

# Research of Parameters in Strategies of the Iterated Continuous Prisoner's Dilemma

Adrian Hossner

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Theoretical Foundations</b>	<b>3</b>
2.1	Prisoner's Dilemma . . . . .	4
2.2	Iterated . . . . .	4
2.3	Continuous . . . . .	5
2.4	Additional variants . . . . .	6
2.5	Current Findings . . . . .	6
<b>3</b>	<b>Methods and Implementation</b>	<b>6</b>
3.1	Study design . . . . .	6
3.2	Implemented Strategies . . . . .	7
3.3	Implementation Details . . . . .	8
<b>4</b>	<b>Results</b>	<b>9</b>
4.1	Random-Discrete . . . . .	12
4.2	Random-Continuous . . . . .	13
<b>5</b>	<b>Discussion</b>	<b>14</b>
5.1	Random-Discrete vs. AlwaysSame . . . . .	14
5.2	Random-Discrete vs. Adapt-Discrete . . . . .	16
5.3	Random-Discrete vs. Adapt-Continuous . . . . .	17
5.4	Random-Continuous vs. Adapt-Discrete . . . . .	19
<b>6</b>	<b>Conclusions and Outlook</b>	<b>20</b>
<b>7</b>	<b>References</b>	<b>20</b>
<b>8</b>	<b>Appendix</b>	<b>20</b>

# Abstract

## 1 Introduction

- Interactions  
implication range is wide:  
from two persons in a relationship to countries in alliance  
from two impala's allogrooming to zebras warning the herd about a lion
- Key hook:  
helping someone with possible cost – cooperation  
being selfish but breaking trust – defection
- Game Theory (optional, not needed for understanding)  
mathematical framework to investigate interactions
- model PD:  
single interaction  
already well analysed  
Nash found solution/equilibrium
- Iterated:  
more interesting, insightful  
strategies
- non-suited variants:  
Axelrod's Tournament  
Evolution
- continuous:  
more complex, accurate
- parameters:  
determines behaviour  
create surface  
the part that has not been experimented
- overall and difference:  
overall: wealth of population  
difference: competence between individuals

## 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 Prisoner's Dilemma (PD) (2.1), the introduction of iterations as an extension (2.2) and the approach of submitting continuous values (2.3). Furthermore, relevant information of published researches are provided as it seems helpful (2.5).

## 2.1 Prisoner's Dilemma

The Prisoner's Dilemma (PD) is a famous model in Game Theory. It was invented by Merrill Flood and Melvin Dresher in 1950. 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 utilises one of its methods. They interrogate them separately without giving them a chance to coordinate beforehand. The convicts receive two options. On 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. Naturally, the police rewards the act of revealing the accomplice's delinquency over exercising the right of solidarity. In case mutual trust and thus no confession took place, the twain separately receives an imprisonment of three years. If both confess, either obtains five years in prison. Nevertheless, by taking this course of action, this outcome is not the only existing. If one chooses to confess while the other remains silent, the informer is released, whereas the betrayed must devote ten years in prison. Remaining silent is considered as cooperation. Whereby defection is the opposite, meaning breaking confidence and therefore state. The following matrix visualises the pay-offs in this game in a comprehensive way:

	C	D
C	3, 3	0, 10
D	10, 0	5, 5

C represents cooperation and D corresponds to defection. The dilemma consists of the following.

The act of confessing appears alluring since there is a chance of eluding imprisonment. Nonetheless, it is possible to receive five years in prison depending on the confederate's decision. If the second criminal also confesses, both get five years in prison. This is the most inferior outcome for both which could have been avoided as either would have stayed silent.

John Nash, a mathematician which made great contributions in Game Theory, has proved the most reasonable course of action to be confessing. The Nash-equilibrium is introduced. It defines a stable state in which both would not reconsider their own judgement although each of them would be aware of the other's decision.

This game composes a one-time interaction excellently in a mathematical and analytical perspective. The entitlement of the variables derives from the pay-off they suggest.

$T$ : Temptation,  $R$ : Reward,  $P$ : Punishment,  $S$ : Sucker's pay-off

The pay-offs in the matrix can be altered as long as these inequalities stay valid.

