

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

High-School Thesis
(Maturaarbeit)

Mathematical-Computational
Analysis

Author:
Adrian Hossner

Supervisor: Geoffrey Ostrin

Submission Date: September 7, 2025

School: Gymnasium Thun

Word Count (Sections 1-6): 8000+

Abstract

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 47th 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 a 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	4
2	Theoretical Foundations	4
2.1	Prisoner's Dilemma	4
2.2	Iterated Prisoner's Dilemma	5
2.3	Iterated Continuous Prisoner's Dilemma	5
2.4	Further Variants	6
2.5	Scientific Findings	6
3	Methods and Implementation	7
3.1	Study design	7
3.2	Implemented Strategies	8
3.3	Implementational Details	10
4	Results	10
4.1	Random-Discrete	14
4.2	Random-Continuous	15
5	Discussion	17
5.1	Shallow Randomness	17
5.2	Random-Discrete vs. AlwaysSame	18
5.3	Random-Discrete vs. Adapt-Discrete	19
5.4	Random-Discrete vs. Adapt-Continuous	21
5.5	Random-Continuous vs. Adapt-Discrete	23
6	Conclusions and Outlook	24
7	Source List	25
A	Appendices	27
A.1	Appendix A. Literature Search	27
A.2	Appendix B.	31
A.3	Appendix C.	31
B	Declaration of Integrity	31

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), appendices (Section A) and a declaration of integrity (Section 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 pay-offs 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. Although this—which could have been avoided if both had remained silent and thus obtained a reward of three easy surmountable years in confinement—is nevertheless 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:

$$T < R < P < S \quad (1)$$

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 will be applied on 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:

$$T > R > P > S \quad (2)$$

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

$$2R > T + S \quad (3)$$

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

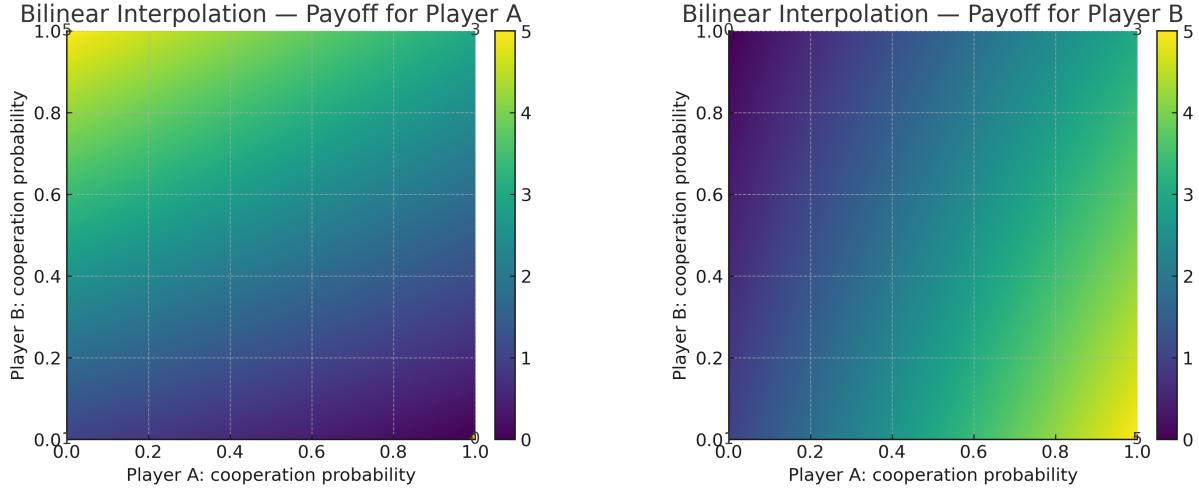


Figure 1: Pay-off heat maps in the ICPD for player A (left) and player B (right).

be represented by colour grades, a heat map as depicted in Figure 1. The equation used to calculate this heat map can be found in the Appendix.

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

$$p_A = y - cx + c \quad (4)$$

$$p_B = x - cy + c \quad (5)$$

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 seemingly 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.1). 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 original tournament organized 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” and 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 than 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

For answering the central research question, it was necessary to develop a overarching study design (3.1), to specify the strategies that will be pursued in the iterated continuous Prisoner’s dilemma (3.2), and to determine details of the implementation in terms of IT (3.3).

3.1 Study design

Erratic behaviour in respect of the PD can be equated to a random conduct. Three strategy-types are proposed by reason that counter such erraticism. Hereunto can a strategy either be unaffected and behold its manner of responding or react to the randomness. In the former case, strategies are identified as rigid. Whereas the latter is characterised as adaptive.

All of these categories of behaviour can algorithmised differently. This particularly concerns the above introduced distinction of the discrete and the continuous form. Applied to the erratic strategies, this means that the investment is a substantial random number between zero and

one. In contrast, the discrete version of this strategy generates a random number that is then rounded to the nearest integer. The same approach hold true in the adaptive strategy. Discrete contributions of rigid strategies are equal to either always cooperating (AlwaysCooperate) or constantly defecting (AlwaysDefect). Whereas continuous strategies 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 investment of a rigid strategy can be defined. On this account, a parameter is incorporated into the determination of the investment and the evaluation of various conditions. It has, incidently, never been undertaken in this manner in the history of researches that concern the ICPD.

To keep the number of comparisons manageable within the realms of a Maturaarbeit (MA), five strategy were chosen. Firstly, the selected strategies in the random class were named to be a discrete (Random-Discrete) and a continuous (Random-Continuous) version. Secondly, the same concept was established in the adaptive category with the discrete (Adapt-Discrete) and the continuous (Adapt-Continuous) behaviour. These strategies hold a parameter which determines the peculiarity level and ranges from zero to ten. So ultimately, the discrete and the continuous implementation are included in one strategy (AlwaysSame) due to this parameter. Thus, the distinction of a discrete and a continuous design would be trivial.

3.2 Implemented Strategies

As justified before, five strategies, namely Random-Discrete, Random-Continuous, Adapt-Discrete, Adapt-Continuous and AlwaysSame, will be presented and specified in the following.

