

# Dynamical instabilities in a simple minority game with discounting

D Challet<sup>¶</sup>, A De Martino<sup>†</sup> and M Marsili<sup>‡</sup>

<sup>¶</sup> Institute for Scientific Interchange (ISI), Viale S. Severo 65, 10133 Torino (Italy) and Département de Physique, Université de Fribourg, Pérolles, 1700 Fribourg, Switzerland

<sup>†</sup> CNR/INFN SMC, Dipartimento di Fisica, Università di Roma “La Sapienza”, p.le A. Moro 2, 00185 Roma (Italy)

<sup>‡</sup> The Abdus Salam International Centre for Theoretical Physics, Strada Costiera 14, 34014 Trieste (Italy)

**Abstract.** We explore the effect of discounting and experimentation in a simple model of interacting adaptive agents. Agents belong to either of two types and each has to decide whether to participate a game or not, the game being profitable when there is an excess of players of the other type. We find the emergence of large fluctuations as a result of the onset of a dynamical instability which may arise discontinuously (increasing the discount factor) or continuously (decreasing the experimentation rate). The phase diagram is characterized in detail and noise amplification close to a bifurcation point is identified as the physical mechanism behind the instability.

PACS numbers: 02.50.Le, 87.23.Ge, 05.70.Ln, 64.60.Ht

E-mail: [challet@isi.it](mailto:challet@isi.it), [andrea.demartino@roma1.infn.it](mailto:andrea.demartino@roma1.infn.it), [marsili@ictp.it](mailto:marsili@ictp.it)

## 1. Introduction

Individual optimizing behavior is sometimes sufficient for a population of agents to coordinate on efficient outcomes. For example, competitive behavior in markets allows buyers and sellers to coordinate their demand and supply, and drivers trying to minimize transit time may avoid crowded routes, thus reducing the chances of traffic jams. Such consequences of individual behavior have been thoroughly studied in economics within an equilibrium framework. In many situations, however, such idealized conditions as those postulated in economics (e.g. perfect information and rationality), may be unrealistic [1]. When individuals are engaged in contexts involving many other individuals, as in financial markets, it is normally more realistic to assume that they acquire information from their environment and from other agents as they learn the relative efficiency of their strategies/actions. Experimentation – i.e. the fact of sometimes playing sub-optimal actions – and flexibility – the ability to change strategy in response to a changing environment – become important aspects of a learning dynamics in an evolving system. The former means that agents’ behavior is best represented by a

stochastic choice model [2]. Flexibility implies instead that agents should discount past payoffs in their learning behavior, as they might not be relevant for the current state of their environment (at the same time, discounting too much past payoffs, may not allow agents to learn the full complexity of their environment [3]). It is thus essential to understand how the efficiency of learning is affected by individual stochastic choice and by payoff discounting.

Here we address these issues in an extremely simple setting that allows us to draw sharp conclusions and to achieve a full understanding of the results. The model we consider is a stylized description of situations where two different groups of agents, taken of equal size for the sake of simplicity, interact, each of them providing a resource for agents of the other type. Examples of this generic setting include buyers and sellers – which is the situation we shall informally refer to henceforth – but also (heterosexual) individuals of opposite sex attending a party: for each of them the party might be most interesting if the majority of the participants are of the opposite sex. In addition to this, individuals may have *a priori* incentives to participate in the game. In our model, individuals take decisions based on an estimate of the payoff, which they compute from the outcomes of the game in the recent past.

The model falls in the class of Minority Games (MGs) [4], which have been studied extensively in the recent past. Most theoretical results have been confined so far to the case of infinite memory, where agents do not discount payoffs. Numerical simulations for more complex versions of the model than that discussed here show that, interestingly, strong fluctuations can arise when the discount factor is introduced [3, 5]. Remarkably, these fluctuations arise after a much longer time than any characteristic time scale in the system. Understanding the origin of such non-trivial fluctuations is indeed one of the motivations of the present work.

Strong fluctuations and dynamical instabilities are persistent and ubiquitous in a multitude of social systems that can be modeled with interacting adaptive agents (like economies, financial markets, traffic, elections, web-communities, etc.; see e.g. [6, 7]), hence the issues raised here have a rather general relevance. Dynamic instabilities in models of interacting agents have been also addressed in the economics literature (see Ref. [8] for a review), but the role of stochastic fluctuations has only been recently recognized [9].