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

$$T > R > P > S$$

## 2.2 Iterated

The Iterated Prisoner's Dilemma (IPD) is a game which extends the PD. As the name suggests, the PD is played a number of times sequentially. To refer to the number of rounds played,  $n$  is used as the denotation. The gained pay-offs accumulated over the last  $n - i$  rounds are summed where  $i$  represents the current round. A strategy is an autonomous player which has its own behaviour. Such a strategy has the whole history of the previous PDs at hand to generate the decision for the next round. These decisions are calculated using conditions and probabilities that will be applied on the given data. In this, and in the following variants, the analogy to the real world is changed. Maximising one owns points is preferred. Since a strategy strives to gain as many points as possible, the pay-off matrix must be altered accordingly.

	C	D
C	5, 5	10, 0
D	0, 10	3, 3

And to generalise this pay-off matrix as well, the numbers can be replaced with variables. The concept is used for  $T$ ,  $R$ ,  $P$ , and  $S$ , provided that the four inequalities are satisfied.

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

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

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

Inequality (2) is necessary to avoid receiving greater pay-offs by alternating cooperation and defection. The Nash-equilibrium in the IPD differs from the that in the PD. This implies that mutual cooperation yields long-term benefits. Playing cooperation increases trust whereas defection allows exploitation. Maximum pay-off during all the games is achieved only by consistent exploitation ( $n \cdot T$ ). Since persistent defection leads to mutual defection as a response, the practical attainment of the maximum points is severely limited. To conclude, both strategies benefit the most with mutual cooperation.

A problem arises in the last round. Since cooperation is only played to conserve trust, the reasonable action in the last round is to defect as trust has no use beyond the game. If both strategies act rational, mutual defection will be played in the last round. And because both are aware of this fact, there is no logical reason to cooperate the second last round. This will end up in an inductive defection behaviour for both strategies the whole game. So, to prevent constant mutual defection, the number of rounds played is an information that needs to be hidden from both strategies. As a consequence of that,  $n$  is a random number whose limits can be defined.

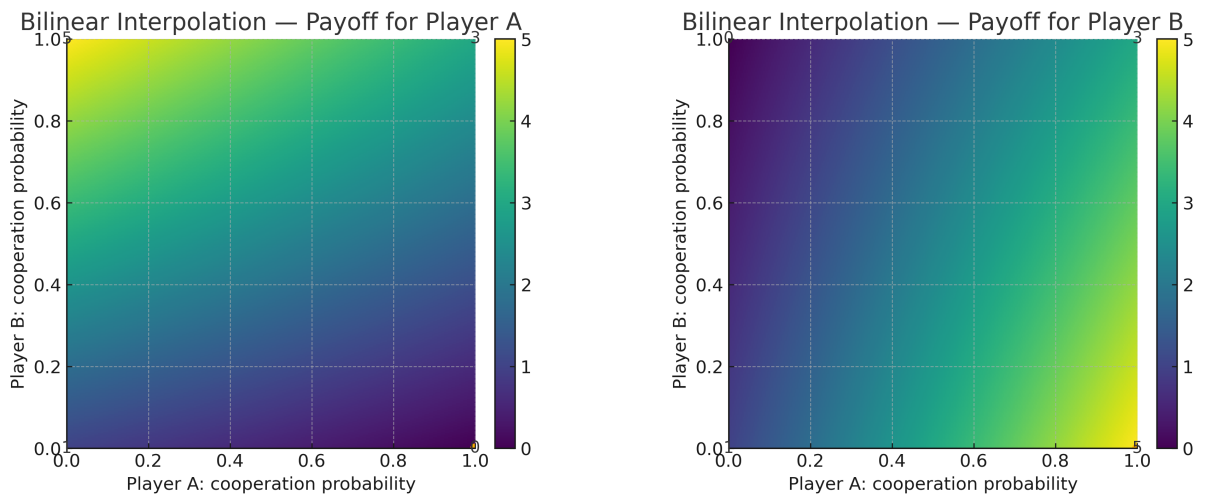
This extension of the PD allows to simulate a relationship of two individuals.

### 2.3 Continuous

In this variant, strategies are not limited to either cooperate or defect; instead, they can chose any level of cooperation between 0 and 1. It is thus named Iterated Continuous Prisoner's Dilemma (ICPD) This cooperation level, called investment or contribution, ranges from zero (defect) to one (cooperate). The continuous form offers enhanced precision and mathematical applicability.

Two main approaches will be presented and evaluated for practicality in this thesis. First, an intuitive extension of the iterated form. And second, a way that simplifies mathematically analysability.