Adapt-Continuous

Adapt-Continuous (AdpC) is an, as the name suggests, adaptive and thus responsive strategy. It starts with full cooperation to offer a constructive relationship within this simulation. After having played one round, it will adapt to the submissions of the opponent by shifting its own investment towards the opponent's.

$$i_k(\theta_{\text{AdpC}}) = i_{k-1} + s(\theta_{\text{AdpC}}, k)$$

$$s(\theta_{\text{AdpC}}, k) = \frac{\theta_{\text{AdpC}}}{5} \cdot (\bar{i}_{k-1} - i_{k-1})$$

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 of the past rounds. The shift being applied to its own previous investment is notated with the function $s(\theta_{\text{AdpC}})$. This strategy with the parameter 0 is identical to AlwaysCooperate since the difference and thus the shift is multiplied by 0. Dividing θ_{AdpC} by 5 implies the fact that this coefficient to the difference is equal to one if the parameter equals five. 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. $\theta_{\text{AdpC}} = 10$ also shows an interesting nature of the strategy. By having the parameter set to ten, it means Adapt will add the shift to its own previous investment twice. Of course, one strategy is not allowed to submit any investment that exceeds the limits that is defined to be from 0 to 1. The implementation of this strategy will simply set its investment to the reached limit if surpassed.

Adapt-Discrete

As mentioned, there is one continuous and one discrete strategy. This means there must be an Adapt-Discrete (AdpD) implemented. The simplest implementation of this strategy is to take the same investment function as Adapt-Continuous and round the result.

$$i_k(\theta_{\text{AdpC}}) = \lfloor i_{k-1} + s(\theta_{\text{AdpC}}, k) \rfloor$$

$$s(\theta_{\text{AdpC}}, k) = \frac{\theta_{\text{AdpC}}}{5} \cdot (\bar{i}_{k-1} - i_{k-1})$$

Something that is not seen in these equations is that the strategy will start with full cooperation, as Adapt-Continuous. If $i(\theta_{\text{AdpD}}) \geq 0.5$, the ultimate investment will be 1. Otherwise it will be equal to 0. Once the strategy is at full defection, it can only get to cooperation again if $\theta_{\text{AdpD}} \geq 3$. Here is why:

The maximum difference there can be is when this strategy submits 1 and the opponent submits a 0, so 1. Looking at the function, this threshold of 0.5 can only be overstepped if the shift is at least equal to 0.5. This only can be achieved if the coefficient is greater or equal to 0.5 ($\frac{\theta_{\text{AdpD}}}{5} \geq 0.5$). This inequation is fulfilled if $\theta_{\text{AdpD}} \geq 3$.

$\theta_{\text{AdpD}} = 0$ means, as in the continuous version, 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, it is most likely to jump from 1 to 0 back to 1 and so on. This only doesn't happen if the difference is less or equal to 0.25.

AlwaysSame

AlwaysSame's (AlwS) parameter calculates the investment very simply. The investment follows the equation:

$$i(\theta_{\text{AlwS}}) = \frac{\theta_{\text{AlwS}}}{10}$$

This means that after each incrementation of the parameter, the investment increases by 0.1. Parameter 0 is equivalent to AlwaysDefect and parameter 1 has the same behaviour as AlwaysCooperate. If the parameter is set to five, the strategy will always submit the investment 0.5 which is identical to Neutral.

Random-Continuous

The other strategy in the group of the random is oppositely continuous, meaning it can additionally submit any number between 0 and 1. The name of this strategy is logically Random-Continuous (RndC). This strategy has a base being at 0.5. The parameter defines the shift being applied to the base, calculated by the equation:

$$i(\theta_{\text{RndC}}) = 0.5 + \cdot s(\theta_{\text{RndC}})$$

$$s(\theta_{\text{RndC}}) = \epsilon \frac{\theta_{\text{RndC}}}{20}$$

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

s means the shift. It is divided by 20 so the maximum shift, parameter being 10, can not exceed the range of 0 and 1. There is a 50% probability of subtracting or adding this shift, indicated by the symbol ϵ .

Random-Discrete

Random-Discrete (RndD) is a strategy in the group of the 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 of submitting full defection is described in the following equation.

$$i(\theta_{\text{RndD}}) = \Pr(i = 1 \mid \theta_{\text{RndD}}) = \frac{\theta_{\text{RndD}}}{10}$$

or

$$i(\theta_{\text{RndD}}) = \Pr(i = 0 \mid \theta_{\text{RndD}}) = 1 - \frac{\theta_{\text{RndD}}}{10}$$

Parameter 0 is equivalent to AlwaysDefect since the probability of submitting 0 is 0%. And on the other hand, parameter 10 is equivalent to AlwaysCooperate since the probability of submitting 1 is 100%.

	$\theta = 0$	$\theta = 5$	$\theta = 10$
Random-Discrete	AlwaysDefect	-	AlwaysCooperate
Random-Continuous	AlwaysNeutral	-	Random-Discrete ($\theta = 5$)
Always-Same	AlwaysDefect	AlwaysNeutral	AlwaysCooperate
Adapt-Discrete	AlwaysCooperate	-	-
Adapt-Continuous	AlwaysCooperate	Tit-For-Tat	-

3.3 Implementational Details

Regarding the initial problem 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 100 times to flatten non-recurring abnormalities. To avoid gratuitous huge numbers, the mean of the ultimate points of these 100 repetitions is taken. This comes down to $100 \cdot 20 \cdot 1210 = 2420000$ executions of the continuous PD.

One encounter of the $11 \cdot 11$ parameter-based ICPD (pbICPD) is displayed in form of a surface. There are two surfaces for each pbICPD, both of which represent the gained points of the corresponding strategy. Additive, two further surfaces can be established. These ones are calculated by subtracting the gained-points surfaces and depict the relative gained, i.g. how much better was strategy A against strategy B. Ultimately, a fifth surface will be generated by adding the two gained-points surfaces. They illustrate the overall gain, meaning the points that have been gained in the game.

The simulation, data generation and visualisation is completely written in Python (version 3.13.5). For the visualisation, as in generating the surface plots, a python library was used, namely Plotly. More information about the python scripts and data visualisation files that were used can be found in the Appendix.

