

## Econophysics Review: II. Agent-based models

Anirban Chakraborti<sup>a)</sup> and Ioane Muni Toke<sup>b)</sup>

*Laboratoire de Mathématiques Appliquées aux Systèmes, Ecole Centrale Paris, 92290 Châtenay-Malabry, France*

Marco Patriarca<sup>c)</sup>

*National Institute of Chemical Physics and Biophysics, Rūvala 10, 15042 Tallinn, Estonia and*

*Instituto de Física Interdisciplinaria y Sistemas Complejos (CSIC-UIB), E-07122 Palma de Mallorca, Spain*

Frédéric Abergel<sup>d)</sup>

*Laboratoire de Mathématiques Appliquées aux Systèmes, Ecole Centrale Paris, 92290 Châtenay-Malabry, France*

(Dated: 3 November 2010)

This article is the second part of a review of recent empirical and theoretical developments usually grouped under the term *Econophysics*. In the first part, we have reviewed statistical properties of financial times series, statistics exhibited on order books and discussed some studies of correlations of assets. This second part deals with models in Econophysics through the point of view of agent-based modelling. Amongst a large number of multi-agent-based models, we have identified three representative areas. First, using previous work originally presented in the fields of behavioural finance and market microstructure theory, econophysicists have developed agent-based models of order-driven markets that are extensively presented here. Second, kinetic theory models designed to explain some empirical facts on wealth distribution are reviewed. Third, we briefly summarize game theory models by reviewing the now classic minority game and related problems.

Keywords: Econophysics; Stylized facts ; Financial time series; Correlations; Order book models ; Agent-based models; Wealth distributions; Game Theory; Minority Games; Pareto Law; Entropy maximization; Utility maximization.

PACS Nos.: 05.45.Tp, 02.50.Sk, 05.40.-a, 05.45.Ra, 89.75.Fb

### I. INTRODUCTION

In the first part of the review, empirical developments in Econophysics have been studied. We have pointed out that some of these widely known “stylized facts” are already at the heart of financial models. But many facts, especially the newer statistical properties of order books, are not yet taken into account. As advocated by many during the financial crisis in 2007-2008 (see e.g. Bouchaud (2008); Lux and Westerhoff (2009); Farmer and Foley (2009)), agent-based models should have a great role to play in future financial modelling. In economic models, there is usually the representative agent, who is “perfectly rational” and uses the “utility maximization” principle while taking actions. Instead the multi-agent models that have originated from statistical physics considerations have allowed to go beyond the prototype theories with the “representative” agent in traditional economics. In this second part of our review, we present recent developments of agent-based models in Econophysics.

There are, of course, many reviews and books already published in this areas (see e.g. Bouchaud *et al.*

(2009), Lux and Westerhoff (2009), Samanidou *et al.* (2007), Yakovenko and Rosser (2009), Chatterjee and Chakraborti (2007), Challet *et al.* (2004), Coolen (2005), etc.). We will present here our perspectives in three representative areas.

### II. AGENT-BASED MODELLING OF ORDER BOOKS

#### A. Introduction

Although known, at least partly, for a long time – Mandelbrot (1963) gives a reference for a paper dealing with non-normality of price time series in 1915, followed by several others in the 1920’s – “stylized facts” have often been left aside when modelling financial markets. They were even often referred to as “anomalous” characteristics, as if observations failed to comply with theory. Much has been done these past fifteen years in order to address this challenge and provide new models that can reproduce these facts. These recent developments have been built on top of early attempts at modelling mechanisms of financial markets with agents. For example, Stigler (1964), investigating some rules of the SEC<sup>1</sup>, or Garman (1976), investigating double-auction microstruc-

