



UvA-DARE (Digital Academic Repository)

Evolution of repeated prisoner's dilemma play under logit dynamics

Ochea, M.I.

Publication date

2012

Document Version

Submitted manuscript

[Link to publication](#)

Citation for published version (APA):

Ochea, M. I. (2012). *Evolution of repeated prisoner's dilemma play under logit dynamics*. (CeNDEF Working Paper; No. 12-10). Universiteit van Amsterdam.
<http://www1.fee.uva.nl/cendef/publications/>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Evolution of Repeated Prisoner's Dilemma Play under Logit Dynamics

Marius-Ionut Ochea*

CeNDEF University of Amsterdam

19th December 2012

Abstract

In an evolutionary set-up, we append an ecology of iterated Prisoner's Dilemma (IPD) game strategies, consisting of unconditional cooperators (AllC), unconditional defectors (AllD) and reactive players (TFT) with two repeated strategies that have received less attention in the evolutionary IPD game literature: the error-proof, "generous" tit-for-tat (GTFT) which, with a certain probability, re-establishes cooperation after a (possibly by mistake) defection of the opponent and the penitent, "stimulus-response" (WSLS) strategy that resets cooperation after the opponent punished for defection. An abundance of Rock-Paper-Scissors like patterns is discovered in the 3x3 ecologies comprising Pavlovian and "generous" (GTFT) players. Interestingly, the evolutionary success of Pavlov seems to depend on the *absence* of unconditional (AllC) cooperators in the ecologies investigated.

JEL classification: C72, C73, D43

Keywords: Prisoner's Dilemma, repeated games, evolution, Pavlov

*Center for the Nonlinear Dynamics in Economics and Finance, University of Amsterdam, the Netherlands. m.i.ochea@uva.nl, 0031-20-525 4248

1 Introduction

Studying evolutionary dynamics on iterated prisoner's dilemma (IPD) requires a selection out of the vast set of the repeated game strategies. This choice turns out to be important for the outcome of the iterated game and the level of cooperation 'evolved'. For instance, although the Tit-For-Tat strategy was an undisputed winner of two round-robin tournaments, Axelrod (1997) stresses the crucial role of the surrounding ecology of submitted iterated rules for the success of this direct reciprocity norm. However, which strategies to choose out of the large set of repeated PD is a question lacking a definite answer in the literature. Kraines and Kraines (2000) define adaptive dynamics¹ on the space of memory one repeated PD strategies and show that this process selects, as long-run outcome, only three classes of repeated rules: cooperative, alternating and defective. They also claim that, out of the cooperative subset, the Pavlov strategy is the only one that cannot be invaded by a strategy from the same 'cooperative' class (including TFT). Sigmund and Brandt (2006) investigate one such ecology consisting of three strategies: unconditional cooperators (AllC), unconditional defectors (AllD) and reactive players (TFT) and show that Replicator Dynamics exhibits, among others, a rock-scissors-paper pattern of cyclic behavior. However, such replicator cycles are not robust under Replicator Dynamics, as small payoff perturbations drive all but one strategy near to extinction².

In this paper, we contribute to the stream of work on evolution of rules in repeated games by analyzing the repeated Prisoner's Dilemma with a small number of simple, memory-one strategies. In particular, we extend Sigmund and Brandt (2006) ecology with two additional repeated strategies that seem to have received less attention in the evolutionary IPD game literature: the error-proof, "generous" tit-for-tat which, with a certain probability, re-establishes cooperation after a (possibly by mistake) defection of the op-

¹Adaptive dynamics are regularly used in evolutionary biology to model evolution through natural selection on the fitness landscape.

²mathematically, a so-called heteroclinic cycle is born, with Replicator Dynamics trajectories lingering longer and longer near the boundary of the simplex.

ponent and the penitent, "stimulus-response" (thus dubbed Pavlov strategy Kraines and Kraines (1995), Sigmund and Nowak (1993a), Sigmund and Nowak (1993b)) that resets cooperation after the opponent punished for defection. Second, we contrast Replicator Dynamics behavior with a perturbed version of the Best-Reply dynamics, the Logit Dynamics, allowing for an imperfect switching towards a myopic best reply to the existing strategies distribution. Our emphasis is on the possibility of complicated dynamics such as multiplicity of steady states, limit cycles or chaos in the resulting dynamical system. We start with a systematic investigation of all ten 2×2 repeated game rules interactions and then build-up towards more complex ecologies of 3, 4, and 5 strategies, combining theoretical considerations about the best-reply structure of the resulting normal form games with numerical analysis and simulations. A bifurcation analysis with respect to various model parameters is performed in order to reveal qualitative changes in the set of long-run (non)cooperative behaviors. Preliminary results, in particular for the 4×4 and 5×5 IPD ecologies, show that Logit Dynamics displays *stable co-existence* of repeated strategies but subjected to perpetual oscillations or even chaotic patterns in the distribution of IPD strategies.

The paper is structured as follows: Section 2 introduces the selection of iterated PD strategies and the resulting evolutionary IPD game. The 2, 3 and 4–types ecologies are discussed in Sections 3, 4 and 5, respectively, while Section 6 investigates, mainly via computer simulations, the full 5×5 ecology. The final section is reserved for concluding remarks and future research directions.

2 An Evolutionary Iterated PD game

We consider a standard 2x2 Prisoner's Dilemma stage game where players can either cooperate (C) or defect (D). The payoff matrix of the stage game is given by:

$$\begin{bmatrix} C/D & C & D \\ C & b-c, b-c & -c, b \\ D & b, -c & 0, 0 \end{bmatrix}; b > c > 0 \quad (1)$$

where b stands for the benefits of cooperation and c for the costs associated with cooperative behavior. At each time t the state of the play between two opponents is given by an element belonging to set : $\Omega = \{CC, CD, DC, DD\}$. For the iterated PD game we restrict the choice of iterated Prisoner's Dilemma meta-rules to a set of stochastic *memory-one* strategies (see, for instance, Kraines and Kraines (2000)). Our motivation for such a simple subset of IPD strategies resides, on the one hand, on our attempt to model a particular form of boundedly rational players (namely limited memory agents or forgetting effects) and, on the other hand, to obtain analytical tractability of the resulting Markov chain. The IPD game starts with a first random move C or D and then proceeds with playing C with probability (r, s, t, p) conditional on the realized state at time $t - 1$ being CC, CD, DC, DD , respectively. Deterministic strategies are particular limits in this stochastic strategy space. In particular we focus on 5 well-known strategies:

- unconditional cooperators $AllC : (1, 1, 1, 1)$;
- unconditional defectors $AllD : (0, 0, 0, 0)$;
- conditional cooperators "Tit-for-Tatters" $TFT : (1, 0, 1, 0)$;
- generous cooperators "Generous-Tit-for-Tat" $GTFT : (1, m, 1, n)$ ³;

³ m stands for the probability of cooperating after the opponent defected and n is the probability of playing cooperate after mutual defection.

- penitent or Pavlov players "WinStayLoseShift" $WLSL : (1, 0, 0, 1)$.

Stochastic strategies are ε -perturbations of the deterministic ones, where ε has the natural interpretation of a probability of mistakes or errors in implementation/execution of the deterministic strategies. Following Kraines&Kraines (2000) the iterated Prisoner's Dilemma game between two stochastic players $S_1 = (r, s, t, p)$ and $S_2 = (x, y, z, w)$ leads to a Markov chain on states CC, CD, DC, DD with transition probabilities given by:

$$T = \begin{bmatrix} \text{state} & CC & CD & DC & DD \\ CC & rx & sz & ty & pw \\ CD & r(1-x) & s(1-z) & t(1-y) & p(1-w) \\ DC & (1-r)x & (1-s)z & (1-t)y & (1-p)w \\ DD & (1-r)(1-x) & (1-z)(1-s) & (1-t)(1-y) & (1-p)(1-w) \end{bmatrix} \quad (2)$$

Each column of the matrix T in (2) contains transition probabilities to states CC, CD, DC, DD , respectively and naturally, its entries add up to one. For instance, in the first column, entry t_{11} gives the probability of next period state staying in CC : row player S_1 cooperates with probability r after the realized state was CC while column player S_2 cooperates with probability x after the same realization, and thus:

$$t_{11} = \text{prob}(CC | CC) = rx.$$

Similarly,

$$\begin{aligned} t_{21} &= \text{prob}(CD | CC) = \text{prob}(S_1 \text{ cooperates after } CC) \times \text{prob}(S_2 \text{ defects after } CC) \\ &= r \times (1 - \text{prob}(S_2 \text{ cooperates after } CC)) = r(1-x). \end{aligned}$$

One can show that, for strictly positive perturbation parameters, this Markov process is ergodic, i.e. there is positive probability of escaping from any of the states in Ω . Therefore,

it has a stationary invariant distribution given by the eigenvector of T corresponding to an eigenvalue 1 (Kemeny and Snell (1975)). The invariant distribution represents the average time the play between two stochastic strategies spends in each state in Ω and, without discounting future payoffs⁴, it enables computation of the *average* expected payoff resulting from the interaction of two *repeated* game strategies (Appendix A contains detailed calculations of the invariant distributions and average payoffs). One such average payoff matrix is constructed below for an ecology consisting of the following stochastic versions of the five strategies described above:

- $AllC - (1 - \varepsilon, 1 - \varepsilon, 1 - \varepsilon, 1 - \varepsilon)$;
- $AllD - (\varepsilon, \varepsilon, \varepsilon, \varepsilon)$;
- $TFT - (1 - \varepsilon, \varepsilon, 1 - \varepsilon, \varepsilon)$;
- $GTFT - (1 - \varepsilon, m, 1 - \varepsilon, n^5)$;
- $WSLS - (1 - \varepsilon, \varepsilon, \varepsilon, 1 - \varepsilon)$.

$$M = \begin{bmatrix} & AllD & TFT & GTFT & WSLS & AllC \\ AllD & \varepsilon(b - c) & m_{12} & m_{13} & \frac{1}{2}b - c\varepsilon & b - b\varepsilon - c\varepsilon \\ TFT & m_{21} & \frac{1}{2}b - \frac{1}{2}c & m_{23} & \frac{1}{2}b - \frac{1}{2}c & m_{25} \\ GTFT & m_{31} & m_{32} & \frac{n}{n+\varepsilon}(b - c) & m_{34} & m_{35} \\ WSLS & b\varepsilon - \frac{1}{2}c & \frac{1}{2}b - \frac{1}{2}c & m_{43} & m_{44} & b - \frac{1}{2}c - b\varepsilon \\ AllC & b\varepsilon - c + c\varepsilon & m_{52} & m_{53} & m_{54} & (1 - \varepsilon)(b - c) \end{bmatrix} \quad (3)$$

⁴Sigmund and Brandt (2006) do allow for discounting in their 3x3 ecology of behaviors, and construct the iterated game matrix in a different manner, by computing discounted sums of future payoffs, under stochastic strategies; thus, the first move in the game becomes relevant and a repeated game strategy must be characterized, in addition, by the probability of playing C in the first round.

⁵Throughout the paper we assume that the ‘generosity’ probabilities n and m are an order of magnitude higher than the probability of an error in the strategy execution ε (e.g. $\varepsilon = 0.01, n = 0.1$).

where m'_{ij} s are complicated algebraic expressions in the stage game parameters (b, c) , the errors (ε) and the 'generosity' parameters n (see Appendix A).

Players are assumed to switch their repeated game strategies based on realized, past *average* performance. Thus, at time t the repeated strategy i will be played with the logistic probability:

$$x_{i,t} = \frac{e^{\beta(M\mathbf{x})_{i,t-1}}}{\sum_{i=1}^5 e^{\beta(M\mathbf{x})_{i,t-1}}}, \quad \sum_{i=1}^5 x_{i,t-1} = 1 \quad (4)$$

with β denoting the responsiveness to payoff differences between alternative strategies and $x_{i,t}$ s the time t fractions of *ALLD*, *TFT*, *GTFT*, *WSLS*, *ALLC* players in the population, respectively. Each entry in the payoff matrix (3) defines the *long-run*⁶ contribution m_{ij} to the fitness of strategy S_i due to a particular encounter with strategy S_j . This long-run contribution to fitness is used first to approximate the specific performance of strategy S_i in a *finitely* repeated PD game and then to update the behavior for the next IPD encounter.

In order to disentangle the rich behavior of the full 5-dimensional system (4) we start from simple pairwise interactions and then build larger ecologies of behaviors to capture the contribution of each type to the evolution of (non) cooperative outcomes.

3 2×2 Ecologies

In this section we discuss in more detail four out of the $\binom{5}{2} = 10$ cases of an ecology with only two repeated strategies⁷. Depending on the stage game payoff matrix parameterization and of the error parameter ε , the following best-response structures emerge for the reduced IPD game: Coordination, Hawk-Dove, Prisoner's Dilemma and dominance solvable games.

⁶i.e. in the invariant distribution of the Markov chain

⁷For the complete analysis of all 10 cases we refer the reader to Chapter V in PhD thesis Ochea (2010).

3.1 AllD vs. TFT

In a 2×2 ecology of unconditional defectors (AllD) and reciprocators (TFT) the reduced payoff matrix has the following form:

$$\begin{bmatrix} & \text{AllD} & \text{TFT} \\ \text{AllD} & \varepsilon(b-c) & -\varepsilon(c-2b+2b\varepsilon) \\ \text{TFT} & \varepsilon(b-2c+2c\varepsilon) & \frac{1}{2}b - \frac{1}{2}c \end{bmatrix} \quad (5)$$

For $b > c/(1-2\varepsilon)$ and $\varepsilon > 0$ the following inequalities hold:

$$\begin{aligned} \varepsilon(b-c) &> \varepsilon(b-2c+2c\varepsilon) \\ \frac{1}{2}b - \frac{1}{2}c &> -\varepsilon(c-2b+2b\varepsilon) \end{aligned}$$

Thus, the IPD game (5) is a Coordination game⁸ with two pure strategy equilibria (*AllD*, *AllD*) and (*TFT*, *TFT*) and, depending on the initial population mixture, the Best-Reply limit ($\beta \rightarrow \infty$) of the Logit Dynamics converges to either of them. When implementation errors are very small ($\varepsilon = 0.01$), the (*AllD*, *AllD*) equilibrium basin of attraction is much smaller compared to the other equilibrium basin. This is due to the fact that *TFT* against itself performs much better than *AllD* against itself, as can be seen from the payoff matrix (5): $m_{22} = 1/2(b-c) > m_{11} = \varepsilon(b-c)$. It turns out that, for large β and small ε , *TFT* can invade and take over the entire population, irrespective of the initial mixture (Fig. (1), Panel (a) show time series of the fractions for highly asymmetric starting conditions of 99% *AllD* and only 1% *TFT*). Increasing the probability of mistakes ε , increases the payoff of defectors against themselves and, in the β large limit, the system displays co-existence of stable steady states. A certain large initial critical mass

⁸In the $\varepsilon \rightarrow 0$ limit TFT weakly dominates AllD and it is selected by the logit dynamics given when the sensitivity to payoffs difference is high enough (large β). However, taking this limit is problematic as the invariant distribution of the Markov chain is no longer unique, but history-dependent (more precisely the long-run payoff matrix now depends on the first move in the game). Still, for $\varepsilon = 0$, the payoff matrix could be regarded as the outcome of a particular unfolding of the game history.

of defectors (fraction of *AllD* 99% and *TFT* 1%) can eliminate *TFT* with the population ending up in a monomorphic *AllD* state (Fig. 1b). The multiple steady states are created via saddle-node bifurcations: e.g. (Fig. 1cd) there is an unique interior fixed point for low values of $\beta(\varepsilon)$ but as the intensity of choice(mistake probability) reaches certain thresholds $\beta \approx 140(\varepsilon \approx 0.05)$ two additional steady states are created: one stable(observed in the panels) and one unstable.

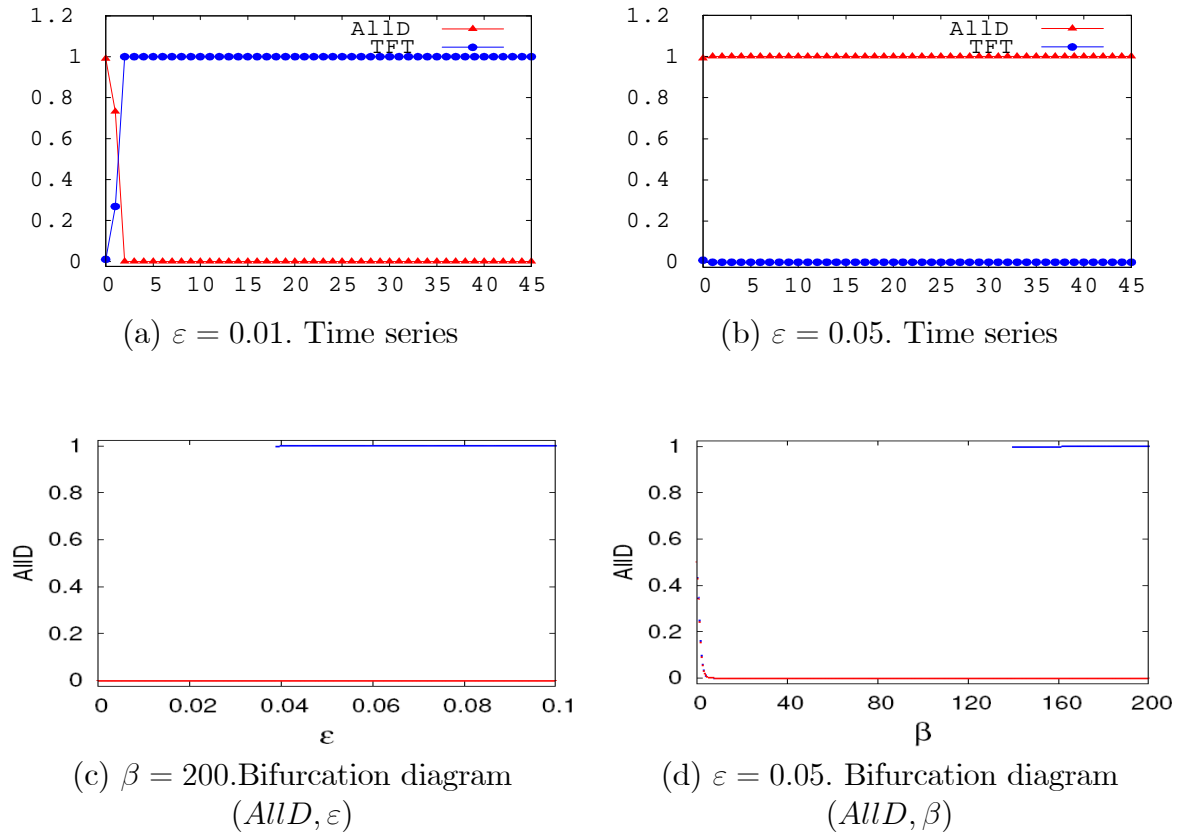


Figure 1: Unconditional defectors (*AllD*) vs. Reciprocators (*TFT*). Different errors in the implementation lead to different long-run steady states: a population of reciprocators (Panel (a)) or a population of defectors (Panel (b)). Panels (c)-(d) display co-existence of stable steady states for large values of β and ε . Remaining game parameters: $b = 4, c = 1, \beta = 200$.