The aim of the present paper is to present a model which is simple enough to allow us to understand the origin of strong fluctuations in a system of boundedly rational interacting agents. In particular we will focus on the interplay between discounting and the stochastic nature of the learning process.

We find that, in a region of parameter space, increasing the discount factor, the system crosses a discontinuous transition and enters a region where two distinct dynamical steady state solutions exist. The emergence of strong fluctuations is then an activated process and, as such, these materialize (and dematerialize) after times which can be extremely long. This same transition arises continuously, decreasing the randomness (or experimentation) in the choice behavior, at a critical threshold. At a still higher

threshold a stable solution reappears, coexisting with the strongly fluctuating one.

## 2. The model

We consider a system of  $N$  buyers and  $N$  sellers. We denote by  $B$  and  $S$  the sets of buyers and sellers, respectively. At time step  $t$ , agent  $i$  can either play the game ( $n_i(t) = 1$ ) or abstain ( $n_i(t) = 0$ ). We assume that he chooses to play with probability

$$\text{Prob}\{n_i(t) = 1\} = \frac{1 + \tanh[\Gamma U_i^\pm(t)]}{2} \quad (1)$$

where  $\Gamma \geq 0$  is the (uniform) learning rate of agents, while the functions  $U_i^\pm(t)$ , representing the cumulated score of buyers and sellers, respectively, evolve according to

$$U_i^\pm(t+1) = (1 - \lambda)U_i^\pm(t) \mp A(t) - \epsilon \quad (2)$$

Here  $\epsilon$  is a cost (or incentive if  $\epsilon < 0$ ) which agents incur for participation. This represents outside opportunities, such as investing in a risk free asset instead of buying.  $A(t)$  stands for the participation imbalance (or ‘excess demand’), defined as

$$A(t) = \frac{1}{N} \left[ \sum_{i \in B} n_i(t) - \sum_{i \in S} n_i(t) \right]. \quad (3)$$

Finally the constant  $\lambda$  ( $0 \leq \lambda \leq 1$ ) plays the role of a discount factor, introducing a finite memory on the score. In words, agents estimate a score for participating in the game taking an exponential moving average of a payoff. The payoff  $\mp A(t) - \epsilon$  depends on how many more agents of the other type had recently participated (i.e. on  $A(t)$ ) and on a constant cost  $\epsilon$ . Agents base their decision on whether to play or not by learning how profitable was it to play in the past. Note that the reinforcement term in the learning dynamics is independent of the agent, hence we can drop the index  $i$  from (2). For the analysis which follows, it is reasonable to separate the sum and the difference of  $U^\pm$  by introducing the variables

$$\bar{u}(t) = \Gamma \frac{U^+(t) + U^-(t)}{2} \quad (4)$$

$$y(t) = \Gamma \frac{U^+(t) - U^-(t)}{2} \quad (5)$$

so that  $U^\pm = (\bar{u} \pm y)/\Gamma$ . The dynamics of  $\bar{u}$  is deterministic and it can be easily solved to yield

$$\bar{u}(t) = \left[ \bar{u}(0) + \frac{\Gamma\epsilon}{\lambda} \right] (1 - \lambda)^t - \frac{\Gamma\epsilon}{\lambda} \quad (6)$$

so that  $\bar{u}$  converges over times of the order of  $1/\lambda$  to its asymptotic value  $-\Gamma\epsilon/\lambda$ . We shall henceforth neglect such a transient, and set  $\bar{u} = -\Gamma\epsilon/\lambda$ . The time evolution of  $y$  is instead described by

$$y(t+1) = y(t)(1 - \lambda) - \Gamma A(t). \quad (7)$$

In order to understand the structure of fluctuations, we shall now focus on the above process.

