Recognising and evaluating the effectiveness of extortion in the Iterated Prisoner's Dilemma

Vincent A. Knight Nikoleta E. Glynatsi

June 21, 2018

Abstract

The Iterated Prisoner's Dilemma is a model for rational and evolutionary interactive behaviour. It has applications both in the study of human social behaviour as well as in biology.

This game is used to understand when and how a rational individual might accept an immediate cost to their own utility for the direct benefit of another.

Much attention has been given to a class of strategies for this game, called Zero Determinant strategies. It has been theoretically shown that these strategies can "extort" any player.

In this work, an approach to identify if observed strategies are playing in a Zero Determinant way is described. Furthermore, experimental analysis of a large tournament with 204 strategies is considered. In this setting the most highly performing strategies do not play in a Zero Determinant way. This suggests that whilst the theory of Zero Determinant strategies indicates that memory is not of fundamental importance to the evolution of cooperative behaviour, this is incomplete.

1 Introduction

2 Recognising Extortion

In [4], given a match between 2 memory-one strategies, the concept of Zeto-Determinant (ZD) strategies is introduced. The main result of that paper showed that given two players $p, q \in \mathbb{R}^2$ a linear relationship between the players scores could be forced.

Assuming the utilities for player p are given by $S_x = (R, S, T, P)$ and for player q by $S_y = (R, T, S, P)$ and that the stationary scores of each player is given by S_X and S_Y respectively. The main result of [4] is that if

$$\tilde{p} = \alpha S_x + \beta S_y + \gamma \tag{1}$$

or

$$\tilde{q} = \alpha S_x + \beta S_y + \gamma \tag{2}$$

where $\tilde{p} = (1 - p_1, 1 - p_2, p_3, p_4)$ and $\tilde{q} = (1 - q_1, 1 - q_2, q_3, q_4)$ then:

$$\alpha S_X + \beta S_Y + \gamma = 0 \tag{3}$$

This work is interested with identifying a test for the reverse problem: given a $p \in \mathbb{R}^4$ how does one identify α, β, γ if they exist.

Note that equation 1 can be expressed linear algebraically as:

$$M x = \tilde{p} \qquad x = (\alpha, \beta, \gamma) \tag{4}$$

with $M \in \mathbb{R}^{4\times 3}$ given by:

$$M = \begin{bmatrix} R & R & 1 \\ S & T & 1 \\ T & S & 1 \\ P & P & 1 \end{bmatrix} \tag{5}$$

Note that in general, equation (4) will not necessarily have a solution. From the Rouch-Capelli theorem []if there is a solution it is unique as $\text{rank}(M) = 3$ which the dimension of the variable x). Furthermore, removing a single row of M would ensure that the corresponding matrix is invertible. This corresponds to the fact that a ZD strategy is defined by only 3 of its values.

As an example, consider the known ZD strategy $p = (8/9, 1/2, 1/3, 0)$ from [5] which is referred to as `Extort-2`. In the standard case of $(R, S, T, P) = (3, 0, 5, 1)$ the inverse of $M_{(4)}$ (removing the last row of M) is given by:

$$M_{(4)}^{-1} = \begin{bmatrix} 1 & -\frac{3}{5} & -\frac{2}{5} \\ 1 & -\frac{2}{5} & -\frac{3}{5} \\ -5 & 3 & 3 \end{bmatrix} \tag{6}$$

This allows us to find the $x = (\alpha, \beta, \gamma)$ corresponding to $\tilde{p}$:

$$x = M_{(4)}^{-1}\tilde{p}_{(4)} = \begin{bmatrix} \frac{1}{18} & -\frac{1}{9} & \frac{1}{18} \end{bmatrix} \tag{7}$$

Using (4) gives that these values lead to the correct value for $p_4 = 0$ confirm that p_4 is a ZD strategy.

This approach could in fact be used to confirm that a given strategy p represents is ZD. However, in practice, if a closed form for p is not known, then due to measurement and/or numerical error this would not work.