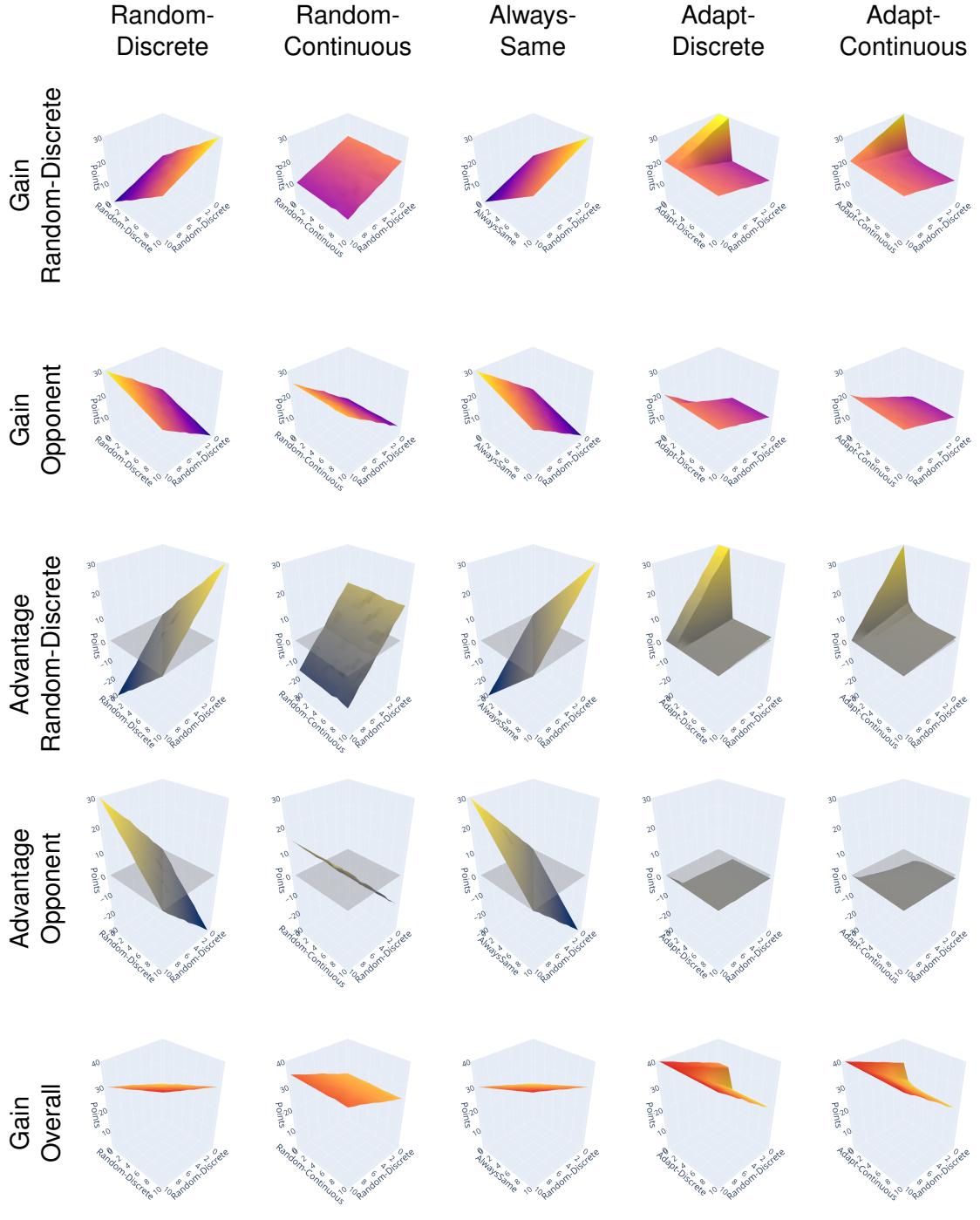
4 Results

In Figures 2 and 3, the surfaces for the Random-Discrete and Random-Continuous simulations are portrayed. There are five surfaces for each interaction that are generated from one pbICPD. The x-axis, the horizontal one in the bottom-right, display the parameter of the main strategy, that is, either Random-Discrete (Figure 3) or Random-Continuous (Figure 4). The y-axis, the other horizontal axis at the bottom of each graph, indicates the parameter of the opponent's strategy. The z-axis always exhibits points gained over all iterations. These dependent values are not only marked by height in the graphs but also by colour (cf. Figure 2). It should be

noted that the two horizontal axis are scaled in such a way that the minimum parameter value of $\theta = 0$ appears at the rear end of the graph and its maximum of $\theta = 10$ 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 Figure 3 and 4, an aggregated version of the explanations provided in Table 1 has been added. [Hier Figure 2 einfügen] Within each column, the first two rows are the most substantial surface plots. They illustrate the absolute gains when applying the respective strategy, first of the random itself and then by the opponent’s strategy, dependent on the column specifying the interaction with a particular other strategy. The highest point that can appear is 30 and the lowest is 0. The maximum can be calculated from the pay-off equations; 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, $P_A \cdot 20 = 30$. So, the boundaries in which the points can vary is from 0 to 30. All the points in a absolute-gain surface with a value of 30 have their correspondent x-y-coordinate in the opponent’s surface, so as to receive 30 points, exploitation is required. This exact same principle holds for the minimum value, merely reversed. Wherever a minimum point can be seen, the counterpart is the maximum point at the same x-y-coordinate and contrariwise. If a point is at the z-coordinate 20, it does not guarantee the incidence of mutual cooperation but portends it. If the correspondent point’s z-coordinate is situated at the same height, mutual cooperation is ensured. This principle is also applied to points at height 10 but with mutual defection.