### 3. Deterministic approximation

For  $N \rightarrow \infty$  a deterministic theory holds. Indeed the quantity  $A(t)$  satisfies the law of large numbers, i.e. it converges almost surely to

$$\langle A|y \rangle = \frac{\tanh(\bar{u} + y) - \tanh(\bar{u} - y)}{2} \quad (8)$$

(with  $\langle \dots |y \rangle$  we denote expected values with respect to the distribution of  $n_i(t)$ , at fixed  $U^\pm$ ). This in turn implies that (7) is well approximated by

$$y(t+1) = y(t)(1 - \lambda) - \Gamma \langle A|y(t) \rangle. \quad (9)$$

so that, if  $y(t) = y$ ,  $A(t)$  has variance

$$\langle \delta A^2 |y \rangle = \frac{2 - \tanh^2(\bar{u} + y) - \tanh^2(\bar{u} - y)}{4N} \quad (10)$$

$$= \frac{\cosh^{-2}(\bar{u} + y) + \cosh^{-2}(\bar{u} - y)}{4N} \quad (11)$$

which vanishes as  $N \rightarrow \infty$ .

Equation (9) is a nonlinear map with multiplicative noise. It possesses a trivial fixed point  $y^* = 0$ , which is stable if  $|1 - \lambda - \Gamma[1 - \tanh(\bar{u})^2]| < 1$  i.e. when

$$\Gamma < (2 - \lambda) \cosh^2(\Gamma\epsilon/\lambda). \quad (12)$$

For  $\epsilon = 0$ , one recovers the stability condition of the original MG,  $\lambda + \Gamma < 2$  with  $2N$  players [10]. The role of  $\epsilon$  is then to stabilize the system, as it decreases the right-hand side of (12), independently of its sign. Note that (12) can be re-written as

$$\epsilon > \frac{\lambda}{\Gamma} \operatorname{acosh} \sqrt{\frac{\Gamma}{2 - \lambda}} \quad (13)$$

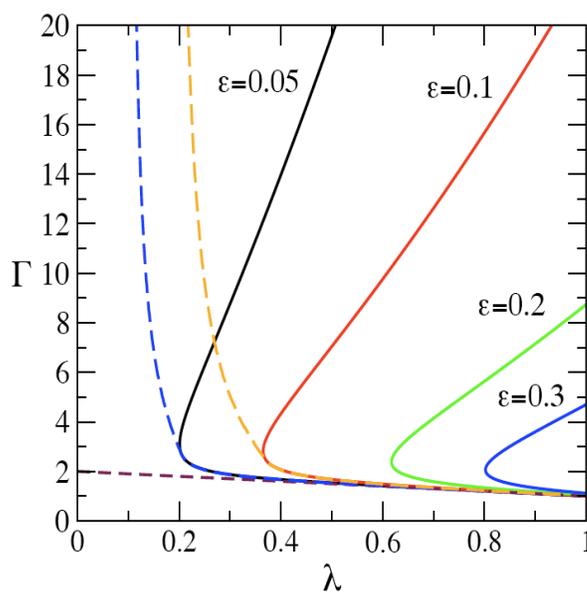
Also note that increasing  $\lambda$  drives the system closer to the instability.

Figure 1 reports the stability condition of the  $y^* = 0$  fixed point in the  $(\lambda, \Gamma)$  plane (full lines). A notable feature of these curves is the existence of two transitions as  $\Gamma$  is varied while  $\lambda$  and  $\epsilon$  are kept constant. This is confirmed by the numerical simulations shown in Fig. 2, where the fluctuations  $\langle y^2 \rangle$  of  $y$  around its fixed point are plotted as a function of  $\Gamma$  at fixed  $\epsilon$  and  $\lambda$ , for different values of  $N$ .

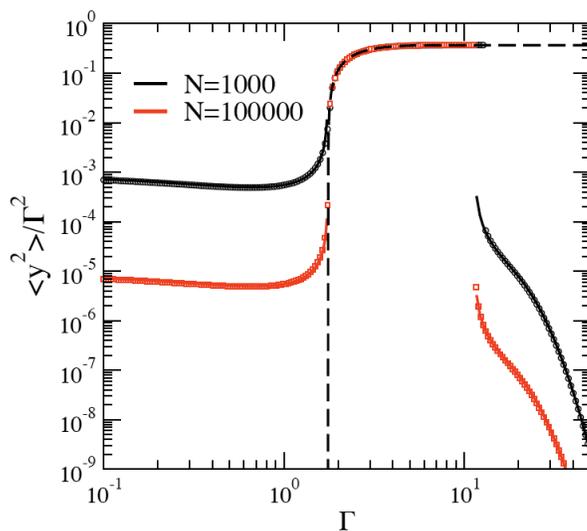
In the region where  $y^* = 0$  is unstable, a different solution must be considered. Indeed, besides the trivial fixed point, there is also a period two solution of the form  $y(t) = (-1)^t z^*$ , where  $z^*$  solves the equation

$$z = \frac{\Gamma}{2 - \lambda} \langle A|z \rangle = \frac{\Gamma}{2 - \lambda} \frac{\sinh(2z)}{\cosh(2\bar{u}) + \cosh(2z)} \quad (14)$$