3.2 TFT vs. AllC

For an interaction between unconditional cooperators (*AllC*) and reciprocators (*TFT*) the payoff matrix (3) reduces to:

$$\begin{bmatrix} \text{TFT} & \frac{1}{2}(b-c) & b-c-b\varepsilon+2c\varepsilon-2c\varepsilon^2 \\ \text{AllC} & b-c-2b\varepsilon+c\varepsilon+2b\varepsilon^2 & (1-\varepsilon)(b-c) \end{bmatrix} \quad (6)$$

It is straightforward to see that, for small ε and $b > c$ we have:

$$\begin{aligned} b-c-2b\varepsilon+c\varepsilon+2b\varepsilon^2 &> \frac{1}{2}(b-c) \\ b-c-b\varepsilon+2c\varepsilon-2c\varepsilon^2 &> (1-\varepsilon)(b-c) \end{aligned}$$

Consequently, the reduced 2×2 game (6) is of Hawk-Dove type and, thus, it has three equilibria: two asymmetric in pure strategies (*TFT, AllC*) and (*AllC, TFT*) and one symmetric in mixed strategies. The two asymmetric equilibria give rise to a 2-cycle under best-reply dynamics (i.e. logit dynamics with large β) while the interior equilibrium is unstable. A two-cycle, with population swinging back and forth between the two asymmetric equilibria, is created via a period-doubling bifurcation when either errors in implementation are small (Fig. 2c) or players choose a best-reply (large β , Panel(d)). The intuition for the two-cycle relates to the best-response correspondences in the payoff matrix above: $TFT \in BR\{AllC\}$ and $AllC \in BR\{TFT\}$. In the population game interpretation, when everyone plays *TFT* (*AllC*) it is better to switch to *AllC* (*TFT*). Still, the amplitude of the cycle fluctuations varies with how responsive players are to payoff differences, with the full-scale two-cycle observed only for large β .

For intermediated values of ε (e.g. $\varepsilon \in (\frac{1}{4}, \frac{1}{2})$, for the $b=2, c=1$ parameterization) matrix (6) defines a dominance solvable (DS) game with (*TFT, TFT*) as unique pure strategy equilibrium while for large $\varepsilon > \frac{1}{2}$ the game is still dominance solvable, but with (*AllC, AllC*) the unique equilibrium. The transition from a Hawk-Dove to a Dominance

Solvable structure in the space of 2×2 games is illustrated in Fig. 2ef for finite and best-reply limit of β , respectively.

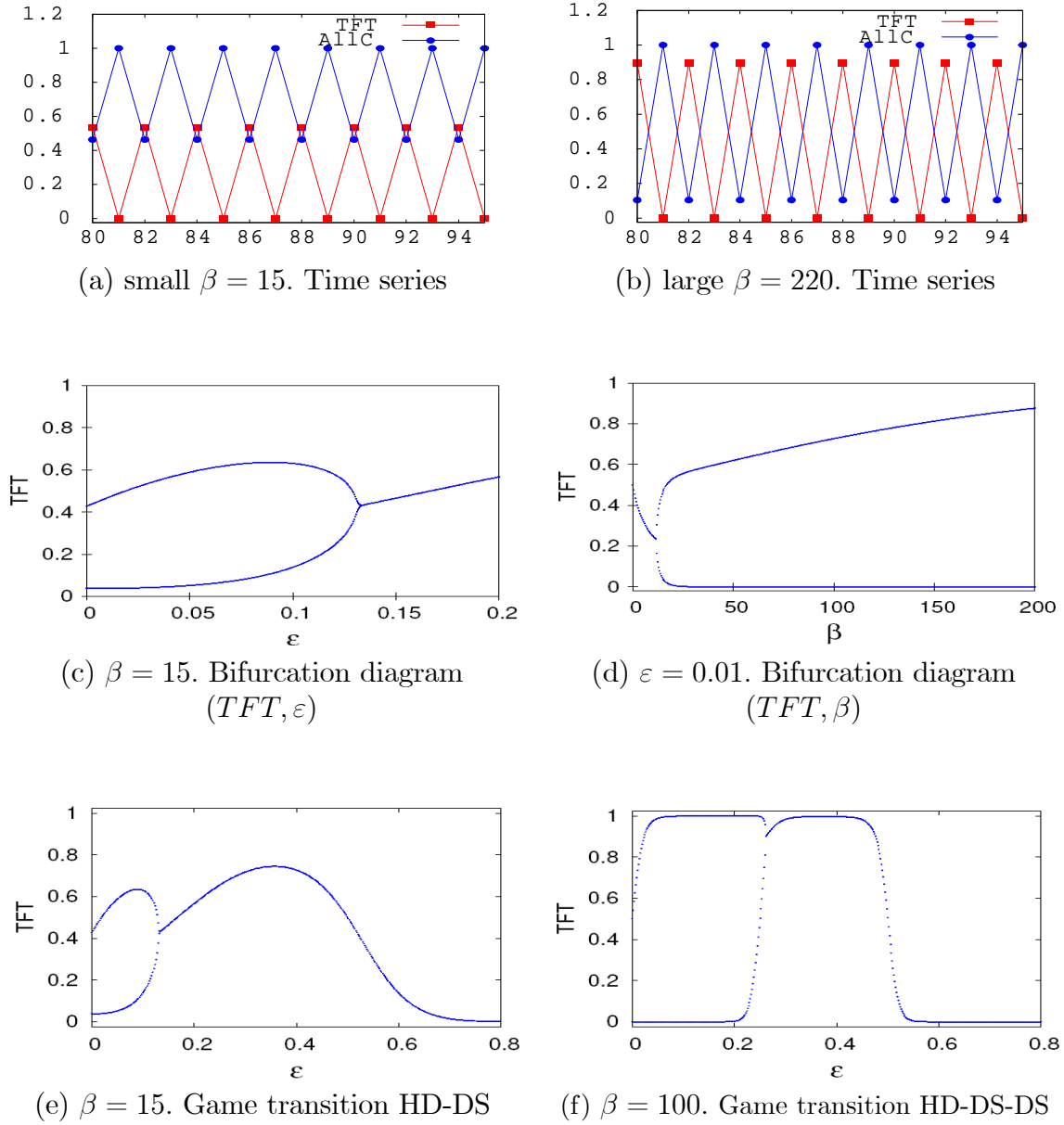


Figure 2: Reciprocators(TFT) vs Unconditional Cooperators(AllC). Time series of fractions of TFT and AllC players for small (Panel (a)) and large (Panel (b)) intensity of choice values, respectively. Panels (c)-(d) display a period-doubling bifurcation with respect to ϵ and β , leading to a 2-cycle. Last, Panels (e)-(f) show transition from a Hawk-Dove to Dominance Solvable pattern. Game parameters: $b = 2, c = 1, \epsilon = 0.01$.

3.3 AllD-WSLS

In this sub-case, the payoff matrix (3) reduces to :

$$\begin{bmatrix} AllD & \varepsilon(b-c) & \frac{1}{2}b - c\varepsilon \\ WSLS & b\varepsilon - \frac{1}{2}c & (b-c)(1 - 4\varepsilon^3 + 6\varepsilon^2 - 3\varepsilon) \end{bmatrix} \quad (7)$$

For high enough benefits of cooperation ($b > b^{*9}$), this payoff matrix represents a Coordination game if:

$$\begin{aligned} \varepsilon(b-c) &> b\varepsilon - \frac{1}{2}c \\ (b-c)(1 - 4\varepsilon^3 + 6\varepsilon^2 - 3\varepsilon) &> \frac{1}{2}b - c\varepsilon \end{aligned}$$

In this case two pure strategy equilibria ($AllD, AllD$) and ($WSLS, WSLS$) exist. One could notice first that, for small implementation errors ε , Pavlov fares much better against itself than $AllD$ does against itself ($(b-c)(1 - 4\varepsilon^3 + 6\varepsilon^2 - 3\varepsilon) > \varepsilon(b-c)$). However, this may not be enough to offset the heavy exploitation Pavlov incurs in mixed encounters when AllD gets, on average, half from the benefits $\lim_{\varepsilon \rightarrow 0}(\frac{1}{2}b - c\varepsilon)$ while $WSLS$ pays half of the costs $\lim_{\varepsilon \rightarrow 0}(b\varepsilon - \frac{1}{2}c)$. This situation is depicted in Fig. (3)a where the population ends up in an all-defectors state. If, in addition to small error probabilities, the benefits of cooperation are relatively large ($b > b^*$), the "own-type" interaction effect may dominate the "cross" interaction effect and the whole population becomes entirely Pavlovian (Fig. 3b). However, this results is dependent on the initial mixture of the population as for a given initial threshold of $AllD$ players, population cannot turn Pavlovian, irrespective of the relative size of the benefits accrued to cooperation b . The co-existence of two stable, monomorphic steady states for high benefits of cooperation is documented in Fig. 3c-d where additional fixed points emerge as the parameters of interest $-\beta$ and ε respectively - are changed. When benefits are low enough, $b < b^*$ the 2×2 game becomes dominance

⁹ $b^* = 2c(1 - 2\varepsilon + 2\varepsilon^2)/(-4\varepsilon + 4\varepsilon^2 + 1)$

solvable (DS) with $(AllD, AllD)$ the unique equilibrium. A bifurcation diagram with respect to the benefits of cooperation b (Fig. 3e) reveals once the threshold b^* is passed, a transition in the space of 2×2 games from a Dominance Solvable game with unique equilibrium to a Coordination game with two equilibria.

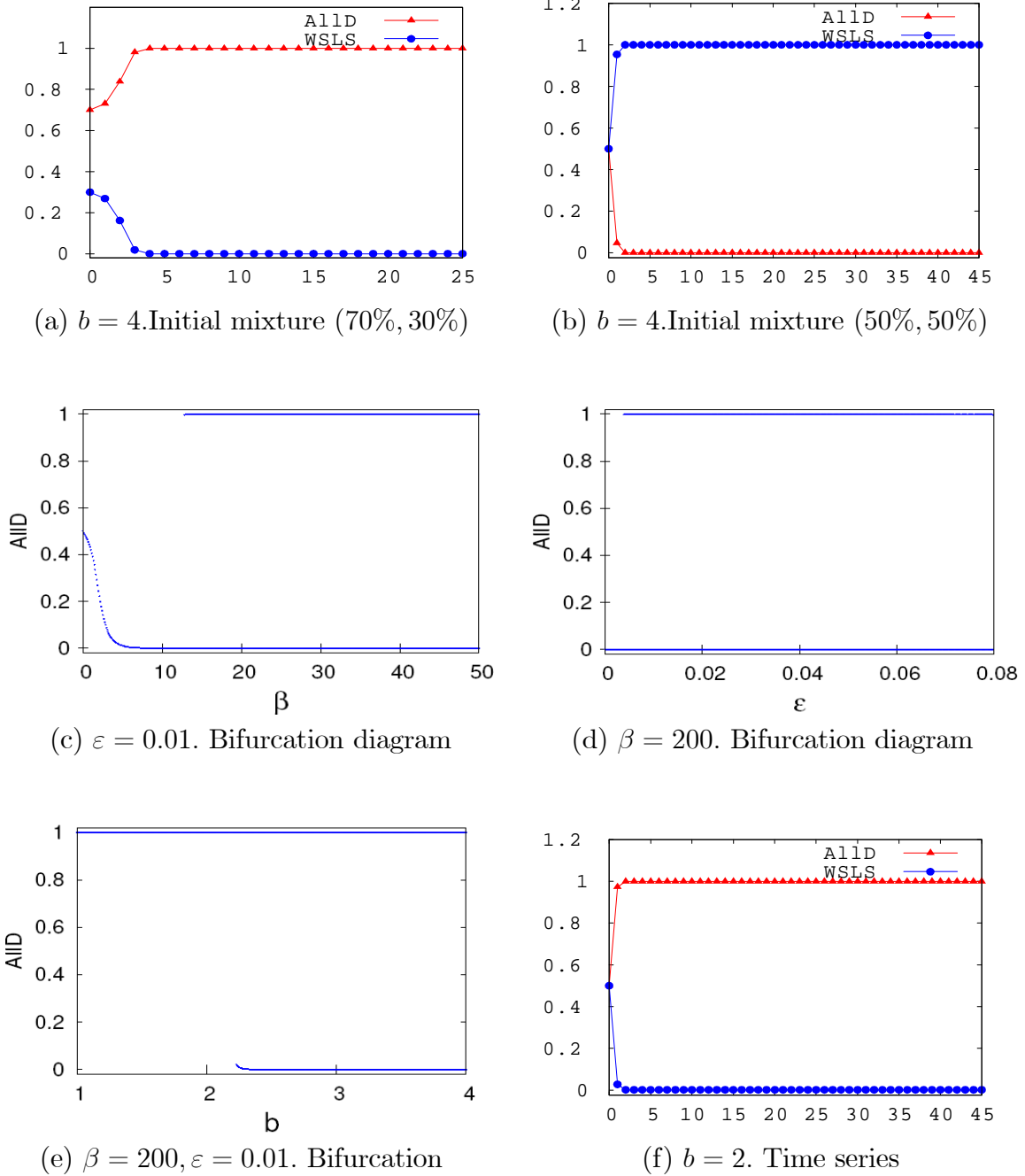


Figure 3: Unconditional defectors(AllD) vs. Pavlov(WSLS). The selected long-run monomorphic state, depends both on the size of the implementation errors ε and on the relative benefit of cooperation b . Panels (c)-(d): fold bifurcations in ε, β spaces and co-existence of steady states. Panel (e) displays the bifurcation diagram with respect to the benefit of cooperation b and show the transition from a dominance solvable game (low b) to a coordination game (high b). Panel (f) reports time series converging to an AllD population as WSLS is a strictly dominated strategy for $b < 2c$. Baseline game parameters: $b = 4, c = 1, \varepsilon = 0.01, \beta = 15$.

3.4 GTFT vs. WSLS

For this pair of repeated strategies the payoff matrix (3) reduces to:

$$\begin{bmatrix} GTFT & \frac{n}{n+\varepsilon}(b-c) & m_{34} \\ WSLS & m_{43} & (b-c)(1-4\varepsilon^3+6\varepsilon^2-3\varepsilon) \end{bmatrix} \quad (8)$$

where m_{43} and m_{34} are complicated expressions of all model parameters (see Appendix (A) for explicit formulas). One can show, by algebraic manipulations, that for $n > \varepsilon$ (by assumption):

$$\begin{aligned} \frac{n}{n+\varepsilon}(b-c) &> m_{43} \\ (b-c)(1-4\varepsilon^3+6\varepsilon^2-3\varepsilon) &> m_{34} \end{aligned}$$

and this 2×2 ecology defines a Coordination game with two pure strategy, symmetric Nash equilibria ($GTFT, GTFT$) and ($WSLS, WSLS$). In the β large limit, the equilibria structure of the coordination game describe the long-run behavior of the logit dynamics. The corresponding basins of attraction relative sizes are significantly biased towards Pavlov players as it takes 70% of $GTFT$ players in the initial population in order to converge to a long-run all $GTFT$ state (Fig.4a), with anything below this threshold reversing fortunes toward a long run Pavlov monomorphic state. Intuitively, a stimulus-response, past-performance driven strategy like Pavlov, is able to take advantage of the generosity built into $GTFT$ which reverts, with large probability n , to "nice" behaviour after an opponent defection *irrespective* of the realized history. This is because $GTFT$ does not discriminate among possible causes of defection (be it after a DD, or a DC history), while Pavlov re-establishes cooperation only after own defection. For finite, small β the population converges to a unique fixed point with co-existing behaviors of the form $(x_{GTFT}^*, x_{WSLS}^*), x_{GTFT}^* + x_{WSLS}^* = 1$, with Pavlov outnumbering generous reciprocators (x_{WSLS}^* slides up in the interval $[1/2, 1]$ as we increase β). Fig. 4c-d depict a saddle-node

bifurcation route towards co-existing monomorphic steady states. For low β (given n) or small n (fixing β close to the "best-reply" limit) there is a unique polymorphic population steady state.