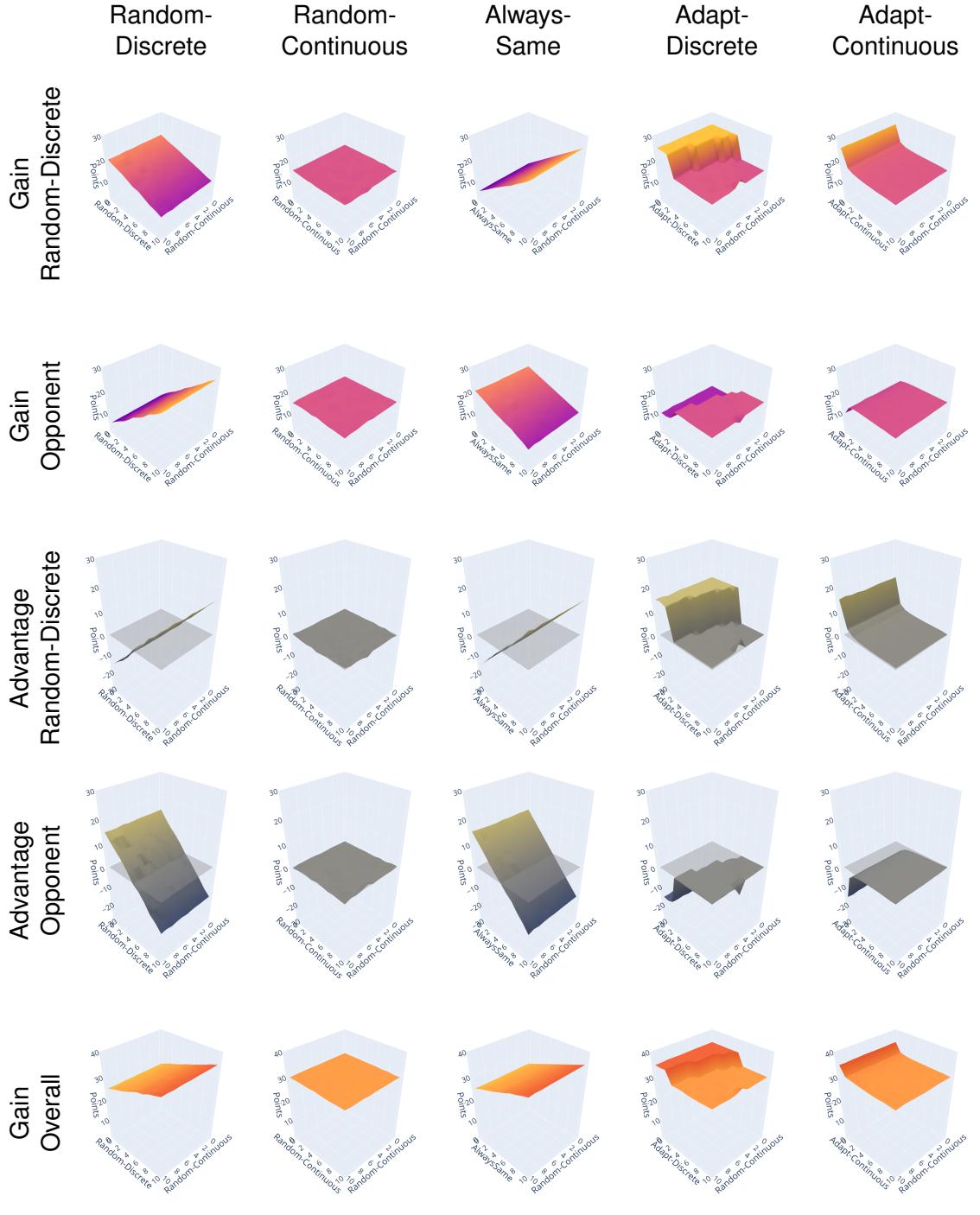
The next two surface plots within a column concern the relative gain, that means the advantage of one strategy over the other. This illustrated how superior both strategies are when compared the to the opponent’s strategy. For the surface of the main strategy, the opponent’s absolute gain surface is 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 (regardless of the parameters). The maximum and the minimum values of the resulting relative-gain surfaces can be calculated by taking the difference of the maximum and the minimum 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 number is always equal to zero. Finally, it should be noted that the two relative-gain surfaces are always mirrored by the zero-plane since every time you subtract $A - B$, it is the same as $(B - A) \cdot (-1)$.

The last row of graphs 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 graphs. Due to the fact that both opponents gain the most points if both constantly cooperate, 1 point is provided for both players after every round and 20 rounds are played in total, the maximum value for these overall-gain surfaces is equal to 40. It should be taken notice of the fact 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 at the same time. 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.



	$\theta = 0$	$\theta = 5$	$\theta = 10$
Random-Discrete	AlwaysDefect	-	AlwaysCooperate
Random-Continuous	AlwaysNeutral	-	Random-Discrete ($\theta = 5$)
Always-Same	AlwaysDefect	AlwaysNeutral	AlwaysCooperate
Adapt-Discrete	AlwaysCooperate	-	-
Adapt-Continuous	AlwaysCooperate	Tit-For-Tat	-

Figure 2: *pbICPD of Random-Discrete.*



	$\theta = 0$	$\theta = 5$	$\theta = 10$
Random-Discrete	AlwaysDefect	-	AlwaysCooperate
Random-Continuous	AlwaysNeutral	-	Random-Discrete ($\theta = 5$)
Always-Same	AlwaysDefect	AlwaysNeutral	AlwaysCooperate
Adapt-Discrete	AlwaysCooperate	-	-
Adapt-Continuous	AlwaysCooperate	Tit-For-Tat	-

Figure 3: *pbICPD* of Random-Continuous.

In the following two pages, the results of each column, meaning the five surface plots, will be presented objectively. The description of each interaction follows a certain pattern. First, the two absolute-gain surfaces are exhibited. Second, both relative-gain surface plots are explained. Ultimately, the overall-gain surface is clarified.

4.1 Random-Discrete

Random-Discrete_2

The left (10, 0) and right (0, 10) corner are the extremes. Random-Discrete is exploited if its parameter is equal to 10, meaning it corresponds to AlwaysCooperate, and Random-Discrete_2's parameter is at 0, i.e. it is equivalent to AlwaysDefect. The hindmost corner at the coordinate (0, 0) is where both defect constantly. This fact can be excogitated by the bespoken feature that this coordinate is located at a z-coordinate of 10. The nearest point to this perspective is at the coordinate (10, 10). Either's value of their parameter is 10, this indicates that both Random-Discrete and Random-Discrete_2 behave equally, viz. as AlwaysCooperate. So, both always cooperate and thus gain 20 points separately.

The peaks of both relative-gain surface plots derive from having the maximum output. Since one can get the maximum output only if the other is exploited, the advantage surface result in the highest z-value possible. When the two parameter's are identical, however, they are not any better than the other. This makes sense since they behave the same when they have the same parameter and can thus not exploit each other.

The overall-gain surface indicates clearly that the pinnacle of the slope is at (10, 10). This point, as already mentioned, has the quality that both strategies are equal to AlwaysCooperate. On the opposite corner, however, it becomes clear this is the lowest point. As already mentioned, the most behind point, being the lowest, results from the fact that both defected constantly. So, for the welfare of the population, it is best if both cooperate interminably.

Random-Continuous

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

The two relative-gain surfaces, if put together in one plot, would intersect exactly in the middle where the parameter of Random-Discrete is equal to 5. Since they range from different limits but both upper and lower limit add up to 30, they intersect in the middle where the parameter of Random-Discrete is 5. That is why the surface of the relative-gain intersects the zero-plane at the same line. Naturally, this surface is only influenced by the parameter of Random-Discrete and not Random-Continuous. They range continuously from -15 to 15.

The overall-gain surface shows a constant slope ranging from 35 to 25. The maximum value, 40, has not been achieved since exploitation did not take place.

