesis would aim at identifying strategies that prove to be effective countermeasures against randomness. In addition, it seems interesting to ask whether gradations of erratic behaviour matter, meaning that a game-theoretical computational analysis would be required to allow parametrisation. Finally—and in sharp contrast to an 'America First' ideology—it appears worthwhile to investigate not only superiority over opponents but the overall gain collectively achieved as well.

The leading question derived in the present introduction (Section 1) will be explored below in five consecutive steps: After relevant theoretical foundations have been laid (Section 2), applied methods and implementational specifications of the computational analysis will be presented (Section 3). Subsequently, the obtained results will be described (Section 4) and discussed (Section 5). The thesis will be rounded off with a conclusion and outlook that refers back to the leading question just formulated (Section 6) and will be supplemented by a source list (Section 7), an appendix (A) and a declaration of integrity (B).

2 Theoretical Foundations

To grasp the game-theoretical approach pursued in this thesis, it is necessary to first understand the underlying key concepts. The most important ones, being the PD (2.1), the introduction of iterations as an extension (2.2) and the approach of submitting continuous values (2.3). Finally, further PD variants will be mentioned (2.4) and relevant information of published researches will be provided (2.5). Excellent overviews are available in standard encyclopedia, from which the information reported in 2.1–2.4 was compiled, in particular the *International Encyclopedia of the Social and Behavioral Sciences* [1] and the *Stanford Encyclopedia of Philosophy* [2].

2.1 Prisoner's Dilemma

The PD is a famous model in game theory. To understand the dilemma, a proper context has to be established: Two persons are arrested and sued for having committed a crime jointly. There is yet, however, insufficient evidence to imprison the two suspects. Despite not having enough prove, the police interrogates them separately without giving them a chance to coordinate beforehand. The suspects receive two options. On the one hand, they can confess that the other suspect was involved in the crime. On the other hand, they have the right to remain silent. In case mutual trust and thus no confession took place, both separately receive an imprisonment of three years. If both confess, either obtains ten years in prison. However, if one chooses to confess while the other remains silent, the informer is sentenced to only one year in prison, whereas the betrayed is punished with 25 years. Remaining silent is considered as cooperation, whereby defection is the opposite, meaning breaking confidence. The following matrix visualises the payoffs in this game in a comprehensive way, with C representing cooperation and D corresponding to defection:

	C	D
C	3, 3	25, 1
D	1, 25	10, 10

The dilemma consists of the following: The act of defecting is tempting since there is a chance of avoiding a long sentence. However, it is then also possible to receive a punishment of ten years in prison depending on the confederate's decision. Nevertheless, this outcome—which could have been avoided if both had remained silent and thus obtained a reward of three easy surmountable years in confinement—is the most likely result, as the temptation to defect is simply too great. Additionally, the possibility of becoming the sucker in this game forces the player to consider trust as the deciding factor. As trust is bound to be unreliable, defection is the more reasonable choice. The PD thus composes a one-time interaction from a mathematical and analytical perspective. In a more general form, the variables can be described as Reward (R), Sucker's pay-off (S), Temptation (T), and Punishment (P) as follows:

	C	D
C	R, R	S, T
D	T, S	P, P

The pay-offs in the matrix can be altered as long as the following inequalities stay valid:

$$T < R < P < S \tag{1}$$

2.2 Iterated Prisoner's Dilemma

The Iterated Prisoner's Dilemma (IPD) extends the PD as the game is played multiple times sequentially. To refer to the number of rounds played, n is used as the denotation. The gained pay-offs are accumulated over the last $n - k$ rounds where k represents the current round. A player's strategy can be based on the entire history of the previous PDs to generate the decision for the next round. These decisions can be calculated using conditions and probabilities that are applied to the given data. The prison analogy is changed from minimising years in prison to maximising one's points. Thus, the pay-off matrix must be altered accordingly. In particular, the following inequalities need to be satisfied:

$$T > R > P > S \tag{2}$$

An additional inequality is necessary to avoid receiving greater pay-offs by coordinated alternation of cooperation and defection ($T + S$) than continued mutual cooperation ($2R$):

$$2R > T + S \tag{3}$$

In the IPD, mutual cooperation yields long-term benefits since playing cooperation increases trust. The maximum pay-off, however, can only be achieved by exploitation ($p_{\max} = nT$). Because persistent defection leads to mutual defection as a response, the attainment of this maximum value is highly unlikely.

2.3 Iterated Continuous Prisoner's Dilemma

In the Iterated Continuous Prisoner's Dilemma (ICPD), strategies are not limited to either cooperate or defect. Instead, the players can choose any level of cooperation between 0 (defect) and 1 (cooperate). This cooperation level is called investment or contribution. The continuous form offers enhanced precision and mathematical applicability.

To include continuous values in the ICPD, a mathematical procedure can be applied to the discrete version of the matrix, namely bilinear interpolation. This will establish a field that can be represented by colour grades, a heat map as depicted in Figure 1.

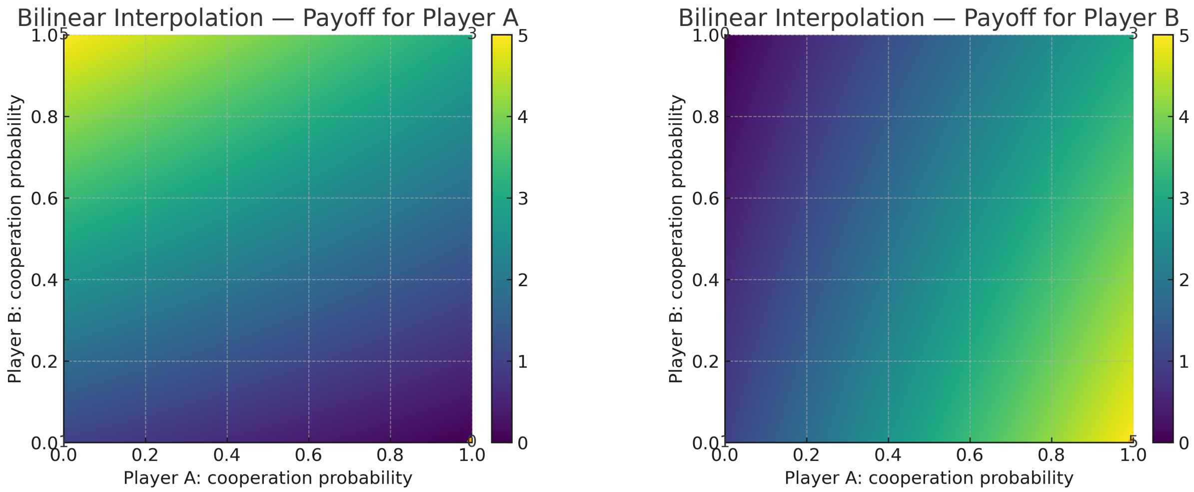


Figure 1: *Pay-off heat maps in the ICPD for player A (left) and player B (right).*

For reason of simplifying analysability, this pay-off system can be expressed in the following equations [cf. 3, p. 259]:

$$p_A = y - cx + c \tag{4}$$

$$p_B = x - cy + c \tag{5}$$

In these equations, p_A and p_B are the pay-offs of strategy A and B, respectively. x is the investment of strategy A and y is that of strategy B. The coefficient c ranges, as well as x and y, from 0 to 1 (inclusively) and describes the cost of cooperation. The pay-off is highly affected by the opponent's investment. Only a fraction of one's own contribution is subtracted, c being the coefficient. Consequently, the minimum pay-off would then be $-c$. However, to avoid negative pay-offs and enhance comprehensibility of results achieved, a third addend is appended in the present thesis, namely c. So, the least points that can be gained is 0 while the maximal points that can be earned by exploitation is $1 + c$.

2.4 Further Variants

There exist several extensions of the PD, one of which is noise, meaning that investments are altered with a certain probability. This noise is introduced to simulate potential misunderstandings. Another advancement is the alternating form. In this variant, investments are contributed by the players in succession rather than concurrently as in the conventional simultaneous model. Finally, the evolutionary PD game surveys the success of strategies in a world where natural selection takes place. This implies that losing strategies vanish, while successful ones replicate and ultimately survive a PD tournament. However, since the ICPD—as introduced above—offers a promising and apparently sufficient framework for approaching the leading question, those further variants can be disregarded in the present thesis.

2.5 Scientific Findings

In the context of the PD, erratic behaviour might be characterised best as 'random strategy'. To retrieve the most important scientific findings on this topic, a literature search was conducted on `scholar.google.com` with the search terms 'Prisoner's Dilemma' and 'random strategy' (for details see Appendix A). A criterion-based process of stepwise reduction of the initially found 650 hits finally resulted in eight publications of particular relevance.

Interestingly, a random strategy was already included in the first conducted PD tournament organised by Axelrod and Hamilton [4]. Their main finding is that a simple Tit-for-Tat strategy (TFT: "start with a C, and then use your co-players previous move", [5]) won over all other strategies. In his review on PD studies, Brembs [6, p. 14] highlights as one of the "most important advancements in the game-theoretical work on different aspects of the game" the turn from deterministic to stochastic variants. In this context, Nowak and Sigmund [5, p. 56] argue that "occasional mistakes between two TFT players cause long runs of mutual backbiting." They thus introduce a so-called Pavlov strategy, which "cooperates if and only if both players opted for the same alternative in the previous move (...). Thus Pavlov (...) can correct mistakes." Building on this, Kraines and Kraines [7, p. 112] demonstrate that "Pavlov will exploit (...) random strategies" in the discrete IPD game. Even further optimisations were suggested by Hauert and Stenull [8] in terms of adaptive strategies and by Hao, Li and Zhou [9] for ways of determining the opponent's maximum and minimum gain. Jurišić, Kermek and Konecki [10] report in their review on repetitions of Axelrod's tournament in 2004 and 2005 on a successful Omega TFT strategy that always defects as soon as a certain randomness threshold is crossed by the opponent. Finally, Mathieu and Delahaye [11, pp. 4, 5] introduce new strategies that perform better than probabilistic or random strategies, namely soft_majo ("begins by cooperating and cooperates as long as the number of times the opponent has cooperated is greater that or equal to the number of times it has defected") and gradual ("cooperates on the first move, then defect n times after n^{th} defections of its opponent, and calms down with 2 cooperations").