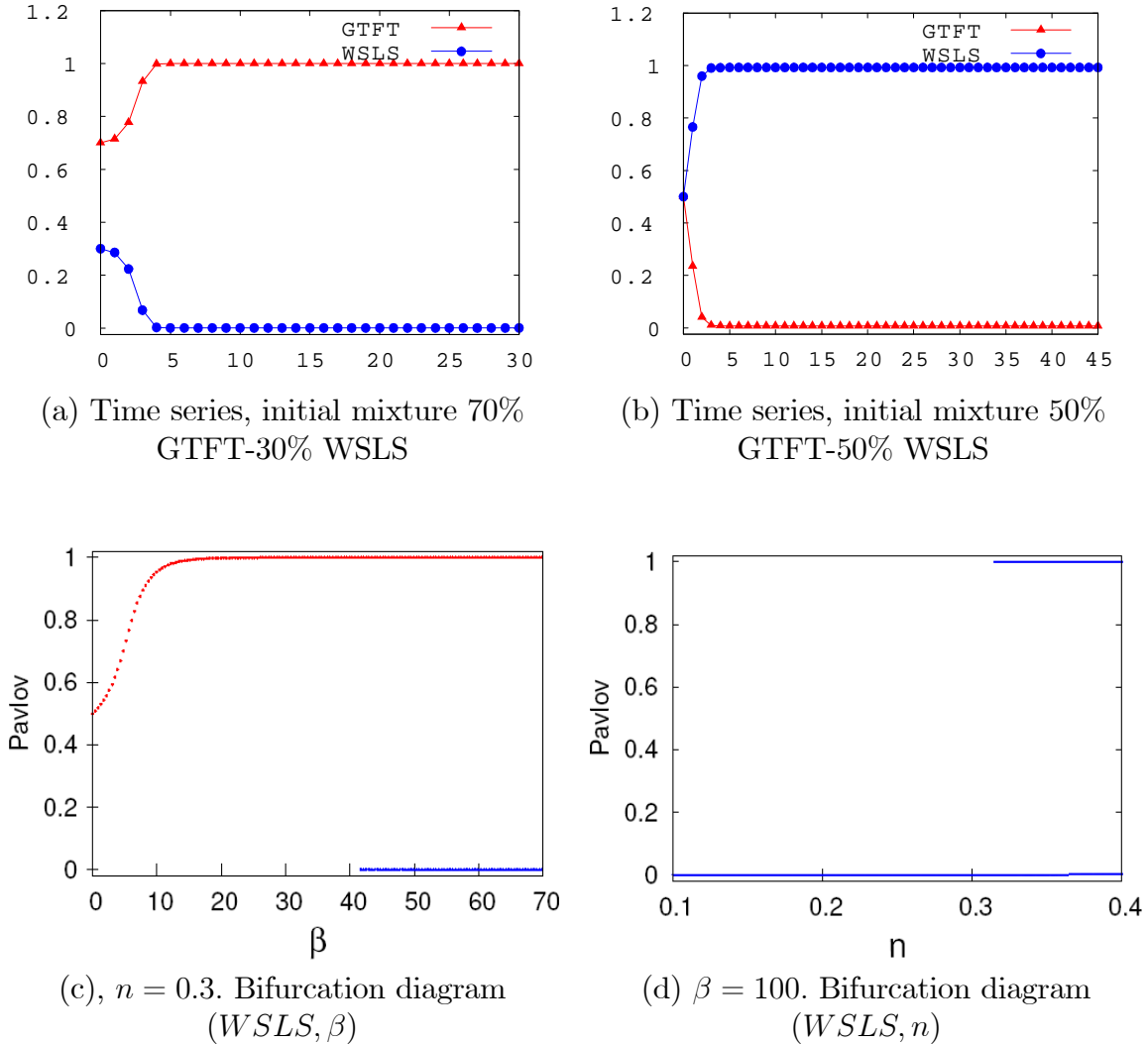


Figure 4: Generous reciprocators (GTFT) vs. Pavlovian(WSLS). Depending on the initial fractions in the populations, in the β large limit, the system converges to a monomorphic state with only one type of behavior surviving. Panels (c)-(d): creation of two co-existing steady states via a fold bifurcation when the degree of generosity/intensity of choice is varied. Game parameters: $b = 4, c = 1, \varepsilon = 0.01, n = 0.3, \beta = 100$.

3.5 Summary

Table 1 summarizes the normal form games resulting from 10 different the strategic pairwise interaction of the five repeated PD strategies together with the long-run attractors of the corresponding evolutionary game under the perturbed best-reply dynamics. In the limiting $\beta \rightarrow \infty$ case, these long-run outcomes coincide with the attracting sets of the unperturbed best-reply dynamics on the respective 2×2 game.

No.	2x2	Game	Bifurcation	Attractors	Path-Dependence
				β large	
2.1	AllD-TFT	WDS,C	SN	multiple SS	yes, $\varepsilon = 0.05$
2.2	TFT-AllC	HD	PD	2-cycle	no
2.3	AllD-AllC	DS	none	unique SS	no
2.4	AllD-GTFT	C	SN	multiple SS	yes, $\varepsilon = 0.01$
2.5	AllD-WSLS	DS,C	SN	multiple SS	yes, $b = 4$
2.6	GTFT-AllC	HD	PD	2-cycle	yes
2.7	GTFT-WSLS	C	SN	multiple SS	yes
2.8	TFT-GTFT	DS, C	PD	2-cycle	no
2.9	TFT-WSLS	WDS	none	unique SS	no
2.10	WSLS-AllC	DS,HD	PD	2-cycle	no

Table 1: Summary of the type of game and dynamical behaviour for all 10 pairwise interaction of iterated Prisoner’s Dilemma strategies under logit dynamics.

We have encountered the following typical pairwise interactions between iterated Prisoner’s Dilemma strategies together with transition from one game form to another as the payoff matrix parameters (b, c, ε, n) change:

- Coordination game (C): multiplicity of monomorphic(only one type of behaviour surviving) steady states with strong path-dependence
- Hawk-Dove game (HD): asymmetric equilibria; 2-cycle with continuous population switching from one monomorphic state to the other
- Prisoner’s Dilemma (PD): uniquely selected long run symmetric, Pareto-inferior, equilibrium (i.e. monomorphic population *state*)

- Dominance Solvable game (DS), non-PD : unique, monomorphic long-run steady state
- Games with a weakly dominated strategy (WDS): unique, monomorphic long-run steady state

The co-existing steady states and 2-cycles emerging under logit dynamics are usually created via fold/ saddle-node (SN) and period-doubling (PD) bifurcations, respectively. Although the qualitative changes in the space of 2×2 games do not make the object of our study, some interesting game transitions were also revealed and it would be worth studying them systematically, as future research.

4 3×3 Ecologies

In this section we will discuss four¹⁰ out of the $\binom{5}{3} = 10$ sub-ecologies of three iterated PD strategies, focusing the analysis on the resulting 3×3 normal form games and on the attracting sets and long-run behavior of the logit evolutionary dynamics.

4.1 AllD-GTFT-WSLS

The 3×3 payoff matrix reads:

$$\left[\begin{array}{ccc} & \textit{AllD} & \textit{GTFT} & \textit{WSLS} \\ \textit{AllD} & \varepsilon(b-c) & m_{13} & \frac{1}{2}b - c\varepsilon \\ \textit{GTFT} & m_{31} & \frac{n}{n+\varepsilon}(b-c) & m_{34} \\ \textit{WSLS} & b\varepsilon - \frac{1}{2}c & m_{43} & (b-c)(1 - 4\varepsilon^3 + 6\varepsilon^2 - 3\varepsilon) \end{array} \right] \quad (9)$$

Using the best-response functions for the reduced 2×2 games 2.4, 2.5 and 2.7 in table (1), and given that $n > \varepsilon$ (by assumption), we can easily derive the best-response correspondences in game form (9) as follows: $\{\textit{AllD}\} = BR\{\textit{AllD}\}$, $\{\textit{GTFT}\} = BR\{\textit{GTFT}\}$ and $\{\textit{WSLS}\} = BR\{\textit{WSLS}\}$. Thus, the game is a 3×3 coordination game, and, in the best-reply limit of our logit dynamic (i.e. $\beta \rightarrow \infty$) each of these three pure strategy symmetric equilibria could be obtained as long-run outcome if the population starts out with the appropriate distribution of fractions. See Fig.5f for a set of initial conditions from which the $(\textit{WSLS}, \textit{WSLS})$ is selected. For β finite, the situation becomes more intricate as the Logit dynamics appear to break the typical coordination game best-response structure described above and generate Rock-Paper-Scissors patterns of cyclical oscillations irrespective of the initial mixture of the population (Fig.5a-d).

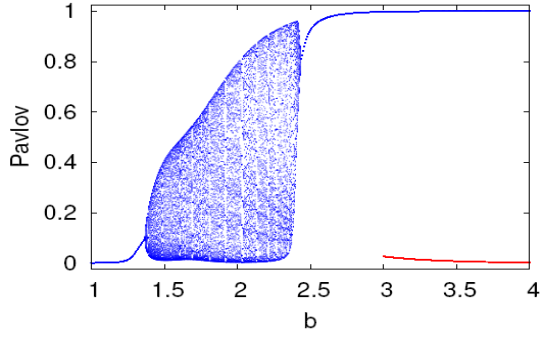
As the bifurcation diagram with respect to β (Fig.5e) suggests, a unique, fully mixed

¹⁰The reader is referred to the corresponding chapter in Ochea (2010) for the complete analysis of the remaining 6 ecologies of repeated rules.

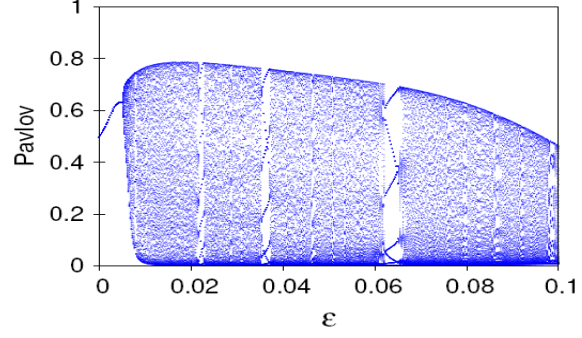
steady state for low values, destabilizes, for moderate values of the intensity of choice, via a Neimark-Sacker (NS) bifurcation, and RSP-like stable, limit cycles appear. For high values of β a sequence of saddle-node (SN) bifurcations occurs, conducting to three¹¹ co-existing stable steady states, in the $\beta \rightarrow \infty$ limit. This bifurcation scenario is confirmed in Fig. 6 where we continue a barycentric fixed point $(1/3, 1/3, 1/3)$ in the benefit of cooperation/error in strategy execution space (top panels) and detect the conjectured Neimark-Sacker (NS)-Period-Doubling (PD)-Saddle-Node (LP) sequence of bifurcations along the equilibrium curve. The middle and bottom panels of Fig. 6 ‘continue’ each detected codimension I singularity with respect to another game or behavioral parameter and displays the resulting bifurcation curves. When a NS/PD/LP curve of codimension I bifurcations is crossed from below a limit cycle/2-cycle/co-existing stable and unstable steady states emerge, respectively. Also, it is worthwhile pointing out the rich selection of codimension II bifurcations occurring along the curves of codimension I bifurcations in this 3x3 ecology of repeated Prisoner’s Dilemma rules under the Logit Dynamics.

The intuition for the emergence, within a Coordination game, of the Rock-Paper-Scissors cycles "*GTFT* beats *AllD*, *WSLS* beats *GTFT* and *AllD* beats *WSLS*" is not immediately clear. We know that in the $\beta \rightarrow \infty$ limit the interaction is of Coordination game type with the three co-existing stable steady states created via a sequence of saddle-node bifurcations as shown before. Still, for finite values of the intensity of choice β (Fig.5e) the typical Coordination game (each strategy E_i is a best-reply to itself) behaves dynamically *as if* a RSP game (each strategy E_i is a best reply to its successor E_{i+1}). Nevertheless, even for finite β the Coordination game dynamic behavior can be recovered if the benefits accrued to cooperation b are high enough (Fig.5a).

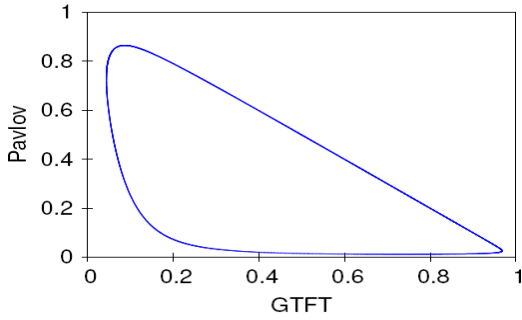
¹¹Panel (e) displays, for large β only two values of Pavlov fractions steady states, the top one with 100% share of WSLS and the bottom one with 0%. The third steady state supersedes the bottom curve, with zero WSLS fraction, too.



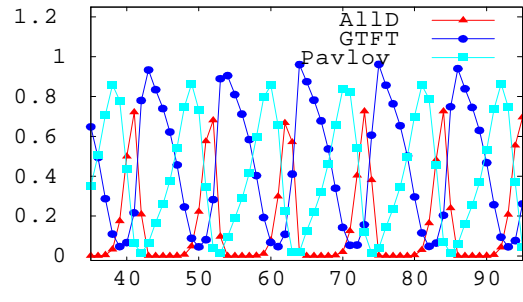
(a) $\beta = 15$. Bifurcation diagram, b



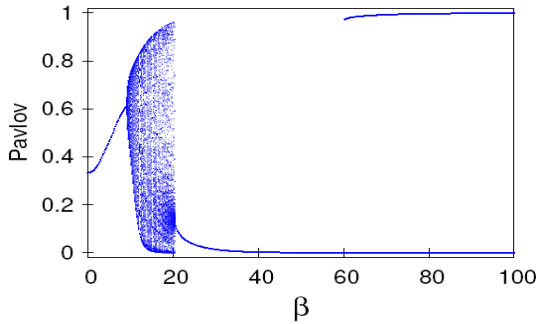
(b) $\beta = 15$. Bifurcation diagram, ϵ



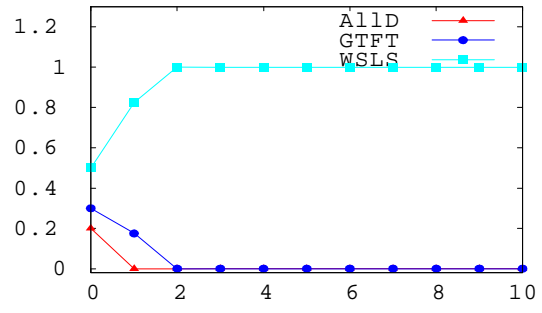
(c) $b = 2.16$. Phase plot



(d) $b = 2.16$. Time series



(e) $b = 2.2$. Bifurcation diagram



(f) $\beta = 100$. Initial fractions (0.2,0.3,0.5)

Figure 5: AllD vs. GTFT vs. WSLs. Bifurcation diagrams with respect to the benefit of cooperation parameter b (Panel (a)) and probability of mistakes (Panel (b)). Phase plots in GTFT-WSLS space (Panel (c)) and time series (Panel (d)), displaying Rock-Paper-Scissors type of long-run behaviour for moderate β . Bifurcation diagram with respect to β and time series (Panel (f)) showing convergence, in the β large limit to the all Pavlov steady state (for an appropriate initial population mixture). Baseline game parameters: $b = 2, c = 1, \epsilon = 0.01, n = 0.1, \beta = 15$.

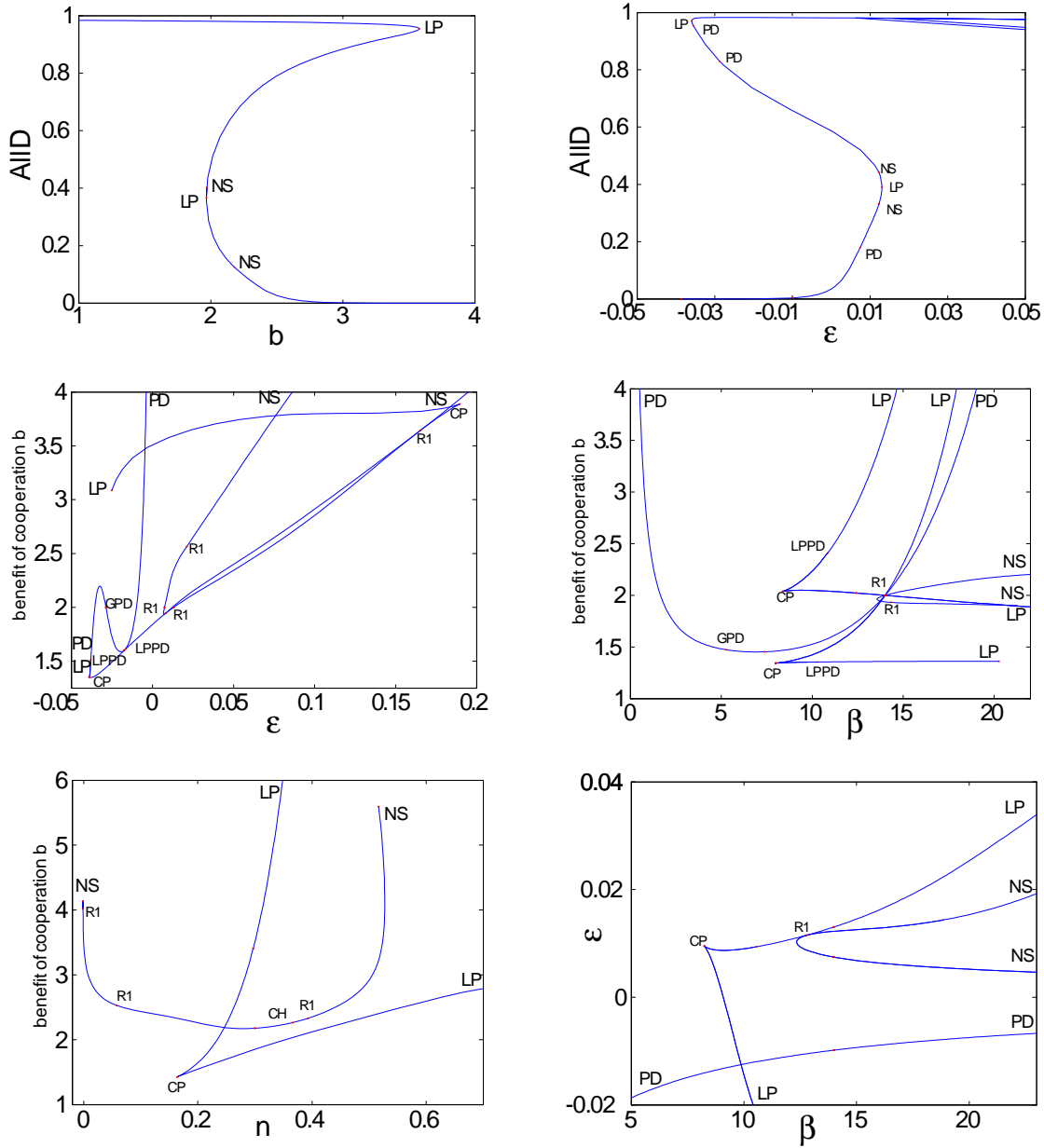


Figure 6: AllID vs. GTFT vs. WSLs. Top panels: continuation of an equilibrium in the benefit of cooperation and error probability space. Middle and bottom panels: continuation of the Period-Doubling (PD), Saddle-Node (LP) and Neimark-Sacker (NS) singularities detected in the top two panels with respect to a second parameter: intensity of choice β , degree of generosity n , benefit of cooperation b and error in implementation ϵ . Baseline game parameters: $b = 2, c = 1, \epsilon = 0.01, n = 0.1, \beta = 15$.

