Iterated Prisoner's Dilemma among mobile agents performing 2D random walk

Jurica Hižak[1]

[1] *University North, University Center Varaždin, J.Križanića 31b, Varaždin, Croatia*
E-mail: ⟨jurica.hizak@unin.hr⟩

Abstract. When Iterated Prisoner's Dilemma takes place on a two-dimensional plane among mobile agents, the course of the game slightly differs from that one in a well-mixed population. In this paper we present a detailed derivation of the expected number of encounters required for Tit-for-tat strategy to get even with Always-Defect strategy in a Brownian-like population. It will be shown that in such an environment Tit-for-Tat can perform better than in a well-mixed population.

Keywords: game theory, payoff, prisoner's dilemma, random walk, tit-for-tat

---

1. Introduction

Prisoner's Dilemma (PD) is a well-known game which is used as a mathematical model for studying the emergence of cooperation in a population of selfish agents without central authority [3]. PD can provide a framework for many real world situations involving a conflict between competitive and cooperative behavior, such as food sharing between vampire bats [4]. Besides biology, PD can be used in economics for understanding how do firms set prices under perfect competition, in the environmental studies to model a fail of current climate negotiations among countries hesitant to curb carbon-dioxide emissions etc. [5, 13, 19].

In the PD game two players can cooperate with or betray each other, with the outcome for each player varying according to the combined moves of both. When both players cooperate, they are awarded at an equal but intermediate level [2, 3, 11]. If only one player defects, s/he wins the highest possible payoff (the temptation, T) while the other player gets the lowest possible payoff (the sucker, S). When both players defect, they receive the penalty P which is not as bad as S. If the player's choices are denoted by C and D for cooperation and defection respectively, one could state that PD game has four possible outcomes (CC, CD, DC, DD). These outcomes are usually represented in the form of the bimatrix

$$
\begin{array}{cc}
 & \text{Player 2} \\
 & \begin{array}{cc} C & \quad D \end{array} \\
\text{Player 1} \quad \begin{array}{c} C \\ D \end{array} &
\begin{array}{|c|c|}
\hline
(R,R) & (S,T) \\
\hline
(T,S) & (P,P) \\
\hline
\end{array}
\end{array}
$$

where R represents payoff in the case of mutual cooperation (reward), P represents payoff in the case of mutual defection (punishment), S and T represent payoffs in the case when one player cooperates and the other one defects and vice versa. The relative sizes of the payoffs are arranged in such way that

$$S < P < R < T.$$

In a symmetric game, such as Prisoner's dilemma, both players have the same payoffs and the same moves at their disposal (C or D). Therefore, by switching the identities of the players nothing changes essentially. If the first player gets payoff a_{ij}, the second one gets $b_{ij} = a_{ji}$ and vice versa. In other words, a matrix with single values (a_{ij}) can describe payoffs for both players. Therefore, a classical matrix can be used instead of a bimatrix.

$$\begin{bmatrix} (a_{11}, b_{11}) & (a_{12}, b_{12}) \\ (a_{21}, b_{21}) & (a_{22}, b_{22}) \end{bmatrix} \Longleftrightarrow \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

i.e.

$$\begin{bmatrix} (R, R) & (S, T) \\ (T, S) & (P, P) \end{bmatrix} \Longleftrightarrow \begin{bmatrix} R & S \\ T & P \end{bmatrix}.$$

If only one round of PD is played it can be easily shown that defection is the only stable strategy [14]. However, when agents are allowed to play multiple rounds, cooperative strategies can survive and even take over the population under certain circumstances. Such repeated PD game is called the Iterated Prisoner's Dilemma (IPD)[4]. Modelling the iterated kind of game makes sense as it enables description of a real world where individuals will interact more than once.

Since individuals in IPD can play repeatedly, there is a whole spectrum of strategies that could be used. Besides ALLD that always defects and ALLC that always cooperates, there are strategies that combine cooperation with defection. Players who offer cooperation do not have to play strictly cooperatively; they can alternate their moves or even punish the players who have cheated on them in the previous rounds. One such strategy is Tit-for-tat (TFT), which is proven to be the most resilient of all cooperative strategies [3, 4, 15]. TFT is not in itself the strongest strategy, but in a mixed surrounding it can win a tournament: it always scores a lot in duels with cooperative strategies and does not lose much from selfish strategies since it does not allow exploitation [4]. TFT is a very simple strategy; it initially cooperates and then simply mimics the previous move of the opponent.

Agents that play TFT strategy are named TFT agents. TFT agents always play cooperatively with the unknown players in their first encounter. If the unknown player cooperates as well, a TFT agent with whom he has interacted will continue to cooperate in the next round. However, if tricked, TFT agent will memorize, recognize and punish the defector in the next encounter. Hence, the payoff matrix for the game between TFT and the unknown ALLD in the *first encounter* may be presented in the following form:

$$E_1 = \begin{array}{c} \\ TFT \\ ALLD \end{array} \begin{array}{c} TFT \quad ALLD \\ \begin{bmatrix} R & S \\ T & P \end{bmatrix} \end{array} \tag{1}$$

Obviously, in the first encounter TFT agent recieves the smallest payoff $S = 0$, while ALLD agent wins the highest award $T = 5$ (if we use Axelrod's payoff parameters $S = 0, P = 1, R = 3, T = 5$).