AlwaysSame

It has to be mentioned that Random-Discrete behaves like AlwaysSame. This is due to the fact that this simulation has been run 100 times to smooth out one-time abnormalities. This

is notable by looking at the column where Random-Discrete plays against itself and against AlwaysSame. Since the two strategies behave exactly in the same way, these two columns of surface plots are duplicates. The differential between Random-Discrete and AlwaysSame is later defined in the Discussion.

Adapt-Discrete

In the surface of Random-Discrete against Adapt-Discrete, there is a stripe at the parameter of Adapt-Discrete being from 0 to 2. This stripe ranges from absolute exploitation at (0, 0) to mutual cooperation at parameter of Random-Discrete being equal to 10. Abruptly, this stripe does not follow its pattern in the y-direction any more. Rather a new inclination is visible. This sudden alternation will be analysed in the Discussion. This slope's lowest points are at a height of 10. But, both slopes range from where they begin at the y-z-plane to 20 at the other y-z-plane. Respecting the fact that the height of 20 is an indicator of mutual cooperation, it can be verified by regarding the height of the correspondent points. Since they are indeed at the same height, mutual cooperation takes place where the parameter of Random-Discrete is equal to 10. Also the maximum points of the first plot indicate that the same points on the other surface must be at height 0.

The relative-gain surface of Adapt-Discrete hides its most interesting element, the inverted stripe that leads ends at absolute exploitation and thus at height 0. The rest of the surface is very near to the zero-plane which means it was as good as Random-Continuous in this area. Because the slopes look pretty similar, there are practically no advantages in this scope. At the coordinates of the stripe, however, is a great advantage which only Random-Continuous has over Adapt-Discrete.

As a matter of course, the overall-gain surface is the highest where only mutual cooperation occurred, which is where the parameter of Random-Discrete is equal to 10. The stripe breaks off at height 30 which indicates exploitation. And the broader slope finishes at height 20 which suggests mutual defection.

Adapt-Continuous

The absolute-gain surfaces look akin to the ones against Adapt-Discrete. The difference concerns the angularity. Curves appear to replace the edges.

Since this phenomenon differs from that in the interaction between Adapt-Discrete, also this occurrence will be investigated further in the Discussion.

The same applies to the relative-gain and overall-gain surface plots.

4.2 Random-Continuous

Random-Discrete

Certain interactions appear twice, as this one. Random-Discrete has already played against Random-Continuous. There is thus no need for any explanation.

Random-Continuous_2

As already mentioned, Random-Continuous behaves like AlwaysNeutral by having played several times. This results to having a completely flat surface. We can calculate at which height the surface is. Plugging in 0.5 for both x and y in the pay-off equations, we get $P_A = 0.5 - c \cdot 0.5 + c$ where $c = 0.5$. The pay-off of strategy A is $P_A = 0.75$ and having played the ICPD 20 rounds, it is equal to $P_A \cdot 20 = 15$. Both strategies A and B get the same pay-off since they submitted the same investment. $P_B = P_A \rightarrow P_B \cdot 20 = 15$. So, both surfaces are at the same height at every point.

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

Adding these even two surfaces which are both at height 15, a surface at height $15 + 15 = 30$ follows.

AlwaysSame

These surfaces are identical to the one's of the Random-Continuous vs. Random-Discrete interaction. I will not describe them as they already have been.

Adapt-Discrete

There are two main levels on this surface. The upper level in the Random-Continuous surface is at z-coordinate equal to 25 and lower to 10. There is a little peak at the coordinate (5, 10) with an approximate height of 17, 12 respectively.

It seems as if the two absolute-gain surfaces are congruent. This lower level is at $z = 10$ and the upper one at $z = 15$. Also, the peak is pointing downwards making it to a lacuna.

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 nearly the same z component. Notably, the relative-gain surface of Random-Continuous is above the zero-plane. This means that Random-Continuous was better in every single game.

According to this surface, the upper level of the absolute-gain surface of Random-Continuous added to the lower level of the one of Adapt-Discrete is greater than adding the two other levels. The peak pointing upwards indicates that the peak is higher than the lacuna is profound.

Adapt-Continuous

The absolute-gain surfaces resemble a wave or a curve in the y-direction. This derives from the fact that, as already mentioned, Random-Continuous behaves like AlwaysNeutral. The upper and lower limits conform to that of the correspondent absolute-gain surfaces against Adapt-Discrete. Interestingly, this curve gives the impression that the lower limit has been reached once the parameter of Adapt-Continuous arrives at a certain value, estimative 4.

The same properties can be seen in the relative-gain surface plots. Upper and lower limits correspond to that of the relative-gain surfaces against Adapt-Discrete. Also, the relative-gain surface of Random-Continuous is completely above the zero-plane.

This bespoken relation between the interaction against Adapt-Discrete is indifferent in the overall-gain surface plot. The upper limit is at $z = 35$ and the lower boundary is at z-coordinate equal to 30.

5 Discussion

The results suggest certain effects of the PD. Many additional findings have been identified and were promised to be analysed in the Discussion since they need mathematical context for better understanding.

5.1 Shallow Randomness

Random-Discrete and Random-Continuous, as a mixed strategies, behave equitably as pure strategies if the average of all the generated surfaces is taken. The mathematical framework used to describe this phenomenon is called the Law of Large Numbers (LLN). 100 repetitions were named to be sufficient. Random-Discrete acts identically as AlwaysSame and Random-Continuous as AlwaysNeutral.

Random-Discrete

To clarify the resemblance between Random-Discrete and AlwaysSame, the equations to calculate the investment have to analysed.

$$i \sim \begin{cases} 1 & \text{with probability } \frac{1}{2} \\ -1 & \text{with probability } \frac{1}{2} \end{cases} \quad (6)$$

Regarding the investment calculation of Random-Discrete (6), we can generate n i's (7).

$$i_1, i_2, i_3, \dots, i_n \sim \text{Bernoulli}\left(\frac{\theta_{\text{RndD}}}{10}\right) \quad (7)$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n i_k = \frac{\theta_{\text{RndD}}}{10} \text{ almost surely} \quad (8)$$

$$\frac{1}{100} \sum_{k=1}^{100} i_k \approx \frac{\theta_{\text{RndD}}}{10}$$

Substituting n with 100 in equation (8) would be an approximation. But, for the scope of this project, 100 repetitions were named to be sufficient.

The fact that the surfaces of AlwaysSame seem identical to those against itself is now proven and justified.