Despite the scientific progress that becomes obvious with this short summary, it should be first noted that all of the retrieved articles refer to an overall success in a tournament with multiple strategies rather than a specific advantage over a random opponent. Second, success is understood as outperforming the opponent without also considering the overall gain of both players. Third, the discrete variant of the IPD is used, which does not include certain gradations of randomness. The last-mentioned point is strikingly corroborated by the fact that an additionally conducted scholar.google.com search with the terms 'continuous Prisoner's Dilemma' and 'random strategy' delivered three hits only, all of which are irrelevant for the present topic [12–14]. It seems thus fair to conclude that a mathematical simulation of a competition of continuously parametrised strategies against strategies of graded randomness with a particular focus also on both players' overall gain is worth pursuing further in the present thesis.

3 Methods and Implementation

To answer the leading question, it is necessary to develop an overarching study design (3.1), to specify the parametrised strategies that will be pursued in the ICPD (3.2), and to determine details of the implementation in terms of IT (3.3).

3.1 Study Design

Erratic behaviour with respect to the PD can be equated to a random conduct. Two strategy types are proposed by reason to counter such erraticism: a strategy remains unaffected and beholds its manner of responding and a strategy that actively reacts to the randomness. In the former case, strategies are identified as rigid whereas the latter are characterised as adaptive.

Behavioural categories can be algorithmised differently. This particularly concerns the above introduced distinction of the discrete and the continuous form of the IPD. Applied to the erratic strategies in their continuous form, this means that the investment is a random number between 0 and 1. In contrast, the discrete version of this strategy generates a random number that is then rounded to the nearest integer. The same approach applies to the adaptive strategies. Finally, discrete contributions of rigid strategies are equal to either always cooperating (AlwaysCooperate) or constantly defecting (AlwaysDefect) whereas continuous versions of the rigid category contribute continuous investments invariably (e.g. 0.5; AlwaysNeutral).

Furthermore, it is possible to specify the extremeness of each behaviour. This means that the randomness of an erratic strategy, the adaptiveness of an adaptive strategy and the constant investment of a rigid strategy can be specified. On this account, a parameter (θ) is incorporated into the determination of the investment and the evaluation of various conditions (which—to the best of the authors knowledge—has not been undertaken in PD research so far).

To keep the number of comparisons manageable within the realms of a high-school thesis, five strategies were finally chosen: a discrete and a continuous version of the random category (Random-Discrete, Random-Continuous), as well as of the adaptive category (Adapt-Discrete, Adapt-Continuous), complemented by a rigid strategy (Always-Same) in which the extreme values of the parameter correspond to the discrete form and the in-between parametrisations implement the continuous form. According to the leading question, the two random strategies will play against all strategies, including themselves.

3.2 Implemented Strategies

As justified before, five strategies, namely Random-Discrete, Random-Continuous, Always-Same, Adapt-Discrete, and Adapt-Continuous will be presented and specified in the following paragraphs. The implemented parameter takes an integer value ranging from 0 to 10 (inclusively). As the discrete variant of the adaptive strategy is derived from its continuous variant, the continuous implementation will be described first. Special features of the parametrisations are summarised in advance in Table 1.

Table 1: *Implemented strategies, illustrated by parameter values meanings (TFT = Tit-for-Tat).*

	$\theta = 0$	$\theta = 5$	$\theta = 10$
Random-Discrete	AlwaysDefect	(50% 0.00|1.00)	AlwaysCooperate
Random-Continuous	AlwaysNeutral	(50% 0.25|0.75)	(50% 0.00|1.00)
Always-Same	AlwaysDefect	AlwaysNeutral	AlwaysCooperate
Adapt-Discrete	AlwaysCooperate	(rounded TFT)	(rounded overshot TFT)
Adapt-Continuous	AlwaysNeutral	TFT	(overshot TFT)

Random-Discrete

Random-Discrete (RD) is a strategy in the group of random. It is also discrete, meaning it can only submit either 0 or 1. The parameter in this strategy determines the likelihood of submitting 0 and 1, respectively. The probability (Pr) of submitting either full cooperation or full defection is described in the following equations:

$$\Pr(i = 1) = \frac{1}{10}\theta_{\mathrm{RD}} \tag{6}$$

$$\Pr(i = 0) = 1 - \frac{1}{10}\theta_{\mathrm{RD}} \tag{7}$$

Please note that parameter 0 is equivalent to AlwaysDefect since the probability of submitting 0 is 0%. Parameter 10, on the contrary, is equivalent to AlwaysCooperate since the probability of submitting 1 is 100%. Parameter 5 thus means that the two possible investments equally occur in random order.

Random-Continuous

The other strategy in the random group is Random-Continuous (RC), meaning it can additionally submit any number between 0 and 1. The investment of this strategy is anchored around the value of 0.5. The parameter specifies the shift (s) upwards or downwards from that anchor depending on ϵ, as calculated by the following equations:

$$i(\theta_{\mathrm{RC}}) = 0.5 + \epsilon \cdot s(\theta_{\mathrm{RC}}) \tag{8}$$

$$s(\theta_{\mathrm{RC}}) = \frac{1}{20}\theta_{\mathrm{RC}} \tag{9}$$

$$\Pr(\epsilon = 1) = \Pr(\epsilon = -1) = \frac{1}{2} \tag{10}$$

Parameter 0 results in a behaviour corresponding to AlwaysNeutral. As the maximum value of the parameter is 10, due to the division by 20, the maximum shift cannot exceed the limits of 0 and 1. The probability of subtracting or adding this shift is 50%. Consequently, parameter 10 leads to contributing a discrete value of either 0 or 1 with a chance of 50%.

Always-Same

Always-Same (AS) calculates the investment as a function of the parameter according to the following equation:

$$i(\theta_{\mathrm{AS}}) = \frac{1}{10}\theta_{\mathrm{AS}} \tag{11}$$

This means that after each incrementation of the parameter, the investment increases by 0.1. Consequently, parameter 0 is equivalent to AlwaysDefect and parameter 10 to AlwaysCooperate. If the parameter is set to 5, the strategy will always submit the investment 0.5 which is identical to AlwaysNeutral.

Adapt-Continuous

Adapt-Continuous (AC) starts with full cooperation to offer a constructive relationship as recommended in fundamental PD literature [1, 2]. After the first round, it will adapt to the submissions of the other player by shifting its own investment towards the opponent's on the basis of the following equations:

$$i_0 = 1 \tag{12}$$

$$i(\theta_{\mathrm{AC}}, k) = i_{k-1} + s(\theta_{\mathrm{AC}}, k) \tag{13}$$

$$s(\theta_{\mathrm{AC}}, k) = \frac{1}{5}\theta_{\mathrm{AC}} \cdot (\tilde{i}_{k-1} - i_{k-1}) \tag{14}$$

Please note that k indicates the current round and is used to notate the equations recursively. k only appears in the adaptive strategies as only they require information about the past rounds. The shift being applied to one's own previous investment is defined by the function $s(\theta_{\mathrm{AC}})$. Parameter 0 is identical to AlwaysCooperate since the difference and thus the shift is multiplied by 0. Dividing θ_{AC} by 5 implies that this coefficient to the difference is 1 if the parameter equals 5. This means that the strategy will shift its next investment to exactly the previous investment of the opponent. This behaviour corresponds to Tit-for-Tat (cf. 2.5). Parameter 10 means that the strategy will add the shift to its own previous investment twice, which comes down to an overshooting Tit-for-Tat. However, since strategies are not allowed to submit any investments exceeding the limits of 0 to 1, the implementation of this strategy will simply set its investment to the reached limit if surpassed.

Adapt-Discrete

Adapt-Discrete (AD) is insofar derived from Adapt-Continuous as the same investment function is taken and the results simply rounded to the nearest integer.

$$i_0 = 1 \tag{15}$$

$$i(\theta_{\mathrm{AD}}, k) = \begin{cases} 1 & \text{if } i_{k-1} + s(\theta_{\mathrm{AD}}, k) \geq 0.5 \\ 0 & \text{if } i_{k-1} + s(\theta_{\mathrm{AD}}, k) < 0.5 \end{cases} \tag{16}$$

$$s(\theta_{\mathrm{AD}}, k) = \frac{1}{5}\theta_{\mathrm{AD}} \cdot (\tilde{i}_{k-1} - i_{k-1}) \tag{17}$$

The strategy will start with full cooperation. In the following rounds, the submitted investment will be 1 if the calculated investment is greater or equal to 0.5, otherwise it will be equal to 0. Parameter 0 thus means that the strategy corresponds to AlwaysCooperate because the shift is equal to 0 and thus always stays at its first investment, being full cooperation. If the parameter is equal to 10, due to the rounding of the overshot Tit-for-Tat, the investment is most likely to change from 1 to 0 or vice versa.

3.3 Implementational Details

For the pay-off system of the ICPD, a value of 0.5 has been chosen for c (see Equations 4 and 5 in 2.3). Regarding the leading question of this thesis, both Random-Discrete and Random-Continuous play against every strategy (including themselves). As there are eleven parameter values for each strategy, there will be $2 \cdot 5 \cdot 11 \cdot 11 = 1210$ encounters and thus comparisons. 20 iterations of the ICPD has been decided to suffice. In addition, this game of 20 iterations is played a hundred times to flatten random peaks or valleys. This comes down to $100 \cdot 20 \cdot 1210 = 2420000$ executions of the continuous PD (CPD). Therefore, only the means of the 100 repetitions will be plotted.

One encounter of the $11 \cdot 11$ parameter-based ICPD (pbICPD) will be displayed in form of a surface. There are two surfaces for each pbICPD, both of which represent the gained points of the corresponding strategies (absolute-gain). Additionally, two further surfaces can be established. These ones are calculated by subtracting the absolute-gain values and thus depict the advantage of strategy A over strategy B and vice versa (relative-gain). Finally, a fifth surface will be generated by adding the two absolute-gain surfaces, which represents the points both players gained together (overall-gain).