A similar concept can be used to calculate the pay-offs given the investments of both strategies. This concept is referring to a matrix-like pay-off system. To include continuous values, a mathematical procedure is 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. Please see Figure 1. The equation used to calculate this heat map can be found in the Appendix.



**Figure 1:** *Pay-off heat maps in the ICPD.*

The second approach utilises only equations instead. This pay-off system of [sciencedirect.com] is formed by:

$$p_A = y - c \cdot x + c$$

$$p_B = x - c \cdot y + c$$

$p_A$  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 and describes the cost-to-benefit ratio. The pay-off is highly influenced by the opponent's investment. Only a fraction of one owns contribution is subtracted,  $c$  being the coefficient. The minimum pay-off one strategy can collect is then  $-c$ . To avoid negative pay-offs and enhance subsequent convenience, a third addend is appended,  $c$ . So, the least points that can be gained is 0. The maximal points, which can be earned by exploitation, is  $1 + c$ .

## 2.4 Additional variants

There exist several extensions of the PD that will remain untouched in this thesis, one of which is noise, meaning that investments are altered with a certain probability. This noise is introduced to simulate misunderstandings.

Another advancement being the alternating form. In this variant, investments are contributed in succession rather than concurrently as in the conventional simultaneous model.

The evolutionary research domain surveys the success of strategies in a world where natural selection takes place. This implies that losing strategies vanish, while successful ones replicate.

## 2.5 Current Findings

Axelrod's Tournament In the simpler variants, e.g. the IPD without noise, it has been proved many times over that the strategy Tit-For-Tat is the most successful one. [sources] The success of strategies, however, can vary depending on certain conditions such as the influence of noise or different pay-off systems where continuous investments can be submitted. So, the very specific area of the simultaneous ICPD, has not been very well explored. ...

# 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 strategies 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 (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.

$$i(\theta_{\text{AdpC}}) = i_1 + s(\theta_{\text{AdpC}})$$

$$s(\theta_{\text{AdpC}}) = \frac{\theta_{\text{AdpC}}}{5} \cdot (\bar{i}_1 - i_1)$$

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  $\theta_{\text{AdpC}}$  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.

- As I 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.

$$i(\theta_{\text{AdpD}}) = \lfloor i_1 + s(\theta_{\text{AdpD}}) \rfloor$$

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

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 of 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'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.

- 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:

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

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

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

$s$  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  $\epsilon$ .

- 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 Implementation 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  $100 \cdot 20 \cdot 1210 = 2420000$  procedures of the continuous PD.

One encounter of the  $11 \cdot 11$  parameter-based ICPD (pbICPD) is displays 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 farther 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.