<sup>a)</sup> Electronic mail: anirban.chakraborti@ecp.fr

<sup>b)</sup> Electronic mail: ioane.muni-toke@ecp.fr

<sup>c)</sup> Electronic mail: marco.patriarca@kbfi.ee

<sup>d)</sup> Electronic mail: frederic.abergel@ecp.fr

<sup>1</sup> Security Exchange Commission

ture, belong to those historical works. It seems that the first modern attempts at that type of models were made in the field of behavioural finance. This field aims at improving financial modelling based on the psychology and sociology of the investors. Models are built with agents who can exchange shares of stocks according to exogenously defined utility functions reflecting their preferences and risk aversions. LeBaron (2006b) shows that this type of modelling offers good flexibility for reproducing some of the stylized facts and LeBaron (2006a) provides a review of that type of model. However, although achieving some of their goals, these models suffer from many drawbacks: first, they are very complex, and it may be a very difficult task to identify the role of their numerous parameters and the types of dependence to these parameters; second, the chosen utility functions do not necessarily reflect what is observed on the mechanisms of a financial market.

A sensible change in modelling appears with much simpler models implementing only well-identified and presumably realistic “behaviour”: Cont and Bouchaud (2000) uses noise traders that are subject to “herding”, i.e. form random clusters of traders sharing the same view on the market. The idea is used in Raberto *et al.* (2001) as well. A complementary approach is to characterize traders as fundamentalists, chartists or noise traders. Lux and Marchesi (2000) propose an agent-based model in which these types of traders interact. In all these models, the price variation directly results from the excess demand: at each time step, all agents submit orders and the resulting price is computed. Therefore, everything is cleared at each time step and there is no structure of order book to keep track of orders.

One big step is made with models really taking into account limit orders and keeping them in an order book once submitted and not executed. Chiarella and Iori (2002) build an agent-based model where all traders submit orders depending on the three elements identified in Lux and Marchesi (2000): chartists, fundamentalists, noise. Orders submitted are then stored in a persistent order book. In fact, one of the first simple models with this feature was proposed in Bak *et al.* (1997). In this model, orders are particles moving along a price line, and each collision is a transaction. Due to numerous caveats in this model, the authors propose in the same paper an extension with fundamentalist and noise traders in the spirit of the models previously evoked. Maslov (2000) goes further in the modelling of trading mechanisms by taking into account fixed limit orders and market orders that trigger transactions, and really simulating the order book. This model was analytically solved using a mean-field approximation by Slanina (2001).

Following this trend of modelling, the more or less “rational” agents composing models in economics tends to vanish and be replaced by the notion of flows: orders are not submitted any more by an agent following a strategic behaviour, but are viewed as an arriving flow whose properties are to be determined by empirical observa-

tions of market mechanisms. Thus, the modelling of order books calls for more “stylized facts”, i.e. empirical properties that could be observed on a large number of order-driven markets. Biais *et al.* (1995) is a thorough empirical study of the order flows in the Paris Bourse a few years after its complete computerization. Market orders, limit orders, time of arrivals and placement are studied. Bouchaud *et al.* (2002) and Potters and Bouchaud (2003) provide statistical features on the order book itself. These empirical studies, that have been reviewed in the first part of this review, are the foundation for “zero-intelligence” models, in which “stylized facts” are expected to be reproduced by the properties of the order flows and the structure of order book itself, without considering exogenous “rationality”. Challet and Stinchcombe (2001) propose a simple model of order flows: limit orders are deposited in the order book and can be removed if not executed, in a simple deposition-evaporation process. Bouchaud *et al.* (2002) use this type of model with empirical distribution as inputs. As of today, the most complete empirical model is to our knowledge Mike and Farmer (2008), where order placement and cancellation models are proposed and fitted on empirical data. Finally, new challenges arise as scientists try to identify simple mechanisms that allow an agent-based model to reproduce non-trivial behaviours: herding behaviour in Cont and Bouchaud (2000), dynamic price placement in Preis *et al.* (2007), threshold behaviour in Cont (2007), etc.

In this part we review some of these models. This survey is of course far from exhaustive, and we have just selected models that we feel are representative of a specific trend of modelling.

## B. Early order-driven market modelling: Market microstructure and policy issues

The pioneering works in simulation of financial markets were aimed to study market regulations. The very first one, Stigler (1964), tries to investigate the effect of regulations of the SEC on American stock markets, using empirical data from the 20’s and the 50’s. Twenty years later, at the start of the computerization of financial markets, Hakansson *et al.* (1985) implements a simulator in order to test the feasibility of automated market making. Instead of reviewing the huge microstructure literature, we refer the reader to the well-known books by O’Hara (1995) or Hasbrouck (2007), for example, for a panorama of this branch of finance. However, by presenting a small selection of early models, we here underline the grounding of recent order book modelling.

