Zero-Determinant Strategies and their performance on Evolutionary Scenarios

75657 - Paulo Gouveia 75694 - Daniel Figueira 79758 - Luisa Santo

***Abstract*— The goal of this project is to understand what Zero-Determinant Strategies are, what makes them different from other strategies, and how well they perform. Our focus will be on the Iterated Prisoner's Dilemma and more specifically in trying to understand how well do Zero-Determinant Strategies fare in evolutionary scenarios where players are allowed to evolve. In order to test their performance we intend to create a tournament similar to Axelrod's Tournaments (described in section [5]). For that purpose, we will be using the "Axelrod" python library.**

I. Introduction

A. The Prisoner's Dilemma

The Prisoner's Dilemma is a famous decision analysis paradox that has been used as a study case since the 1950's, in which two individuals acting in their own self-interest end up choosing a course of action that leads to a suboptimal outcome [4].

This paradox got its name from the scenario where two criminals (A and B) are being interrogated in separate rooms, with no means of communication, and are both given the chance to implicate the other in the crime, in order reduce their own crime sentence. Each of them must then choose to either stay silent (cooperate with his partner) or sell his partner to the cops (defect) without knowing what his partner will choose to do. The consequences, which are known to them "a priori", would then be as are shown in Fig 1.

Given some thought, prisoner A would realize that "if prisoner B confesses, the rational choice would be to confess as well and get 5 years; if prisoner B stays silent the rational choice (which would get me less years in prison) would also be to confess and walk way free", thus concluding that no matter what choice the other prisoner would make, it would still be best for him to sell its partner to the cops (defect), making defecting the dominant strategy in this case (the best strategy regardless of the opponent's choice).

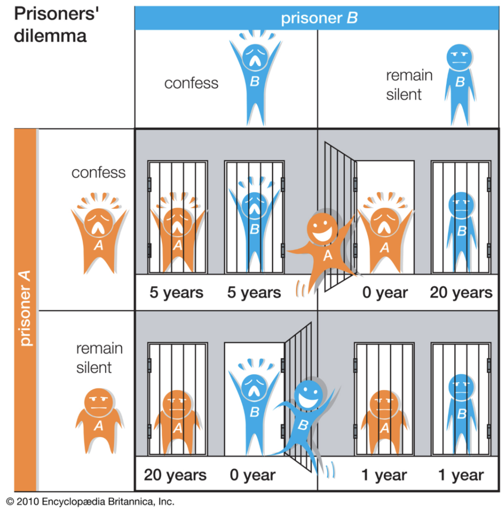

Fig. 1. Possible outcomes of the provisioner's dilemma

Assuming that the other prisoner would also choose to confess, which is the "rational" choice, each of them would get to spend 5 years in jail. The paradox lies in the fact that, although both of them made their choices with their own self interest in mind, in the end the result that they got wasn't the best for them, since they could have both gotten way with only 1 year if both of them had chosen to stay silent (cooperate).

B. Iterated Prisoner's Dilemma

While the Iterated Prisoner's Dilemma is simply a normal prisoner's dilemma played successively by the same participants, it differs from the original scenario because each participant can take into account the opponent's previous choices when making its decision.

In this case, the knowledge that you'll be facing the same opponent again in the future becomes an incentive for cooperation. Theoretically, in an infinite Iterated Prisoner's Dilemma cooperating is considered the rational choice. However, since in reality there are no infinite games, there will eventually be a last round where defecting should lead to a better result. The question then becomes ”How soon should I stop cooperating and start to defect?”, which ultimately leads us back to the same paradox, where two individuals that acted with their self-interest in mind end up not achieving the optimal possible result.

C. **Axelrods Tournaments**

Taking the Iterated Prisoner’s Dilemma into consideration, and using points instead of years spent in jail, it’s easy to imagine it as game between two players where each player would try to amass the biggest amount of points, over a finite number of rounds.