Random-Continuous

Also in this case, the investment calculations need to be investigated to understand the appearance of AlwaysNeutral.

$$i(\theta_{\text{RndC}}) = 0.5 + s(\theta_{\text{RndC}}) \quad (9)$$

$$s(\theta_{\text{RndC}}) = \epsilon \cdot \frac{\theta_{\text{RndC}}}{20} \quad (10)$$

$$\epsilon \sim \begin{cases} 1 & \text{with probability } \frac{1}{2} \\ -1 & \text{with probability } \frac{1}{2} \end{cases} \quad (11)$$

Only the equation (10) is considerable since the probability or randomness singly appears there. The same procedure is applied.

$$\epsilon_1, \epsilon_2, \epsilon_3, \dots, \epsilon_n \sim \begin{cases} 1 & \text{with probability } \frac{1}{2} \\ -1 & \text{with probability } \frac{1}{2} \end{cases} \quad (12)$$

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \epsilon_k &= 0 \text{ almost surely} \\ \bar{\epsilon} = \frac{1}{100} \sum_{k=1}^{100} \epsilon_k &\approx 0 \end{aligned} \quad (13)$$

As a consequence of that, the investment is always equal to 0.5, as shows equation (14).

$$i(\theta_{\text{RndC}}) = 0.5 + \bar{\epsilon} \cdot \frac{\theta_{\text{RndC}}}{20} = 0.5 \quad (14)$$

The non-existent influence of Random-Continuous' parameter is traced back to this fact.

5.2 Random-Discrete vs. AlwaysSame

Random-Discrete, as a mixed strategy, behaves equitably as AlwaysSame when taking the average of all the generated surfaces. In the simulation, all games were played 100 times and taken the average. Having the surface only generated once, results in a different-looking surface as Figure 5 reveals.



(a) Absolute-gain surface of Random-Discrete.

(b) Absolute-gain surface of AlwaysSame.

Figure 4: One iteration of the pbICPD.

Figure 5a and 5b show the gained points of Random-Discrete and AlwaysSame when playing the pbICPD solely once. A more coarse surface is visible. This joltiness is due to the fact that Random-Discrete uses probabilities after all. Where as AlwaysSame submits always a determined investment.

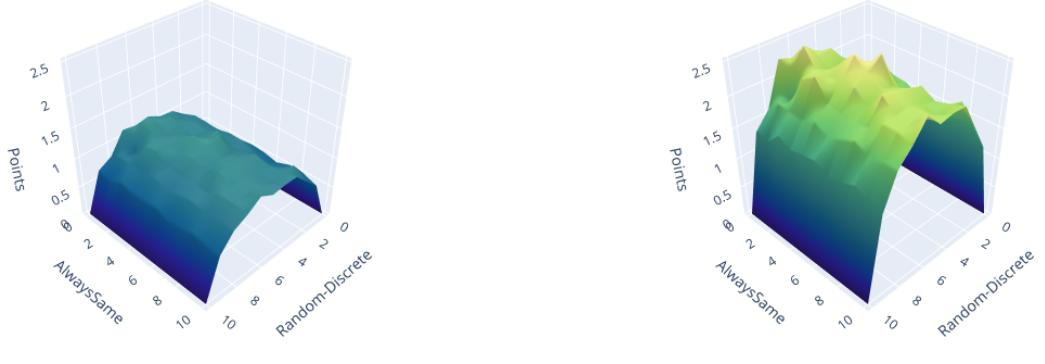
The surface of Random-Discrete is a bit evener. This fact can be explained by examining the pay-off equations: $P_A = y - c \cdot x + c$ and $P_B = x - c \cdot y + c$
 x being the previous investment of Random-Discrete and y is equal to the previous investment

of AlwaysSame. x is relatively random, depending on Random-Discrete's parameter and y is a constant over one game. Subtracting only a fraction of the random investment pays off a more constant amount of points.

To calculate how much Random-Discrete digresses from its constant-behaviour opponent, we can use the standard deviation for every game. This means we can display a new surface plot which will describe the deviation of the investments compared to the average. Using equation (15), the Standard Deviation (SD), Figures 5a and 5b are generated.

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (15)$$

To conclude, choosing AlwaysSame, especially parameter from 2 to 8, as one's own behaviour



(a) Standard Deviation surface of Random-Discrete.

(b) Standard Deviation surface of AlwaysSame.

Figure 5: Standard Deviation surfaces.

means a high risk of loosing or winning as shows Figure 5b. This deviation from the average is at every point twice as much with AlwaysSame than Random-Discrete.

This can be seen in these two following surface plots.

The SD defines the boundaries in which 68.2% of the samples lie within. So, the chance of getting a pay-off within this volume is equal to 68.2%.

Remarkably, the probability of loosing or winning is higher with AlwaysSame than Random-Discrete. This means that the Random-Discrete' pay-offs are more foreseeable than those of AlwaysSame.

5.3 Random-Discrete vs. Adapt-Discrete

Glancing at the surface against Adapt-Discrete, this stripe is remarkable. The existence of this stripe and the factor of Adapt-Discrete's parameter from range 0 to 2 are clarified.

$$i(\theta_{AdpD}) \sim \begin{cases} 1 & \text{if } i_1 + s(\theta_{AdpD}) \geq 0.5 \\ 0 & \text{if } i_1 + s(\theta_{AdpD}) < 0.5 \end{cases} \quad (16)$$

$$s(\theta_{AdpD}) = \frac{\theta_{AdpD}}{5} \cdot (\bar{i}_1 - i_1) \quad (17)$$



(a) Standard Deviation \pm Average surface of Random-Discrete.

(b) Standard Deviation \pm Average surface of AlwaysSame.

Figure 6: Standard Deviation \pm Average surfaces.

As Adapt-Discrete is defined, its investment is rounded to the nearest integer as expression (16) demonstrates. To get from investment 0 to 1 or vice versa, its investment has to get over a certain threshold being equal to 0.5. In the first round, Adapt-Discrete contributes a 1. I.e. to surpass the threshold, $s(\theta_{\text{AdpD}})$ must be under -0.5. To calculate a general formula that calculates the opponent's investment that is necessary to exceed the threshold in the first round given the parameter, it is required to solve for \bar{i}_1 .

$$\begin{aligned}
 s(\theta_{\text{AdpD}}) &< -0.5 \\
 \frac{\theta_{\text{AdpD}}}{5} \cdot (\bar{i}_1 - 1) &< -0.5 & | \cdot \frac{5}{\theta_{\text{AdpD}}} \\
 \bar{i}_1 - 1 &< -\frac{2.5}{\theta_{\text{AdpD}}} & | + 1 \\
 \bar{i}_1 &< -\frac{2.5}{\theta_{\text{AdpD}}} + 1
 \end{aligned}$$

Figure (7) exhibits this the graph of this formula. The curve of this figure defines the area under it. Every point on this curve is not included. It implies every possible investment of the opponent in order Adapt-Discrete to surpass the threshold.