The simulation, data generation and visualisation is completely written in *Python* (Version 3.13.5). For generating the surface plots, a *Python* library was used, namely *Plotly*. The *Python* scripts and data visualisation files can be found here: `https://github.com/adho08/Prisoner-s-Dilemma`

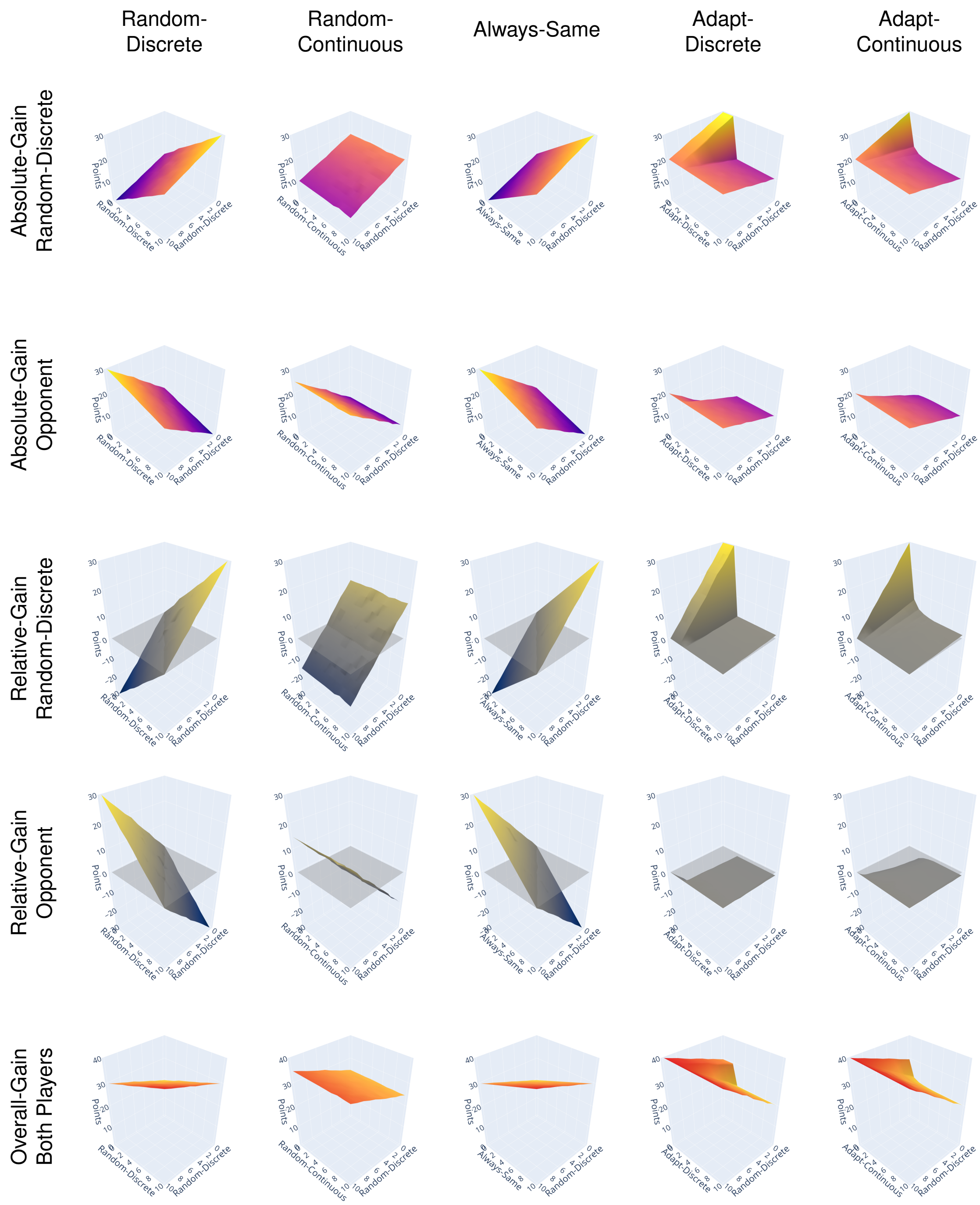
4 Results

In Figures 2 and 3, the surfaces for the Random-Discrete and Random-Continuous simulations are shown, respectively. There are five surfaces for each interaction that are generated from one pbICPD. The x-axis—the horizontal axis at the bottom right of each plot—displays the parameter value of the main strategy, i.e., either Random-Discrete (Figure 2) or Random-Continuous (Figure 3). The y-axis—the horizontal axis at the bottom left of each plot—indicates the parameter value of the opponent's strategy. The z-axis always exhibits points gained in the ICPD. These dependent values are not only marked by height in the plots but also by colour. It should be noted that the two horizontal axes are scaled in such a way that the minimum parameter value of 0 appears at the rear end of the plot and its maximum of 10 at its front end, meaning that the left axis needs to be read 'from the left to the right' and the right axis 'from the right to the left', respectively. To help the readers to relate these values to meaningful content, at the bottom of Figures 2 and 3, the explanations provided in Table 1 are repeated.

Within each column, the first two rows are the most substantial surface plots. They illustrate the absolute gains of the random strategy itself and then of the opponent's strategy. The opponent's strategy is specified in the five columns of the figures. The maximum of absolutely gained points can be calculated from the pay-off Equations 4 and 5; they are gained by exploitation. So, $p_A = y - c \cdot x + c$ becomes by substitution $p_A = 1 - 0.5 \cdot 0 + 0.5 \rightarrow p_A = 1.5$. And since the CPD is played 20 times, $p_A \cdot 20 = 30$. Consequently, the boundaries in which the points can vary are 0 and 30. All the points in an absolute-gain surface with a value of 30 have their correspondent x-y-coordinate of 0 in the opponent's surface since to receive 30 points, exploitation is required. The exact same principle holds for the minimum value, merely reversed since $p_B = 0 - 0.5 \cdot 1 + 0.5 \rightarrow p_B = 0$. Mutual cooperation results in both players receiving a pay-off of 1 point because if y and x are equal to 1, $p_A = p_B = 1 - 0.5 \cdot 1 + 0.5 \rightarrow p_A = p_B = 1$. Thus, if the players constantly mutually cooperate, both receive an end result of 20. However, the z-coordinate of 20 does not guarantee mutual cooperation because the CPD is not a zero-sum game. Finally, mutual defection leads to both receiving 0.5 points as $p_A = p_B = 0 - 0.5 \cdot 0 + 0.5 \rightarrow p_A = p_B = 0.5$. Therefore, if both decide to defect constantly, the resulting absolute-gain is 10.

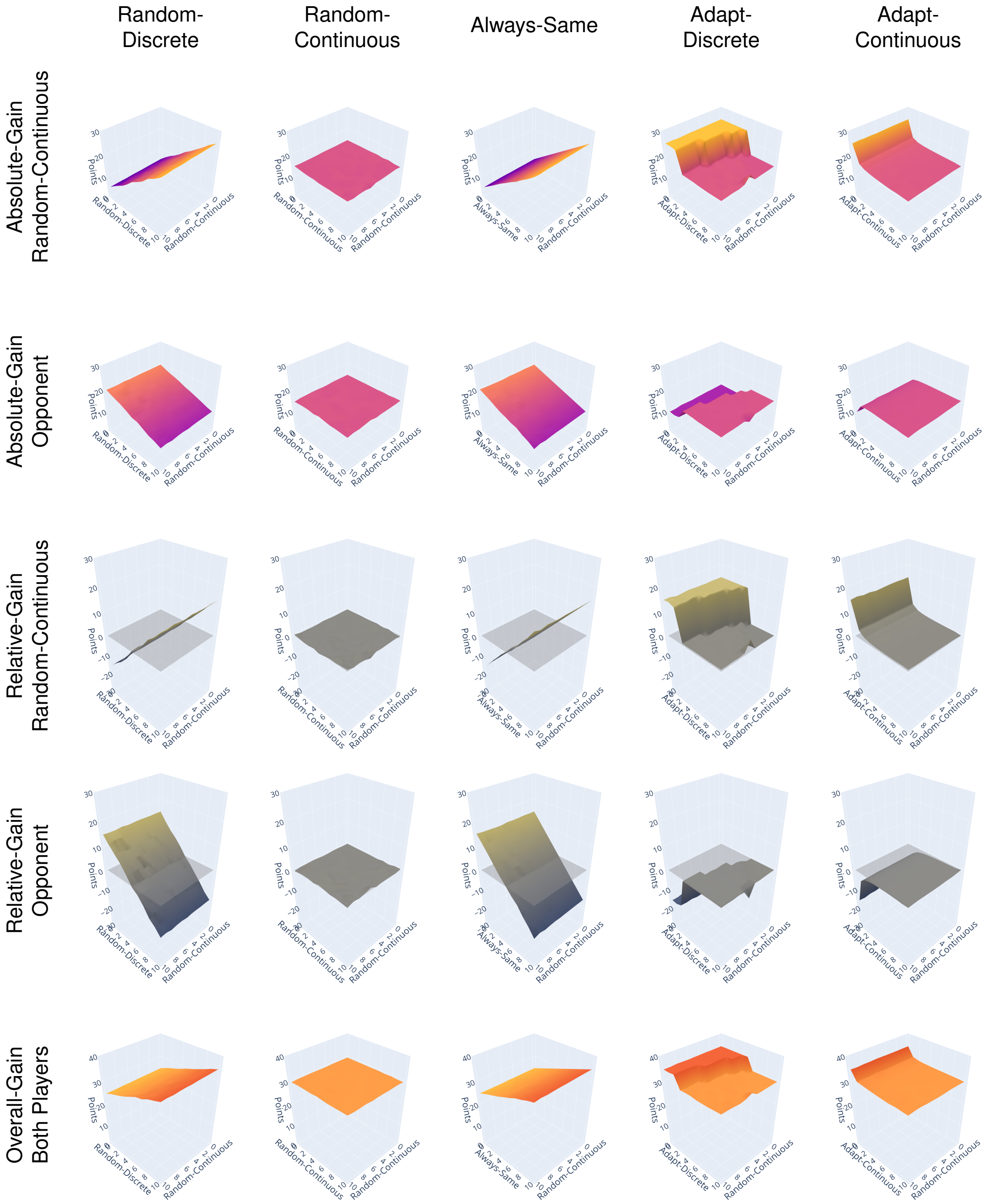
The next two surface plots within a column concern the relative gain, that means the advantage of one strategy over the other. For the surface of the main strategy, the opponent's absolute gain surface is subtracted from its own absolute-gain surface and vice versa. Consequently, if the resulting surface plot happens to be entirely above the zero-plane, it means that the displayed strategy has won in every single game of the pbICPD regardless of the parameters. The maximum and the minimum values of the resulting relative-gain surfaces can be calculated from the difference of the maximum and the minimum values of the first two surfaces. Accordingly, the points can vary between -30 and 30. Since the two absolute-gain surfaces would intersect if they were put into the same coordinate system, the intersection points can be found in the relative-gain plots where the surfaces intersect with the zero-plane because the difference of two same numbers is always equal to zero. Finally, it should be noted that the two relative-gain surfaces are always mirrored by the zero-plane since $A - B = (B - A) \cdot (-1)$.