In 1980’s Robert Axelrod took this idea and organized a tournament where each participant would submit an algorithm that implemented a specific strategy. Each strategy would then play against all the other strategies for 200 rounds each, and the strategy that accumulated the highest score in the end would be declared the winner. In total 14 strategies (plus a ”random” strategy) took part in this tournament, ranging from simple strategies such as ”always cooperate” or ”always defect”, to more complex strategies with many lines of code, and in the end the strategy ”Tit-for-Tat” came out on top achieving 504 out of 600 points [6].

Later Axelrod organized a second tournament, this time with a slight change of rules: instead of each game taking exactly 200 rounds there was a small chance that the game would end after each round. Although the expected number of rounds was still around 200, this change prevented each algorithm to know exactly when each game would end. This time, a total of 62 strategies were submitted, yet despite everyone knowing the results of the first tournament, ”Tit-for-Tat” still came out on top.

II. **Classical Strategies**

These so called ”classical strategies” are the most basic or most popular strategies used for the Iterated Prisoner’s Dilemma. Some of them are just ”common sense” with no complexity at all, while others became known for the positive results they obtained [1].

A. Random

A simple coin throw, with 50% chance of cooperating and 50% chance of defecting. It’s often used as a baseline strategy: every other strategy must be at least better than ”Random” to be even worth considering.

B. Always Defect

As the name implies, always choose to defect. The so called ”dominant strategy”, which is always makes the most rational choice, but not always achieves the best results.

C. Always Cooperate

The inverse of always defect - always choose to cooperate. Unless paired with another generous strategy, it does not achieve good results.

D. Tit-for-Tat

The strategy that won both Axelrod’s tournaments. It starts by cooperating and from then on it replicates the opponent’s last move. Although simple, it’s surprisingly effective.

E. Generous Tit-For-Tat

One of the variations of Tit-for-Tat, where there is a small chance to return an opponent’s defect with cooperation. This is useful in preventing cycling defections when playing against another Tit-for-Tat or one of its variations, specially if there is noise involved (small of chance of playing the wrong move).

F. Win-Stay-Lose-Shift

Just like Tit-for-Tat it’s considered a memory-one strategy (choice is dependent on the result of the last round), where it will keep picking the same choice as long as the result corresponds to the maximum payoff possible, and change its choice if it ends up on the losing side.

III. **Zero-Determinant Strategies**

Although the Iterated Prisoner’s Dilemma has been used as a model to study cooperation for a long time, and it’s thus a topic that received a lot of attention throughout the years, in 2012 there was yet another significant discovery on this field [7].

From the set of memory-one strategies, Press and Dyson managed to identify a new subset of strategies that, due to a vanishing determinant in their mathematical proof, were named Zero-Determinant Strategies.

Surprisingly, these new strategies were able to ensure an unfair share of the rewards between both participants, more specifically, they were able to ensure a linear relationship between both player's payoffs (P_1 and P_2):

$$\alpha P_1 + \beta P_2 + \gamma = 0, \tag{1}$$

where α, β and γ are constants, as long as the strategy followed by the player verifies:

$$p(C|CC) = \alpha R + \beta R + \gamma + 1 \tag{2}$$

$$p(C|CD) = \alpha S + \beta T + \gamma + 1 \tag{3}$$

$$p(C|DC) = \alpha T + \beta S + \gamma \tag{4}$$

$$p(C|DD) = \alpha P + \beta P + \gamma \tag{5}$$

where R, S, T, P correspond to the payoffs for each of the possible situations:

Player A \ Player B	Cooperate	Defect
Cooperate	**R**eward	**S**ucker
Defect	**T**emptation	**P**unishment

Fig. 2. Payoff matrix

A. ZD Extortion Strategies

An obvious approach to the Iterated Prisoner's Dilemma with the knowledge of these ZD strategies would be for a player to set the ratio between her and her opponent's payoff as high as possible. This idea lead to the conception of Extortion Strategies which ensure that either the extortioner receives a higher payoff than her opponent, or that both players receive the payoff for mutual defection [9].