It can be seen that the opponent's investment must be below zero to pass over the threshold value if $0 \leq \theta_{\text{AdpD}} \leq 2$. A contribution below 0 is impossible. To conclude, if Adapt-Discrete's parameter is either 0, 1, or 2, the shift required to overstep the marginal value cannot be achieved. Thus, Adapt-Discrete is equivalent to AlwaysCooperate in cases where its parameter is equal to 0, 1, or 2.

This explains the gainfulness of this stripe in the absolute-gain surface of Random-Discrete and thus the detriment of that stripe in the other absolute-gain surface. Adapt-Discrete is exploited since it cannot react because it only submits full cooperation in this area of the surface.

Coming to the next attribute of this plot; the evenness of the rest of the surface. The other part of the surface is first, not influenced by the parameter of Adapt-Discrete and second, has a little inclination. The first point can be answered with the already proven fact that Random-Discrete parallels AlwaysSame. Adapt-Discrete mimics the opposing strategy if it belongs to

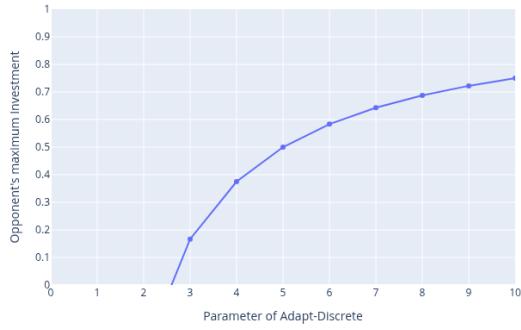


Figure 7: *he*

the rigid category since the shift is equal to zero after the first few round, as will be discussed later.

The second quality can be justified by becoming clear about the fact that Random-Discrete will range its contributions from 0 to 1 as its parameter increases. But, a little detail can be observed concerning the inclinations of the two areas which do not coincide with each other. The subarea of the absolute-gain surface of Adapt-Discrete is a smidgen steeper than the other. This item substantiated by reason. Adapt-Discrete's first investment is always full cooperation. Random-Discrete with its parameter set to 0 acts like AlwaysDefect. Adapt-Discrete is thus completely exploited in the first round. Random-Discrete's parameter equal to 10, in contrary, will let Random-Discrete act like AlwaysCooperate. Hence, Adapt-Discrete is not at all exploited in the first round, and will never be during the game. To deduce, the nearer Random-Discrete's parameter is to zero, the more Adapt-Discrete is exploited in the first round. As a consequence of that, Adapt-Discrete will have a minor point deficit at most of the games.

5.4 Random-Discrete vs. Adapt-Continuous

A same idea can be used to explain the similar appearance of the surfaces against Adapt-Continuous compared to the ones against Adapt-Discrete. The following line chart demonstrates the investments of Adapt-Continuous in one ICPD. I assumed that Random-Discrete behaves exactly the same as AlwaysSame due to discussed reasons (see sections 5.1 and 5.2). Adapt-Continuous pursues getting on the same level of investment as the opponent.

Random-Discrete's parameter equals 5 in Figure 8.

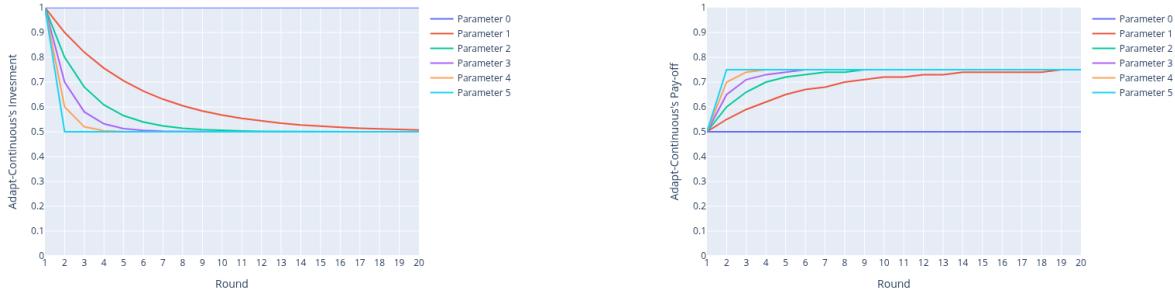
The function to calculate each graph can be derived from the equation of Adapt-Continuous that calculates the investment. Equation (18) describes this process recursively. k indicates the current round and θ is the parameter of Adapt-Continuous. \bar{i} has no numbering because it is constant.

$$i_k = i_{k-1} + \frac{\theta}{5} \cdot (\bar{i} - i_{k-1}) \quad (18)$$

It is not possible to calculate a graph with the recursive form of this equation. The explicit version is required as equation (19) implies.

$$i_k = \bar{i} + (1 - \frac{\theta}{5})^k \cdot (i_0 - \bar{i}) \quad (19)$$

The higher the parameter of Adapt-Continuous, the faster it can adjust its investment. This means that Adapt-Continuous, when the parameter is small, will be more exploited in the first



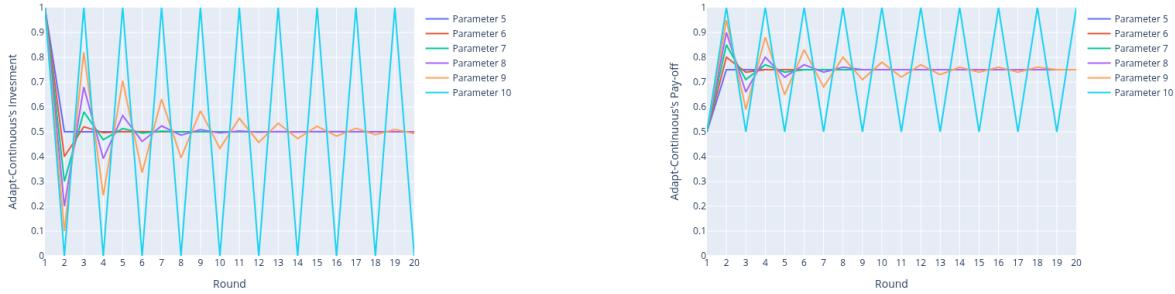
(a) Investments of Adapt-Continuous with parameters ranging from 0 to 5.

(b) Pay-offs of Adapt-Continuous with parameters ranging from 0 to 5.