In the next encounter with the same ALLD agent, TFT will play defection, which means that both of them will receive the same payoff $P = 1$. If the other player persists on defection

(as ALLD does), TFT will play defection consistently in every following encounter. Therefore, the payoff matrix must be rewritten as:

$$E_2 = E_3 = ... = \begin{array}{c} \\ TFT \\ ALLD \end{array} \begin{array}{cc} TFT & ALLD \\ \left[\begin{array}{cc} R & P \\ P & P \end{array} \right] \end{array} \tag{2}$$

where E_2 and E_3 represent the payoff matrices for the second and the third round between the same two players. After two consecutive rounds against the same defector, TFT player will get $\sum E = S + P$. After three rounds s/he will get $\sum E = S + 2P$ and so on. Therefore, in the case of the iterated game against the same ALLD player, TFT will get $\sum E = S + (m-1)P$. Both payoff matrices (1) and (2) may be unified so that the payoff matrix after m rounds takes the following form:

$$\begin{array}{c} \\ TFT \\ ALLD \end{array} \begin{array}{cc} TFT & ALLD \\ \left[\begin{array}{cc} mR & S+(m-1)P \\ T+(m-1)P & mP \end{array} \right] \end{array} \tag{3}$$

where m is the number of rounds.

Such a matrix may be used to calculate the expected payoffs of the strategies even if more than two players are involved as well. For example, it perfectly well describes the payoffs in a round-robin tournament where each player encounters all other players equal number of times in every round.

However, when Prisoner's Dilemma takes place in a population with random interactions, the payoff matrix is usually expressed as a function of the average number of rounds

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \bar{m}R & S+(\bar{m}-1)P \\ T+(\bar{m}-1)P & \bar{m}P \end{bmatrix}. \tag{4}$$

Here $\bar{m}$ is the average number of rounds, while a, b, c and d represent the elements of the payoff matrix in general. For example $b = S + (\bar{m}-1)P$ represents the payoff that TFT strategy will earn in $\bar{m}$ rounds against ALLD strategy. Using $\bar{m}$ in the payoff matrix is a common practice when dealing with well-mixed populations [14, 10, 20, 1, 6]*.

2. Well-mixed population

There are two main approaches to model IPD; one is to allow the individuals to interact with neighbors in predefined network which is usually called *structured population* while the other approach assumes that interaction links between any two individuals exist at all times, which is usually refered to as *well-mixed population*. Payoff matrix (4) describes the latter approach - it is assumed that all-to-all interactions are possible at all time so that spatial structure is not relevant for the game. The existence of all-to-all interactions is the main reason why scholars regard a well-mixed population as an equivalent to a round-robin tournament. This may be justified if every agent plays the same number of games on average and if every pair of rivals is randomly selected in a way that all agents meet each other by the time that is comparable with a time needed for one round of game.

To calculate the expected payoffs of the strategies in a well-mixed population, each matrix element has to be multiplied with the corresponding probability of the outcome. This probability depends on the number of agents in the population. For example, in a very large population

---

*Some of the authors referenced above use *probability of another round* w which is actually the equivalent of the average number of rounds since $\bar{m} = 1/(1-w)$

whereby opponents are randomly selected, the probability that a TFT agent will meet another TFT agent is equal to relative frequency x of TFT strategy. The probability that a TFT agent will meet an ALLD agent is $(1 - x)$. Hence, the expected payoffs of the strategies may be calculated as follows:

$$E_{TFT} = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ 1 - x \end{bmatrix} \tag{5}$$

$$E_{ALLD} = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ 1 - x \end{bmatrix} \tag{6}$$

Therefore, the expected payoffs are the following:

$$E_{TFT} = ax + b(1 - x) \tag{7}$$

$$E_{ALLD} = cx + d(1 - x) \tag{8}$$

However, if IPD takes place in a finite population, the probabilities are slightly different. Assuming that population is a two-component mixture that consists of $N = N_T + N_A$ individuals (agents) whereby $N_T = i$ is the number of individuals carrying the TFT strategy and $N_A = N - i$ is the number of individuals carrying the ALLD strategy, there are $N - 1$ individuals in the population with whom a particular individual can play. For each TFT player there are $(i - 1)$ other TFT individuals and $(N - i)$ ALLD individuals s/he can play with. Therefore, the probability that a TFT agent interacts with another TFT agent is:

$$p_{TT} = \frac{i - 1}{N - 1}, \tag{9}$$

whereas the probability that TFT agent interacts with an ALLD agent is as follows:

$$p_{TA} = \frac{N - i}{N - 1}. \tag{10}$$

The corresponding probabilities that an ALLD agent interacts with a TFT and ALLD agents are as follows:

$$p_{AT} = \frac{i}{N - 1}, \tag{11}$$

$$p_{AA} = \frac{N - i - 1}{N - 1}. \tag{12}$$

Therefore, the expected payoffs in a finite, well-mixed population are as follows:

$$E_{TFT} = a\frac{i - 1}{N - 1} + b\frac{N - i}{N - 1}, \tag{13}$$

$$E_{ALLD} = c\frac{i}{N - 1} + d\frac{N - i - 1}{N - 1}, \tag{14}$$

Each round in a well-mixed population would include approximately $N(N - 1)/2 \approx N^2/2$ of encounters with $(N - 1)$ encounters per agent on average. Hence the average number of encounters per agent is

$$\mu = \bar{m}(N - 1). \tag{15}$$

By equating the expected payoffs $E_{TFT} = E_{ALLD}$ one can determine how many rounds are required for TFT to catch up with ALLD

$$a\frac{i - 1}{N - 1} + b\frac{N - i}{N - 1} = c\frac{i}{N - 1} + d\frac{N - i - 1}{N - 1} \tag{16}$$