The last row of plots displays how many points have been gained by both opponents together. To generate these surfaces for the overall-gain, one must simply add the values from the two absolute-gain plots. Due to the facts that both opponents together gain the most points if both constantly cooperate, namely 1 point per round, the maximum value for the overall-gain surface is equal to $2 \cdot 20 \cdot 1 = 40$. It should be noticed that maximum values in the relative-gain surfaces do not necessarily lead to maximum values in the overall-gain surface due to potential losses of the opponent. Therefore, the overall gain should be understood as a marker for population welfare, since the higher the points, the most has been gained by both players together.



<table>
<tr><th></th><th>θ = 0</th><th>θ = 5</th><th>θ = 10</th></tr>
<tr><td>Random-Discrete</td><td>AlwaysDefect</td><td>(50% 0.00|1.00)</td><td>AlwaysCooperate</td></tr>
<tr><td>Random-Continuous</td><td>AlwaysNeutral</td><td>(50% 0.25|0.75)</td><td>(50% 0.00|1.00)</td></tr>
<tr><td>Always-Same</td><td>AlwaysDefect</td><td>AlwaysNeutral</td><td>AlwaysCooperate</td></tr>
<tr><td>Adapt-Discrete</td><td>AlwaysCooperate</td><td>(rounded TFT)</td><td>(rounded overshot TFT)</td></tr>
<tr><td>Adapt-Continuous</td><td>AlwaysNeutral</td><td>TFT</td><td>(overshot TFT)</td></tr>
</table>

Figure 2: *pbICPD surface plots for Random-Discrete. Absolute-, Relative-, and Overall-Gain values (rows) are shown for each implemented strategy (columns). Parameter value meanings for each strategy are repeated to enhance comprehensibility (Table 1).*



	θ = 0	θ = 5	θ = 10
Random-Discrete	AlwaysDefect	(50% 0.00|1.00)	AlwaysCooperate
Random-Continuous	AlwaysNeutral	(50% 0.25|0.75)	(50% 0.00|1.00)
Always-Same	AlwaysDefect	AlwaysNeutral	AlwaysCooperate
Adapt-Discrete	AlwaysCooperate	(rounded TFT)	(rounded overshot TFT)
Adapt-Continuous	AlwaysNeutral	TFT	(overshot TFT)

Figure 3: *pbICPD surface plots for Random-Continuous. Absolute-, Relative-, and Overall-Gain values (rows) are shown for each implemented strategy (columns). Parameter value meanings for each strategy are repeated to enhance comprehensibility (Table 1).*

In the following two subsections (4.1, 4.2), the results of each column of each figure, meaning the five surface plots, will be described. The description follows a certain patter: First, the two absolute-gain, second, both relative-gain, and third, the overall-gain surfaces are explained in more detail.

4.1 Random-Discrete

Random-Discrete

The left (10, 0) and right (0, 10) corner are the extremes of the surface. Random-Discrete is exploited if its parameter equals 10, meaning it corresponds to AlwaysCooperate, and the parameter of Random-Discrete as an opponent is 0, i.e., equivalent to AlwaysDefect. The rearmost corner at the coordinate (0, 0) is where both players defect constantly, which relates to the already clarified fact that this point is located at a height of 10 for both players. When both parameters equal 10, both points have a z-component of 20. This indicates that both strategies behave equally, i.e., as AlwaysCooperate, which leads to a gain of 20 points separately.

The peaks of both relative-gain surfaces derive from the maximum outputs of the pay-off system. Since one can get the maximum output only if the other is exploited, the relative-gain surfaces entail the extreme outcomes of the ICPD. When the two parameters are identical, however, no strategy wins.

The overall-gain surface clearly indicates that the peak of the surface is at (10, 10). This point has the quality that both strategies are equivalent to AlwaysCooperate. The rearmost point, being the lowest, results from the fact that both defected constantly. So, for the welfare of the population, it is best if both always cooperate. In sum, this is the standard outcome of the PD game.

Random-Continuous

The absolute-gain surfaces concerning the interaction of Random-Discrete and Random-Continuous continuously range from 20 to 10 and from 5 to 25, respectively. Both surfaces are not influenced by the parameter of Random-Continuous.

The two relative-gain surfaces, show that Random-Discrete wins if the parameter values are lower than 5, which corresponds to defecting with a higher probability than 50%. For the parameter values higher than 5, i.e., cooperating more than defecting, the contrary is true. The relative-gain surfaces range continuously from -15 to 15.

The overall-gain surface shows a constant slope with points ranging from 35 to 25. The maximum value of 40 is not achieved since full exploitation does not take place. The fact that the surfaces are not affected by the parameter value of Random-Continuous will be analysed in the discussion section 5.

Always-Same

Remarkably, Always-Same behaves exactly like Random-Discrete. This becomes visible when looking at the columns where Random-Discrete plays against itself and against Always-Same. The surprising absence of any difference between these two strategies needs further explanation and will thus be revisited in the discussion section 5.

Adapt-Discrete

In the absolute-gain surface of Random-Discrete against Adapt-Discrete, a stripe appears at Adapt-Discrete parameter values from 0 to 2. This stripe ranges from absolute exploitation at (0, 0) to mutual cooperation where the parameter of Random-Discrete is equal to 10. Abruptly, this stripe does not continue to follow its pattern in the y-direction, rather a new inclination is visible. Both surfaces' lowest points are at a height of 10, which means that—after adaptation of Adapt-Discrete—mutual defection was played. However, if Random-Discrete always cooperates, the same effect appears, due to the adaptation of Adapt-Discrete, mutual cooperation can be observed resulting in 20 points for both players. Additionally, the maximum points of the first plot indicate that the same points on the corresponding surface of the opponent must be at height 0.

The relative-gain surfaces—with the exception of the stripe highlighted above—is close to 0, meaning that there are practically no advantages. At the coordinates of the stripe, however, a remarkable advantage for Random-Discrete over Adapt-Discrete stands out, which needs further discussion (5).

As a matter of course, the overall-gain surface shows highest points where only mutual cooperation occurs, which is where the parameter of Random-Discrete is equal to 10 and Adapt-Discrete adapts to this value. The rear end of the stripe is at a height of 30 which indicates full exploitation. And the broader subsurface ends at a height of 20 which suggests mutual defection. Once again, the reason for this discontinuity will be analysed in the discussion (5).

Adapt-Continuous

The absolute-gain surfaces considerably resemble the ones of the Adapt-Discrete interaction. The difference mainly concerns the angularity since curves appear to replace the edges. So on the one hand, the similarity can be attributed to the description already delivered for Adapt-Discrete. On the other hand, the rounding off effect needs further investigation in the discussion (5). The same applies to the relative-gain and overall-gain surface plots.

4.2 Random-Continuous

Random-Discrete

The interaction of Random-Discrete and Random-Continuous has already been described above (4.1). So, no further explanation is required.

Random-Continuous

The interaction of Random-Continuous with itself results in completely flat surfaces. The absolute-gain surfaces are at a height of 15. This can be ascribed to the fact that Random-Continuous on average delivers 0.5 as an investment. Plugging in 0.5 for both x and y in the pay-off equations (see 2.3), we get $p_A = p_B = 0.5 - 0.5 \cdot 0.5 + 0.5 \rightarrow p_A = p_B = 0.75$. After 20 rounds of the CPD, this comes down to 15 points for both players.

Subtracting two flat surfaces at the same height results in an also flat surface which corresponds to the zero-plane. No strategy was better than the other at any point.

Adding these two even surfaces—which are both at height 15—an overall surface at height 30 follows. The flatness of all the surfaces deserves to be revisited in the discussion (5).

Always-Same

As already observed in the subsection on the Random-Discrete strategy (4.1), Always-Same behaves identically to Random-Discrete. Once again, this effect requires further examination in the discussion section (5).

Adapt-Discrete

For the absolute-gain plots, there are two main levels of the surfaces. The upper level in the Random-Continuous surface is at a height of 25 and the lower one at a height of 10.

Additionally, there is a little peak at the coordinate (5, 10) with an approximate height of 17 and 12, respectively. The lower level in the Adapt-Discrete surface is located at a height of 10 and the upper one at a height of 15 while the peak is pointing downwards.

The upper level of the absolute-gain surface of Random-Continuous is approximately 15 points higher than the lower level of the other absolute-gain plot. Additionally, the lower level of the surface of Random-Continuous and the upper one of the surface of Adapt-Discrete have almost the same z value. Notably, the relative-gain surface of Random-Continuous is completely above the zero-plane, which pronouncedly applies for small parameter values of Adapt-Discrete, i.e., corresponding to AlwaysCooperate. This means that Random-Continuous won the entire pbICPD.

In terms of overall-gain, the upper level of the absolute-gain surface of Random-Continuous added to the lower level of the one of Adapt-Discrete results in higher values than adding the two levels of similar height. Furthermore, a small positive peak remains. Again, the reason for this discontinuity will be discussed later (5).