Thus, an approach based on least squares [1] is proposed. This approach finds the best fitting $\bar{x} = (\bar{\alpha}, \bar{\beta}, \bar{\gamma})$ which minimises:

$$\text{SSError} = \|Mx - \tilde{p}\|_2^2 = \sum_{i=1}^{4} \left((M\bar{x})_i - \tilde{p}_i\right)^2 \tag{8}$$

Note that SSError which is the square of the Frobenius norm [1] becomes a measure of how close p is to being a ZD strategy.

In [4] a particular type of ZD is defined, if:

$$P(\alpha + \beta) + \gamma = 0 \tag{9}$$

then the player can ensure they get a score at χ times better than the opponent with. This extortion coefficient is given by:

$$\chi = \frac{-\beta}{\alpha} \tag{10}$$

Thus, if (9) holds and $\chi > 1$ a player is said to extort their opponent. Recalling (5), equation (9) corresponds to the last row of M, thus a player p extorts their opponent if and only if $p_4 = 0$ and the estimated $\chi > 1$. Using the method of least squares, the strategy p can be considered ZD for a given threshold of SSError and χ can be approximated. Figure 1 illustrates this for potential strategies: for all of these we see that for the strategy to be extortionate p_2 and p_3 seem to have a linear relationship.

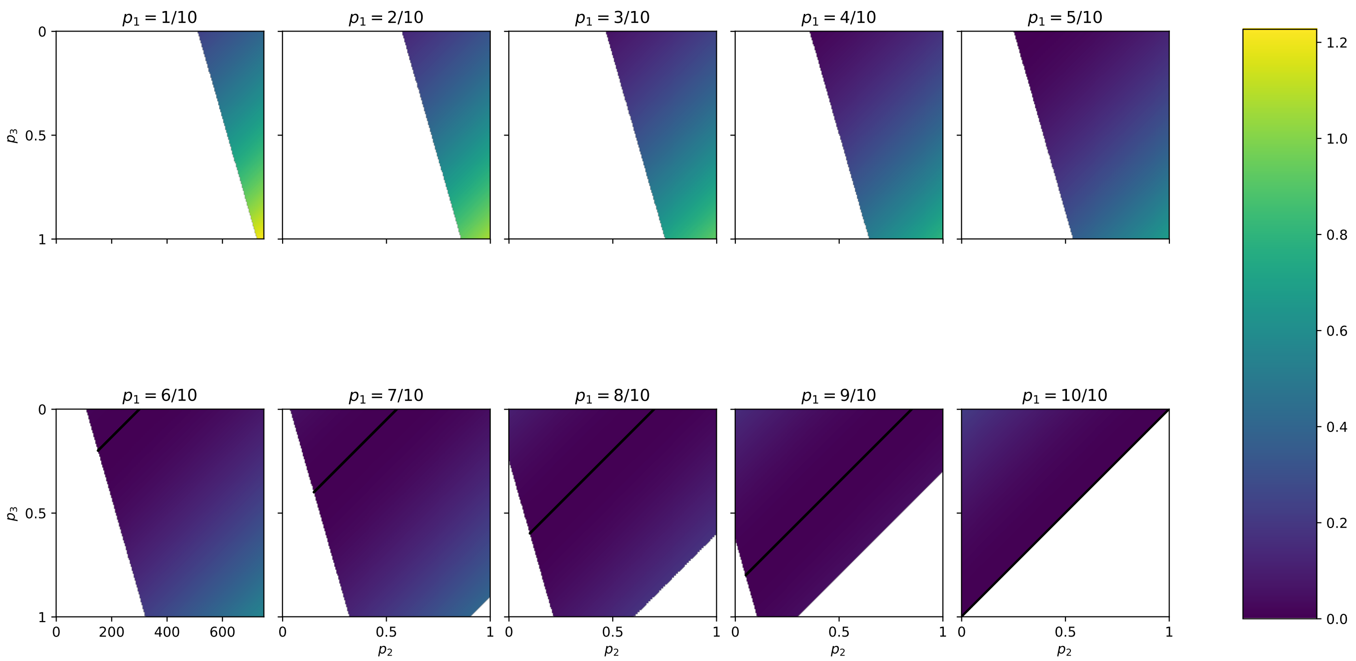

Figure 1: Strategies of the form: $p = (p_1, p_2, p_3, 0)$. The solid line shows the values for which SSError $< 10^{-6}$. Only values for which $\chi > 1$ are displayed, strategies outside of this region can never be extortionate.