Although this kind of strategies as proven to fare distinctly well in winning head-to-head matches, that does not necessarily mean they'll do well in terms of overall score. This because, while extortionate strategies are concerned with receiving a higher payoff than their opponents, non-extortionate strategies would be concerned with maximizing the payoff through cooperation, thus leading to an overall better result.

A particular situation where Extortion Strategies shine it's when they are pitted against an "evolutionary player", who tries to adjust its strategy to maximize its score in response to its opponent's actions. This results in the "evolutionary player" being extorted, and greatly benefiting its opponent while trying to increase its score [8].

B. ZD Generous Strategies

In the other end of Zero-Determinant Strategies are the Generous Strategies, that ensure that both players receive the payoff for mutual cooperation, or that the generous player otherwise receives a lower payoff than its opponent.

Although they may seem to be too generous to succeed in a competitive environment, they have proved to do surprisingly well in round-robin tournaments, were they outperform all the other leading strategies such as Tit-for-Tat, WSLS, etc. Contrary to Extortion Strategies, they never aim to win any game and forgive opponents for defecting, but in the long run they seem to end up amassing a greater payoff than any other strategy [3].

Just like Extortion Strategies, and in fact all Zero-Determinant Strategies, they fare particularly well against "evolutionary players" where they can convince their opponents to cooperate benefiting both players' results.

IV. Evolution

From the application of game theory to biology, a new area of game theory was born called Evolutionary Game Theory. Simply put, this particular area of game theory tries to understand what makes a strategy thrive in a population of players where the players can evolve (change their strategy or aspects of it) [10].

In the case of the Iterated Prisoner's Dilemma an evolutionary scenario is normally achieved by running several iterations of a tournament and at the end of each selecting some players to change their strategies for ones that are achieving better results. As for how this selection is done, it can vary from picking players at random to more complex models. And, since this idea originally came from biology, things like "mutated strategies" can even be taken into equation.

A strategy p is considered evolutionary stable (or robust) if for any two strategies p and q, one of the two relations is satisfied:

$$f(p,p) > f(p,q) \tag{6}$$

$$f(p,p) = f(p,q) \land f(p,q) = f(q,q) \tag{7}$$

where $f(x,y)$ is the long-term payoff of x, pitted against player y.

A. Extortion and Generous Strategies in Evolutionary Scenarios

In an evolutionary scenario where the population can evolve and give up on their strategies in favor of more successful ones, it is proven that Extortion Strategies are not evolutionarily stable since they don't achieve good long-term results, and because they fare particularly bad against other strategies that also try to extort their opponents [2]. Therefore, the greater the ratio of Extortion Strategies on a population the harder it will be for them to thrive.

On the other hand, in this kind of scenario, Generous Strategies are considered to be evolutionarily robust, meaning they won't get replaced by other strategies and will in fact end up replacing other non-cooperative strategies. When it comes to interactions between Generous Strategies, that seems to work particularly well, since both strategies are determined to cooperate. This means that as the ratio of generous strategies in a population goes up, the harder it gets for the other strategies to compete against them.

V. Code Architecture

We wrote our program in Python3 using the Axelrod-Python library which has a comprehensive collection of strategies and tournament variations for reproducible research into the Iterated Prisoner's Dilemma and game theory in general [5]. The developed program supports, with some customization, three modes of operation: a regular tournament, an evolutionary scenario and an interactive mode for a human player.

A. Players and strategies

The developed code has a default list of ten strategies which we considered to be the most interesting to pitch against each other (six well known non-ZD strategies and four Zero-Determinant: two extortionate and two generous). However, it is also possible to add any strategy from the library or remove any from the list.