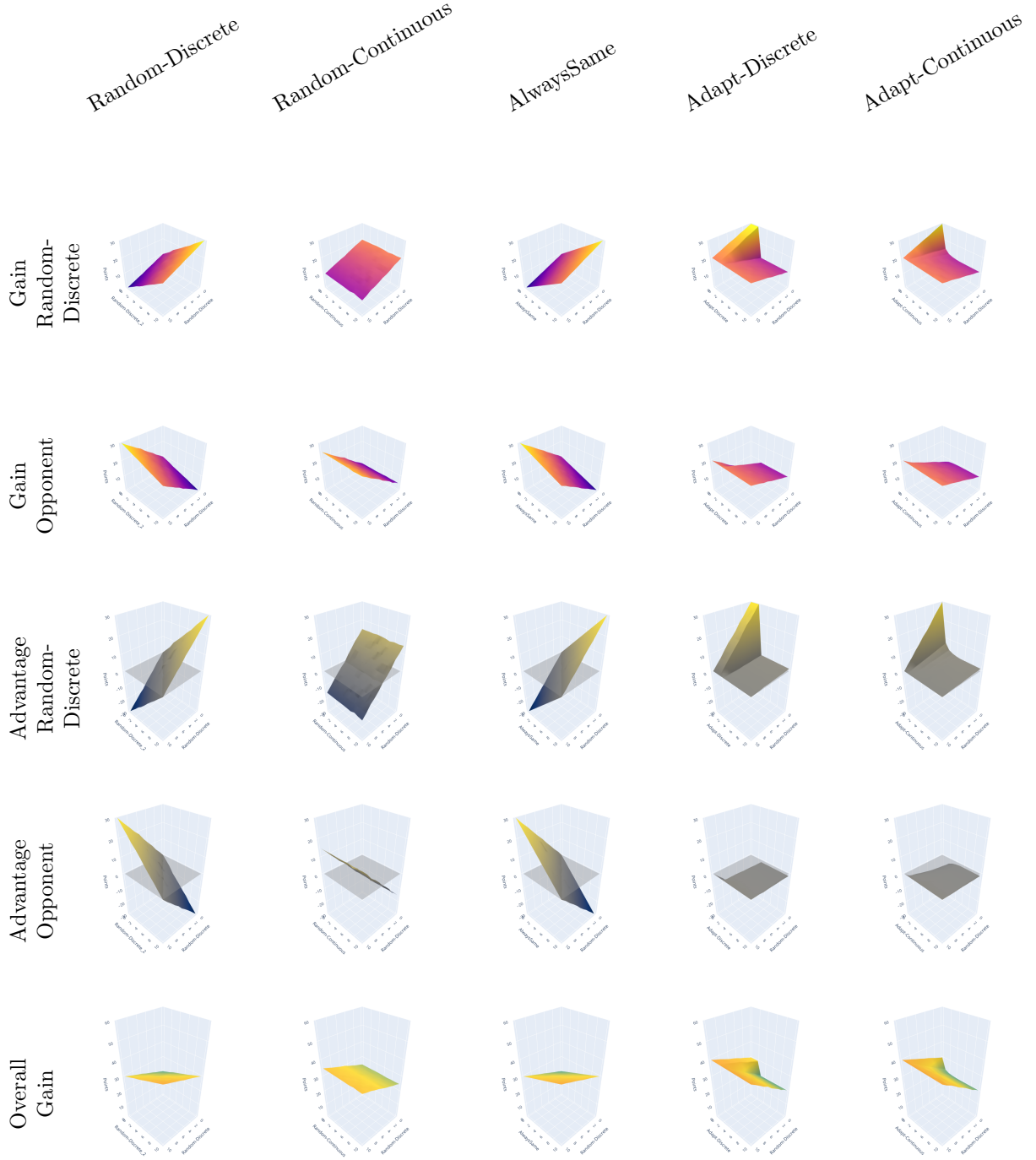


## 4 Results

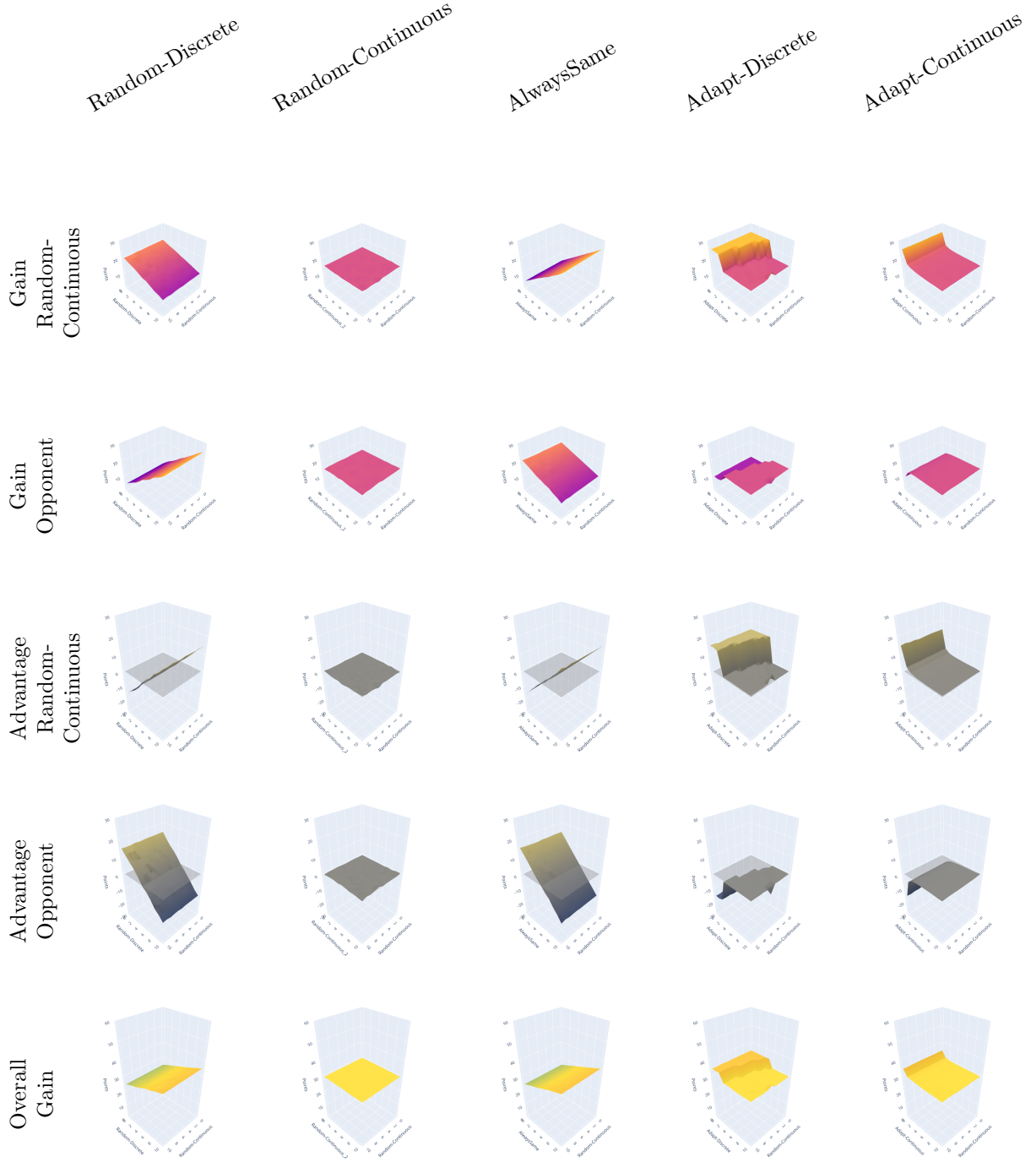
In Figures 3 and 4, the surfaces for the Random-Discrete and Random-Continuous simulations are portrayed. There are five surfaces for each interaction that are generated from one pbiCPD. The x-axis, the horizontal one in the bottom-right, display the parameter of the main strategy, that is, either Random-Discrete (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.