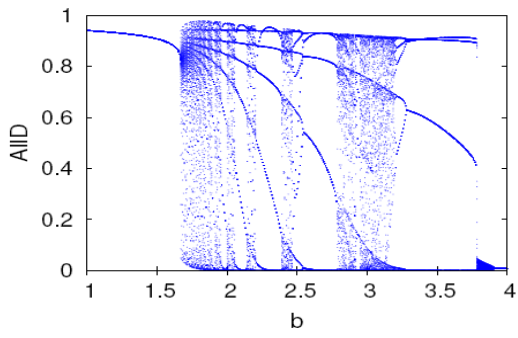
4.2 AllD-GTFT-AllC

This ecology is a generalization of the one investigated thoroughly in Sigmund and Brandt (2006) under Replicator Dynamics¹². The restriction of the full payoff matrix (3) to this 3×3 sub-ecology, can be arranged, after the normalization of diagonal payoffs into the following form:

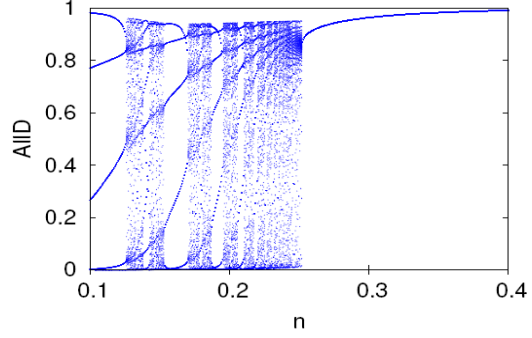
$$\left[\begin{array}{ccc} & \textit{AllD} & \textit{GTFT} & \textit{AllC} \\ \textit{AllD} & 0 & -\frac{1}{n+\varepsilon} (n - \varepsilon^2 - n\varepsilon) (b - c - bn - b\varepsilon) & c(1 - 2\varepsilon) \\ \textit{GTFT} & -c(n - \varepsilon^2 - n\varepsilon) & 0 & c\varepsilon(1 - n - \varepsilon) \\ \textit{AllC} & -c(1 - 2\varepsilon) & \frac{\varepsilon}{n+\varepsilon} (1 - n - \varepsilon) (b - c - bn - b\varepsilon) & 0 \end{array} \right] \quad (10)$$

It is easy to see that, for $\varepsilon, n \rightarrow 0$ this normalized payoff matrix has a generalized Rock-Paper-Scissors structure, i.e. *AllD* outcompetes *AllC*, *AllC* outcompetes *GTFT* and *GTFT* outcompetes *AllD*. In this case the logit dynamics has a limit cycle as an attracting set for moderate β and a three cycle with perpetual switching between the three monomorphic steady states in the β large limit. For strictly positive perturbations ε, n the game is not RSP *per se* but, nevertheless, the Logit Dynamics generates, for low to intermediate values of the intensity of choice, periodic behavior resembling the described RSP cycles. Reasoning for the emergence of such limit cycles in Fig. 7c-d goes along the best-reply correspondences of the reduced 3x3 game 4.1 in table (2). Unlike the *AllD* – *TFT* – *AllC* ecology, the shape of the curve, as well as the oscillations bounds, are distorted, with the original ones recovered in the limit $n \rightarrow \varepsilon$ when *GTFT* approaches the *TFT* strategy. However, in the $\beta \rightarrow \infty$ limit system converges to a monomorphic *AllD* state (Fig. 7e-f). The presence of indiscriminating cooperators leads to the extinction of the generous reciprocators.

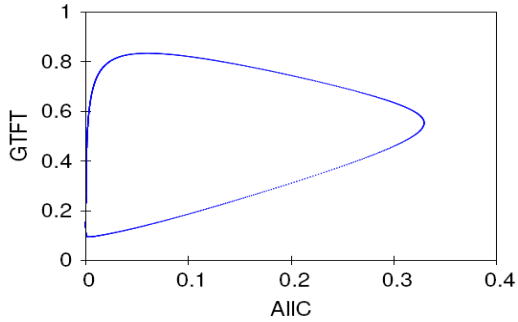
¹²Sigmund and Brandt (2006) also discover, when mistakes probability $\varepsilon \rightarrow 0$, a *RSP* pattern under Replicator Dynamics, for a similar ecology of *IPD* game rules, with discounting.



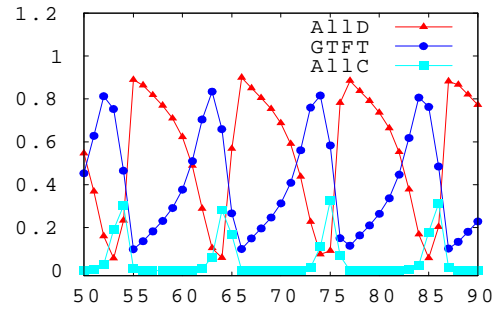
(a) $n = 0.2$. Bifurcation diagram



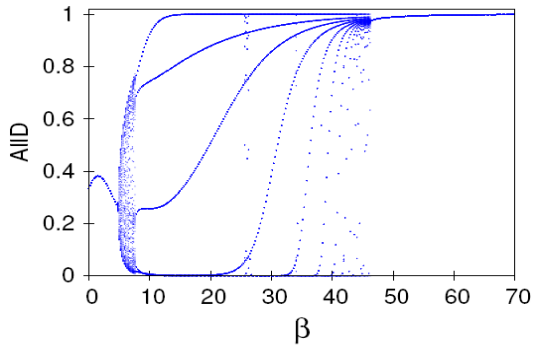
(b) $\varepsilon = 0.01$. Bifurcation diagram



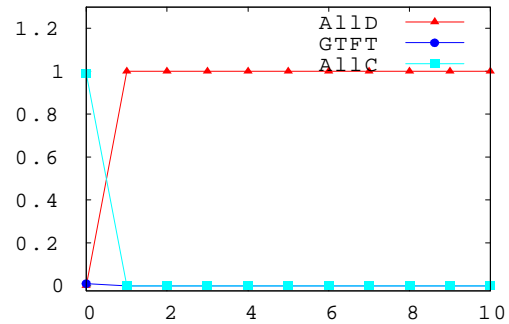
(c) $b = 1.72$. Phase plot



(d) $b = 1.72$. Time series



(e) $\varepsilon = 0.01$. Bifurcation diagram



(f) $\beta = 100$. Time series

Figure 7: AllD vs. GTFT vs. AllC. Bifurcation diagrams with respect to the benefit of cooperation parameter b (Panel (a)) and degree of generosity n (Panel (b)). Phase plots in AllC-GTFT space together with Rock-Paper-Scissors time series are reported in Panels (c) and (d), respectively. Panels (e),(f) display the long run behaviour in the β large limit. Baseline game parameters: $b = 2, c = 1, \varepsilon = 0.01, n = 0.1, \beta = 15$,

4.3 AllD-TFT-WSLS

The payoff matrix for this sub-ecology takes the following form:

$$\begin{bmatrix} & \textit{AllD} & \textit{TFT} & \textit{WSLS} \\ \textit{AllD} & 0 & (\frac{1}{2} - \varepsilon)(c - b + 2b\varepsilon) & (\varepsilon - \frac{1}{2})(b - 2c - 4b\varepsilon + 4c\varepsilon + 4b\varepsilon^2 - 4c\varepsilon^2) \\ \textit{TFT} & c\varepsilon(2\varepsilon - 1) & 0 & \frac{1}{2}(b - c)(2\varepsilon - 1)^3 \\ \textit{WSLS} & c(\varepsilon - \frac{1}{2}) & 0 & 0 \end{bmatrix} \quad (11)$$

which, in the limit $\varepsilon \rightarrow 0$, simplifies to:

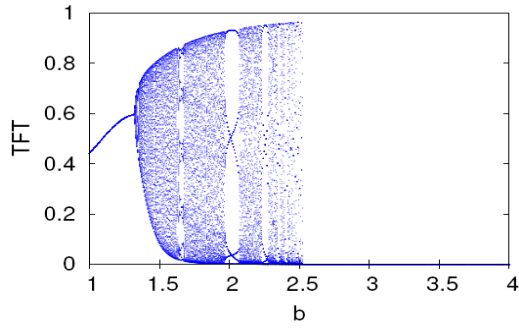
$$\begin{bmatrix} & \textit{AllD} & \textit{TFT} & \textit{WSLS} \\ \textit{AllD} & 0 & -\frac{1}{2}(b - c) & c - \frac{1}{2}b \\ \textit{TFT} & 0 & 0 & -\frac{1}{2}(b - c) \\ \textit{WSLS} & -\frac{1}{2}c & 0 & 0 \end{bmatrix}$$

For relatively small benefits of cooperation ($c < b \leq 2c$) we see that *AllD* loses to *TFT* ($(\frac{1}{2} - \varepsilon)(c - b + 2b\varepsilon) < 0$). Unconditional defectors *AllD* win against *WSLS* ($(\varepsilon - \frac{1}{2})(b - 2c - 4b\varepsilon + 4c\varepsilon + 4b\varepsilon^2 - 4c\varepsilon^2) > 0$) and tie against own-type. Thus, this matrix has a degenerate generalized Rock-Paper-Scissors structure and limit cycles are born via a Hopf bifurcation, for low¹³ to moderate values of the intensity of choice parameter β (Fig. 8a-d) with a 3-cycle arising in the $\beta \rightarrow \infty$ limit. For strictly positive mistake probability ε there are still RPS cycles born, but only for moderate values of the intensity of choice, while, for large intensity of choice, the population ends up in an *AllD* monomorphism (Panel(e)). The intuition for these oscillations could be extracted from the best-response structure of the corresponding pairwise interactions 2.1, 2.5 and 2.9 in table (1): for small errors of implementation *TFT* weakly dominates *AllD*, in turn, is it weakly dominated by *WSLS* and, finally, *AllD* easily invades a Pavlovian population when be-

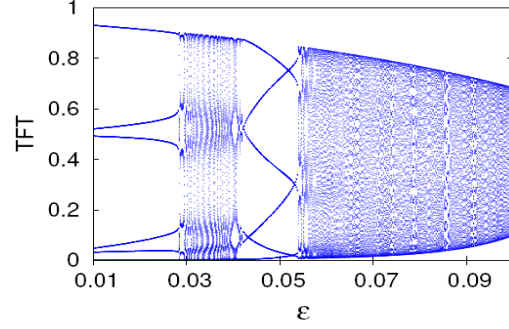
¹³If $\beta = 0$, the repeated strategy choice is random and the unique steady state is the fully mixed population state.

benefits accrued to cooperation are small (see Fig. 3a). For strictly positive ε and $\beta \rightarrow \infty$ the population converges to an *AllD* monomorphic state.

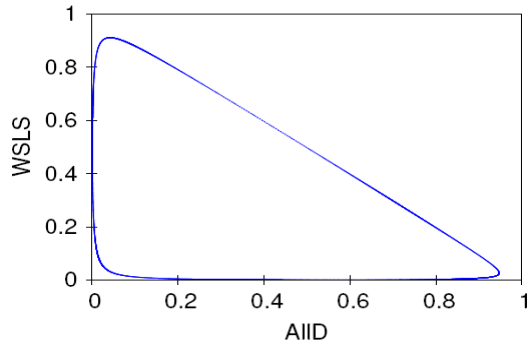
If the relative benefits are large ($b > 2c$), the RSP structure disappears for $\varepsilon \rightarrow 0$, and, for β large the population converges, irrespective of the initial mixture, to a Pavlov-only state (Panel (f)). Notice that, even if the underlying game itself is no longer RSP, there are still oscillations for moderate β due to the evolutionary dynamics.



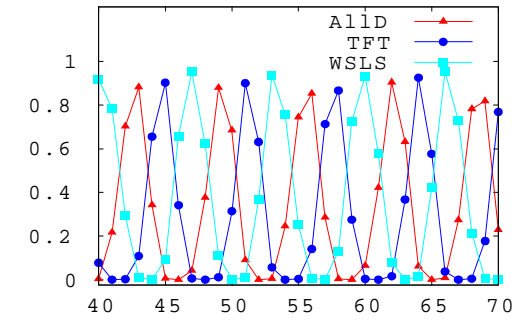
(a) $n = 0.2$. Bifurcation diagram, b



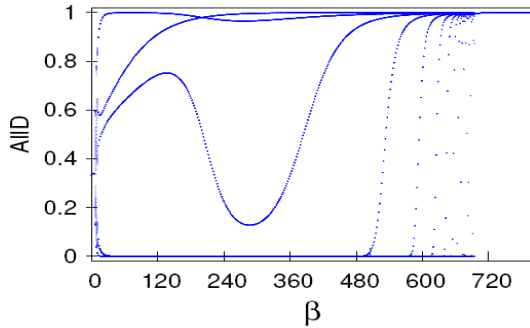
(b) $n = 0.2$. Bifurcation diagram, ε



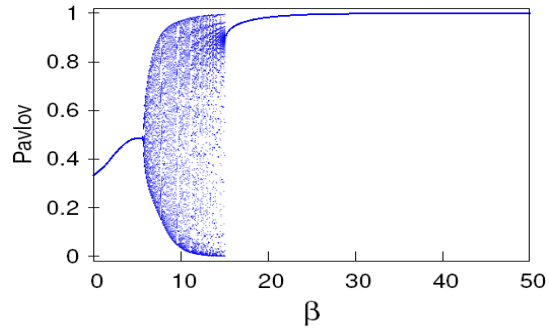
(c) $b = 2.16$. Phase plot



(d) $b = 2.16$. Time series



(e) $b \leq 2c$. Bifurcation diagram



(f) $b > 2c$. Bifurcation diagram

Figure 8: AllD vs. TFT vs. WSLS. Bifurcation diagrams with respect to the benefit of cooperation parameter b (Panel (a)) and probability of mistakes (Panel (b)). Phase plots in AllD-WSLS (Panel (c)) and time series (Panel (d)) displaying Rock-Paper-Scissors type of long-run behaviour. Panels (e)-(f) show long-run behavior and convergence, for large β , to AllD (Pavlov) monomorphism for $b < 2c$ ($b > 2c$, respectively). Baseline game parameters: $b = 2, c = 1, \varepsilon = 0.01, n = 0.2, \beta = 15$,

4.4 AllD-WLSL-AllC

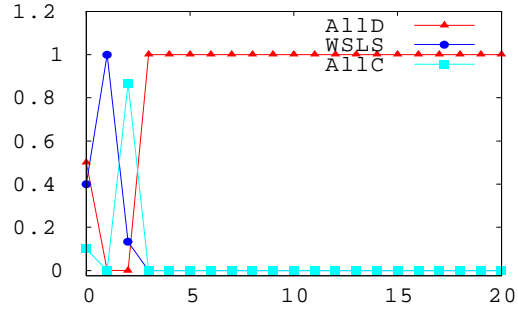
The reduced 3×3 payoff matrix for an ecology of defective, stimulus-response and cooperative players takes the following form:

$$\begin{bmatrix} & \textit{AllD} & \textit{WLSL} & \textit{AllC} \\ \textit{AllD} & \varepsilon(b-c) & \frac{1}{2}b - c\varepsilon & b - b\varepsilon - c\varepsilon \\ \textit{WLSL} & b\varepsilon - \frac{1}{2}c & (b-c)(1 - 4\varepsilon^3 + 6\varepsilon^2 - 3\varepsilon) & b - \frac{1}{2}c - b\varepsilon \\ \textit{AllC} & b\varepsilon - c + c\varepsilon & b - c - 2b\varepsilon + c\varepsilon + 2b\varepsilon^2 & (1 - \varepsilon)(b - c) \end{bmatrix} \quad (12)$$

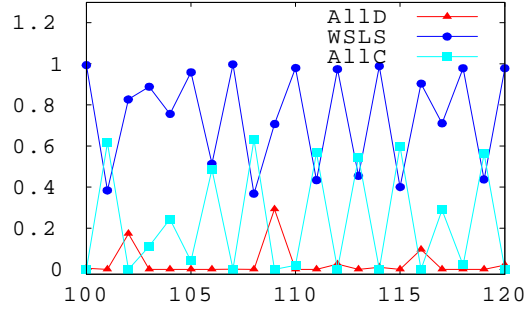
We notice first, that *AllC* is strictly dominated by *WLSL* provided that $b < 2c\frac{1-\varepsilon}{1-2\varepsilon}$, the necessary and sufficient condition for Pavlov to be a best-reply to itself. In this case the equilibria structure is given by the equilibria of the *AllD* vs. *WLSL* game - (*AllD*, *AllD*) and (*WLSL*, *WLSL*) - in reduced game form 2.5. But, in the limit $\varepsilon \rightarrow 0$, the ecology *AllD* - *WLSL* was dominance solvable for $b < 2c$ with *WLSL* the strictly dominated strategy. Combining the two restrictions, we obtain (*AllD*, *AllD*) as the only remaining Nash equilibrium in our 3x3 ecology for small benefits of cooperation $b < 2c$.

If $b > 2c\frac{1-\varepsilon}{1-2\varepsilon}$, then *AllC* is a best-reply to *WLSL*, but *WLSL* is not a best reply to *AllC* (as $b - b\varepsilon - c\varepsilon > b - \frac{1}{2}c - b\varepsilon$) so both the Coordination and the Hawk-Dove patterns of the *AllD* - *WLSL* and *WLSL* - *AllC* ecologies (see restricted game 2.10) are broken. Again we are only left with (*AllD*, *AllD*) as the unique pure strategy equilibrium.(Fig. 9a). We observe the same phenomenon as with the *AllD* - *GFTF* - *AllC* ecology, i.e. the presence of indiscriminating players *AllC* leads to the extinction of the discriminating type Pavlov.

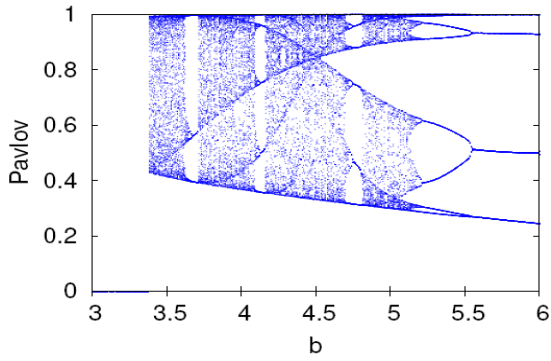
Interestingly, if we relax the $\beta \rightarrow \infty$ assumption, the logit dynamics easily generates periodic (Fig. 9b-c) or chaotic behavior with irregular switching between the three strategies. Fig. (9d) shows a plot of the largest Lyapunov exponent λ , with $\lambda > 0$ implying chaos. In this case complicated dynamics arises through a period-doubling route to chaos.