- Cooperator
- Defector
- Random
- Win-Stay-Lose-Shift
- Tit-For-Tat
- GTFT
- ZD-Extortion
- ZD-Extort2
- ZD-GTFT2
- ZD-Set2

B. Tournaments

In a regular tournament each strategy is considered a player with no repeated strategies. The tournament operates in a round-robin manner: each player plays a match of t turns with every other player and this round-robin process is repeated an r number of iterations.

C. Evolution iterations

For the evolution model, we first create a population consisting o p players per strategy in the list of strategies assuring all strategies start even. Then, for an i number of iterations, a regular tournament is played amongst the population following the same structure described in the previous section and a random player is picked from the population. This player has a probability of 1/PopulationSize to be mutated. If not mutated, a second random player is picked and, if using different strategies, their fitnesses are computed (the payoff normalized per number of turns and matches played). The probability of player one copying player two is given by a Fermi-Dirac distribution calculated with the difference between the two respective payoffs. After the loop of iterations is finished we count the existing strategies in the population.

D. Display of the results

After any run of the program (tournament, evolution or interactive) the user is prompted with the option to save or view the wins distribution, the payoff distribution and/or the payoff matrix.

VI. Results

As expected, in head-to-head matches the extortionate strategy dominates any other strategies that we tested it with except when pitted against other extortionate strategies where they would drag down both of their payoffs. Regarding tournaments, despite their high number of wins, Fig.3, they end up falling behind more cooperative strategies in terms of the overall payoff, Fig.4.

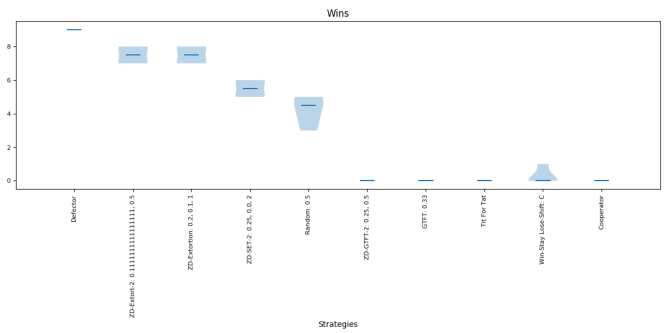

Fig. 3. Wins per strategy in a regular Axelrod tournament.

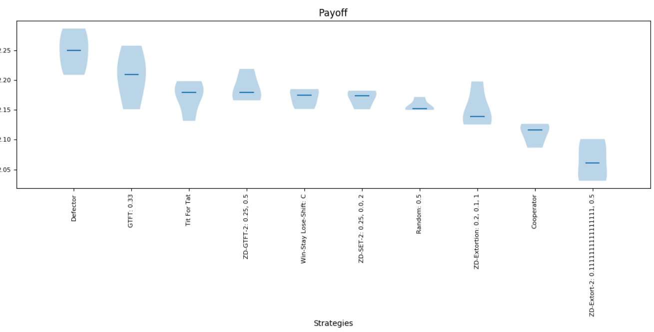

Fig. 4. Wins per strategy in a regular Axelrod tournament.

In order to test the evolutionary performance of the ZD strategies we created a population of 100 players, 10 per strategy and iterated our evolutionary loop 5000 times saving the state of the population every 100 iterations. The tournaments consisted of matches with 100 turns each where everyone played against everyone else.

In the absence of noise (the probability of a player making a mistake and not playing what his strategy dictated) the population quickly evolves into full cooperation eradicating ZD-extortionate strategies and unconditional defectors, Fig.5. ZD-generous prevail longer but they end up eventually perishing as well since they still attempt to get an unfair claim of the rewards despite trying to maintain cooperation.

When noise is introduced the results become more interesting as the defection prone strategies, quickly eradicated in the absence of noise, now prevail longer taking over most of the population, Fig.6. However, the superior robustness of Tit-For-Tat can be verified as continues getting a higher payoff and eventually starts taking over the population, Fig.7.