Adapt-Continuous

The absolute-gain surfaces resemble a wave or a curve in the y-direction. The fact that the Random-Continuous parameter does not affect the surface's form can be traced back to the already discussed phenomenon that Random-Continuous on average behaves like AlwaysNeutral. Strikingly, the upper and lower limits of the surfaces correspond to the absolute-gain surfaces described above for the Adapt-Discrete interaction. For Adapt-Continuous parameter values of higher than three, the curved surface appears to pronouncedly converge to a flat plane at a height of about 15.

The same properties can be seen in the relative-gain surface plots. Upper and lower limits correspond to those of the relative-gain surfaces of the Adapt-Discrete interaction and the relative-gain surface of Random-Continuous is again completely located above the zero-plane. The same applies to the overall-gain surface plot with its upper limit at a height of 35 and the lower limit at a height of 30. The reasons for this curvature will be revisited in the discussion (5).

5 Discussion

In the results section (4), several issues have been identified that deserve a closer inspection. In the following subsections, these issues will be discussed in regard to aspects of shallow randomness (5.1), dispersion and risk (5.2), threshold values for discrete adaptations (5.3), and dynamics of continuous adaptations (5.4). Finally, the interactions in terms of gains against random strategies will be referred back to the leading question of the present thesis (5.5).

5.1 Shallow Randomness

It was found that Random-Discrete acts identically as Always-Same and Random-Continuous as AlwaysNeutral. This means that Random-Discrete and Random-Continuous, as random strategies, behave equitably as deterministic strategies if the average of all the generated surfaces is taken.

To clarify the resemblance between Random-Discrete and Always-Same, the equations to calculate the investment have to be analysed (see Equation 6).

$$\Pr(i = 1) = \frac{1}{10}\theta_{\mathrm{RD}} \tag{18}$$

Regarding the investment calculation of Random-Discrete (Equation 18), we can generate n investments (Equation 19).

$$i_1, i_2, i_3, \ldots, i_n = \begin{cases} 1 & \text{with probability } \frac{1}{10}\theta_{\mathrm{RD}} \\ 0 & \text{with probability } 1 - \frac{1}{10}\theta_{\mathrm{RD}} \end{cases} \tag{19}$$

The average of all the generated investments in Equation 19 converges to a certain value as shown in Equation 20.

$$\bar{i} = \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} i_k = \theta_{\mathrm{RD}} \frac{1}{10} \text{ almost surely} \tag{20}$$

Substituting n with 100—the number of repetitions in the present simulation—in Equation 20 would be an approximation.

$$\bar{i} \approx \frac{1}{100} \sum_{k=1}^{100} i_k \approx \theta_{\mathrm{RD}} \frac{1}{10} \tag{21}$$

Hence, calculating an investment with a certain probability ultimately comes down to investing the same contribution constantly, which in turn corresponds to the behaviour of Always-Same.

A similar effect applies for Random-Continuous and the convergence to the behaviour of the AlwaysNeutral strategy. To investigate this effect, the same procedure is conducted. First, the investment equations of Random-Continuous need to be revisited (see Equations 8-10).

$$i(\theta_{\mathrm{RC}}) = 0.5 + \epsilon \cdot s(\theta_{\mathrm{RC}}) \tag{22}$$

$$s(\theta_{\mathrm{RC}}) = \theta_{\mathrm{RC}} \frac{1}{20} \tag{23}$$

$$\epsilon = \begin{cases} 1 & \text{with probability } \frac{1}{2} \\ -1 & \text{with probability } \frac{1}{2} \end{cases} \tag{24}$$

Only Equation 24 contains a probability We thus form n ϵ values, as Equation 25 implies.

$$\epsilon_1, \epsilon_2, \epsilon_3, \ldots, \epsilon_n = \begin{cases} 1 & \text{with probability } \frac{1}{2} \\ -1 & \text{with probability } \frac{1}{2} \end{cases} \tag{25}$$

The average of these ϵ values cancels out and therefore equals 0 (Equations 26 and 27).

$$\bar{\epsilon} = \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \epsilon_k = 0 \text{ almost surely} \tag{26}$$

$$\bar{\epsilon} \approx \frac{1}{100} \sum_{k=1}^{100} \epsilon_k \approx 0 \tag{27}$$

Consequently, the investment is—over 100 repetitions—approximately equal to 0.5, as shows Equation 28.

$$i(\theta_{\mathrm{RC}}) \approx 0.5 + \bar{\epsilon} \cdot \theta_{\mathrm{RC}} \frac{1}{20} \approx 0.5 \tag{28}$$

The generally non-existent influence of Random-Continuous' parameter values can be traced back to this phenomenon—which is labelled as the Law of Large Number in mathematical statistics.

5.2 Dispersion and Risk

Building on the discussions on shallow randomness (section 5.1), it seems worthwhile to delve deeper into the dispersion of the Random-Discrete and Always-Same surfaces, which on average show the same behaviour. To this end, Figure 4 illustrates the absolute-gain surfaces of the interaction of both strategies if only played once (with 20 rounds).

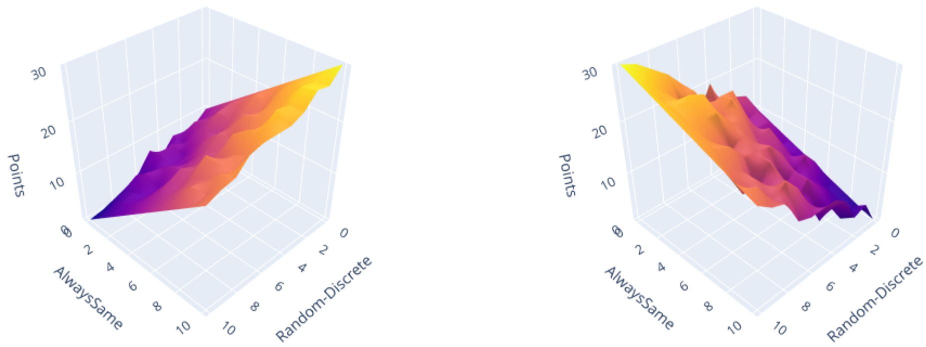


Figure 4: *pbICPD surfaces without repetitions. Absolute-gain surfaces for Random-Discrete (left) and for Always-Same (right).*

Coarser surfaces are visible, which is due to the phenomenon of the Law of Large Numbers as explained above (section 5.1). The surface of Random-Discrete is a bit evener. This fact can be explained by examining the pay-off Equations 4 and 5: $p_A = y - cx + c$ and $p_B = x - cy + c$, where x is the previous investment of Random-Discrete and y the previous investment of Always-Same. Since x is random—depending on Random-Discrete's parameter—and y is a constant over one ICPD, subtracting only a fraction of the random investment pays off a more constant amount of points.

To calculate the dispersion of the average absolute-gain surface, the Standard Deviation (SD) (Equation 29) can be used for every game and thus new surface plots can be generated (see Figure 5).

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} (x_i - \overline{x})^2} \tag{29}$$

As can be seen in Figure 5, the SD surface of Random-Discrete is considerably smaller than that of Always-Same, which refers to the explanation already provided above. Furthermore, it becomes obvious that both SD surfaces are only affected by the parameter value of Random-Discrete with the highest SDs located at a parameter value of 5, which represents the most random version of the strategy. To conclude, on the one hand, there is a more pronounced chance for Always-Same to win or lose if the parameter of Random-Discrete takes medium values. On the other hand, as can be seen in Figure 6, the resulting risk is largely negligible since the average absolute-gain surfaces are not remarkably altered by the SD (which covers 68.2% of the samples).

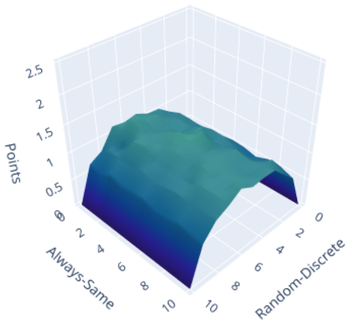


Figure 5: *pbICPD Standard-Deviation surfaces for 100 repetitions. Standard deviations of the absolute-gain surfaces of Random-Discrete (left) and of Always-Same (right).*

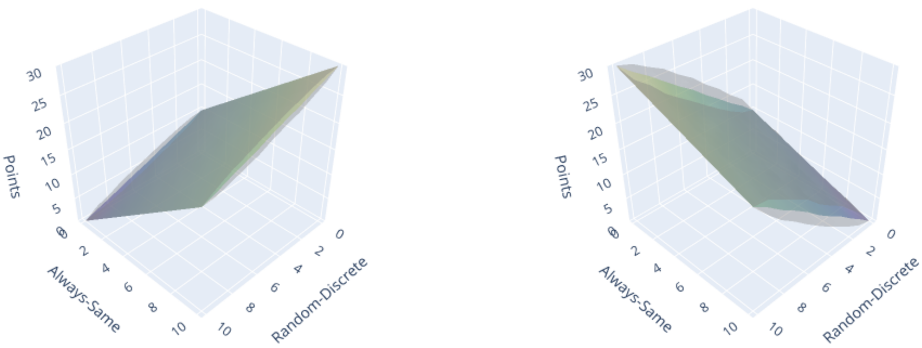


Figure 6: *pbICPD surfaces after 100 repetitions. Absolute-gain surfaces ± SD for Random-Discrete (left) and for Always-Same (right).*

5.3 Threshold Values for Discrete Adaptations

Regarding the surfaces of the interactions of Random-Discrete with Adapt-Discrete, a surprising stripe can be observed. To understand the reason for this effect, the investment calculations of Adapt-Discrete should be recalled (see Equations 15-17).

$$i_0 = 1 \tag{30}$$

$$i(\theta_{\mathrm{AD}}, k) = \begin{cases} 1 & \text{if } i_1 + s(\theta_{\mathrm{AD}}, k) \geq 0.5 \\ 0 & \text{if } i_1 + s(\theta_{\mathrm{AD}}, k) < 0.5 \end{cases} \tag{31}$$

$$s(\theta_{\mathrm{AD}}, k) = \frac{1}{5}\theta_{\mathrm{AD}} \cdot (\tilde{i}_{k-1} - i_{k-1}) \tag{32}$$