### 1. A pioneer order book model

To our knowledge, the first attempt to simulate a financial market was by Stigler (1964). This paper was

a biting and controversial reaction to the Report of the Special Study of the Securities Markets of the SEC (Cohen (1963a)), whose aim was to “study the adequacy of rules of the exchange and that the New York stock exchange undertakes to regulate its members in all of their activities” (Cohen (1963b)). According to Stigler, this SEC report lacks rigorous tests when investigating the effects of regulation on financial markets. Stating that “demand and supply are [...] erratic flows with sequences of bids and asks dependent upon the random circumstances of individual traders”, he proposes a simple simulation model to investigate the evolution of the market. In this model, constrained by simulation capability in 1964, price is constrained within  $L = 10$  ticks. (Limit) orders are randomly drawn, in trade time, as follows: they can be bid or ask orders with equal probability, and their price level is uniformly distributed on the price grid. Each time an order crosses the opposite best quote, it is a market order. All orders are of size one. Orders not executed  $N = 25$  time steps after their submission are cancelled. Thus,  $N$  is the maximum number of orders available in the order book.

In the original paper, a run of a hundred trades was manually computed using tables of random numbers. Of course, no particular results concerning the “stylized facts” of financial time series was expected at that time. However, in his review of some order book models, Slanina (2008) makes simulations of a similar model, with parameters  $L = 5000$  and  $N = 5000$ , and shows that price returns are not Gaussian: their distribution exhibits power law with exponent 0.3, far from empirical data. As expected, the limitation  $L$  is responsible for a sharp cut-off of the tails of this distribution.

## 2. Microstructure of the double auction

Garman (1976) provides an early study of the double auction market with a point of view that does not ignore temporal structure, and really defines order flows. Price is discrete and constrained to be within  $\{p_1, p_L\}$ . Buy and sell orders are assumed to be submitted according to two Poisson processes of intensities  $\lambda$  and  $\mu$ . Each time an order crosses the best opposite quote, it is a market order. All quantities are assumed to be equal to one. The aim of the author was to provide an empirical study of the market microstructure. The main result of its Poisson model was to support the idea that negative correlation of consecutive price changes is linked the microstructure of the double auction exchange. This paper is very interesting because it can be seen as precursor that clearly sets the challenges of order book modelling. First, the mathematical formulation is promising. With its fixed constrained prices, Garman (1976) can define the state of the order book at a given time as the vector  $(n_i)_{i=1,\dots,L}$  of awaiting orders (negative quantity for bid orders, positive for ask orders). Future analytical models will use similar vector formulations that can be cast it into known

mathematical processes in order to extract analytical results – see e.g. Cont *et al.* (2008) reviewed below. Second, the author points out that, although the Poisson model is simple, analytical solution is hard to work out, and he provides Monte Carlo simulation. The need for numerical and empirical developments is a constant in all following models. Third, the structural question is clearly asked in the conclusion of the paper: “Does the auction-market model imply the characteristic leptokurtosis seen in empirical security price changes?”. The computerization of markets that was about to take place when this research was published – Toronto’s CATS<sup>2</sup> opened a year later in 1977 – motivated many following papers on the subject. As an example, let us cite here Hakansson *et al.* (1985), who built a model to choose the right mechanism for setting clearing prices in a multi-securities market.