By observing interactions (human or otherwise), the transition rates of given individuals can be approximated and this approach can be used to recognise extortionate behaviour. Note that if the environment is noisy then no strategy can be considered to be extortionate.

In the next section, this idea will be illustrated by observing the interactions that take place in a computer based tournament of the IPD.

3 Numerical experiments

In [5] a tournament with 19 strategies, was run with specific consideration given to ZD strategies. This tournament is reproduced here using the Axelrod-Python project [2]. To obtain a good measure of the corresponding transition rates for each strategy all matches have been run for 2000 turns and every match has been repeated 60 times.

The transition rate model of the observed strategies can be be computed using the corresponding steady state model for the measured matrix: A [6]. Using the pair wise interactions the transition rates p, q can be measured and the steady state probabilities inferred and compared to the actual probabilities of each state. Figure 2 shows a regression line fitted to these with a reported R^2 value. This serves to validate the approach: a part from some edge cases the relationship is consistent.

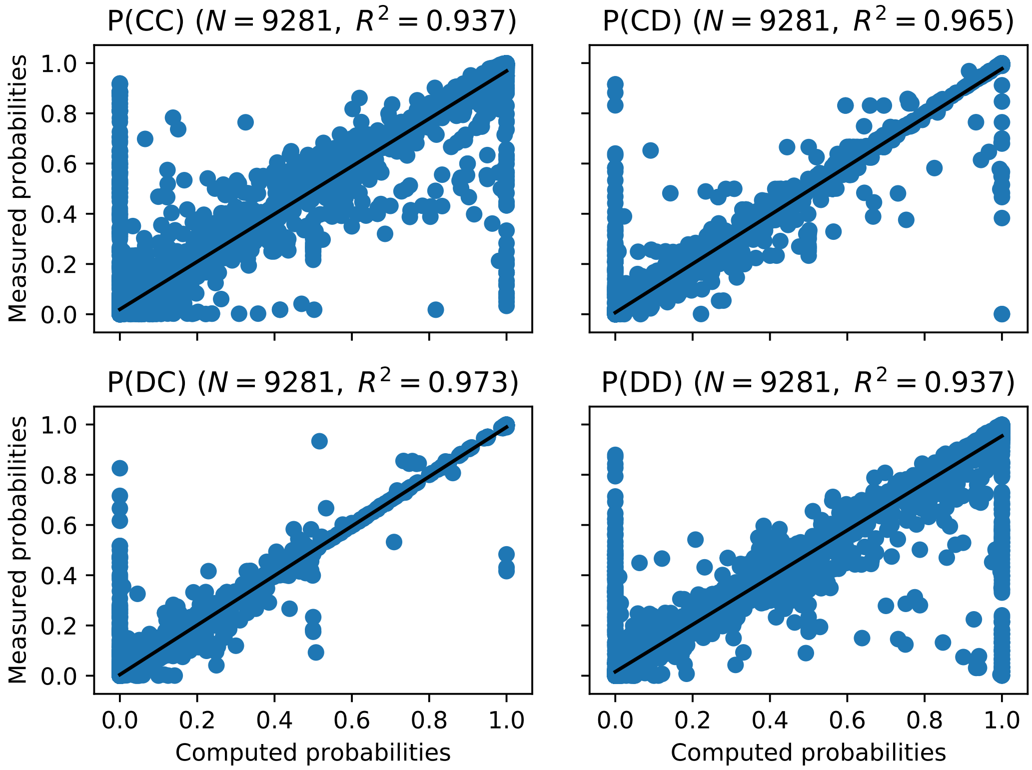

Figure 2: The relationship between the steady state probabilities inferred from the measured transitions and the actual steady state probabilities. A linear regression line is included validating the approach.