This admits again a trivial solution  $z^* = 0$ , which is stable in the same region where the  $y^* = 0$  solution is. In addition, it can also have a solution  $z^* \neq 0$  which is stable in the region where the  $y^* = 0$  solution is unstable. The stability of this solution may be studied precisely in the same way as before, and it reveals that stability extends beyond the region where the  $y^* = 0$  solution is unstable, as shown in Fig. 1. In particular, it extends to the region of large  $\Gamma$  for  $\lambda > 2\epsilon/(1 + \epsilon)$ , where the solution takes the simple



**Figure 1.** Stability in the parameter space. The trivial fixed point is stable on the left of the respective  $\epsilon$  lines (continuous curves) and below the  $2\Gamma + \lambda = 2$  line (dashed line). The stability lines of the oscillating solution (dashed curves) are shown only for  $\epsilon = 0.05$  and  $\epsilon = 0.1$  for sakes of clarity.



**Figure 2.** Fluctuations of  $y(t)$  for  $\epsilon = 0.05$ ,  $\lambda = 0.35$  as a function of  $\Gamma$ . In the unstable region (intermediate  $\Gamma$ ), fluctuations are dominated by the periodic oscillation of  $y(t)$ , giving rise to a  $\langle y^2 \rangle$  that can be computed analytically (dashed line).

asymptotic form  $z \simeq 1/(2 - \lambda)$  ( $\Gamma \gg 1$ ). This region is limited by a dashed curve in Fig. 1, which indicates that the transition from the  $z^* \neq 0$  solution to the  $y^* = 0$  solution is discontinuous in the region where the latter is stable, whereas it is continuous in the region where the fixed point  $y^* = 0$  is unstable. In the region between the dashed and the solid line both solutions coexist.

So when  $\lambda$  is large enough and  $\Gamma$  is increased from a small value, we expect a continuous transition, whereas when  $\Gamma$  is large enough and  $\lambda$  increases from zero across the transition, we find a sharp discontinuous transition and phase coexistence. The latter matches closely the numerical finding of [5] which reported the onset of instabilities occurring for very large times in a more complex version of the present model.

It is worth to comment on the effect of  $\epsilon$  on the dynamics. In the fixed point  $y^* = 0$ ,  $\epsilon$  controls the average fraction of agents, through  $\bar{u}$ :

$$\langle n_i \rangle = \frac{1 + \tanh(\bar{u})}{2} = \frac{1 + \tanh(-\Gamma\epsilon/\lambda)}{2}$$

This decreases with  $\epsilon$  and it vanishes if  $\epsilon \gg \lambda/\Gamma$  is large and positive. On the other hand  $\langle n_i \rangle \rightarrow 1$  if  $\epsilon$  is large and negative, i.e. all agents play if the incentive is large enough. This is the main effect of the sign of  $\epsilon$ . Indeed both the stability condition as well as the value of  $z$  and the fluctuation properties (see below) are independent of the sign of  $\epsilon$ .

#### 4. Fluctuations in the stochastic system

In order to explain this rich behavior, it is necessary to study the fluctuations in the stochastic system.

Let us first notice that  $\langle y^2 \rangle \simeq z^2$  on the periodic solution  $z^* \neq 0$ . On this solution, the fluctuations in  $A$  are of order one. Indeed they are given by

$$\langle A^2 \rangle = \langle A \rangle^2 + \langle \delta A^2 \rangle \simeq \left( \frac{2 - \lambda}{\Gamma} \right)^2 z^2 + O(N^{-1})$$

where we have neglected fluctuations  $\langle \delta A^2 \rangle \sim O(N^{-1})$ . In brief, the fluctuations are dominated by the periodic oscillation and the noise has a negligible effect.