## 3. Zero-intelligence

In the models by Stigler (1964) and Garman (1976), orders are submitted in a purely random way on the grid of possible prices. Traders do not observe the market here and do not act according to a given strategy. Thus, these two contributions clearly belong to a class of “zero-intelligence” models. To our knowledge, Gode and Sunder (1993) is the first paper to introduce the expression “zero-intelligence” in order to describe non-strategic behaviour of traders. It is applied to traders that submit random orders in a double auction market. The expression has since been widely used in agent-based modelling, sometimes in a slightly different meaning (see more recent models described in this review). In Gode and Sunder (1993), two types of zero-intelligence traders are studied. The first are unconstrained zero-intelligence traders. These agents can submit random order at random prices, within the allowed price range  $\{1, \dots, L\}$ . The second are constrained zero-intelligence traders. These agents submit random orders as well, but with the constraint that they cannot cross their given reference price  $p_i^R$ : constrained zero-intelligence traders are not allowed to buy or sell at loss. The aim of the authors was to show that double auction markets exhibit an intrinsic “allocative efficiency” (ratio between the total profit earned by the traders divided by the maximum possible profit) even with zero-intelligence traders. An interesting fact is that in this experiment, price series resulting from actions by zero-intelligence traders are much more volatile than the ones obtained with constrained traders. This fact will be confirmed in future models where “fundamentalists” traders, having a reference price, are expected to stabilize the market (see Wyart and Bouchaud (2007) or Lux and Marchesi (2000) below). Note that the results have been criticized by Cliff and Bruten (1997), who show that

---

<sup>2</sup> Computer Assisted Trading System

the observed convergence of the simulated price towards the theoretical equilibrium price may be an artefact of the model. More precisely, the choice of traders' demand carry a lot of constraints that alone explain the observed results.

Modern works in Econophysics owe a lot to these early models or contributions. Starting in the mid-90's, physicists have proposed simple order book models directly inspired from Physics, where the analogy "order  $\equiv$  particle" is emphasized. Three main contributions are presented in the next section.

### C. Order-driven market modelling in Econophysics

#### 1. The order book as a reaction-diffusion model

A very simple model directly taken from Physics was presented in Bak *et al.* (1997). The authors consider a market with  $N$  noise traders able to exchange one share of stock at a time. Price  $p(t)$  at time  $t$  is constrained to be an integer (i.e. price is quoted in number of ticks) with an upper bound  $\bar{p}$ :  $\forall t, p(t) \in \{0, \dots, \bar{p}\}$ . Simulation is initiated at time 0 with half of the agents asking for one share of stock (buy orders, bid) with price:

$$p_b^j(0) \in \{0, \bar{p}/2\}, \quad j = 1, \dots, N/2, \quad (1)$$

and the other half offering one share of stock (sell orders, ask) with price:

$$p_s^j(0) \in \{\bar{p}/2, \bar{p}\}, \quad j = 1, \dots, N/2. \quad (2)$$

At each time step  $t$ , agents revise their offer by exactly one tick, with equal probability to go up or down. Therefore, at time  $t$ , each seller (resp. buyer) agent chooses his new price as:

$$p_s^j(t+1) = p_s^j(t) \pm 1 \quad (\text{resp. } p_b^j(t+1) = p_b^j(t) \pm 1). \quad (3)$$

A transaction occurs when there exists  $(i, j) \in \{1, \dots, N/2\}^2$  such that  $p_b^i(t+1) = p_s^j(t+1)$ . In such a case the orders are removed and the transaction price is recorded as the new price  $p(t)$ . Once a transaction has been recorded, two orders are placed at the extreme positions on the grid:  $p_b^i(t+1) = 0$  and  $p_s^j(t+1) = \bar{p}$ . As a consequence, the number of orders in the order book remains constant and equal to the number of agents. In figure 1, an illustration of these moving particles is given.

As pointed out by the authors, this process of simulation is similar the reaction-diffusion model  $A + B \rightarrow \emptyset$  in Physics. In such a model, two types of particles are inserted at each side of a pipe of length  $\bar{p}$  and move randomly with steps of size 1. Each time two particles collide, they're annihilated and two new particles are inserted. The analogy is summarized in table I. Following this analogy, it thus can be showed that the variation

TABLE I. Analogy between the  $A + B \rightarrow \emptyset$  reaction model and the order book in Bak *et al.* (1997).

| Physics     | Bak <i>et al.</i> (1997) |
|-------------|--------------------------|
| Particles   | Orders                   |
| Finite Pipe | Order book               |
| Collision   | Transaction              |