VII. Conclusions

Cooperative behavior seems to go against the idea of survival of the fittest, yet cooperation is abundant in nature. Researchers have for years used the Iterated Prisoner's Dilemma game to study the emergence and stability of cooperation. Recent work has discovered a new class of strategies that provide one player a disproportionate payoff when facing an unwitting opponent. Extortion strategies perform very well in head-to-head competitions, but they fare poorly in large, evolving populations. To tackle this drawback, researchers proposed a related set of generous strategies, which cooperate with others and forgive defection, that replace defection-prone players (like extortionists) and dominate in large populations. The results of our work help explain the evolutionary stability of cooperation and demonstrate how an almost 40 year old strategy, Tit-For-Tat, is notably both one of the simplest strategies and one of the most successful in direct competition.

VIII. Acknowledgements

Special thanks to Mr. Gouveia's desktop for staying up all night running 8000 iterations of the evolutionary loop and to Mr. Gouveia himself for publicly hosting the results of the testing at $http://web.ist.utl.pt/ist175657/$.

REFERENCES

[1] The Axelrod project developers . Axelrod-python/axelrod: v3.9.0, April 2016. [Online; accessed 2-December-2017].

[2] Christoph Adami and Arend Hintze. Evolutionary instability of zero-determinant strategies demonstrates that winning is not everything. *Nature communications*, 4:2193, 2013.

[3] Christian Hilbe, Martin A Nowak, and Arne Traulsen. Adaptive dynamics of extortion and compliance. *PloS one*, 8(11):e77886, 2013.

[4] Investopedia. Prisoners dilemma., 2017. [Online; accessed 2-December-2017].

[5] Vincent Knight, Owen Campbell, Marc Harper, Karol Langner, James Campbell, Thomas Campbell, Alex Carney, Martin Chorley, Cameron Davidson-Pilon, Kristian Glass, et al. An open framework for the reproducible study of the iterated prisoners dilemma. *Journal of Open Research Software*, 4(1), 2016.

[6] OpenABM. Understanding games people play: Axelrods tournament, 2017. [Online; accessed 2-December-2017].

[7] William H Press and Freeman J Dyson. Iterated prisoners dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences*, 109(26):10409–10413, 2012.

[8] Alexander J Stewart and Joshua B Plotkin. Extortion and cooperation in the prisoners dilemma. *Proceedings of the National Academy of Sciences*, 109(26):10134–10135, 2012.

[9] Alexander J Stewart and Joshua B Plotkin. From extortion to generosity, evolution in the iterated prisoners dilemma. *Proceedings of the National Academy of Sciences*, 110(38):15348–15353, 2013.

[10] Wikipedia. Evolutionary game theory — wikipedia, the free encyclopedia, 2017. [Online; accessed 2-December-2017].