Figure 3 shows the SSError values for all the strategies in the tournament, as reported in [5] the extortionate strategy (which has an expected SSError approximately 0) gains a large number of wins. Whilst [5] described the performance of each strategies, this approach also informs the behaviour of a strategy when their opponent is not a memory one strategy: so it will not create a Markov chain.

Here, the work of [5] is extended by investigating a tournament with 204 strategies.

The results of this analysis are shown in Figure 4. The top ranking strategies all achieve mutual cooperation and do not extort each other, however they **do** extort a number of the lower ranking strategies. It is also clear that apart from the low ranking strategies no single strategy is able to extort the higher ranking ones: this is due to the fact that they are not playing a memory one strategy, thus not creating a Markov chain. This is a simple technique to avoid being extorted.

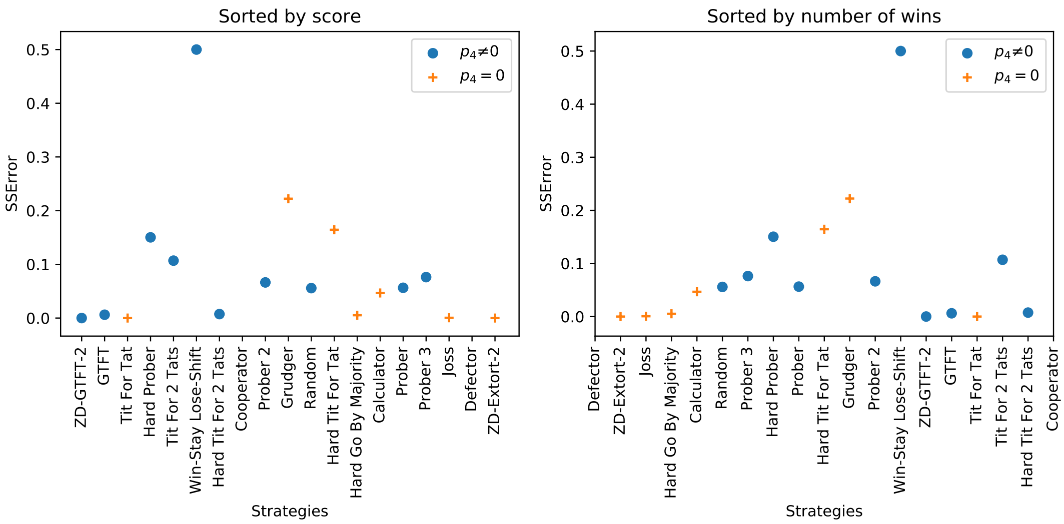

Figure 3: SSError for the strategies of [5], ordered both by number of wins and overall score. Cooperator and Defector are omitted as they do not visit all the states.