Figure 8

several rounds than a Adapt-Continuous with a big value of its parameter.

The question arises what the behaviour looks like if Adapt-Continuous' parameter has a value over 5. Figure 9a demonstrates the answer to this question.



(a) Investments of Adapt-Continuous with parameters ranging from 5 to 10.

(b) Pay-offs of Adapt-Continuous with parameters ranging from 5 to 10.

Figure 9

The investments always overstep the value of the opponent's investment. The lines of parameters from six to nine tend to that value. The ultimate accumulated points they receive is, however, indifferent to that of the behaviour of parameter equal to ten. This is a consequence of the fact that the pay-offs of these investments cancel out. The points that are gained more is the counterpart of the deficit of points in the next round.

The peaks and the lows at parameter from 5 to 10 converge into a common average. This average can be seen in Figure 10.

To summarise, if $0 \leq \theta \leq 5$, their investment converge to the opponent's contribution since they can adapt faster the higher the parameter. If $5 \leq \theta \leq 10$ however, the investments equal the opponent's contribution since their overpassing the value of the opponent's investment results in an average being equal to the opponent's investment.



Figure 10: Accumulated points of Adapt-Continuous. Parameters ranging from 0 to 10.

5.5 Random-Continuous vs. Adapt-Discrete

The existence of two different plateaus is noteworthy.

Figures 11 and 12 attempt to justify this attribute of the absolute-gain surface of Random-Continuous and Adapt-Continuous.

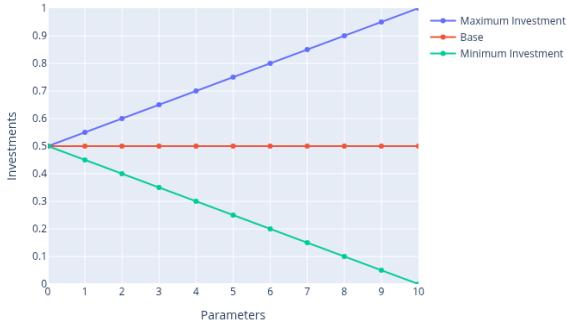


Figure 11: Investments of Random-Continuous. Parameters from 0 to 10.

Every investment of the opponent above the line of 'Maximum Investment' in Figure 12 results in an exploitation of Adapt-Discrete. This comes down to a simple fact. The case of Adapt-Discrete's parameter equal to 3 will be taken as a sample. The same idea applies to parameter of Adapt-Discrete equal to 4 and 5. Parameters equal to either 0, 1, or 2 necessitate a different approach (see 5.3). The height at parameter 3 lies between the lines RndC 7 and RndC 6. The line underneath describes that this and every greater parameter of Random-Continuous does not exploit Adapt-Discrete. Adapt-Discrete can surpass the threshold of 0.5 because the investment of Random-Continuous 7 and every further one submits a lower investment (with a 50% change) and thus defects completely. The line above indicates the first parameter 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 of Adapt-Discrete. To review this fact, please see section 5.3.

The reason for the existence of the peak at coordinates $(5, 10, z)$ has not yet been clarified. Figure 13 demonstrate this aspect. At Adapt-Discrete 10, the lines of Random-Continuous 5

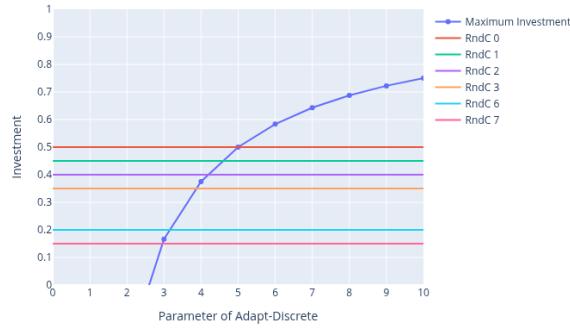


Figure 12: Maximal investments of Random-Continuous to surpass threshold.

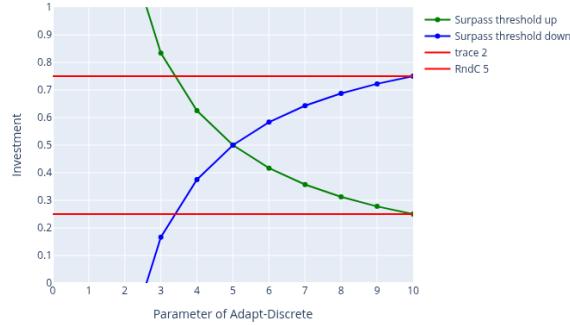


Figure 13: Influence of parameter of Random-Continuous 5 to Adapt-Discrete parameter 10.

intersect perfectly. This means that surpassing the threshold down is happening 50% of the submissions of Random-Continuous since only one line is below the blue point. But, surpassing the threshold upwards is always the case since the two lines are the at the same height or higher than the green line at parameter of Adapt-Discrete equal to 10. So, Adapt-Discrete will go down 50% of the times 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 the probability of Adapt-Discrete cooperating fully is high. And with this comes the high probability of being exploited. Every other combination of parameters of Adapt-Discrete and Random-Continuous prevent this situation from happening.

6 Conclusions and Outlook

- Summary and Quintessence

First, we looked at the features of each surface. To conclude, both Adapt-Discrete and Adapt-Continuous did very bad against Random-Discrete and Random-Continuous. AlwaySame did equally well as bad against both random strategies. The random strategies that played against themselves also scored fairly well.

Some areas in this simulation have not been explored due to factors concerning the time. These areas apply to parameters such as the coefficient c in the pay-off equations. Or, the factor of noise, which changes the investments slightly and thus create a simulation

where misunderstanding come to play. Another limitation is the applicability to the real world. To verify these results, further researches must be made. These further researches concern empirical social or political researches.

To answer the initial question. What behaviour is best when encountering an erratic behaviour. As already mentioned, the interpretation into the real world are questionable. But, from what we can conclude from this simulation is the following. First, responsive/adaptive strategy could beat a random/erratic strategy. The maximum they could get is mutual cooperation which is best for the total welfare but not ideal when one cares about advantage over the other. Second, the rigid strategy could explore every outcome. It could exploit the erratic strategy or lose against it. But, it could also cooperate and defect mutually. Ultimately, the random strategies itself could also get every outcome from the simulation. This means when encountering a erratic strategy, it is best to be rigid or random oneself. Being adaptive is only a loss.

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 Appendices

A.1 Appendix A. 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

A.2 Appendix B.

A.3 Appendix C.

B Declaration of Integrity