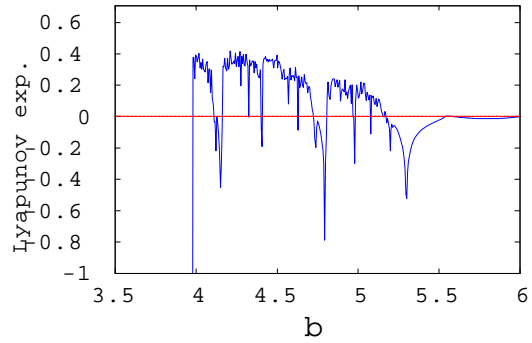
(a) $b > 2c, \beta = 300$. Time series



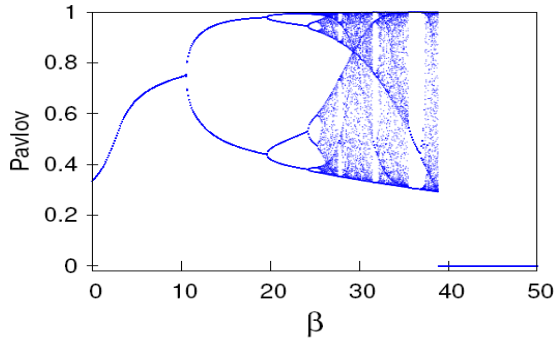
(b) $b > 2c, \beta = 30$. Time series



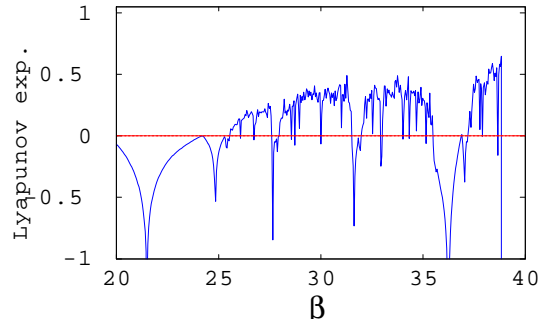
(c) $\beta = 30$. Bifurcation diagram



(d) $\beta = 30$. Largest Lyapunov exponent



(e) $b = 4.4$. Bifurcation diagram



(f) $b = 4.4$. Largest Lyapunov exponent

Figure 9: AllD vs. WSLs vs. AllC. Time series evolution for small and high values of β (Panels (a)-(b)). Bifurcation diagram with respect to the benefit of cooperation b (intensity of choice β) and evidence for chaotic behavior in Panels (c)-(d) and (e)-(f), respectively. Game parameters: $b = 4, c = 1, \varepsilon = 0.01, n = 0.3$.

4.5 Summary

Table 2 summarizes the qualitative long-run behavior of logit dynamics on various 3×3 sub-ecologies of PD repeated strategies. There are two basic routes to periodic behavior: the first one is the limiting case $\varepsilon \rightarrow 0$ when five out of the ten normal form games unveil a (degenerate) generalized Rock-Paper-Scissors structure. The Logit Dynamics on this generalized RSP game gives rise to stable limit cycles created via a supercritical Neimark-Sacker(Hopf) bifurcation with the three monomorphisms connected via a 3-cycle in the $\beta \rightarrow \infty$ best-reply limit. In the second route to oscillating behavior, when we maintain the strictly positive error parameter ε , the game forms are not of RPS type, but, nevertheless, for moderate values of the intensity of choice stable oscillations created via a supercritical NS bifurcation are observed. However, for such non-RPS games, in the limiting case $\beta \rightarrow \infty$, the cyclical structure disappears with one of the strict equilibria selected as the long run outcome.

A somewhat different scenario occurs for the *AllD-GTFT-WLSL* 3×3 coordination game, where a sequence of Neimark-Sacker and Fold bifurcations unfolds as the intensity of choice slides from low to high values: the long-run behavior varies from a unique polymorphic steady state for low β , through limit cycles to three co-existing steady states corresponding to the underlying game Nash equilibria in the best-reply limit.

In the transition to the *AllC* ecologies, in the second half of the table, we have found one 3×3 example of chaotic dynamics created through a period-doubling route to chaos. Chaos appeared for finite intensity of choice β , in the *AllD-Pavlov-AllC* ecology, whose underlying payoff matrix has a *non-RSP* structure even for the limiting case $\varepsilon \rightarrow 0$. The typical RSP cyclical dominance structures seem to preclude the onset of chaos.

In the last four ecologies in Table 2 with no defectors (*AllD*), the long-run, $\beta \rightarrow \infty$ outcomes consist, generally, of simple behavior, steady states or 2-cycle with continuous population swings between one monomorphic state to another, with the third strategy

extinct from the population.

No.	3×3	Bifurcation	Attractors	
			moderate β	$\beta \rightarrow \infty$
3.1	AllD-TFT-AllC	NS	limit cycle	3-cycle
3.2	AllD-GTFT-WLSL	NS,LP	limit cycle	3 steady states
3.3	AllD-GTFT-AllC	NS	limit cycle	1 steady state (AllD)
3.4	AllD-TFT-WLSL	NS	limit cycle	1 steady state (Pavlov)
3.5	AllD-TFT-GTFT	NS	limit cycle	1 steady state (GTFT)
3.6	AllD-WLSL-AllC	PD	2-cycle, chaos	1 steady state (AllD)
3.7	TFT-WLSL-AllC	PD	2-cycle	stable steady state
3.8	TFT-GTFT-WLSL	PD	2-cycle	2 steady states
3.9	GTFT-WLSL-AllC	PD	2-cycle	2-cycle
3.10	TFT-GTFT-AllC	PD	2-cycle	2-cycle

Table 2: Summary of the dynamical behaviour for all 10 3×3 interaction of iterated Prisoner's Dilemma strategies under logit dynamics.

5 4×4 ecologies

Before studying the full 5×5 ecology, in this section we discuss the long-run behavior of typical 4-types interactions under the Logit Dynamics.

5.1 No TFT

This ecology appends the *AllD – GTFT – WLSL* subset (subsection (4.1)) with unconditional cooperators. In the β large limit, *AllC* destroys the Coordination game structure of *AllD – GTFT – WLSL* and the population converges to an *AllD* monomorphism, irrespective of the initial conditions (Fig. 10a). This is due, on the one hand, to the Hawk-Dove nature of the *WLSL – AllC* and *GTFT – AllC* interactions-which breaks the best-response correspondences in (4.1) $\{WLSL\} = BR\{WLSL\}$ and $GTFT = BR\{GTFT\}$ - and, on the other hand, to the unconditional defectors *AllD* "easiness" of invading indiscriminating cooperators *AllC* relative to *WLSL* and *GTFT*. Briefly, adding "cooperative" players leads to the extinction of the two discriminating types (*GTFT* and *WLSL*) and,

consequently, the overall level of long-run cooperation drops. However, for moderate β , Pavlov seems to be favoured by the evolutionary dynamics¹⁴ (see Panel (e) for the evolution of an initially balanced population) and rich dynamics of the *WSLS* fraction unfold if either the mistake probability (ε) or the intensity of choice is altered (Fig. 10b-c). Co-existence of attractors - one chaotic (Panel (d) plots the largest Lyapunov exponent as numerical evidence for chaos) and an *AllD* monomorphic steady state - is detected for intermediate values of β while the entire population becomes *AllD* for large values of the intensity of choice. Panels (e)-(f) show plots of fractions evolution on the two attractors: a fully mixed population converges to the chaotic attractor (Panel (e)) while a population with no reciprocators (Panel (f)) ends up in a defectors *AllD* only state. This also illustrates the critical role *TFT* players have in the emergence and success of Pavlovian (*WSLS*) types.

¹⁴Also, from subsection (3.4) (*WSLS, WSLS*) equilibrium is more likely to emerge because of a larger basin of attraction compared to (*GTFT, GTFT*); still this dependence of initial distribution (even though basins are largely asymmetric) seems to be lost once we investigate the enlarged, no-*TFT* ecology.

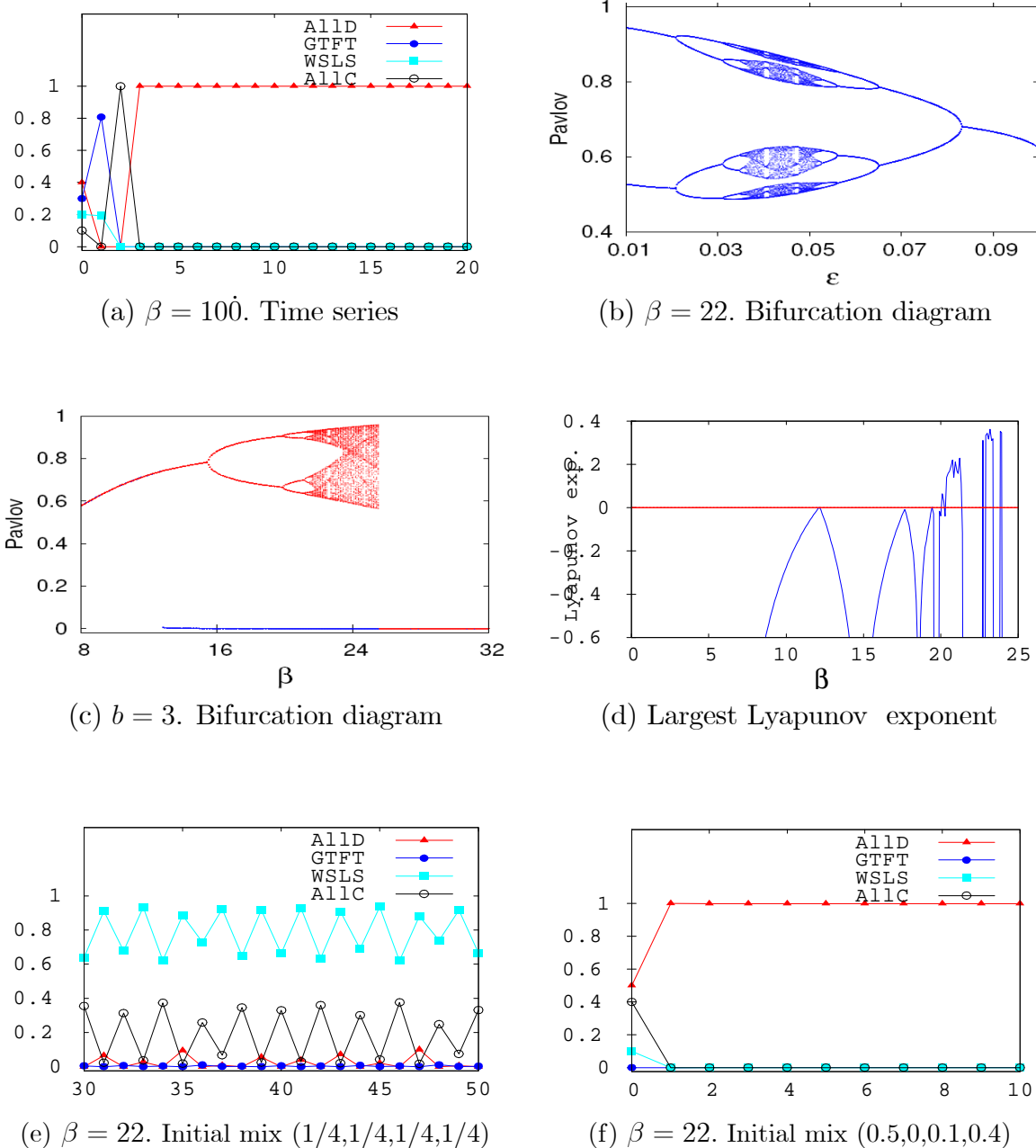


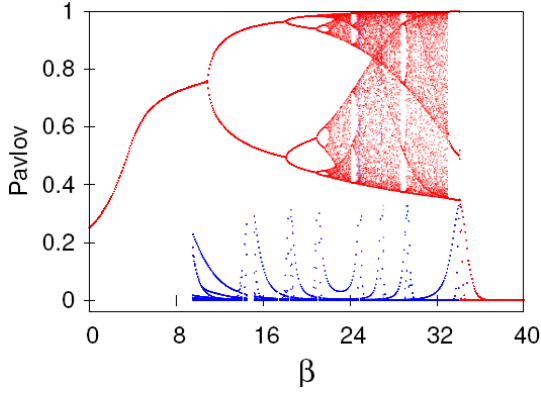
Figure 10: AllD vs. GTFT vs. WSLs vs. AllC. Evolution towards an AllD population for β large limit. Panels (b)-(c) plot bifurcation diagrams with respect to the error in strategy implementation ε and intensity of choice β parameters, while numerical evidence, i.e. positive Lyapunov exponent, for chaos in Panel (d). Last, evolution of fractions on the two co-existing attractors, for the boundedly rational choice, is displayed in Panels (e) and (f). Baseline game parameters: $b = 3, c = 1, \varepsilon = 0.01, n = 0.3$.

5.2 No GTFT

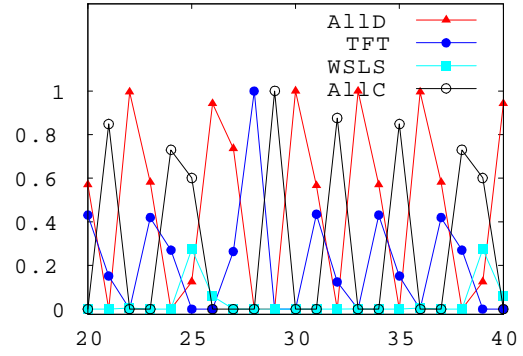
Unlike the previous case, the 4×4 ecology without *GTFT*¹⁵ exhibits path-dependence. Hence, for moderate values of the intensity of choice $\beta \approx 27$, the system displays *co-existence of attractors*. Depending on the initial population mix, the long-run evolutionary outcome can be either the cyclical competition between *AllD*, *TFT* and *AllC* with asymmetric amplitudes (time series Fig. 11b) as in subsection (??), or a 1-piece chaotic attractor (time series in Fig. 11c) with Pavlov having a dominant position, as in, for instance, subsection (4.4). In both cases, the other strategies remain present in the population, but in small numbers.

For high values of β the path-dependence is lost, the only remaining attractor being a full-ranged Rock-Paper-Scissors 3-cycle (Panel (d)) in *AllD*, *TFT* and *AllC*, with stimulus-response players *WSLS* going extinct. The bifurcation diagram in Panel (a) organizes the long-run outcomes of this ecology and displays the two co-existing attractors: the chaotic one in the top and the cyclical one in the bottom curves.

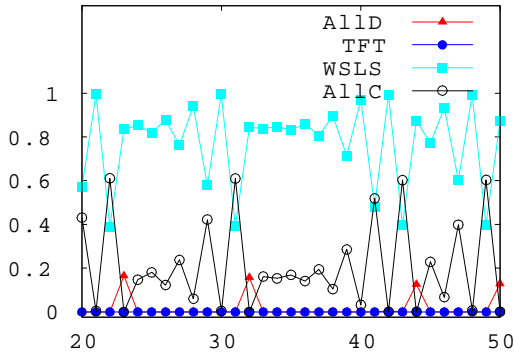
¹⁵Imhof et al. (2007) also discuss, in a *finite* population framework, an ecology of *AllD*, *TFT*, *WSLS* and *AllC* behaviors, under the frequency-dependent, mutation-selection Moran process. They show that *AllD* is selected for stage game payoffs (i.e. b/c ratio) below a certain threshold, while *WSLS* wins the evolutionary competition otherwise.



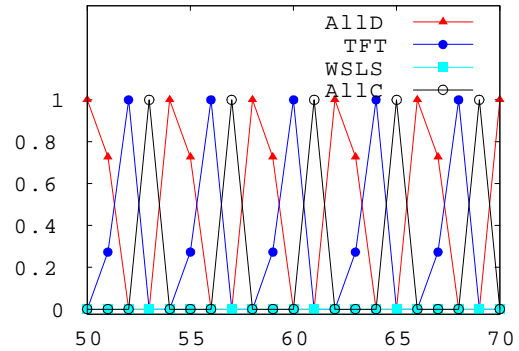
(a) $b = 4$. Bifurcation diagram showing co-existence of attractors



(b) small $\beta = 27$. Time series, initial population mixture: $(1/4, 1/4, 1/4, 1/4)$



(c) small $\beta = 27$. Initial population mixture: $(10\%, 10\%, 50\%, 30\%)$



(d) large $\beta = 100$. Initial population mixture: $(1/4, 1/4, 1/4, 1/4)$

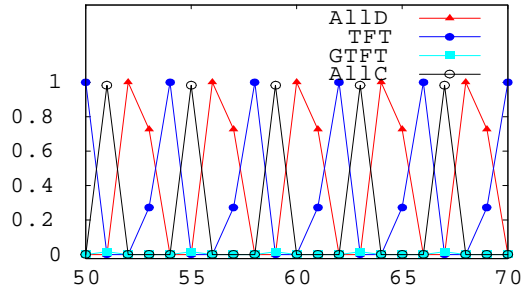
Figure 11: AllD vs. TFT vs. WSLs vs. AllC. Panel (a) bifurcation diagram of Pavlov fractions with respect to the intensity of choice parameter; Panels (b)-(c): depending on initial distribution of fractions the long-run population state may be either ongoing oscillations among AllD, TFT, AllC (Panel (b)) or co-existence, although in a possibly chaotic fashion, of WSLs and AllC players with WSLs largely outnumbering unconditional cooperators (Panel (c)). Last, Panel (d) displays the large β limit of the long-run dynamics of fractions. Remaining game parameters: $b = 4, c = 1, \varepsilon = 0.01, n = 0.3$.

5.3 No WSLs

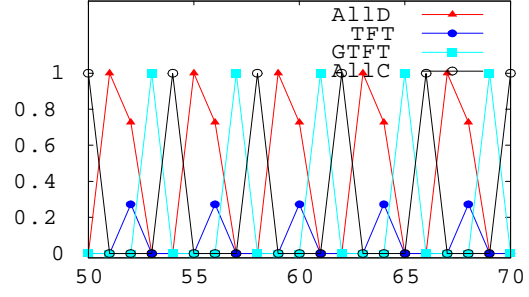
With the stimulus-response *WSLS* rule put aside, the evolutionary dynamics becomes very complicated. In the β large limit, co-existence of the four strategies in a cyclical

fashion, is observed for high benefits of cooperation b (Fig. 12a) while, for low benefits b , $GTFT$ is driven extinct and full-scale cyclical dynamics with the remaining three strategies emerge (Panel (b)).

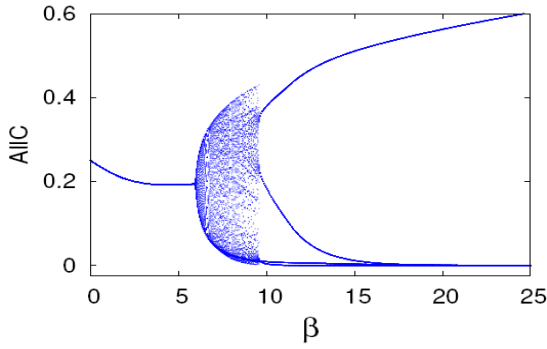
For finite β , a stable limit cycle is born via a Neimark-Sacker bifurcation (Fig. 12a) for small b , while a period-doubling route to chaos emerges for high benefits accrued to cooperation (Panel (d)). Finally, in Fig. 12e a period-doubling route to chaos is depicted (positive Lyapunov exponent plot in Panel (f)) as the generosity parameter n built into the $GTFT$ strategy is fine-tuned.