Using the payoff matrix (4) and (16), we obtain the minimal number of rounds needed to equalize the payoffs

$$\bar{m}_0 = \frac{Ti - S(N - i) + P(N - 2i)}{(R - P)(i - 1)}. \tag{17}$$

With Axelrod's parameters $(S = 0, P = 1, R = 3, T = 5)$ this leads to

$$\bar{m}_0 = \frac{3i + N}{2(i - 1)}. \tag{18}$$

By the time TFT reaches ALLD, each agent would experience μ_0 encounters on average; therefore, equations (15) and (18) lead to

$$\mu_0 = \frac{3i + N}{2(i - 1)}(N - 1) \tag{19}$$

For the sake of simplicity, let us suppose that the number of TFT players is equal to the number of ALLD players, that is $N/2$. In that case, the expected number of encounters per agent required to equalize the payoffs is

$$\mu_0 = \frac{5N(N - 1)}{2(N - 2)} \approx \frac{5N}{2}. \tag{20}$$

3. Brownian-like population

The intuitive notion of random walk often leads to the presumption that a population in which agents move randomly is well-mixed as well [9]. One might think that the equations (4)-(20) would give sufficiently accurate description of such a population. However, our study has shown a significant discrepancy between this approach and the simulation experiments. For example, in a population that consists of 50 TFT and 50 ALLD agents, one would expect that the payoffs get even after $\mu_0 \approx 5N/2 = 250$ interactions per agent. However, the experimental data consistently show that the expected number of encounters is lower than that - the agent experiences around 200 interactions on average. Therefore, TFT is more successful than expected.

Figure 1: IPD among 3 TFT and 3 ALLD. Each round includes 5 encounters per agent on average
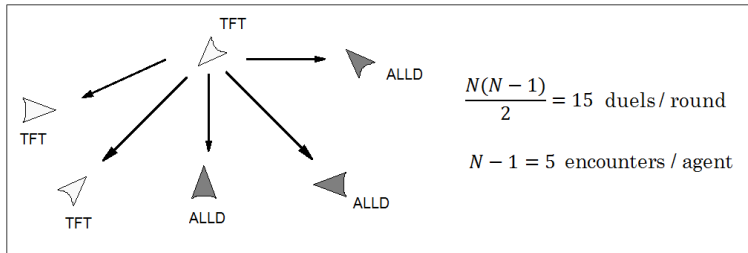
The key to understanding this problem is a gradual decrease of the probability of encounter between a TFT agent and an *unknown* ALLD agent. Let us suppose that population consists of 3 TFT and 3 ALLD agents. In the first round of a round-robin tournament among these 6 agents every TFT agent will meet every individual including ALLD, which means that the first five encounters (one round includes five games per agent) are reserved for playing with the unknown agents. Hence, the probability to meet a new defector in the first round is one hundred percent. Therefore, the probability to meet a new defector in the second round, for example in the 6th encounter is zero. In a Brownian-like population that is not the case. When

agents perform 2D random-walk there is always a probability to meet an agent twice in a row and there is always a chance that some agents will never be encountered.

Let us raise the question: What is the probability for a randomly chosen TFT agent to meet a new (unknown) defector as a function of number of encounters? In the first encounter there is no difference between a Brownian and a round-robin setup: the probability to meet another TFT agent in the first encounter is 2/5, while the probability to meet a defector is 3/5. If the

Figure 2: Probability tree. The probability to meet a new defector decreases gradually with each encounter. For example, the probability to meet a new defector in the second encounter is 12/25.
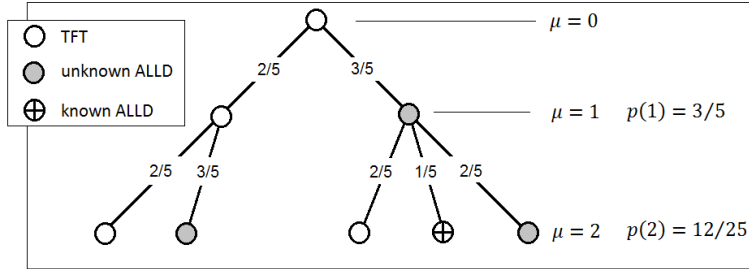
agent has met a defector in the first encounter, then (in the second encounter) the probability to meet a new (unknown) defector in a round-robin setup is 2/4. As opposed to a round-robin, in a Brownian-like population the probability to meet an unknown defector (after a defector) in the second encounter is 2/5. Therefore, the total probability to meet an unknown defector in the second encounter is

$$p_{TA}(2) = \frac{2}{5} \cdot \frac{3}{5} + \frac{3}{5} \cdot \frac{2}{5} = \frac{12}{25}. \tag{21}$$