**Figure 2:** *pbICPD of Random-Discrete*



**Figure 3:** *something*

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 bespoke 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 bespoke 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 Random-Discrete vs. AlwaysSame

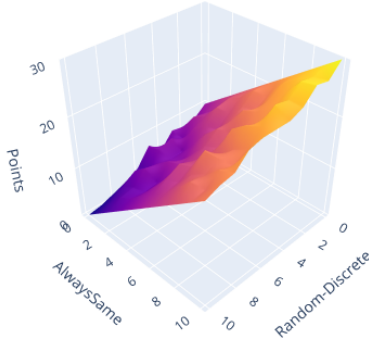
Random-Discrete, as a mixed strategy, behaves equitably as AlwaysSame when taking the average of all the generated surfaces. In the simulation, all games were played 100 times and taken the average to plot the surfaces. Having the surface only generated once, results in a different-looking surface area.

These surfaces show the gained points of Random-Discrete and AlwaysSame when playing the ICPD only once. A rougher surface is visible. This is due to the fact that Random-Discrete uses probabilities after all. Where as AlwaysSame submits always a determined investment.

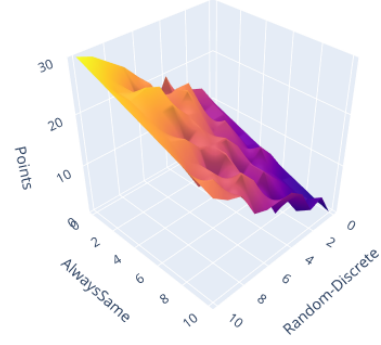
The surface of Random-Discrete is a bit evenner. This fact can be explained by examining the pay-off equations:  $P_A = y - c \cdot x + c$  and  $P_B = x - c \cdot y + c$

$x$  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 in one game always constant. Subtracting only a fraction of the random investment pays off a more constant amount of points.