As the investments of Adapt-Discrete are rounded to the nearest integer (Equation 31), to get from investment 0 to 1 or vice versa, the raw value of the investment has to exceed a certain threshold value which is equal to 0.5. In the first round, Adapt-Discrete contributes full cooperation (Equation 30), so that to surpass the threshold, the shift in Equation 32 must be lower than -0.5. To generate a general formula that calculates the opponent's investment that is necessary to exceed the threshold in the first round given the Adapt-Discrete parameter, it is required to solve for $\tilde{i}_1$.

$$\begin{aligned} s(\theta_{\text{AD}}, k) &< -0.5 \\ \frac{\theta_{\text{AD}}}{5}(\tilde{i}_{k-1} - 1) &< -0.5 && \Big| \cdot \frac{5}{\theta_{\text{AD}}} \\ \tilde{i}_{k-1} - 1 &< -\frac{2.5}{\theta_{\text{AD}}} && \Big| +1 \\ \tilde{i}_{k-1} &< -\frac{2.5}{\theta_{\text{AD}}} + 1 \end{aligned} \tag{33}$$

Figure 7 illustrates this formula graphically. The curve depicts the upper limit of the area of values for which adaptation takes place (exclusively). It can be seen that the opponent's investment would be needed to fall below 0 to exceed the threshold value if $0 \leq \theta_{\text{AD}} \leq 2$. However, a contribution below 0 is impossible. To conclude, if Adapt-Discrete's parameter is either 0, 1, or 2, the shift required to surpass the threshold value cannot be achieved. Thus, Adapt-Discrete is equivalent to AlwaysCooperate in cases where its parameter is lower than 3. This explains the gainfulness of the stripe in the absolute-gain surface of Random-Discrete and thus the corresponding disadvantage of Adapt-Discrete. In this area of the surface, Adapt-Discrete is exploited since it cannot react because of always submitting full cooperation.

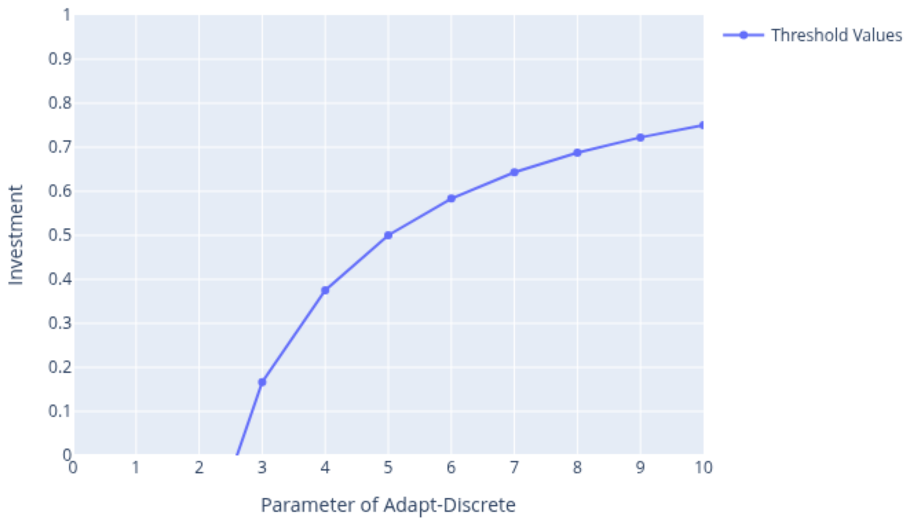


Figure 7: *Threshold values for discrete adaptations as a function of the Adapt-Discrete parameter.*

In the surface plot of the interaction of Random-Continuous with Adapt-Discrete, the existence of two different levels is remarkable and will thus be investigated. To understand the discontinuity of the two absolute-gain surfaces, the investment calculations of Random-Continuous (Equation 8-10) need to be revisited.

$$i(\theta_{\text{RC}}) = 0.5 + \epsilon \cdot s(\theta_{\text{RC}}) \tag{34}$$

$$s(\theta_{\mathrm{RC}}) = \frac{1}{20}\theta_{\mathrm{RC}} \tag{35}$$

$$\Pr(\epsilon = 1) = \Pr(\epsilon = -1) = \frac{1}{2} \tag{36}$$

The investment of Random-Continuous as a function of its parameter can be found as a graph in Figure 8 (left). The lower and upper points of one parameter value are notated as 'RC [parameter value] lower' or 'RC [parameter value] upper', respectively.

Figure 8: *Investment of Random-Continuous as a function of its parameter (left) and intersections of parameter-dependent investments with the* 'Threshold values' *from Figure 7 (right).*

As shown in Figure 8 (right), every investment of Random-Continuous above the line of *'Threshold Values'* results in an exploitation of Adapt-Discrete. Parameters equal to either 0, 1, or 2 necessitate a different approach (as discussed above). To explain this fact, the parameter value of Adapt-Discrete equal to 3 will be taken as an example. The same idea applies to every further value of the parameter. The graph at parameter 3 is at an investment value of $0.1\overline{6}$ and thus lies between the lines *'RC 7 Lower'* and *'RC 6 Lower'*. The line below this point means that this and every greater parameter of Random-Continuous does not exploit Adapt-Discrete. The line above indicates the first version of Random-Continuous in negative direction that exploits Adapt-Discrete. This line can be at the same height as the point and it would still result in a losing situation for Adapt-Discrete.

The appearance of the small peak is also a part of the discontinuity of the surfaces. In Figure 9 it can be seen that the two traces *'RC 5 Lower'* and *'RC 5 Upper'* intersect with *'Surpass Threshold Down'* and *'Surpass threshold Up'* where the parameter value of Adapt-Discrete is equal to 10, respectively. In the first round—as Adapt-Discrete starts with full cooperation—Adapt-Discrete only adapts from 1 to 0 with a probability of 50% since only the investment line *'RC 5 lower'* is located below the point of the *'Surpass threshold down'* curve. However, Adapt-Discrete always surpasses the threshold value of 0.5 from full defection to full cooperation as both *'RC 5 Lower'* and *'RC 5 Upper'* are above the corresponding point of the *'Surpass Threshold Up'* graph. So, Adapt-Discrete will go down 50% of the cases when it is at full cooperation but it will go up to full cooperation when it is at full defection with 100% probability. This implies that Adapt-Discrete is exploited more often as Adapt-Discrete cooperates more often than it defects. Every other combination of parameters of Adapt-Discrete and Random-Continuous prevents this situation from occurring.

5.4 Dynamics of Continuous Adaptations

The absolute-gain surface of Adapt-Continuous in the interaction with Random-Discrete has similar qualities to that of Adapt-Discrete against Random-Discrete. The main difference between these surfaces concerns the angularity. Figure 10 demonstrates the investments of Adapt-

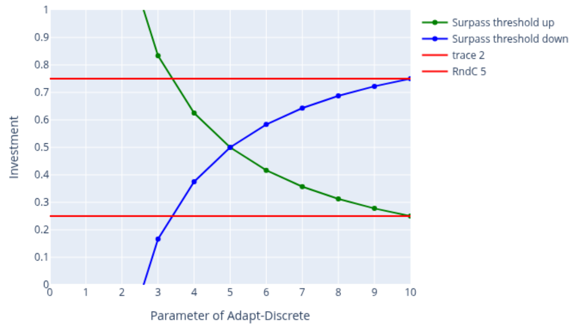


Figure 9: *Intersections of surpass thresholds curves with* 'RC 5 Lower' *and* 'RC 5 Upper'.

Continuous in one ICPD when it plays against Random-Discrete, thereby taking Always-Same as equivalent to Random-Discrete for simplicity (see 5.1 and 5.2). Adapt-Continuous pursuits getting on the same level of investment as the opponent for parameter values from 1 to 5. Parameter 0 forms a special case as its strategy never adapts and thus always cooperates, which can be derived from the fact that the difference to the opponent's investment is multiplied by 0 and thus no adaption is made. The parameter value of Random-Discrete equals 5 in Figure 10.

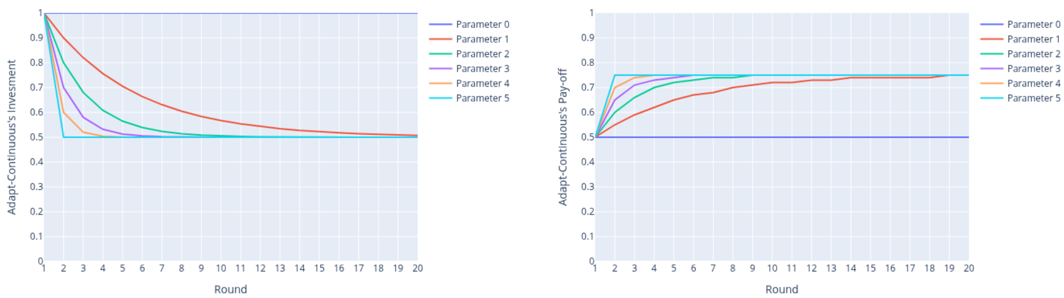


Figure 10: *Investments (left) and pay-offs (right) of Adapt-Continuous with parameters ranging from 0 to 5 against Random-Discrete parameter 5 as a functions of the iteration.*

The function for each graph can be derived from the investment calculation of Adapt-Continuous (Equations 12-14). Equation 37 describes this process recursively. k indicates the current round and θ is the parameter of Adapt-Continuous. $\tilde{i}$ is the opponent's investment and thus constant.

$$i_k = i_{k-1} + \frac{\theta}{5} \cdot (\tilde{i} - i_{k-1}) \tag{37}$$

Since it is not possible to generate a graph using the recursive form, the explicit version is required (Equation 38).

$$i_k = \tilde{i} + (1 - \frac{\theta}{5})^k \cdot (i_0 - \tilde{i}) \tag{38}$$

The higher the parameter of Adapt-Continuous, the faster the strategy can adjust its investment. This means that Adapt-Continuous, when the parameter is small, will be more exploited in the first several rounds than an Adapt-Continuous version with a larger parameter value. For parameter values of Adapt-Continuous greater than or equal to 5, Figure 11 shows the dynamics of the investments and pay-offs.