One can easily show that in a round-robin tournament this probability is as follows:

$$p_{TA}(2) = \frac{2}{5} \cdot \frac{3}{4} + \frac{3}{5} \cdot \frac{2}{4} = \frac{3}{5}. \tag{22}$$

In the following table we can compare the probabilities for a randomly chosen TFT agent to meet an unknown defector for the first 4 encounters:

	Round-robin	2D random walk
$p_{TA}(1)$	3/5	3/5
$p_{TA}(2)$	3/5	12/25
$p_{TA}(3)$	3/5	48/125
$p_{TA}(4)$	3/5	192/625

In order to derive a general formula for the likelihood of meeting a new defector after μ random encounters, we will simplify the problem and assume that there is only one TFT agent in a population. Then the problem is equivalent to the following question: "Suppose there is a deck of playing cards. Let's pick up one card and place it back in the deck. This procedure is repeated μ times. What is the probability of pulling out a card that has not been pulled yet?". If there are $\mu = 3$ elements drawn from the set $S_{52} = \{a_1, a_2, \cdots, a_{52}\}$ there are 52^3 partial permutations with repetition, but there are $(52 - 1)^2 \cdot 52$ ways to pick up a new card in the third selection. Therefore, the probability to draw out a new card in the third selection is $(52 - 1)^2 \cdot 52/52^3$. More generally, if the cards are drawn and returned μ times, the probability to pull out a card that has not been pulled yet is $(\frac{52-1}{52})^{\mu-1}$. Hence, if there is just one TFT

agent among (N-1) ALLD agents, the probability of a TFT encountering a new ALLD agent is as follows

$$p(\mu) = \left(\frac{N-2}{N-1}\right)^{\mu-1}. \tag{23}$$

If there are i TFT agents among $(N-i)$ ALLD agents, the probability p_{TA} of TFT agent encountering a new ALLD agent is as follows

$$p_{TA}(\mu) = \frac{N-i}{N-1}\left(\frac{N-2}{N-1}\right)^{\mu-1}. \tag{24}$$

The expected number of encounters with the unknown defectors from the begining of the random walk till the μth encounter is:

$$\delta_{TA} = i\sum_{k=1}^{\mu} p_{TA}(k), \tag{25}$$

which is the sum of geometrical series. Hence the number of dyadic encounters among TFT and ALLD agents is as follows:

$$\delta_{TA} = i(N-i)\left[1 - \left(\frac{N-2}{N-1}\right)^{\mu}\right]. \tag{26}$$

On the other hand, the probability of a TFT encountering another TFT player is always the same $p_{TT} = (i-1)/(N-1)$ which leads to the expected number of dyadic encounters among TFT agents being

$$\delta_{TT} = i\sum_{k=1}^{\mu} p_{TT}(k) = i\mu\frac{i-1}{N-1}. \tag{27}$$

Probability of a TFT encountering an *old* defector can be found using the probability of opposite event

$$p'_{TA} = 1 - p_{TA} - p_{TT} \tag{28}$$

which leads us to the expected number of dyads with old defectors

$$\delta'_{TA} = i\mu - \delta_{TA} - \delta_{TT}. \tag{29}$$

In a similar way, one can derive the expected number of dyadic encounters for ALLD as well. The formulas are given in the following table:

Dyad	Probability	Expected number of dyads
TT	$\frac{i-1}{N-1}$	$i\mu\frac{i-1}{N-1}$
TA	$\frac{N-i}{N-1}\left(\frac{N-2}{N-1}\right)^{\mu-1}$	$i(N-i)\left[1 - \left(\frac{N-2}{N-1}\right)^{\mu}\right]$
TA'	$1 - p_{TA} - p_{TT}$	$i\mu - \delta_{TA} - \delta_{TT}$
AA	$\frac{N-i-1}{N-1}$	$\mu(N-i)\frac{N-i-1}{N-1}$
AT	$\frac{i}{N-1}\left(\frac{N-2}{N-1}\right)^{\mu-1}$	$i(N-i)\left[1 - \left(\frac{N-2}{N-1}\right)^{\mu}\right]$
AT'	$1 - p_{AT} - p_{AA}$	$(N-i)\mu - \delta_{AT} - \delta_{AA}$

4. Expected payoffs in a Brownian-like population

The average payoff of a strategy can be found by dividing the total payoff of the strategy with the number of individuals using that strategy. In the case of TFT, the average payoff is

$$\langle E_{TFT}\rangle = \frac{1}{i}(\delta_{TT}R + \delta_{TA}S + \delta'_{TA}P). \tag{30}$$

The average payoff of ALLD strategy

$$\langle E_{ALLD} \rangle = \frac{1}{N - i}(\delta_{AT}T + \delta'_{AT}P + \delta_{AA}P). \tag{31}$$

Expected payoff is usually expressed as the average payoff normalized over the number of opponents $(N - 1)$. Therefore the expected payoffs are

$$E_{TFT} = \frac{\delta_{TT}R + \delta_{TA}S + \delta'_{TA}P}{i(N - 1)}. \tag{32}$$

$$E_{ALLD} = \frac{\delta_{AT}T + \delta'_{AT}P + \delta_{AA}P}{(N - i)(N - 1)}. \tag{33}$$

One may notice that both (32) and (33) can be broken into the partial payoffs of the corresponding dyads. For example, partial payoff between TFT and TFT is

$$E_{TT} = \frac{\delta_{TT}R}{i(N - 1)} = \frac{\mu}{N - 1}\frac{i - 1}{N - 1}R. \tag{34}$$