To calculate how much Random-Discrete derived 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



(a) first subfig



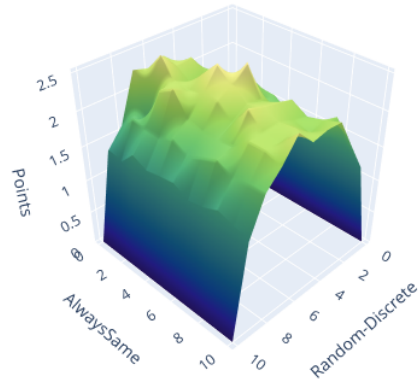
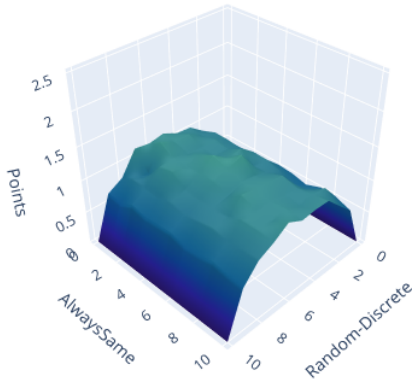
(b) second subfig

**Figure 4:** *here*

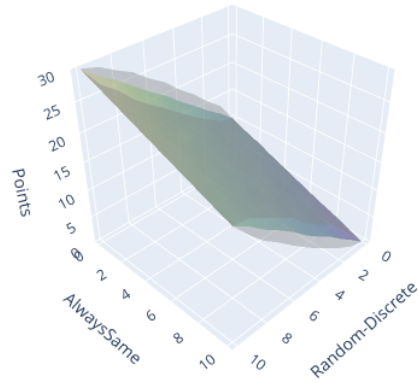
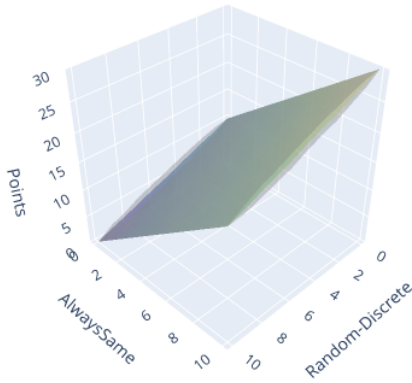
how much off the investments were comparing to the average. This formula is being used, the standard deviation.

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

We get the following surface:



To conclude, choosing Random-Discrete, especially parameter from 2 to 8, as one's own behaviour means a high risk of loosing. But, this risk of deviating from the average is at every point twice as much with AlwaysSame than Random-Discrete. This means that when playing like Random-Discrete, one risks getting  $\pm$  the standard deviation to the average. But, one is applying twice this risk to the opponent which is in this case AlwaysSame. This can be seen in these two following surface plots.



(Why chosen AlwaysSame?: Because AlwS and Rnd-D seem alike, but aren't)

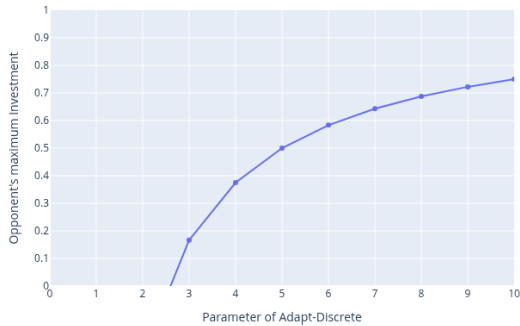
(3D object in which the prob. is 66% of being in there in one match.)

(Why are Standard Deviation surfaces same but different scales?)

## 5.2 Random-Discrete vs. Adapt-Discrete

When you look at the surface, this stripe is very remarkable. We will look at why this stripe exists and why is only exists from AdaptD's parameter 0 to 2.

As Adapt-Discrete is explained, its investment is rounded to the nearest integer. So, in order to get from investment 0 to 1 or vice versa, its investment has to get over the threshold of 0.5. To demonstrate this point, the following graph is used. It shows what the opponent's investment has to be at least or at most to surpass this threshold.



The derivation to calculate this function is the following can be found in the Appendix. It can be seen that if Adapt-Discrete has the parameter from 0 to 2, the opponent's investment must be lower than 0 which is not possible. This explains why Adapt-Discrete with its parameter set to either 0, 1 or 2 behaves exactly like AlwaysCooperate since it cannot surpass the threshold of 0.5.