Things are a bit more complicated for fluctuations around the  $y^* = 0$  fixed point. The starting point of our analysis is (7). We take the square of each side and average over the realizations of the stochastic process. In the stationary state, after rearranging terms, this yields

$$[1 - (1 - \lambda)^2] \langle y^2 \rangle + 2\Gamma(1 - \lambda) \langle yA \rangle - \Gamma^2 \langle A^2 \rangle = 0 \quad (15)$$

Now, the last two averages take into account both the effect of random choice, for a fixed value of  $y$ , as described by (8) and (11), and the effect induced by the fluctuations of  $y$  itself. For the first term, we may write

$$\langle yA \rangle = \int dy P(y) y \langle A|y \rangle \simeq \left. \frac{\partial \langle A|y \rangle}{\partial y} \right|_{y=0} \langle y^2 \rangle + O(\langle y^4 \rangle)$$

where in the second equality we have expanded the conditional average around the fixed point  $y = 0$ . This was to be expected: because the stochastic noise is of order  $1/N$ , also fluctuations  $\delta y \sim 1/\sqrt{N}$  are likely to be small. Therefore for large  $N$  we can safely neglect higher order terms in the above equation. Likewise, we can approximate

$$\langle A^2 \rangle = \int dy P(y) [\langle \delta A^2 | y \rangle + \langle A | y \rangle^2] \quad (16)$$

$$\simeq \langle \delta A^2 | 0 \rangle + \left[ \frac{\partial \langle A | y \rangle}{\partial y} \Big|_{y=0} \right]^2 \langle y^2 \rangle + \dots \quad (17)$$

Here, since  $\langle \delta A^2 | 0 \rangle \sim 1/N$ , we have neglected contributions due to fluctuations in  $y$ , as they only matter to higher orders in  $1/N$ . With these, (15) can easily be solved to yield the fluctuations of  $y$ :

$$\langle y^2 \rangle = \frac{\Gamma^2}{2N \cosh^2(\bar{u}) \left[ 1 - (1 - \lambda - \Gamma / \cosh^2(\bar{u}))^2 \right]} \quad (18)$$

Notice that fluctuations diverge as the instability line is approached, as expected. (18) agrees perfectly with numerical simulations, as shown in Fig. 2.

The approach discussed here can easily be extended, along the lines of Ref. [11], in order to describe the onset of anomalous fluctuations beyond the gaussian regime, close to the instability line.

## 5. Conclusion

We have shown that discounting can have non-trivial consequences on the dynamics of a system of interacting adaptive agents. From a naïve point of view, this sounds counterintuitive as discounting decreases the correlation time, hence the strength of non-linear effects. The interplay with the stochastic nature of the learning process (due to experimentation) is crucial. Indeed the instability does not arise (for  $\lambda < 1$ ) neither if the noise is too strong (small  $\Gamma$ ) or if the dynamics is close to being deterministic ( $\Gamma$  large).

We argue that the dynamic instability discussed here is of the same nature as that which is responsible for the onset of large fluctuations in the Minority Game discussed in Ref. [5]. Indeed, the mechanism discussed here is likely of a very general nature and a similar phase structure may arise in other games with discounting and experimentation.

## Acknowledgements

This work was partially supported by EU-NEST grant 516446 ComplexMarkets.

## References

- [1] Arthur WB 1994 American Economic Review **84** 406
- [2] Luce RD 1959 *Individual choice behaviour: a theoretical analysis* (Wiley)

- [3] Marsili M, Mulet R, Ricci-Tersenghi F and Zecchina R 2001 PRL **87** 208701
- [4] Challet D and Zhang Y-C 1997 Physica A **246** 407
- [5] Challet D, De Martino A, Marsili M and Perez Castillo I 2006 JSTAT P03004
- [6] Wyart M and Bouchaud J-P 2007 J. Econ, Behav. Org. **63** 1
- [7] Raffaelli G and Marsili M 2006 JSTAT L08001
- [8] Hommes CH 2006 In: L. Tesfatsion, K.L. Judd (eds.), *Handbook of Computational Economics, Vol. 2: Agent-Based Computational Economics*, Elsevier Science B.V., 1109-1186
- [9] Alfarano S and Lux T 2007 Macroeconomics Dynamics **11** 80-101
- [10] Mosetti G, Challet D and Zhang Y-C 2005 Physica A **365** 529
- [11] Kravtsov YA and Surovyatkina ED 2003 Phys. Lett. A **319** 348