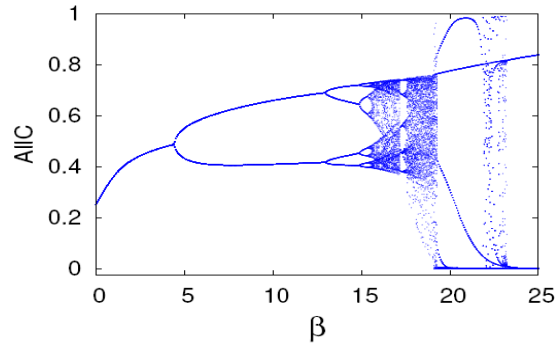
(a) $b = 2$, large $\beta = 100$. Time series



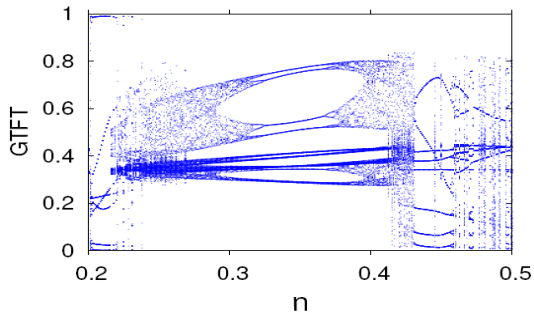
(b) $b = 6$, large $\beta = 100$. Time series



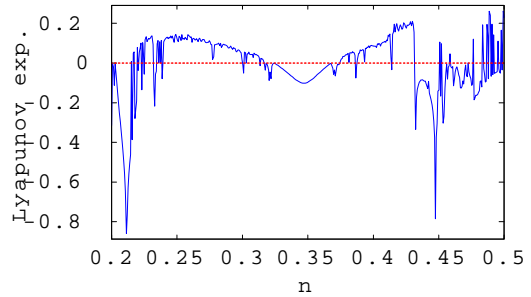
(c) $b = 2$. Bifurcation diagram β



(d) $b = 6$. Bifurcation diagram β



(e) $b = 6, \beta = 15$. Bifurcation diagram n



(f) $\beta = 15$. Largest Lyapunov exponent

Figure 12: AllD vs. TFT vs. GTFT vs. AllC. Time series of the corresponding fractions, for large β , and large/small benefit of cooperation (Panel (a) and (b), respectively) . Bifurcations diagrams for high and low b are also shown in Panel (c)-(d). Last, Panel (e) displays a bifurcation diagram of GTFT fraction with respect to the "generosity" parameter n while Panel (f) shows some numerical evidence for chaos. Remaining game parameters: $c = 1, \varepsilon = 0.01, n = 0.3$.

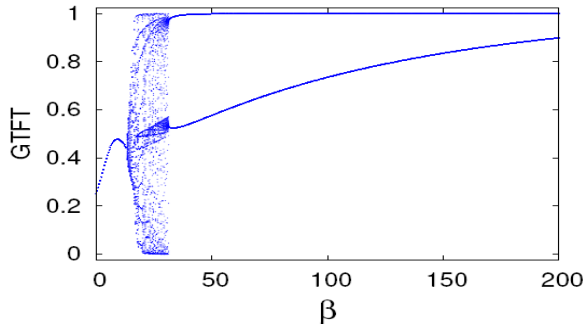
5.4 No AllC

There are two cases to be distinguished for an *AllD* – *TFT* – *GTFT* – *WSLS* ecology. These two situations emerge from the discussion of the corresponding 3×3 sub-ecologies (see subsections (4.1), (4.3) and (??), namely low ($b \leq 2c$) and high benefits of cooperation ($b > 2c$).

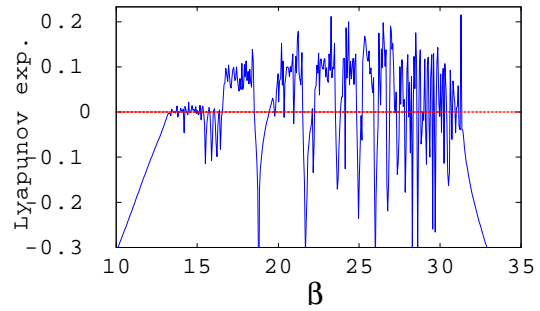
For low $b \leq 2c$ and in the $\beta \rightarrow \infty$ limit, the system converges to a unique monomorphic *GTFT* state (Fig. 13a). However, all four behavioral rules co-exist, either in the cyclical (Panel (a)) or even chaotic (Panels (b)-(d)) manner, for intermediate values of the intensity of choice.

As the responsiveness to payoffs differentials increases, quasi-periodic behavior emerges through a Neimark-Sacker bifurcation, which, if β is pushed even further, breaks into a chaotic attractor (see Fig. 13b for a plot of the positive largest Lyapunov exponent). The strange attractor depicted in Panel (d) emerges via one such Neimark-Sacker, "breaking of the invariant circle" route to chaos.

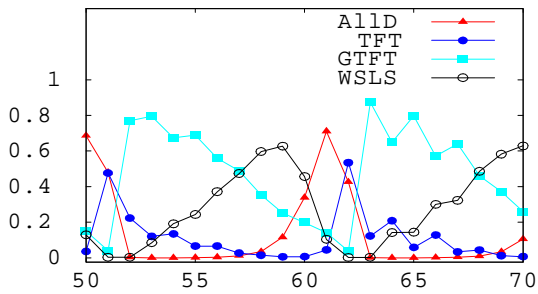
When benefits of cooperation are relative large ($b > 2c$) there is path-dependence producing the co-existence of two monomorphisms *GTFT* and *WSLS*, respectively. Panel (e) shows, for $b = 4$, the bifurcation diagram of *GTFT* fraction with respect to β and we can see the two steady states created for large responsiveness to payoff differences (at the top and bottom of the panel). For instance, an initial fractions distribution biased towards Pavlov converges to a Pavlov-only state (Panel (f)). Alternatively, an even set of initial conditions favors the *GTFT* as the long-run, good-fated type of behavior.



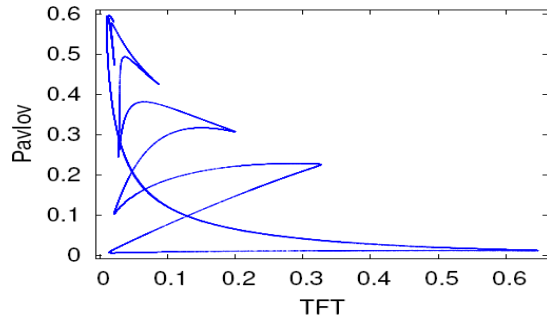
(a) $b = 2$. Bifurcation ($\beta, GTFT$)



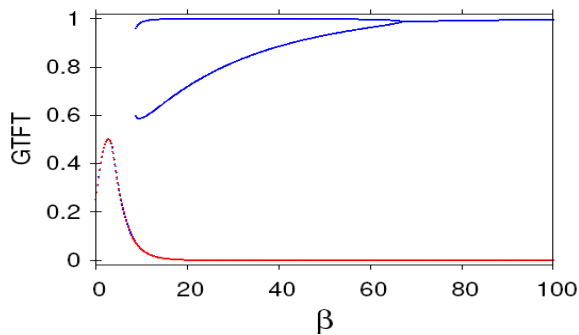
(b) $b = 2$. Lyapunov exponent



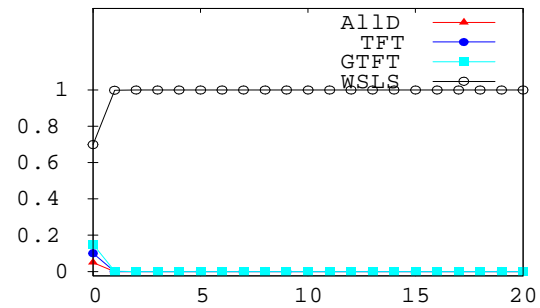
(c) $b = 2, \beta = 18$. Time series



(d) $b = 2, \beta = 17.6$. Strange attractor



(e) $b = 4$. Bifurcation diagram



(f) $\beta = 100$. Initial mix: (0.1, 0.1, 0.1, 0.7)

Figure 13: AllD vs. TFT vs. GTFT vs. WSLs. Panels (a): bifurcation diagram with respect to the intensity of choice showing quasi-periodic behavior. Panels (c)-(d): time series, phase portrait and Lyapunov exponent plot as numerical evidence for chaotic time series when the intensity of choice β increases. Finally, Panels (e)-(f) show the long-run behavior, in the best-reply limit, for large benefits of cooperation: this is, coexistence of monomorphic GTFT and Pavlov WSLs states. Baseline game parameters: $c = 1, \varepsilon = 0.01, n = 0.3$.

5.5 Summary

Table 3 summarizes the long-run behavior of all 4×4 ecologies both the best-reply (β large) and in the boundedly rational (β small) choice cases. In comparison to the 3×3 ecologies reviewed in the previous section, having more strategies not only enlarges the set of possible outcomes but also makes the behavior much more complicated. In particular there is now strong path-dependence in most of the ecologies, leading to the co-existence of complicated attractors, periodic and chaotic ones. We identified two routes to complex dynamics in these 4×4 ecologies: the period-doubling and breaking of the invariant circle route to chaos. However, the path dependence is lost for some ecologies (reduced 4x4 games 4.1, 4.2 and 4.3 in table (3)) when β approaches the "full rationality" limit. *WSLS* fares well in an *AllD* environment with no *AllC* players, but poorly, almost going extinct, in an *AllC* environment with *AllD* strategists around.

No.	4x4	Bifurcations	Attractors	
			small β	large β
4.1	No AllD	PD	2-cycle	2-cycle/unique SS
4.2	No TFT	PD, NS	co-existence SS and chaos	unique SS (AllD)
4.3	No GTFT	PD, NS	co-existence RSP and chaos	unique SS (AllD)
4.4	No WSLS	NS	limit cycles, chaos	4-cycle
4.5	No AllC	NS, PD	limit cycles, chaos	multiple SS

Table 3: Summary of the dynamical behaviour for all 4x4 interaction of iterated Prisoner's Dilemma strategies under logit dynamics.

6 5×5 Ecology

In this section, we report simulations about the behavior of Logit Dynamics on the full 5×5 ecology $AllD - TFT - GTFT - WSLS - AllC$ that generated the game matrix (3). Depending on the relative size of cooperation benefits we distinguish again two situations.

First, for small benefits, $b \leq 2c$ there is co-existence of all five strategies, albeit in a fluctuating manner, for low responsiveness to payoffs differences (Fig. 14a). When β reaches the best-response limit the long-run outcome is an $AllD-TFT-AllC$ 3-cycle (Fig. 14b).

If the benefits of cooperation are large, Pavlov wins the evolutionary competition for small β (Panel (c)) but goes extinct in the $\beta \rightarrow \infty$ limit, with a 4-cycle in the remaining strategies as long run behavior (Panel (d)). Last, Panels (e), (f) display bifurcation diagrams of $AllC$ and Pavlov fractions respectively with respect to the intensity of choice. It is apparent from the $AllC$ bifurcation diagram in Panel (e) that an initially stable steady state for very low β , destabilizes for moderate values, with continuous swings between high and low fractions emerging for large β . As far as Pavlov is concerned the corresponding bifurcation diagram shows its way to extinction when players choose a best-reply (Fig. 14f) .

The chaotic patterns in the two diagrams for intermediate values of β are confirmed in Fig. 15a-f that displays two dimensional projections of a chaotic attractor together with the fractions evolution on the attractor for $b = 2.4805$ and $\beta = 15$. Finally, Fig. 16c-d illustrates the "breaking of an invariant circle" route to chaos: a quasiperiodic attractor for $\beta = 0.05$ breaks into a 6-piece quasi-periodic and then chaotic attractor (Panels (d)-(e)) at $\beta \approx 10.05$ and $\beta = 10.2$. Eventually these pieces join together to form a 1-piece strange attractor (Fig. 16f) when the intensity of choice β reaches the threshold $\beta = 10.8$.

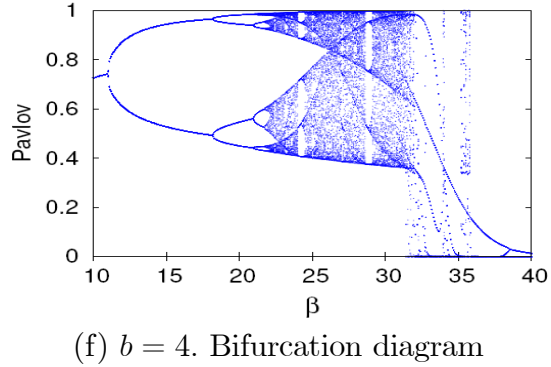
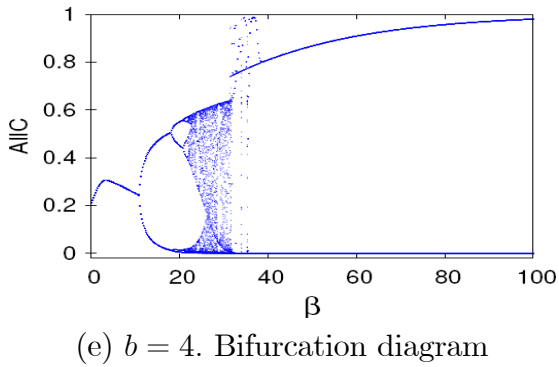
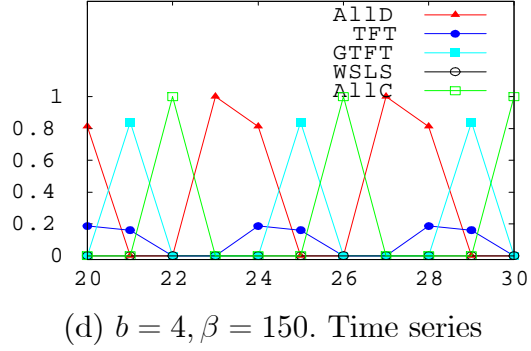
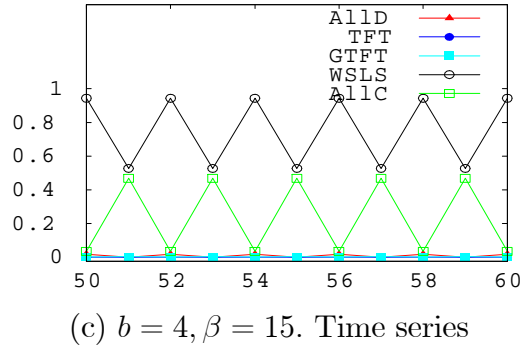
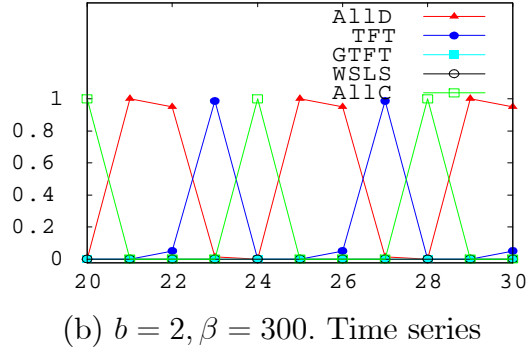
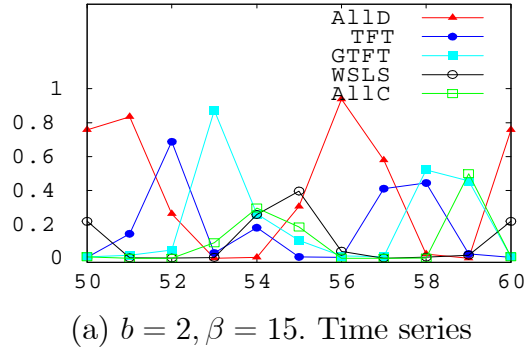
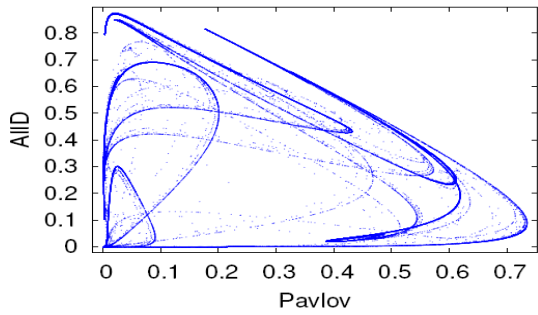
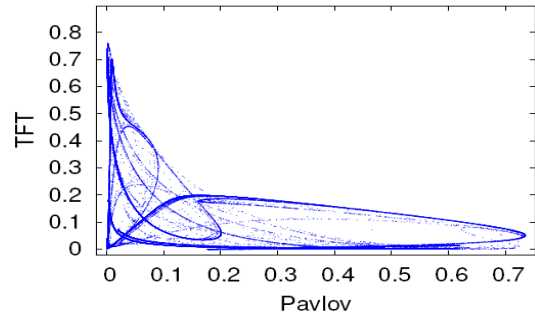


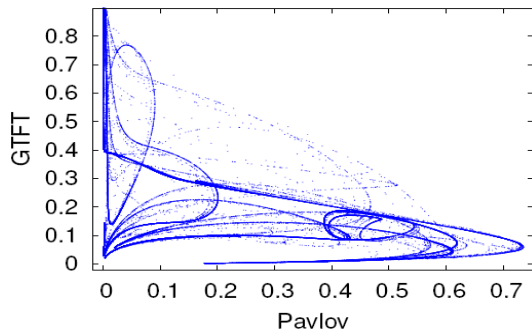
Figure 14: AllD vs. TFT vs. GTFT vs. WSLs vs. AllC. Panels (a)-(b) show the long-run outcome for low benefits of cooperation b in the β low and high limit, respectively. Similar time series are displayed in Panels (c)-(d) for the high benefit b case. Panels (e)-(f) show, for high b , bifurcation diagrams of AllC and Pavlov fractions. Baseline game parameters: $c = 1, \varepsilon = 0.01, n = 0.3$.