In the following expressions we shall use the following supstitution

$$\xi = \left[1 - \left(\frac{N - 2}{N - 1}\right)^{\mu}\right], \tag{35}$$

therefore the partial payoffs are as follows

$$E_{TT} = \frac{\mu}{N - 1}\frac{i - 1}{N - 1}R, \tag{36}$$

$$E_{TA} = \frac{N - i}{N - 1}\xi S, \tag{37}$$

$$E'_{TA} = \frac{N - i}{N - 1}\left(\frac{\mu}{N - 1} - \xi\right)P, \tag{38}$$

$$E_{AT} = \frac{i}{N - 1}\xi T, \tag{39}$$

$$E'_{AT} = \frac{i}{N - 1}\left(\frac{\mu}{N - 1} - \xi\right)P, \tag{40}$$

$$E_{AA} = \frac{\mu}{N - 1}\frac{N - i - 1}{N - 1}P, \tag{41}$$

Using the equations (36)-(41) and expression (15), in the end we obtain the expected payoffs as a function of the average number of rounds

$$E_{TFT} = \bar{m}R\frac{i - 1}{N - 1} + \left[S\xi + (\bar{m} - \xi)P\right]\frac{N - i}{N - 1}, \tag{42}$$

$$E_{ALLD} = \left[T\xi + (\bar{m} - \xi)P\right]\frac{i}{N - 1} + \bar{m}P\frac{N - i - 1}{N - 1}. \tag{43}$$

Expected payoffs (42) and (43) that have been derived differ from the classical equations (for a well-mixed population). When compared to the expressions (13) and (14), equations (42) and (43) point to the conclusion that elements b and c of the payoff matrix ought to be

replaced. Therefore the payoff matrix for the IPD between TFT and ALLD in a Brownian-like population takes a new form

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \bar{m}R & S\xi + (\bar{m} - \xi)P \\ T\xi + (\bar{m} - \xi)P & \bar{m}P \end{bmatrix} \tag{44}$$

where ξ stands for

$$\xi = \left[1 - \left(\frac{N-2}{N-1} \right)^{\mu} \right]. \tag{45}$$

By equating the expected payoffs (42) and (43) we can derive how many rounds TFT needs to reach ALLD. Using the Axelrod's payoff parameters, we get

$$\bar{m} = \frac{3i + N}{2(i-1)} \xi \tag{46}$$

which is exactly the classical result (18) multiplied by factor ξ. By combining the equations (44) and (18), and by using (35) instead of ξ it follows that the expected number of encounters μ required for TFT and ALLD to get even satisfies the following equation

$$\mu = \frac{(3i + N)(N-1)}{2(i-1)} \left[1 - \left(\frac{N-2}{N-1} \right)^{\mu} \right] \tag{47}$$

or in a simpler form

$$\mu = \mu_0 \left[1 - \left(\frac{N-2}{N-1} \right)^{\mu} \right], \tag{48}$$

where μ_0 is the expected number of encounters in the round-robin case given by equation (19).

Equation (48) cannot be solved analytically, but one can use Lambert W function which is defined as inverse function of xe^x. It can be shown that solution of the equation (48) can be written as

$$\mu = \mu_0 \left[1 - \frac{W_{-1}(a \ln a)}{\ln a} \right], \tag{49}$$

where $a = \left(\frac{N-2}{N-1} \right)^{\mu_0}$.

It must be emphasized that for $x < 0$ Lambert W function has two solutions, positive and negative, which are usually denoted as branch W_0 and branch W_{-1}. In our case the equation (48) has two solutions as well; first one is trivial and gives $\mu = 0$ due to a fact that TFT and ALLD have equal payoffs at the beginning of the game. The second solution (49) is nontrivial and describes the moment when TFT gets even with ALLD. In Sec.5 we shall compare it to the experimental results.

5. Simulation experiments

In order to check the validity of the expression (49) we have run the number of simulations in the multi-agent NetLogo modeling environment. In our simulations TFT and ALLD agents perform two-dimensional one-step random walk on a two-dimensional surface with periodic boundary conditions, which may be considered a toroidal space. The surface consists of a finite number of discrete fields called patches. The agents can only change their direction of motion, but not their speed which is constant in time and equal for all the agents. The agents are initially assigned a random position on the surface and a random direction of motion θ. At each time step, they update their directions according to the rule:

$$\theta(t+1) = \theta(t) + \alpha, \tag{50}$$

Figure 3: Expected number of encounters required for payoffs to equalize in the population with equal ratio of TFT and ALLD. Dotted line represents the values given by the classical equation $\mu_0 \approx 5N/2$, while the full line represents the values calculated via equation (49) using the Lambert function
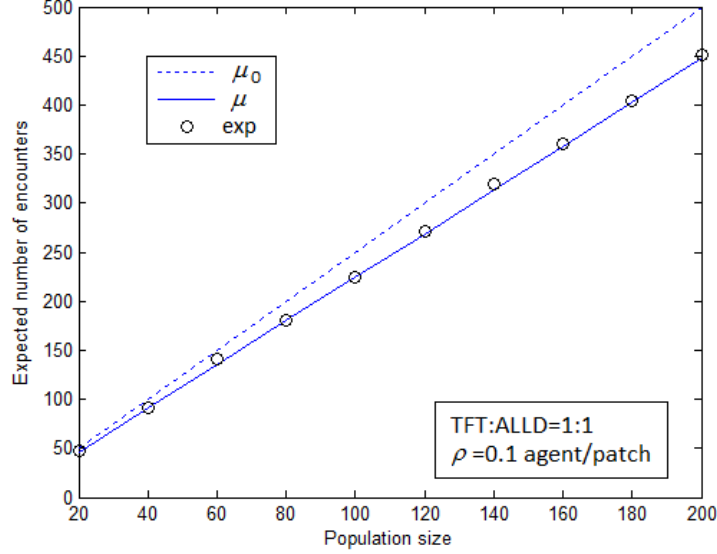
where α is a random number chosen with uniform probability in the interval $[-\pi/2, \pi/2]$. After direction is updated, agent makes one step forward. When two agents meet at the same patch, they play one round of the PD game. After such an encounter they deflect away from each other and carry on with the random walk.

We have run simulations with different parameters of density in a range from 0.05 to 0.5 agents per patch. Number of patches varies from 21×21 to 65×65 in such a way that the total population size N and density ρ are conserved. With a fixed density parameter, as the number of agents increases, the number of patches is automatically adjusted. Such a modification keeps the average number of collisions per unit time constant. A simulation stops when payoffs of the strategies get even. The measured number of encounters are averaged (five simulations on every parameter setup) and presented on FIG.3-5.

As one can notice, the results of the simulation experiments are in a remarkable accordance with theoretical predictions given by the equation (49) and quite robust with respect to the surface density of agents as well. The difference between a round-robin and a Brownian-like population can be seen more vividly when the number of encounters is shifted into the average number of rounds (FIG.6). If the population consists of TFT and ALLD with equal number of individuals, classical (round-robin) result for the average number of rounds is given by the expression

$$\bar{m}_0 = \frac{\mu_0}{N-1} = \frac{5N}{2(N-2)} \tag{51}$$

which is represented by dotted line in FIG.6. On the other hand, the full line represents the average number of rounds which arises from the following equation:

$$\bar{m} = \frac{\mu}{N-1} = \bar{m}_0 \left[1 - \frac{W_{-1}(a \ln a)}{\ln a} \right]. \tag{52}$$

where $a = \left(\frac{N-2}{N-1} \right)^{\mu_0}$.

Figure 4: Expected number of encounters required for payoffs to equalize in the population with $i = N/3$ and $N - i = 2N/3$
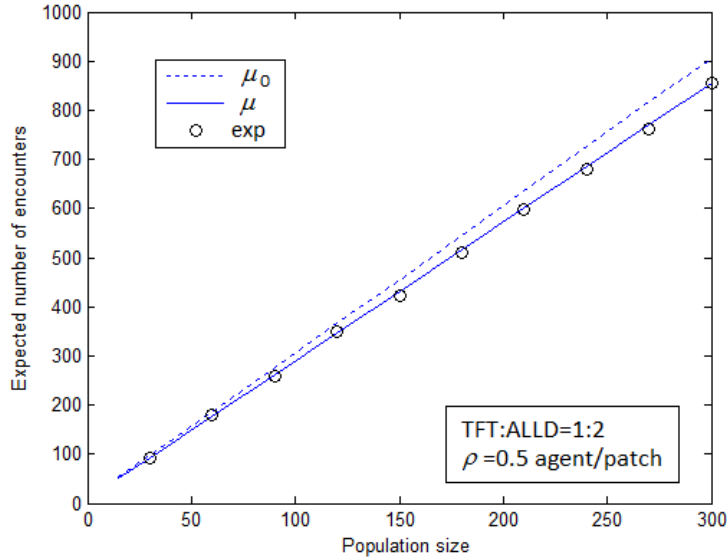