Coming to the next attribute of this surface; 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. This question can be answered with the fact that Random-Discrete behaves, as mentioned many times, like AlwaysSame. The two slopes look nearly identical because Adapt-Discrete will also behave rigidly against a rigid strategy. A little detail is that the slope of Adapt-Discrete has a very little larger inclination. This means this slope is a little bit steeper. This mysterious can be answered by the fact that Adapt-Discrete always starts off the game with full cooperation. So, the first time, Adapt-Discrete gets

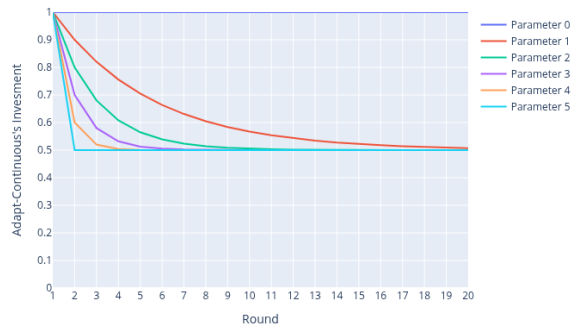


a little bit more exploited the nearer the parameter of Random-Discrete gets to 0. Because at parameter 0, Random-Discrete is like AlwaysDefect and gets thus fully exploited. Random-Discrete's parameter 10, in contrary, it like full cooperation so it get not exploited. This means that it gets more exploited the nearer Random-Discrete's parameter gets to 0.

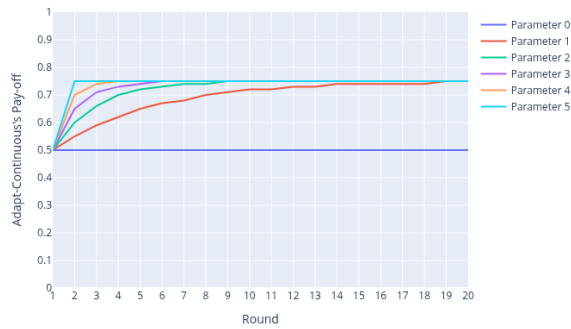
### 5.3 Random-Discrete vs. Adapt-Continuous

A same idea can be used to explain why this surface has a similar appearance like the one before. The following line chart demonstrates the investments of Adapt-Continuous in one game in the ICPD. I assumed that Random-Discrete behaves exactly the same as AlwaysSame. So, Adapt-Continuous pursuits to get to the same level of the parameter of Random-Discrete.

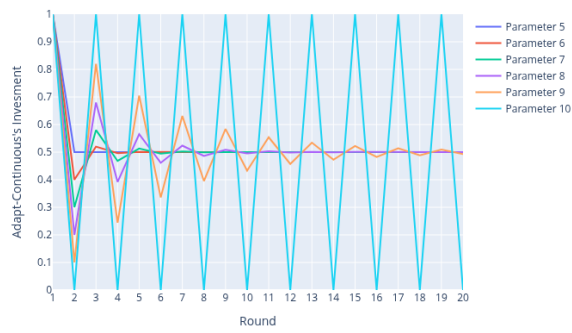
Random-Discrete is, in this chart, at parameter equal to 5.



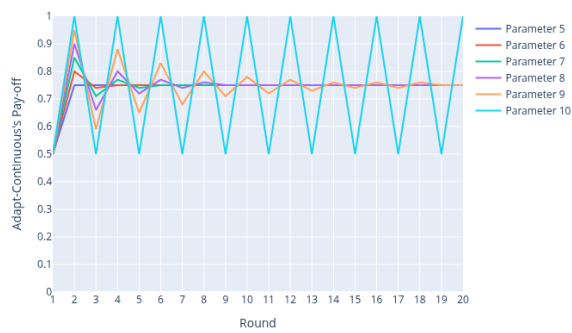
You can see that the higher Adapt-Continuous's parameters are, the faster they can adapt. This means that an Adapt-Continuous with a low parameter gets more exploited in the first few rounds than a Adapt-Continuous with a high parameter.



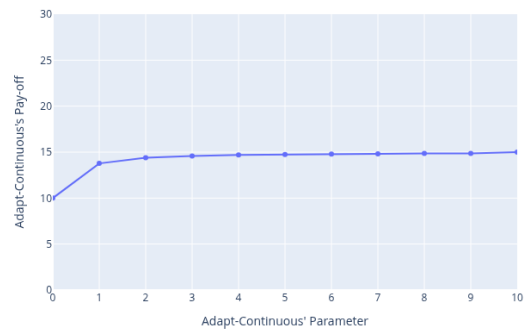
The question arises what the behaviour looks like if Adapt-Continuous has its parameter over 5. The next plot shows exactly this.