IX. Appendix

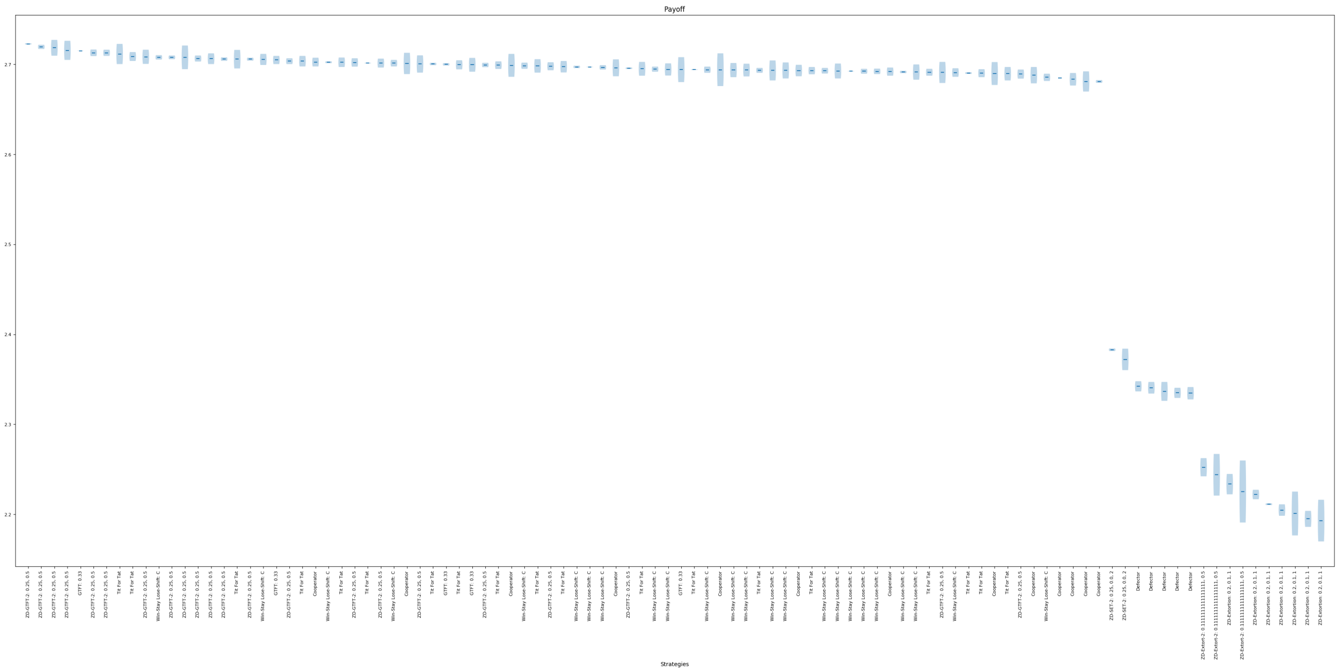

Fig. 5. Payoff distribution of the 1000th iteration on a population of 100 players.