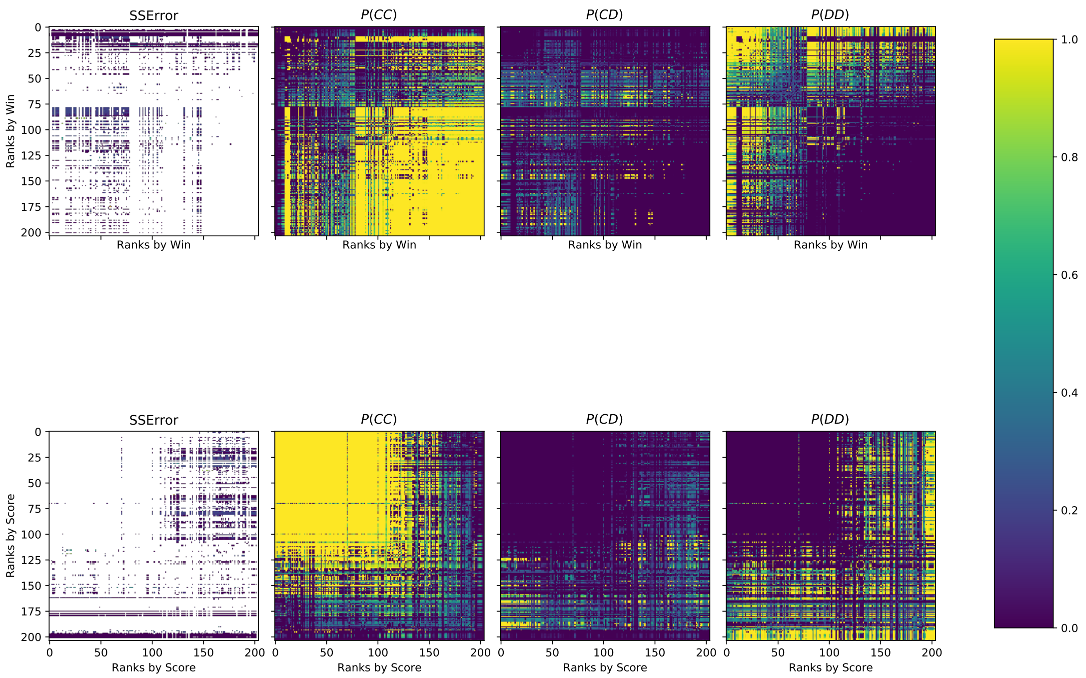

Figure 4: SSError for the strategies for the full tournament. Only strategy interactions for which $p_4 = 0$ and $\chi > 1$ are displayed.

4 Conclusion

This work defines an approach to measure whether or not a player is playing a strategy that corresponds to an extortionate strategy as defined in [4]. This is done through a classic linear algebraic approach for reverting the underlying system of linear equations. Using this, a very large number of pair wise interactions is simulated and in fact very few strategies are found to act extortionately.

The work of [4], whilst showing a clever approach to taking advantage of another memory one strategy exists: this is incomplete. Whilst the elegance of this result is very attractive, just as the simplicity of the victory of Tit For Tat was, it is incomplete. Extortionate strategies are easily beaten in complex environments. This confirms the findings of [3] in which sophisticated strategies resist evolutionary invasion of shorter memory strategies.

References

[1] Gene H. Golub. *Matrix Computations*. J. Hopkins Uni. Press, Jan. 7, 2013. ISBN: 1421407949. URL: `https://www.ebook.de/de/product/20241149/gene_h_golub_matrix_computations.html`.

[2] Vincent Knight et al. "An Open Framework for the Reproducible Study of the Iterated Prisoner's Dilemma". In: *Journal of Open Research Software* 4 (Aug. 2016). DOI: `10.5334/jors.125`.

[3] Vincent Knight et al. "Evolution Reinforces Cooperation with the Emergence of Self-Recognition Mechanisms: an empirical study of the Moran process for the iterated Prisoner's dilemma". In: 1707.

[4] W. H. Press and F. J. Dyson. "Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent". In: *Proceedings of the National Academy of Sciences* 109.26 (May 2012), pp. 10409–10413. DOI: `10.1073/pnas.1206569109`.

[5] A. J. Stewart and J. B. Plotkin. "Extortion and cooperation in the Prisoner's Dilemma". In: *Proceedings of the National Academy of Sciences* 109.26 (June 2012), pp. 10134–10135. DOI: `10.1073/pnas.1208087109`.

[6] William J. Stewart. *Probability, Markov Chains, Queues, and Simulation*. Princeton Univers. Press, July 11, 2009. 760 pp. ISBN: 0691140626. URL: `https://www.ebook.de/de/product/8052317/william_j_stewart_probability_markov_chains_queues_and_simulation.html`.