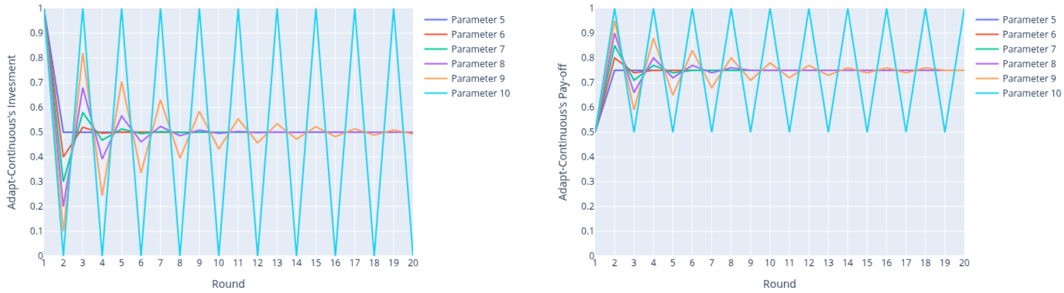


Figure 11: *Investments (left) and pay-offs (right) of Adapt-Continuous with parameters ranging from 5 to 10 against Random-Discrete parameter 5 as a functions of the iteration.*

For the depicted parameter values, Adapt-Continuous versions generally overstep the value of the opponent's investment. However, whilst parameter 5 immediately adapts, the traces of parameters from 6 to 9 converge to a certain pay-off value of 0.75, which can be ascribed to specifics of the applied pay-off system again. On the contrary, the parameter value of 10 always surpasses the opponent's contribution. This results in an alternating cooperating and defecting behaviour.

5.5 Gains Against Random Strategies

When evaluating the surface plots depicted in Figure 2 and 3, random strategies always win against adaptive strategies. This particularly applies to low parameter values of the adaptive strategy, meaning that the adaptation dynamics from the initial state of full cooperation is limited and thus exploitation by the opponent is offered. Positive relative-gain values can only be achieved if the opponent's investment is greater than one's own. Ultimately, it is thus logical to fully defect against random strategies if the aim is to maximise the personal advantage. This can happen by always defecting in a Always-Same strategy or cooperating only with a small probability in a Random-Discrete strategy. If the opponent reacts on this behaviour, the fundamental concepts of the PD surface, meaning that both player end up in mutual defection and thus receiving the least overall-gain.

When the aim is to maximise this overall-gain, cooperating always should be preferred, or at least with a high probability. When pursuing an adaptive strategy, however, another option would be to offer exploitation, meaning being altruistic. This results in a guaranteed, although lower overall-gain (at a height of 30) at the price of accepting negative relative-gain values.

Overall a superior for successfully counteracting a random strategy in terms of maximising relative-gain as well as overall-gain values, could not be identified in the present simulation.

6 Conclusions and Outlook

When coming back to the leading question, one needs to conclude that no strong strategy counteracting erratic behaviour was found in the PD simulations. It was rather confirmed that if the personal advantage is emphasised, cooperation should be completely refused. However, complete erratic behaviours in the real world are practically not found—which is a considerable limitation of the pursued game-theoretical approach. It is reasonable to expect that interaction partners behave rationally at a certain degree at least. When being confronted with an apparent erratic behaviour, the resulting issue then comes ultimately down to the question whether the other person is actually erratic or at least partially constructive. In the first case, harshly finding back is advisable whereas in the second case, some cooperation should be offered. Coming back to the introductory political example of this thesis: Whether Donald J. Trump is more on erratic or on the rational side seems to be an open question...

7 Source List

References

[1] D. Ríos and E. Cresto. “Prisoner’s dilemma, one shot and iterated”. In: *International Encyclopedia of the Social & Behavioral Sciences*. Ed. by J. D. Wright. Second Edition. Elsevier, 2015, pp. 930–937. DOI: https://doi.org/10.1016/B978-0-08-097086-8.63065-9.

[2] S. Kuhn. “Prisoner’s Dilemma”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta and Uri Nodelman. Fall 2025. Metaphysics Research Lab, Stanford University, 2025. URL: https://plato.stanford.edu/archives/fall2025/entries/prisoner-dilemma/.

[3] S. Le and R. Boyd. “Evolutionary dynamics of the continuous iterated prisoner’s dilemma”. In: *Journal of Theoretical Biology* 245.2 (2007), pp. 258–267. DOI: https://doi.org/10.1016/j.jtbi.2006.09.016.

[4] R. Axelrod and W. D. Hamilton. “The evolution of cooperation”. In: *Science* 221.4489 (1981), pp. 1390–1396. DOI: https://doi.org/10.1126/science.7466396.

[5] M. Nowak and K. Sigmund. “A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner’s dilemma game”. In: *Nature* 364 (1993), pp. 56–58. DOI: https://doi.org/10.1038/364056a0.

[6] B. Brembs. “Chaos, cheating and cooperation: Potential solutions to the prisoner’s dilemma”. In: *Oikos* 76 (1996), pp. 14–24. DOI: https://doi.org/10.2307/3545744.

[7] D. Kraines and V. Kraines. “Learning to cooperate with Pavlov an adaptive strategy for the iterated prisoner’s dilemma with noise”. In: *Theory and Decision* 35 (1993), pp. 107–150. DOI: https://doi.org/10.1007/BF01074955.

[8] C. Hauert and O. Stenull. “Simple adaptive strategy wins the prisoner’s dilemma”. In: *Journal of Theoretical Biology* 218.3 (2002), pp. 261–272. DOI: https://doi.org/10.1006/jtbi.2002.3072.

[9] D. Hao, K. Li, and T. Zhou. “Payoff control in the iterated prisoner’s dilemma”. In: *Proceedings of the 27th International Joint Conference on Artificial Intelligence* (2018), pp. 296–302. DOI: https://doi.org/10.48550/arXiv.1807.06666.

[10] M. Jurišić, D. Kermek, and M. Konecki. “A review of iterated prisoner’s dilemma strategies”. In: *MIPRO: 2012 Proceedings of the 35th International Convention* (2012). URL: https://www.researchgate.net/publication/261424948_A_review_of_iterated_prisoner's_dilemma_strategies.

[11] P. Mathieu and J. P. Delahaye. “New winning strategies for the iterated prisoner’s dilemma”. In: *Journal of Artificial Societies and Social Simulation* 20.4 (2017). URL: https://www.jasss.org/20/4/12.html.

[12] W. Ashlock. “Using prisoner’s dilemma fingerprints to analyze evolved strategies”. MA thesis. University of Guelph., 2008. URL: https://atrium.lib.uoguelph.ca/server/api/core/bitstreams/d86079e1-b4ca-4a5c-867e-c6cf536ccee7/content.

[13] M. Hannah, H. M. Lewis, and A. J. Dumbrell. “Evolutionary games of cooperation: Insights through integration of theory and data”. In: *Ecological Complexity* 16 (2013), pp. 20–30. DOI: https://doi.org/10.1016/j.ecocom.2013.02.007.

[14] F. P. Leite. “Larger reward values alone are not enough to entice more cooperation”. In: *Thinking & Reasoning* 17.1 (2011), pp. 82–103. DOI: https://doi.org/10.1080/13546783.2010.537521.

[15] S. Van Segbroeck et al. “Reacting differently to adverse ties promotes cooperation in social networks”. In: *Physical Review Letters* 102.058105 (2009). DOI: https://doi.org/10.1103/PhysRevLett.102.058105.

[16] L. R. Izquierdo, N. M. Gotts, and J. G. Polhill. “Case-based reasoning, social dilemmas, and a new equilibrium concept”. In: *Journal of Artificial Societies and Social Simulation* 7.3 (2004). URL: https://www.jasss.org/7/3/1.html.

[17] N. Hayashi and T. Yamagishi. “Selective play: Choosing partners in an uncertain world”. In: *Personality and Social Psychology Review* 2.4 (1998), pp. 276–289. DOI: https://doi.org/10.1207/s15327957pspr0204_4.

[18] X. Han, X. Zhao, and H. Xia. “Hybrid learning promotes cooperation in the spatial prisoner’s dilemma game”. In: *Chaos, Solitons & Fractals* 164.112684 (2022). DOI: https://doi.org/10.1016/j.chaos.2022.112684.

[19] I. McLean. “The social contract in Leviathan and the prisoner‘s dilemma supergame”. In: *Political Studies* 29.3 (1981), pp. 339–351. DOI: https://doi.org/10.1111/j.1467-9248.1981.tb00500.x.

[20] I Back and A. Flache. “The viability of cooperation based on interpersonal commitment”. In: *Journal of Artificial Societies and Social Simulation* 9.1 (2006). URL: https://www.jasss.org/9/1/12.html.

[21] X. Chen and F. Fu. “Outlearning extortioners: Unbending strategies can foster reciprocal fairness and cooperation”. In: *PNAS Nexus* 2.6 (2023). pgad176. DOI: https://doi.org/10.1093/pnasnexus/pgad176.

A Appendix: Literature Search

To retrieve the most important scientific findings on the topic of PD and erratic behaviour, a literature search was conducted on `scholar.google.com` with the search terms 'Prisoner's Dilemma' and 'random strategy'. To narrow down the search to scientific findings, books (54), mere citations (20) and papers not written in English (6) were excluded from the 650 hits as well as items with no more than ten citations (362). On the basis of the titles of the remaining 208 articles, papers were eliminated that dealt with the evolutionary (55) or other PD variants (14; e.g., with more players), the role of noise or information use (12), implementational or analytical aspects (42) or special applications (34; e.g., in human-robot interaction). From the remaining 51 items, 50 could be retrieved via the *University of Bern* library (access via the GymThun's and University's joint talent promotion programme).

The screening of the abstracts led to a further exclusion of items due to their actual focus on the evolutionary PD variant (10; e.g. [15]), the introduction of more than two players (5; e.g. [16]), other PD variants (8; e.g. [17]), strategy learning (10; e.g. [18]), or specific aspects like philosophical considerations (7; e.g. [19]). For the remaining ten articles, pdf files were downloaded from the *University of Bern* library and studied in more detail. After deleting two more items, in which either a different game was researched ([20]: exchange dilemma) or "random strategy" was used as randomly choosing a certain strategy from a set of strategies [21], eight publications finally remained, which were considered as being most relevant for the present topic. These publications are briefly summarised in subsection 2.5.