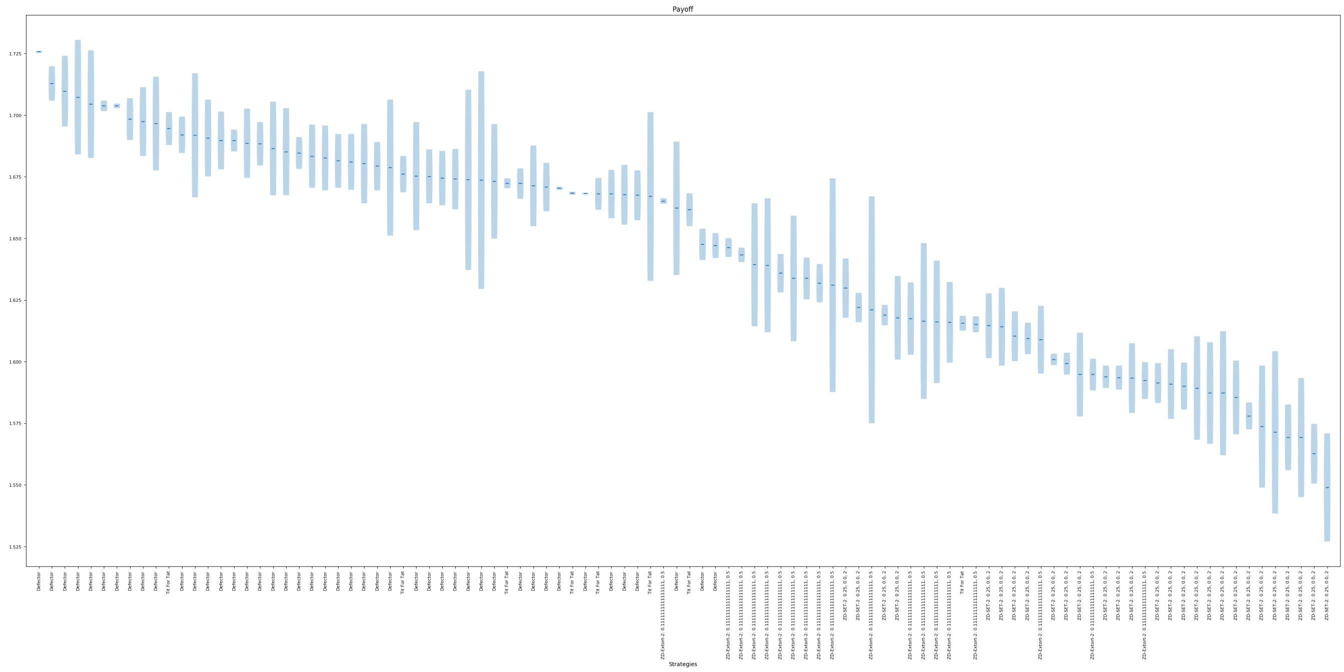

Fig. 6. Payoff distribution of the 4000th iteration on a population of 100 players with noisy tournaments.

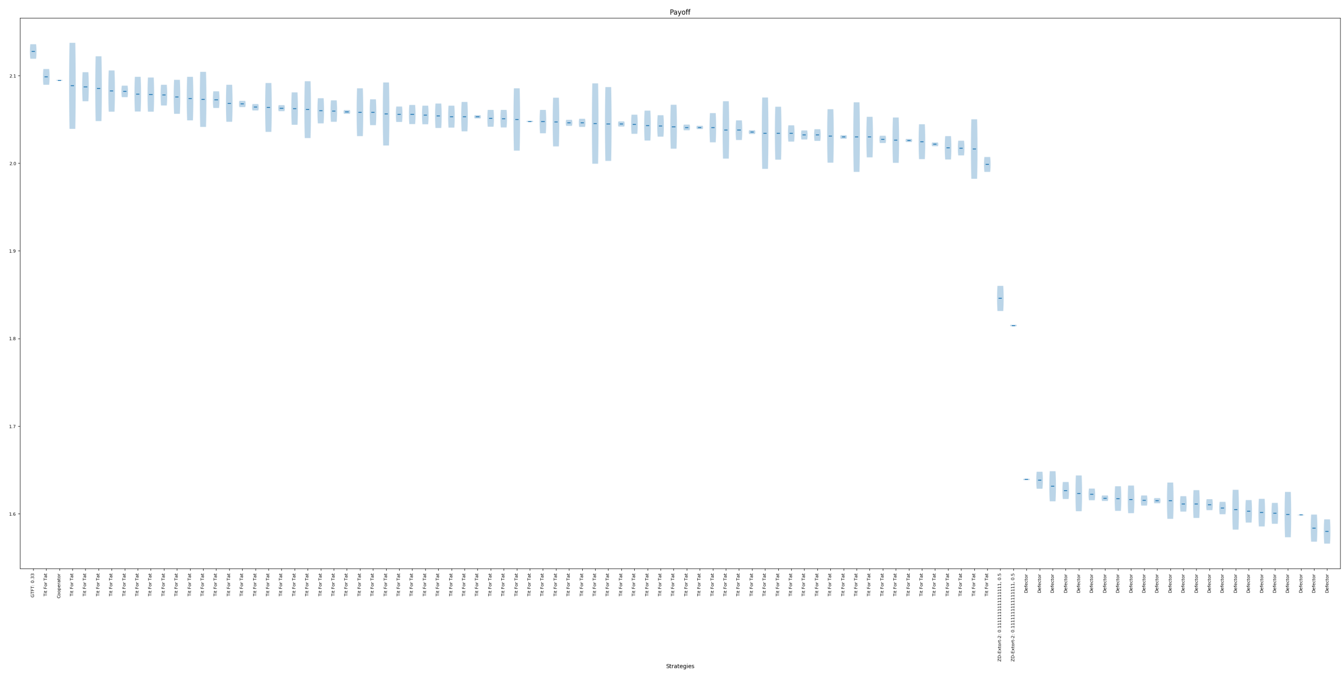

Fig. 7. Payoff distribution of the 7900th iteration on a population of 100 players with noisy tournaments.