Figure 5: Expected number of encounters required for payoffs to equalize with respect to relative frequency of TFT strategy.
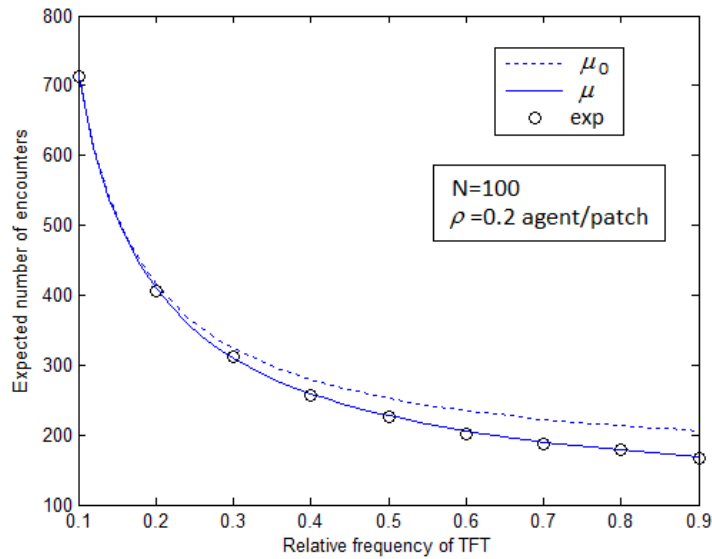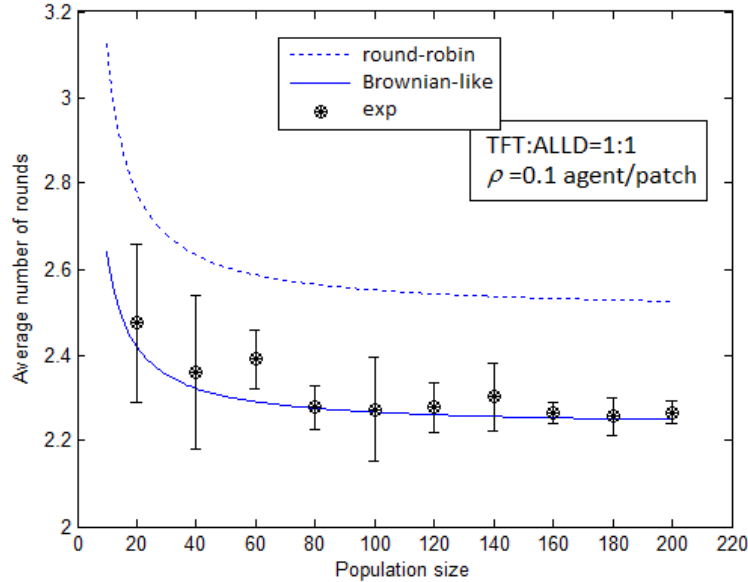

FIG.6. shows the number of rounds required for payoffs to equalize in the environment with the same density, but with different number of agents in the populations. Therefore the number of encounters per time is constant. This is achieved by changing the number of patches. Whenever the size of the population is changed, the number of patches increases in a way that density remains constant. In order to show the general applicability of the formula, the simulations were carried out with a different density parameter setup.

Figure 6: Average number of rounds required for payoffs to equalize in the population with equal ratio of TFT and ALLD.

6. Conclusion

The number of rounds required for one strategy to get even with a rival strategy is one of the most important values in almost every analysis of the Prisoner's Dilemma. For example, in the case of the infinitely large well-mixed population, TFT is stable against ALLD if the average number of rounds is at least $\bar{m} = (T - P)/(R - P) = 2$, assuming that population consists mainly of TFT. This conclusion may be obtained by equating the payoffs with the assumption that relative frequency of TFT is much higher than that one of ALLD.

In finite populations the average number of rounds is an important issue as well, although the mathematical analysis is much more complicated. The requirement for the constant size of the population means that IPD must be analyzed within the framework of the Moran process [12, 7, 16]. One must notice that each step of the Moran process (which consists of birth and death at the same time) is preceded with IPD. The way that IPD is conducted affects the rate of change in difference between the payoffs of the strategies. In a realistic setup where agents move randomly, our simulations show that the difference in payoffs between the strategies decreases faster than in a round-robin setup. This may reflect in the increase of the fixation probability of the TFT strategy[6].

As we have mentioned earlier the intuitive notion of random walk leads to the misconception that Brownian-like population can be described with the classical payoff matrices for a well-mixed population. This might be justified in the case of the unconditioned strategies whereby individuals do not have to be recognized. However, one must be very careful when applying the mathematics onto the strategies that can discern unfamiliar individuals from those with whom they are already acquainted with. In the classical approach it is assumed that TFT agents can meet all of the defectors in the early stage of the game, but as it has been shown, in a Brownian-like population this is not the case. In a large Brownian-like population with 200 individuals, the payoffs will get even although an average TFT agent is never going to encounter at least five ALLD agents. Obviously, this kind of IPD is far from IPD tournament in a well-mixed population. In this paper, the discrepancy between the classical approach and the experiments

has been described. We have seen that probabilities (9)-(12) have to be calculated in a radically different way. In Section 3 of this paper we have investigated this problem more thoroughly; the exact formulas for the probabilities have been derived as well as the formulas for the expected number of dyadic encounters. On the basis of those calculations, in Section 4 the expected payoffs have been calculated. Finally, by equalizing those payoffs the exact formula for the expected number of encounters has been found. In Section 5, the predictions calculated by this formula have been compared with the experimental data. The fact that the formulas were derived rather simplistically using a 'balls and boxes' type probabilistic approach points to the fact that the dimension of space in which the agents live does not play a role. Indeed, it can be shown that the results of the game are more or less the same even if the agents perform 3D random walk. In other words, "balls and boxes" approach is quite suitable since it surpasses the differences between 2D and 3D.

Social structure is particularly important for the application of IPD in social cooperation since people never interact with all the people in a given population. Previous studies have shown that mobility can vastly improve promotion of cooperation in IPD [8]. It has been shown that both randomness and heterogeneity in structured populations can enhance cooperation [17, 18]. In this research it has been shown that mobility may enhance cooperation in a non-structured population as well. The evolution of cooperation is still an important topic on the research agenda of many scientists (not only social scientists). Models so far have shown different circumstances that make this cooperative behaviour easier or more likely to happen (e.g. simultaneous mutations or spatial clusters). This paper has investigated another setting in which agents play iterated prisoners' dilemmas in a population in Brownian motion. It shows that the number of encounters that are necessary for cooperation to survive is smaller than in the same game with a round robin matching scheme. The theoretical result is supported by simulations. The results are not spectacular, but they complement well to the research on cooperation which is out there. The paper supports the basic message that an increase in the likelihood to meet agents again makes cooperation easier to emerge.