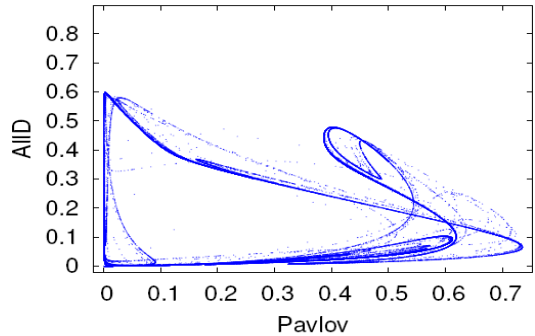
(a) $b = 2.48, \beta = 15$. Strange attractor



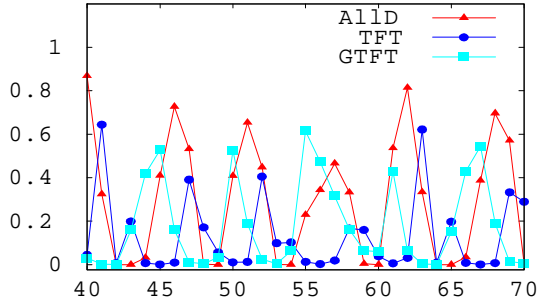
(b) $b = 2.48, \beta = 15$. Strange attractor



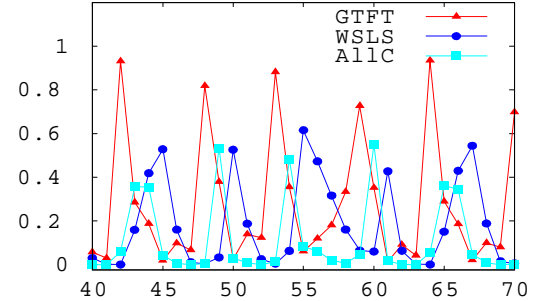
(c) $b = 2.48, \beta = 15$. Strange attractor



(d) $b = 2.48, \beta = 15$. Strange attractor

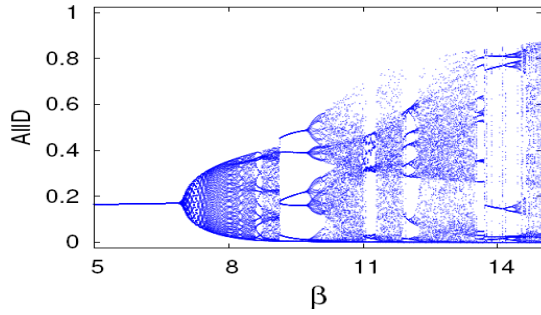


(e) $b = 2.48, \beta = 15$. Time series

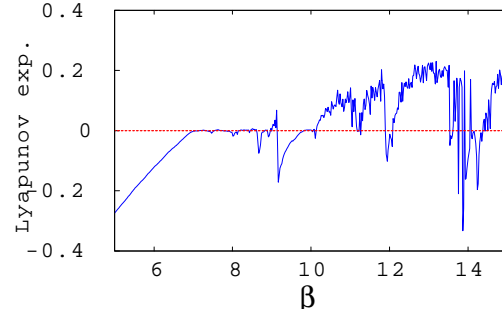


(f) $b = 2.48, \beta = 15$. Time series

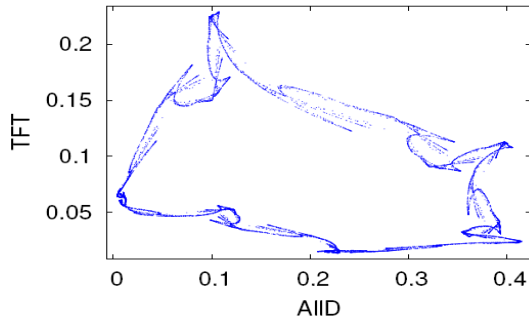
Figure 15: AllD vs. TFT vs. GTFT vs. WLSL vs. AllC. Panels (a)-(d): two-dimensional projections of a strange attractor for $b = 2.4805$ and $\beta = 15$. Panels (e)-(f) evolution of fractions on the strange attractor. Remaining game parameters: $c = 1, \varepsilon = 0.01, n = 0.3$.



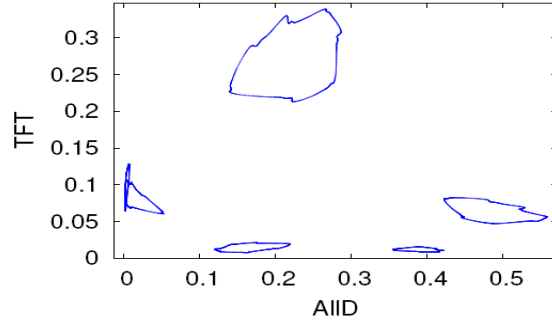
(a) $b = 2.48$. Bifurcation diagram β



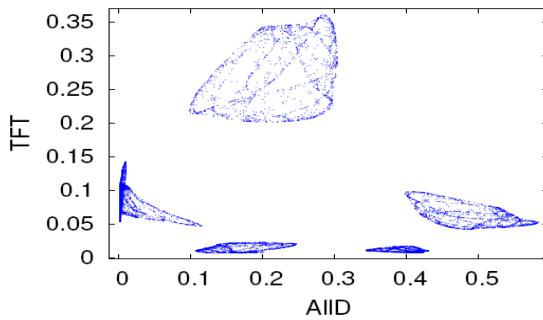
(b) $b = 2.48$. Lyapunov exponent



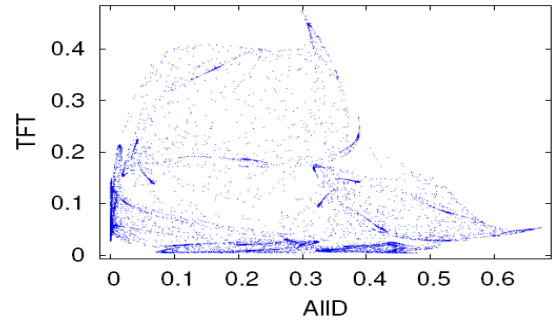
(c) $\beta = 9.05$. 1-piece attractor



(d) $\beta = 10.05$. 6-piece quasiperiodic



(e) $\beta = 10.2$. 6-piece chaotic



(e) $\beta = 10.8$. 1-piece chaotic

Figure 16: AllD vs. TFT vs. GTFT vs. WSLs vs. AllC. Panels (a)-(b): bifurcation diagram with respect to intensity of choice β and numerical evidence of chaos (plot of the largest Lyapunov exponent). Panels (c)-(f) show the evolution of a quasiperiodic attractor into a strange attractor as the intensity of choice increases. Remaining game parameters: $b = 2.48, c = 1, \varepsilon = 0.01, n = 0.3$.

6.1 Numerical Bifurcation Curves

The visual analysis of the bifurcation diagram in Fig. 14 can be confirmed by a rigorous continuation procedure, i.e. the computation of curves of equilibria along with their detected codimension I bifurcations as one parameter is varied using the Cl_Matcont bifurcation software package¹⁶ (Dhooge et al. (2003)).

One such fixed point - (0.4972065; 0.4178546; 0.06911757; 0.01521537; 0.01521537) for the game parameterization $b = 1.1; c = 1; \varepsilon = 0.01; n = 0.3; \beta = 14$ - is continued in Fig. 17 top left panel, as the benefit of cooperation b is increased. Three singularities are detected by the continuation package in the meaningful region [$b > c = 1$] of b : two Neimark-Sacker (NS) points at $b = 1.270046$ and $b = 2.636476$ ¹⁷ and one Period-Doubling (PD) point at $b = 3.158022$. These two points are next 'continued' with respect to another parameter as depicted with codimension II singularities¹⁸ detected along them. The bifurcation curves describe in a systematic way the qualitative changes in the behavior of the dynamical system in a certain parameters space. For instance, in the two bottom panels in Fig. 17 the plotted curves of Neimark-Sacker singularities partition the parameters space in regions with qualitatively similar behavior: when crossing a NS curve from below the system loses stability of the steady state and a stable limit cycles arises around the unstable steady state. Such an analysis reveals that the behavior envisaged by simulations is *robust* to perturbations of the repeated Prisoner's Dilemma game (as parametrized by the b/c ratio), the probability of mistake in implementing a particular strategy ε , and the generosity parameter n . Note that, compared to the 3×3 ecologies, there is no fold bifurcation as coordination on equilibria is more difficult to achieve with more strategies. Instead, increasing the strategy space generates oscillatory behavior more easily.

¹⁶Cl_Matcont for maps, see package documentation at <http://www.matcont.ugent.be/>

¹⁷normal form coefficients are -2.741700 and -5.440686 meaning that the NS bifurcations are supercritical, i.e. the limit cycles are born stable.

¹⁸e.g. a Chenciner CH codim II bifurcation gives rise to much more complicated patterns of behavior (see Kuznetsov (1995) pp. 482).

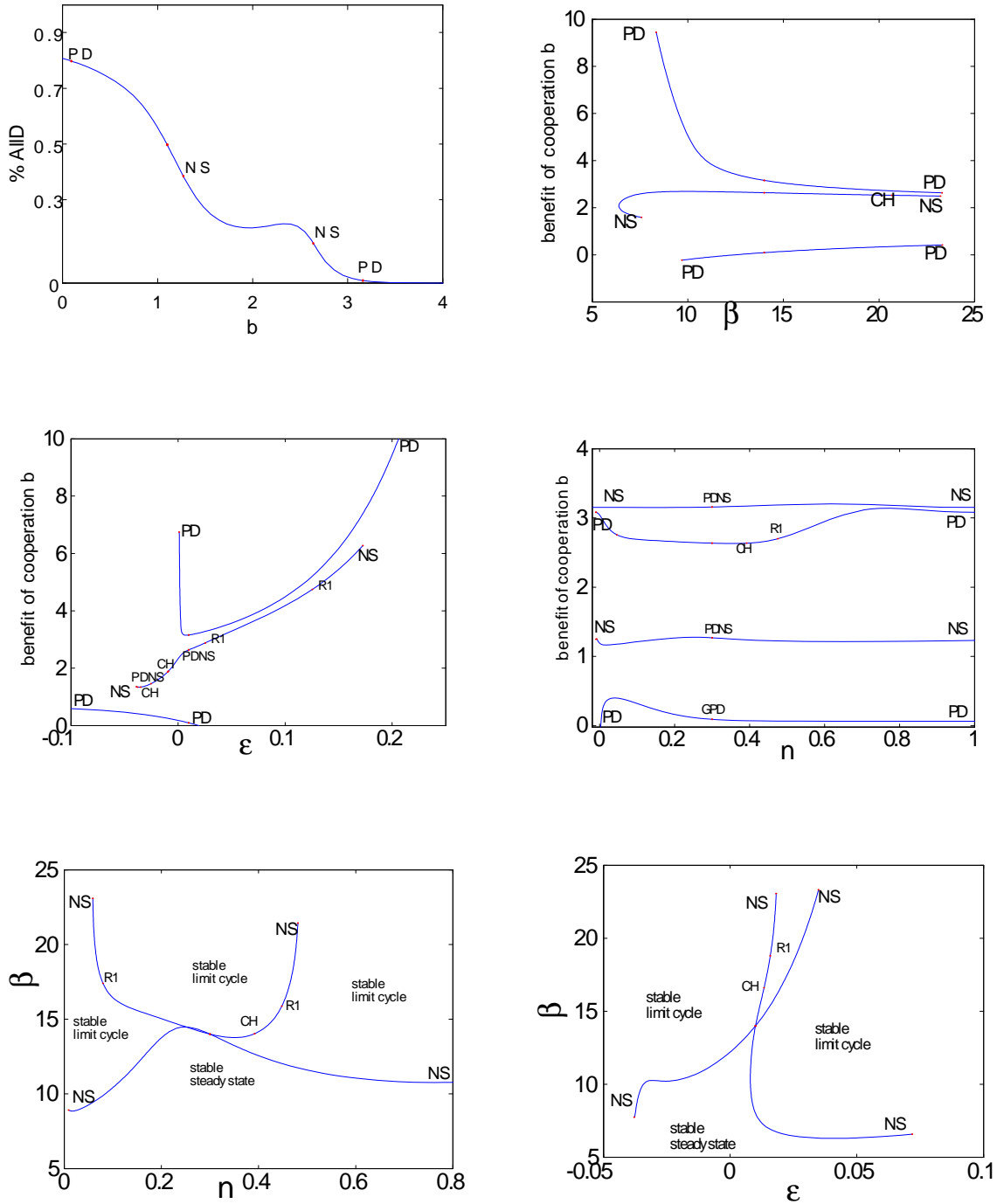


Figure 17: AllD vs. TFT vs. GTFT vs. WSLs vs. AllC. Top left Panel continues a numerically computed fixed point, in the benefit of cooperation space. Detected codimension I singularities are then continued with respect to another game/behavioral parameter. The resulting curves of Period-Doubling (PD) and Neimark-Sacker (NS) points, with respect to the intensity of choice β , probability of error in strategy execution ε , and degree of generosity n , are plotted in the top-right, middle and bottom panels, respectively. The occurrence of codimension II singularities can be observed along all these curves. Unless free to float, game parameters set to: $b = 2, c = 1, \varepsilon = 0.01, n = 0.3, \beta = 15$

7 Conclusions

In an evolutionary set-up, we append an ecology of iterated Prisoner's Dilemma (IPD) game strategies, consisting of unconditional cooperators (*AllC*), unconditional defectors (*AllD*) and reactive players (*TFT*) with two repeated strategies that have received less attention in the evolutionary IPD game literature, the error-proof, "generous" tit-for-tat (*GTFT*) which, with a certain probability, re-establishes cooperation after a (possibly by mistake) defection of the opponent and the penitent, "stimulus-response" (*WSLS*) strategy that resets cooperation after the opponent punished for defection. Stable oscillations in the frequency of both the forgiving (*GTFT*) and repentant (Pavlov) strategy along with chaotic behavior emerge under a perturbed version of best-response dynamics, the logit dynamics. We have performed a detailed analysis of the relatively simple 2×2 and 3×3 case with the dynamic behaviour now fairly well understood. However, the enlarged 4×4 and 5×5 exhibit complicated behavior via period-doubling and breaking of an invariant circle routes to chaos. Finite intensity of choice β leads to rich dynamics, path-dependence and co-existence of cyclical and chaotic attractors for a wide selection of 4×4 ecologies. If we turn to the best-reply limit of the logit dynamics, various 3- and 4-cycles are exhibited in specific 4×4 and in the 5×5 ecologies. Last, in terms of individual strategies performance, we discovered first that the presence of unconditional cooperators turns out detrimental to the discriminating types (*TFT*, *GTFT* and *WSLS*) in some 4×4 interactions leading the population to an *AllD* monomorphism. Second, there is mixed evidence for a Pavlovian meta-rule in an evolutionary repeated PD with a population of rational players: on the one hand, the stimulus-response strategy does very well and wins the evolutionary competition in a 4×4 environment with hard defectors (*AllD*) but no indiscriminating cooperators (*AllC*), but it almost goes extinct in the full 5×5 repeated game. Nevertheless, with boundedly rational players, the fractions of Pavlov maintain high values even within the complete ecology of five rules.

A Iterated Prisoner's Dilemma: stationary distributions and average payoffs

For each pairwise interaction of the iterated PD strategies $S_i \times S_j$, the stationary distribution $\Omega = \{\tau_{CC}, \tau_{CD}, \tau_{DC}, \tau_{DD}\}$ gives the average fraction of time system spends in each of the four states CC, CD, DC, DD , respectively. Average expected payoffs are computed as weighted average of the stage game matrix payoffs with weights obtained from the stationary distribution:

$$\pi_i(S_i, S_j) = \begin{bmatrix} b - c \\ -c \\ b \\ 0 \end{bmatrix}^T \begin{bmatrix} \tau_{CC} \\ \tau_{CD} \\ \tau_{DC} \\ \tau_{DD} \end{bmatrix}$$

$$i, j \in \{AllD, TFT, GTFT, WSLS, AllC\}$$

Using the transition probability matrix (2) we can derive the stationary distribution together with the respective average expected payoffs as summarized in the following table (some expression m_{ij} are too long and given at the end of the table):

Interaction	Stationary Distribution	Average Payoff ¹⁹
AID vs. AID	ε^2 $-\varepsilon(\varepsilon - 1)$ $-\varepsilon(\varepsilon - 1)$ $(\varepsilon - 1)^2$	$m_{11} = \varepsilon(b - c)$
AID vs TFT	$-2\varepsilon^2(\varepsilon - 1)$ $2\varepsilon^3 - 2\varepsilon^2 + \varepsilon$ $2\varepsilon(\varepsilon - 1)^2$ $-2\varepsilon^3 + 4\varepsilon^2 - 3\varepsilon + 1$	$m_{12} = -\varepsilon(c - 2b + 2b\varepsilon)$
AID vs GFT	$\frac{\varepsilon(1-\varepsilon)(n+\varepsilon)}{n-m+m\varepsilon-n\varepsilon+1}$ $\varepsilon \frac{m\varepsilon-\varepsilon-m+\varepsilon^2+1}{n-m+m\varepsilon-n\varepsilon+1}$ $\frac{(\varepsilon-1)^2(n+\varepsilon)}{n-m+m\varepsilon-n\varepsilon+1}$ $\frac{(1-\varepsilon)(m\varepsilon-\varepsilon-m+\varepsilon^2+1)}{n-m+m\varepsilon-n\varepsilon+1}$	m_{13}
AID vs WSLS	$\frac{1}{2}\varepsilon$ $\frac{1}{2}\varepsilon$ $\frac{1}{2} - \frac{1}{2}\varepsilon$ $\frac{1}{2} - \frac{1}{2}\varepsilon$	$m_{14} = \frac{1}{2}b - c\varepsilon$
AID vs AIC	$-\varepsilon(\varepsilon - 1)$ ε^2 $(\varepsilon - 1)^2$ $-\varepsilon(\varepsilon - 1)$	$m_{15} = b - b\varepsilon - c\varepsilon$

¹⁹Stationary distributions are valid for the general formulation of $GTFT - (1, m, 1, n)$. However, expressions for the average payoffs become very complicated and are therefore computed under the restriction $m = n$ in the general form of $GTFT$, with the impact of the asymmetry ($m \neq n$) in the two probabilities of restoring cooperation (i.e. after a CD or DD history) left for further research.