The pay-off's of these investments, however, cancel out.



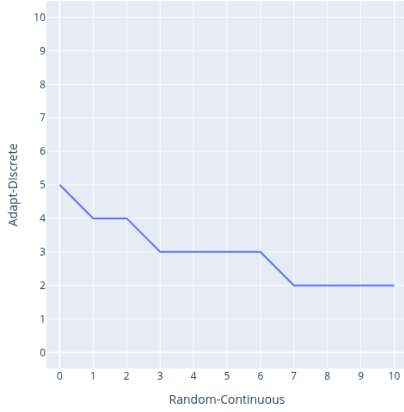
The peaks and the lows at parameter from 5 to 10 converge into an average. These averages can be seen in the next plot.



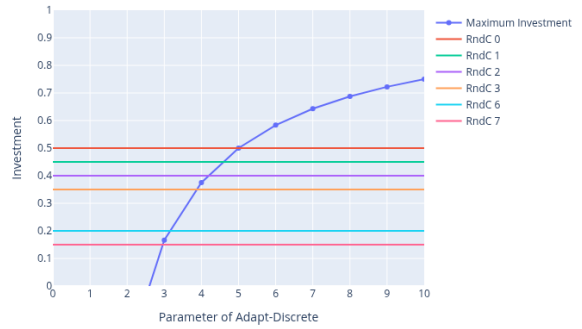
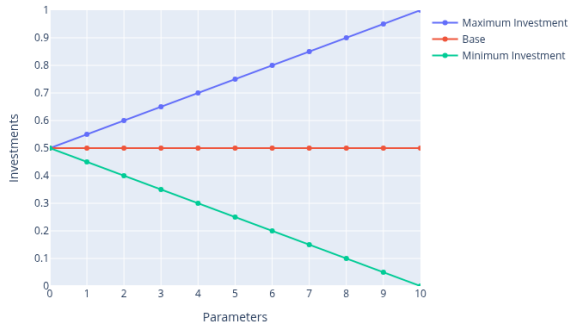
To summarise, from parameter 0 to 5, the pay-offs get better since they can adapt better. But, from parameter 5 onwards, the pay-offs stay exactly the same since their pay-offs converge to an average.

## 5.4 Random-Continuous vs. Adapt-Discrete

It is very interesting why such two plateaus exist.

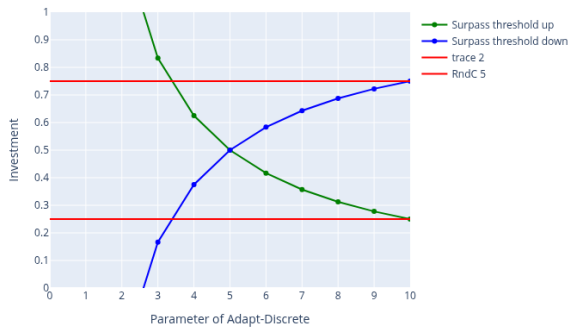


These plots describe exactly this.



Everything under the line of the first chart is where Adapt-Discrete gets exploited. This means that Adapt-Discrete only cooperates and Random-Continuous, naturally, only submits 0.5. This comes down to a simple fact. We look at the case of Adapt-Discrete parameter 3. The same idea applies to parameters 4 and 5 and technically 1 and 2 as well. The height at parameter 3 is between the lines RndC 7 and RndC 6. The lower line describes that this parameter of Random-Continuous and every further one does not exploit Adapt-Discrete. Adapt-Discrete can surpass the threshold of 1.5 because the investment of Random-Continuous 1 and every further one submits a lower investment (with a 50% change) and thus also defects. The line above describes the first parameter that exploits Adapt-Discrete. The investment of Random-Continuous must be below the point in order Adapt-Discrete to be able to jump from 1 to 0. So, every parameter that is lower than 6, exploits Adapt-Discrete.

Only one aspect has not been clarified. It has not been discussed why there is a one-time peak at the coordinates (5, 10)  $\rightarrow$  (Random-Continuous 5, Adapt-Discrete = 10)



This graph is only used to demonstrate that point. At Adapt-Discrete 10, the lines of Random-Continuous

5 intersect perfectly. This means that surpassing the threshold down is happening 50% of the submits 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 at the same height or above the green line at 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 100% of the times. This constellation is only the case with these specific parameters.

## 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. AlwaysSame 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 References

## 8 Appendix