References

[1] R Axelrod and W. Hamilton. "The evolution of cooperation". In: *Science* 211.4489 (Mar. 1981), pp. 1390–1396. DOI: 10.1126/science.7466396. URL: https://doi.org/10.1126/science.7466396.

[2] Robert Axelrod. "Effective Choice in the Prisoner's Dilemma". In: *Journal of Conflict Resolution* 24.1 (Mar. 1980), pp. 3–25. DOI: 10.1177/002200278002400101. URL: https://doi.org/10.1177/002200278002400101.

[3] Robert Axelrod. "The Emergence of Cooperation among Egoists". In: *American Political Science Review* 75.02 (June 1981), pp. 306–318. DOI: 10.2307/1961366. URL: https://doi.org/10.2307/1961366.

[4] Robert Axelrod. *The evolution of cooperation*. New York: Basic Books, 2006. ISBN: 978-0465005642.

[5] Gunnar Breivik. "Doping Games A Game Theoretical Exploration of Doping". In: *International Review for the Sociology of Sport* 27.3 (Sept. 1992), pp. 235–253. DOI: 10.1177/101269029202700303. URL: https://doi.org/10.1177/101269029202700303.

[6]   Mark Broom. *Game-Theoretical Models in Biology*. Hoboken: CRC Press, 2013. ISBN: 978-1-4398-5322-1.

[7]   C.Taylor et al. "Evolutionary game dynamics in finite populations". In: *Bulletin of Mathematical Biology* 66.6 (Nov. 2004), pp. 1621–1644. DOI: 10.1016/j.bulm.2004.03.004. URL: https://doi.org/10.1016/j.bulm.2004.03.004.

[8]   Regis Ferriere and Richard E Michod. "Invading wave of cooperation in a spatial iterated prisoner's dilemma". In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 259.1354 (1995), pp. 77–83.

[9]   Francisco Herrerías-Azcué, Vicente Pérez-Muñuzuri, and Tobias Galla. "Stirring does not make populations well mixed". In: *Scientific Reports* 8.1 (Mar. 2018). DOI: 10.1038/s41598-018-22062-w. URL: https://doi.org/10.1038/s41598-018-22062-w.

[10] L. A. Imhof, D. Fudenberg, and M. A. Nowak. "From The Cover: Evolutionary cycles of cooperation and defection". In: *Proceedings of the National Academy of Sciences* 102.31 (July 2005), pp. 10797–10800. DOI: 10.1073/pnas.0502589102. URL: https://doi.org/10.1073/pnas.0502589102.

[11] Shashi Mittal and Kalyanmoy Deb. "Optimal Strategies of the Iterated Prisoner's Dilemma Problem for Multiple Conflicting Objectives". In: *IEEE Transactions on Evolutionary Computation* 13.3 (June 2009), pp. 554–565. DOI: 10.1109/tevc.2008.2009459. URL: https://doi.org/10.1109/tevc.2008.2009459.

[12] P.A.P Moran. *The Statistical Processes of Evolutionary Theory*. Oxford: Oxford: Clarendon Press, 1962.

[13] Walter Nicholson. *Intermediate microeconomics and its application*. Boston, MA: Cengage Learning, 2016. ISBN: 978-1-133-18903-9.

[14] M. A. Nowak. *Evolutionary dynamics : exploring the equations of life*. Cambridge, Mass: Belknap Press of Harvard University Press, 2006. ISBN: 0674023382.

[15] Martin A. Nowak and Karl Sigmund. "Tit for tat in heterogeneous populations". In: *Nature* 355.6357 (Jan. 1992), pp. 250–253. DOI: 10.1038/355250a0. URL: https://doi.org/10.1038/355250a0.

[16] Martin A. Nowak et al. "Emergence of cooperation and evolutionary stability in finite populations". In: *Nature* 428.6983 (Apr. 2004), pp. 646–650. DOI: 10.1038/nature02414. URL: https://doi.org/10.1038/nature02414.

[17] Jie Ren, Wen-Xu Wang, and Feng Qi. "Randomness enhances cooperation: A resonance-type phenomenon in evolutionary games". In: *Physical Review E* 75.4 (Apr. 2007). DOI: 10.1103/physreve.75.045101. URL: https://doi.org/10.1103/physreve.75.045101.

[18] F. C. Santos, J. M. Pacheco, and T. Lenaerts. "Evolutionary dynamics of social dilemmas in structured heterogeneous populations". In: *Proceedings of the National Academy of Sciences* 103.9 (Feb. 2006), pp. 3490–3494. DOI: 10.1073/pnas.0508201103. URL: https://doi.org/10.1073/pnas.0508201103.

[19] Rory Smead et al. "Addendum: A bargaining game analysis of international climate negotiations". In: *Nature Climate Change* 4.9 (Aug. 2014), pp. 840–840. DOI: 10.1038/nclimate2367. URL: https://doi.org/10.1038/nclimate2367.

[20] T.A.Han, L.M. Pereira, and F.C.Santos. "The Role of Intention Recognition in the Evolution of Cooperative Behavior." In: *Proceedings of the 22nd International Joint Conference on Artificial Intelligence* (2011).