Interaction	Stationary Distribution	Payoffs
TFT vs ALLD	$-2\varepsilon^2(\varepsilon - 1)$ $2\varepsilon(\varepsilon - 1)^2$ $2\varepsilon^3 - 2\varepsilon^2 + \varepsilon$ $-2\varepsilon^3 + 4\varepsilon^2 - 3\varepsilon + 1$	m_{21}
TFT vs TFT	$1/4$ $1/4$ $1/4$ $1/4$	$m_{22} = \frac{1}{2}b - \frac{1}{2}c$
TFT vs GTFT	$\frac{(1-\varepsilon)(n+\varepsilon)(m+\varepsilon-2m\varepsilon)}{m\varepsilon+5n\varepsilon+9\varepsilon^2-12\varepsilon^3+4\varepsilon^4-6m\varepsilon^2+4m\varepsilon^3-10n\varepsilon^2+4n\varepsilon^3+mn-4mn\varepsilon+4mn\varepsilon^2}$ $\frac{2\varepsilon^2\frac{\varepsilon-1}{2\varepsilon-2\varepsilon^2}(m+\varepsilon-2)(\varepsilon+n-2n\varepsilon)}{m\varepsilon+5n\varepsilon+9\varepsilon^2-12\varepsilon^3+4\varepsilon^4-6m\varepsilon^2+4m\varepsilon^3-10n\varepsilon^2+4n\varepsilon^3+mn-4mn\varepsilon+4mn\varepsilon^2}$ $\frac{2\varepsilon(\varepsilon-1)^2(n+\varepsilon)}{m\varepsilon+5n\varepsilon+9\varepsilon^2-12\varepsilon^3+4\varepsilon^4-6m\varepsilon^2+4m\varepsilon^3-10n\varepsilon^2+4n\varepsilon^3+mn-4mn\varepsilon+4mn\varepsilon^2}$ $\frac{2\varepsilon^2(\varepsilon-1)(m+\varepsilon-2)}{m\varepsilon+5n\varepsilon+9\varepsilon^2-12\varepsilon^3+4\varepsilon^4-6m\varepsilon^2+4m\varepsilon^3-10n\varepsilon^2+4n\varepsilon^3+mn-4mn\varepsilon+4mn\varepsilon^2}$	m_{23}
TFT vs WSLS	$1/4$ $1/4$ $1/4$ $1/4$	$m_{24} = \frac{1}{2}b - \frac{1}{2}c$
TFT vs ALLC	$-(\varepsilon - 1)(2\varepsilon^2 - 2\varepsilon + 1)$ $-2\varepsilon^2\frac{\varepsilon-1}{2\varepsilon-2\varepsilon^2}(2\varepsilon^2 - 2\varepsilon + 1)$ $2\varepsilon(\varepsilon - 1)^2$ $-2\varepsilon^2(\varepsilon - 1)$	m_{25}

Interaction	Stationary Distribution	Payoffs
GTFT vs. AllD	$-\varepsilon(\varepsilon - 1) \frac{n+\varepsilon}{n-m+m\varepsilon-n\varepsilon+1}$ $(\varepsilon - 1)^2 \frac{n+\varepsilon}{n-m+m\varepsilon-n\varepsilon+1}$ $\varepsilon \frac{m\varepsilon-\varepsilon-m+\varepsilon^2+1}{n-m+m\varepsilon-n\varepsilon+1}$ $-(\varepsilon - 1) \frac{m\varepsilon-\varepsilon-m+\varepsilon^2+1}{n-m+m\varepsilon-n\varepsilon+1}$	m_{31}
GTFT vs. TFT	$\frac{(1-\varepsilon)(n+\varepsilon)(m+\varepsilon-2m\varepsilon)}{m\varepsilon+5n\varepsilon+9\varepsilon^2-12\varepsilon^3+4\varepsilon^4-6m\varepsilon^2+4m\varepsilon^3-10n\varepsilon^2+4n\varepsilon^3+mn-4mn\varepsilon+4mn\varepsilon^2}$ $\frac{2\varepsilon(\varepsilon-1)^2(n+\varepsilon)}{m\varepsilon+5n\varepsilon+9\varepsilon^2-12\varepsilon^3+4\varepsilon^4-6m\varepsilon^2+4m\varepsilon^3-10n\varepsilon^2+4n\varepsilon^3+mn-4mn\varepsilon+4mn\varepsilon^2}$ $\frac{2\varepsilon^2 \frac{\varepsilon-1}{2\varepsilon-2\varepsilon^2} (m+\varepsilon-2)(\varepsilon+n-2n\varepsilon)}{m\varepsilon+5n\varepsilon+9\varepsilon^2-12\varepsilon^3+4\varepsilon^4-6m\varepsilon^2+4m\varepsilon^3-10n\varepsilon^2+4n\varepsilon^3+mn-4mn\varepsilon+4mn\varepsilon^2}$ $\frac{2\varepsilon^2(\varepsilon-1)(m+\varepsilon-2)}{m\varepsilon+5n\varepsilon+9\varepsilon^2-12\varepsilon^3+4\varepsilon^4-6m\varepsilon^2+4m\varepsilon^3-10n\varepsilon^2+4n\varepsilon^3+mn-4mn\varepsilon+4mn\varepsilon^2}$	m_{32}
GTFT vs. GTFT	$-n \frac{2m-2m\varepsilon+n\varepsilon-mn}{\varepsilon^3-2\varepsilon^2-4n\varepsilon+mn^2+m\varepsilon^2+2n\varepsilon^2+n^2\varepsilon-2mn+2mn\varepsilon}$ $n\varepsilon \frac{n+\varepsilon-2}{\varepsilon^3-2\varepsilon^2-4n\varepsilon+mn^2+m\varepsilon^2+2n\varepsilon^2+n^2\varepsilon-2mn+2mn\varepsilon}$ $n\varepsilon \frac{n+\varepsilon-2}{\varepsilon^3-2\varepsilon^2-4n\varepsilon+mn^2+m\varepsilon^2+2n\varepsilon^2+n^2\varepsilon-2mn+2mn\varepsilon}$ $\varepsilon^2 \frac{m+\varepsilon-2}{\varepsilon^3-2\varepsilon^2-4n\varepsilon+mn^2+m\varepsilon^2+2n\varepsilon^2+n^2\varepsilon-2mn+2mn\varepsilon}$	m_{33}
GTFT vs. WSLs	$\frac{(n+\varepsilon-2n\varepsilon-\varepsilon^2+2n\varepsilon^2-mn+2mn\varepsilon)(\varepsilon-1)}{4m\varepsilon-7\varepsilon-n+4n\varepsilon+15\varepsilon^2-12\varepsilon^3+4\varepsilon^4-8m\varepsilon^2+4m\varepsilon^3-8n\varepsilon^2+4n\varepsilon^3+mn-4mn\varepsilon+4mn\varepsilon^2}$ $\frac{(n\varepsilon-2\varepsilon+\varepsilon^2+1)2\varepsilon(\varepsilon-1)}{4m\varepsilon-7\varepsilon-n+4n\varepsilon+15\varepsilon^2-12\varepsilon^3+4\varepsilon^4-8m\varepsilon^2+4m\varepsilon^3-8n\varepsilon^2+4n\varepsilon^3+mn-4mn\varepsilon+4mn\varepsilon^2}$ $\frac{-(2\varepsilon-2m\varepsilon-n\varepsilon-3\varepsilon^2+\varepsilon^3+4m\varepsilon^2-2m\varepsilon^3+2n\varepsilon^2+mn\varepsilon-2mn\varepsilon^2)}{4m\varepsilon-7\varepsilon-n+4n\varepsilon+15\varepsilon^2-12\varepsilon^3+4\varepsilon^4-8m\varepsilon^2+4m\varepsilon^3-8n\varepsilon^2+4n\varepsilon^3+mn-4mn\varepsilon+4mn\varepsilon^2}$ $\frac{(m\varepsilon-\varepsilon-m+\varepsilon^2+1)2\varepsilon(\varepsilon-1)}{4m\varepsilon-7\varepsilon-n+4n\varepsilon+15\varepsilon^2-12\varepsilon^3+4\varepsilon^4-8m\varepsilon^2+4m\varepsilon^3-8n\varepsilon^2+4n\varepsilon^3+mn-4mn\varepsilon+4mn\varepsilon^2}$	m_{34}
GTFT vs. AllC	$-\frac{\varepsilon-1}{n\varepsilon-m\varepsilon+1} (n\varepsilon - 2\varepsilon + \varepsilon^2 + 1)$ $\frac{\varepsilon}{n\varepsilon-m\varepsilon+1} (n\varepsilon - 2\varepsilon + \varepsilon^2 + 1)$ $\varepsilon(\varepsilon - 1) \frac{m+\varepsilon-2}{n\varepsilon-m\varepsilon+1}$ $-\varepsilon^2 \frac{m+\varepsilon-2}{n\varepsilon-m\varepsilon+1}$	m_{35}

Interaction	Stationary Distribution	Payoffs
WSLS vs. AID	$\frac{1}{2}\varepsilon$ $\frac{1}{2} - \frac{1}{2}\varepsilon$ $\frac{1}{2}\varepsilon$ $\frac{1}{2} - \frac{1}{2}\varepsilon$	m_{41}
WSLS vs. TFT	$1/4$ $1/4$ $1/4$ $1/4$	m_{42}
WSLS vs. GTFT	$\frac{(n+\varepsilon-2n\varepsilon-\varepsilon^2+2n\varepsilon^2-mn+2mn\varepsilon)(\varepsilon-1)}{4m\varepsilon-7\varepsilon-n+4n\varepsilon+15\varepsilon^2-12\varepsilon^3+4\varepsilon^4-8m\varepsilon^2+4m\varepsilon^3-8n\varepsilon^2+4n\varepsilon^3+mn-4mn\varepsilon+4mn\varepsilon^2}$ $\frac{-(2\varepsilon-2m\varepsilon-n\varepsilon-3\varepsilon^2+\varepsilon^3+4m\varepsilon^2-2m\varepsilon^3+2n\varepsilon^2+mn\varepsilon-2mn\varepsilon^2)}{4m\varepsilon-7\varepsilon-n+4n\varepsilon+15\varepsilon^2-12\varepsilon^3+4\varepsilon^4-8m\varepsilon^2+4m\varepsilon^3-8n\varepsilon^2+4n\varepsilon^3+mn-4mn\varepsilon+4mn\varepsilon^2}$ $\frac{(n\varepsilon-2\varepsilon+\varepsilon^2+1)2\varepsilon(\varepsilon-1)}{4m\varepsilon-7\varepsilon-n+4n\varepsilon+15\varepsilon^2-12\varepsilon^3+4\varepsilon^4-8m\varepsilon^2+4m\varepsilon^3-8n\varepsilon^2+4n\varepsilon^3+mn-4mn\varepsilon+4mn\varepsilon^2}$ $\frac{(m\varepsilon-\varepsilon-m+\varepsilon^2+1)2\varepsilon(\varepsilon-1)}{4m\varepsilon-7\varepsilon-n+4n\varepsilon+15\varepsilon^2-12\varepsilon^3+4\varepsilon^4-8m\varepsilon^2+4m\varepsilon^3-8n\varepsilon^2+4n\varepsilon^3+mn-4mn\varepsilon+4mn\varepsilon^2}$	m_{43}
WSLS vs. WSLS	$-4\varepsilon^3+7\varepsilon^2-4\varepsilon+1$ $-\varepsilon(\varepsilon-1)$ $-\varepsilon(\varepsilon-1)$ $4\varepsilon^3-5\varepsilon^2+2\varepsilon$	m_{44}
WSLS vs. AIC	$\frac{1}{2} - \frac{1}{2}\varepsilon$ $\frac{1}{2}\varepsilon$ $\frac{1}{2} - \frac{1}{2}\varepsilon$ $\frac{1}{2}\varepsilon$	m_{45}

Interaction	Stationary Distribution	Payoffs
AIC vs. AID	$\frac{-\varepsilon(\varepsilon-1)}{(\varepsilon-1)^2}$ $\frac{-\varepsilon(\varepsilon-1)}{\varepsilon^2}$	$m_{51} = b\varepsilon - c + c\varepsilon$
AIC vs. TFT	$-2\varepsilon^3 + 4\varepsilon^2 - 3\varepsilon + 1$ $\frac{2\varepsilon(\varepsilon-1)^2}{2\varepsilon^3 - 2\varepsilon^2 + \varepsilon}$ $-2\varepsilon^2(\varepsilon-1)$	$m_{52} = b - c - 2b\varepsilon + c\varepsilon + 2b\varepsilon^2$
AIC vs. GTFT	$-\frac{\varepsilon-1}{n\varepsilon-m\varepsilon+1}(n\varepsilon-2\varepsilon+\varepsilon^2+1)$ $\frac{\varepsilon(\varepsilon-1)}{n\varepsilon-m\varepsilon+1}\frac{m+\varepsilon-2}{n\varepsilon-m\varepsilon+1}$ $\frac{\varepsilon}{n\varepsilon-m\varepsilon+1}(n\varepsilon-2\varepsilon+\varepsilon^2+1)$ $-\varepsilon^2\frac{m+\varepsilon-2}{n\varepsilon-m\varepsilon+1}$	$m_{53} = b - c - 2b\varepsilon + c\varepsilon + b\varepsilon^2 + bn\varepsilon$
AIC vs. WSLS	$-2\varepsilon^3 + 4\varepsilon^2 - 3\varepsilon + 1$ $\frac{2\varepsilon(\varepsilon-1)^2}{2\varepsilon^3 - 2\varepsilon^2 + \varepsilon}$ $-2\varepsilon^2(\varepsilon-1)$	$m_{54} = b - c - 2b\varepsilon + c\varepsilon + 2b\varepsilon^2$
AIC vs. AIC	$\frac{(\varepsilon-1)^2}{-\varepsilon(\varepsilon-1)}$ $\frac{-\varepsilon(\varepsilon-1)}{\varepsilon^2}$	$m_{55} = (1-\varepsilon)(b-c)$

Table 4: Stationary distributions and average expected payoffs for an iterated PD game with an ecology of repeated rules consisting of AID, TFT, GTFT, WSLS and AIC

where,

$$m_{13} = b\varepsilon - c\varepsilon - b\varepsilon^2 + bn - bn\varepsilon$$

$$m_{21} = \varepsilon(b - 2c + 2c\varepsilon)$$

$$m_{23} = -\frac{c\varepsilon - b\varepsilon + b\varepsilon^2 - bn + cn + bn\varepsilon - 2cn\varepsilon}{n + 3\varepsilon - 2n\varepsilon - 2\varepsilon^2}$$

$$m_{25} = b - c - b\varepsilon + 2c\varepsilon - 2c\varepsilon^2$$

$$m_{31} = b\varepsilon - c\varepsilon + c\varepsilon^2 - cn + cn\varepsilon$$

$$m_{32} = \frac{b\varepsilon - c\varepsilon + c\varepsilon^2 + bn - cn - 2bn\varepsilon + cn\varepsilon}{n + 3\varepsilon - 2n\varepsilon - 2\varepsilon^2}$$

$$m_{34} = -m_{341}/m_{342}, \text{ with}$$

$$m_{341} = 3b\varepsilon - 3c\varepsilon - bn^2 + cn^2 - 5b\varepsilon^2 + 2b\varepsilon^3 + 8c\varepsilon^2 - 7c\varepsilon^3 + 2c\varepsilon^4 + bn - cn - 4bn^2\varepsilon^2 + 2cn^2\varepsilon^2 - 6bn\varepsilon + 3cn\varepsilon + 10bn\varepsilon^2 + 4bn^2\varepsilon - 4bn\varepsilon^3 - 6cn\varepsilon^2 - 3cn^2\varepsilon + 4cn\varepsilon^3$$

$$m_{342} = 4n^2\varepsilon^2 - 4n^2\varepsilon + n^2 + 8n\varepsilon^3 - 16n\varepsilon^2 + 8n\varepsilon - n + 4\varepsilon^4 - 12\varepsilon^3 + 15\varepsilon^2 - 7\varepsilon$$

$$m_{35} = b - c - b\varepsilon + 2c\varepsilon - c\varepsilon^2 - cn\varepsilon$$

$$m_{41} = b\varepsilon - \frac{1}{2}c$$

$$m_{42} = \frac{1}{2}(b - c)$$

$$m_{43} = m_{431}/m_{342}$$

$$m_{431} = 3c\varepsilon - 3b\varepsilon + bn^2 - cn^2 + 8b\varepsilon^2 - 7b\varepsilon^3 + 2b\varepsilon^4 - 5c\varepsilon^2 + 2c\varepsilon^3 - bn + cn + 2bn^2\varepsilon^2 - 4cn^2\varepsilon^2 + 3bn\varepsilon - 6cn\varepsilon - 6bn\varepsilon^2 - 3bn^2\varepsilon + 4bn\varepsilon^3 + 10cn\varepsilon^2 + 4cn^2\varepsilon - 4cn\varepsilon^3$$

$$m_{44} = (b - c)(1 - 4\varepsilon^3 + 6\varepsilon^2 - 3\varepsilon)$$

$$m_{45} = b - \frac{1}{2}c - b\varepsilon$$

References

- Axelrod, R. (1997). *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*. Princeton University Press.
- Dhooge, A., W. Govaerts, Y. Kuznetsov, W. Mestrom, and A. Riet (2003). Clmatcont: a continuation toolbox in matlab. *Symposium on Applied Computing*, Proceedings of the 2003 ACM symposium on applied computing, Melbourne, Florida, 161–166.
- Imhof, L., D. Fudenberg, and M. Nowak (2007). Tit-for-tat or win-stay, lose-shift? *Journal of Theoretical Biology* 247(3), 574–580.
- Kemeny, J. G. and J. L. Snell (1975). *Denumerable Markov chains*. New-York: Springer-Verlag.
- Kraines, D. and V. Kraines (1995). Evolution of learning among pavlov strategies in a competitive environment with noise. *Journal of Conflict Resolution* 39(3), 439–466.
- Kraines, D. and V. Kraines (2000). Natural selection of memory-one strategies for the iterated prisoner’s dilemma. *Journal of Theoretical Biology* 203(4), 335–355.
- Kuznetsov, Y. A. (1995). *Elements of applied bifurcation theory*. Berlin and New York: Springer-Verlag.
- Ochea, M. (2010). *Essays on Nonlinear Evolutionary Game Dynamics*. Ph. D. thesis, University of Amsterdam.
- Sigmund, K. and H. Brandt (2006). The good, the bad and the discriminator-errors in direct and indirect reciprocity. *Journal of Theoretical Biology* 239(2), 183–194.
- Sigmund, K. and M. Nowak (1993a). Chaos and the evolution of cooperation. *PNAS* 90(11), 5091–5094.
- Sigmund, K. and M. Nowak (1993b). Win-stay, lose-shift outperforms tit-for-tat. *Nature* 364, 56–58.