For the retrieved publications finally included in the reference list, pdf files are available under `https://github.com/adho08/Prisoner-s-Dilemma`. The following items were not included in the reference list because they were either not accessible or excluded during abstract screening (sorted by number of citations):

Tit for tat in heterogeneous populations
MA Nowak, K Sigmund - Nature, 1992 - nature.com
Speichern Zitieren Zitiert von: 1448 Ähnliche Artikel Alle 9 Versionen

Reciprocity and the induction of cooperation in social dilemmas.
SS Komorita, CD Parks, LG Hulbert - Journal of Personality and ..., 1992 - psycnet.apa.org
Speichern Zitieren Zitiert von: 238 Ähnliche Artikel Alle 6 Versionen

"An eye for an eye leaves everyone blind": Cooperation and accounting systems
P Kollock - American Sociological Review, 1993 - JSTOR
Speichern Zitieren Zitiert von: 214 Ähnliche Artikel Alle 6 Versionen

Reacting differently to adverse ties promotes cooperation in social networks
S Van Segbroeck, FC Santos, T Lenaerts, JM Pacheco - Physical review letters, 2009 - APS
Speichern Zitieren Zitiert von: 200 Ähnliche Artikel Alle 16 Versionen

Is strategic decision making chaotic?
D Richards - Behavioral science, 1990 - Wiley Online Library
Speichern Zitieren Zitiert von: 164 Ähnliche Artikel Alle 6 Versionen

[HTML] Prisoner's dilemma
S Kuhn - 1997 - plato.stanford.edu
Speichern Zitieren Zitiert von: 159 Ähnliche Artikel Alle 6 Versionen

Aspiration dynamics of multi-player games in finite populations

J Du, B Wu, PM Altrock, L Wang - Journal of the Royal . . . , 2014 - royalsocietypublishing.org
Speichern Zitieren Zitiert von: 141 Ähnliche Artikel Alle 12 Versionen

Selective play: Choosing partners in an uncertain world
N Hayashi, T Yamagishi - Personality and Social Psychology . . . , 1998 - journals.sagepub.com
Speichern Zitieren Zitiert von: 138 Ähnliche Artikel Alle 7 Versionen

Prisoner's dilemma and the pigeon: Control by immediate consequences
L Green, PC Price. . . - Journal of the experimental . . . , 1995 - Wiley Online Library
Speichern Zitieren Zitiert von: 129 Ähnliche Artikel Alle 12 Versionen

[HTML] Communication and cooperation
JH Miller, CT Butts, D Rode - Journal of Economic Behavior & Organization, 2002 - Elsevier
Speichern Zitieren Zitiert von: 126 Ähnliche Artikel Alle 19 Versionen

Multiple strategies in structured populations
CE Tarnita, N Wage, MA Nowak - Proceedings of the National Academy of . . . , 2011 - pnas.org
Speichern Zitieren Zitiert von: 120 Ähnliche Artikel Alle 15 Versionen

Resolving the iterated prisoner's dilemma: theory and reality
NJ Raihani, R Bshary - Journal of Evolutionary Biology, 2011 - academic.oup.com
Speichern Zitieren Zitiert von: 118 Ähnliche Artikel Alle 8 Versionen

Self-organization and emergence in social systems: Modeling the coevolution of social environments and cooperative behavior
D Helbing, W Yu, H Rauhut - The Journal of Mathematical . . . , 2011 - Taylor & Francis
Speichern Zitieren Zitiert von: 96 Ähnliche Artikel Alle 13 Versionen

Relationship between cooperation in an iterated prisoner's dilemma game and the discounting of hypothetical outcomes
R Yi, MW Johnson, WK Bickel - Learning & behavior, 2005 - Springer
Speichern Zitieren Zitiert von: 79 Ähnliche Artikel Alle 14 Versionen

Skill in games
P Larkey, JB Kadane, R Austin. . . - Management . . . , 1997 - pubsonline.informs.org
Speichern Zitieren Zitiert von: 66 Ähnliche Artikel Alle 18 Versionen

Finite automata play repeated prisoner's dilemma with information processing costs
TH Ho - Journal of economic dynamics and control, 1996 - Elsevier
Speichern Zitieren Zitiert von: 59 Ähnliche Artikel Alle 6 Versionen

Development and the origin of behavioral strategies
TD Johnston - Behavioral and Brain Sciences, 1984 - cambridge.org
Speichern Zitieren Zitiert von: 58 Ähnliche Artikel Alle 2 Versionen

The social contract in Leviathan and the Prisoner's Dilemma supergame
I McLean - Political Studies, 1981 - journals.sagepub.com
Speichern Zitieren Zitiert von: 58 Ähnliche Artikel Alle 5 Versionen

Social and strategic imitation: the way to consensus
D Vilone, JJ Ramasco, A Sánchez, MS Miguel - Scientific reports, 2012 - nature.com

Speichern Zitieren Zitiert von: 53 Ähnliche Artikel Alle 18 Versionen

[HTML] Modeling cooperation among self-interested agents: a critique
H Gintis - The journal of socio-economics, 2004 - Elsevier
Speichern Zitieren Zitiert von: 45 Ähnliche Artikel Alle 7 Versionen

Chaos in cooperation: continuous-valued Prisoner's Dilemmas in infinite-valued logic
G Mar, PS Denis - International journal of bifurcation and chaos, 1994 - World Scientific
Speichern Zitieren Zitiert von: 43 Ähnliche Artikel Alle 5 Versionen

Invasion Dynamics of the Finitely Repeated Prisoner's Dilemma
MA Nowak, K Sigmund - Games and Economic Behavior, 1995 - Elsevier
Speichern Zitieren Zitiert von: 38 Ähnliche Artikel Alle 4 Versionen

[HTML] Strategies and evolution in the minority game: A multi-round strategy experiment
J Linde, J Sonnemans, J Tuinstra - Games and economic behavior, 2014 - Elsevier
Speichern Zitieren Zitiert von: 32 Ähnliche Artikel Alle 16 Versionen

[PDF] Using the iterated prisoner's dilemma for explaining the evolution of cooperation in open source communities
D Eckert, S Koch, J Mitloehner - Proceedings of the first . . . , 2005 - researchgate.net
Speichern Zitieren Zitiert von: 29 Ähnliche Artikel Alle 3 Versionen
[no access to the abstract via *University of Bern* library]

A framework for learning and planning against switching strategies in repeated games
P Hernandez-Leal, E Munoz de Cote. . . - Connection Science, 2014 - Taylor & Francis
Speichern Zitieren Zitiert von: 27 Ähnliche Artikel Alle 5 Versionen

An experimental comparison of negotiation strategies for siting NIMBY facilities
CP Chiu, SK Lai - . . . and Planning B: Planning and Design, 2009 - journals.sagepub.com
Speichern Zitieren Zitiert von: 24 Ähnliche Artikel Alle 8 Versionen

[HTML] Introspection dynamics: a simple model of counterfactual learning in asymmetric games
MC Couto, S Giaimo, C Hilbe - New Journal of Physics, 2022 - iopscience.iop.org
Speichern Zitieren Zitiert von: 24 Ähnliche Artikel Alle 7 Versionen

[HTML] Case-based reasoning, social dilemmas, and a new equilibrium concept
LR Izquierdo, NM Gotts. . . - Journal of Artificial . . . , 2004 - jasss.soc.surrey.ac.uk
Speichern Zitieren Zitiert von: 23 Ähnliche Artikel Alle 4 Versionen

Why some representations are more cooperative than others for prisoner's dilemma
W Ashlock - 2007 IEEE Symposium on Foundations of . . . , 2007 - ieeexplore.ieee.org
Speichern Zitieren Zitiert von: 19 Ähnliche Artikel Alle 2 Versionen

[HTML] Hybrid learning promotes cooperation in the spatial prisoner's dilemma game
X Han, X Zhao, H Xia - Chaos, Solitons & Fractals, 2022 - Elsevier
Speichern Zitieren Zitiert von: 19 Ähnliche Artikel Alle 5 Versionen

On some winning strategies for the Iterated Prisoner's Dilemma, or, Mr. Nice Guy and

the Cosa Nostra
W Slany, W Kienreich - The iterated prisoners' dilemma, 2007 - books.google.com
Speichern Zitieren Zitiert von: 19 Ähnliche Artikel Alle 9 Versionen

The life game: Cognitive strategies for repeated stochastic games
KL Cheng, I Zuckerman, D Nau… - 2011 IEEE Third …, 2011 - ieeexplore.ieee.org
Speichern Zitieren Zitiert von: 18 Ähnliche Artikel Alle 12 Versionen

Risk and interaction aversion: Screening mechanisms in the prisoner's dilemma game
GA Canova, JJ Arenzon - Journal of Statistical Physics, 2018 - Springer
Speichern Zitieren Zitiert von: 17 Ähnliche Artikel Alle 7 Versionen

[PDF] Game theory and agents
SJ Johansson - Licentiate Thesis, University of Karlskrona …, 1999 - researchgate.net
Speichern Zitieren Zitiert von: 15 Ähnliche Artikel Alle 3 Versionen

Artificial agents play the" Mad Mex trust game": a computational approach
DJ Wu, SO Kimbrough, F Zhong - Proceedings of the 35th …, 2002 - ieeexplore.ieee.org
Speichern Zitieren Zitiert von: 15 Ähnliche Artikel Alle 3 Versionen

The rationality of prejudices
T Chadefaux, D Helbing - PloS one, 2012 - journals.plos.org
Speichern Zitieren Zitiert von: 14 Ähnliche Artikel Alle 18 Versionen

Game theory without rationality
A Rapoport - Behavioral and Brain Sciences, 1984 - cambridge.org
Speichern Zitieren Zitiert von: 13 Ähnliche Artikel Alle 2 Versionen

An exploration of differential utility in iterated prisoner's dilemma
D Ashlock, W Ashlock, G Umphry - 2006 IEEE Symposium on …, 2006 - ieeexplore.ieee.org
Speichern Zitieren Zitiert von: 12 Ähnliche Artikel Alle 2 Versionen

B Declaration of Integrity