$\Delta p(t)$  of the price  $p(t)$  verifies :

$$\Delta p(t) \sim t^{1/4} (\ln(\frac{t}{t_0}))^{1/2}. \quad (4)$$

Thus, at long time scales, the series of price increments simulated in this model exhibit a Hurst exponent  $H = 1/4$ . As for the stylized fact  $H \approx 0.7$ , this sub-diffusive behavior appears to be a step in the wrong direction compared to the random walk  $H = 1/2$ . Moreover, Slanina (2008) points out that no fat tails are observed in the distribution of the returns of the model, but rather fits the empirical distribution with an exponential decay. Other drawbacks of the model could be mentioned. For example, the reintroduction of orders at each end of the pipe leads to unrealistic shape of the order book, as shown on figure 2. Actually here is the main drawback of the model: "moving" orders is highly unrealistic as for modelling an order book, and since it does not reproduce any known financial exchange mechanism, it cannot be the base for any larger model. Therefore, attempts by the authors to build several extensions of this simple framework, in order to reproduce "stylized facts" by adding fundamental traders, strategies, trends, etc. are not of interest for us in this review. However, we feel that the basic model as such is very interesting because of its simplicity and its "particle" representation of an order-driven market that has opened the way for more realistic models.

#### 2. Introducing market orders

Maslov (2000) keeps the zero-intelligence structure of the Bak *et al.* (1997) model but adds more realistic fea-

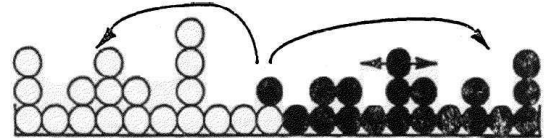


FIG. 1. Illustration of the Bak, Paczuski and Shubik model: white particles (buy orders, bid) moving from the left, black particles (sell orders, ask) moving from the right. Reproduced from Bak *et al.* (1997).



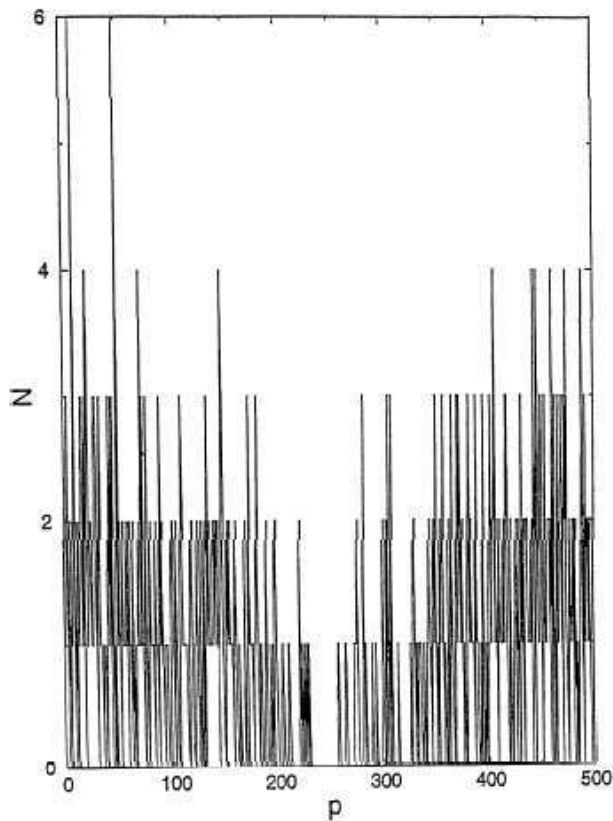


FIG. 2. Snapshot of the limit order book in the Bak, Paczuski and Shubik model. Reproduced from Bak *et al.* (1997).

tures in the order placement and evolution of the market. First, limit orders are submitted and stored in the model, without moving. Second, limit orders are submitted around the best quotes. Third, market orders are submitted to trigger transactions. More precisely, at each time step, a trader is chosen to perform an action. This trader can either submit a limit order with probability  $q_l$  or submit a market order with probability  $1 - q_l$ . Once this choice is made, the order is a buy or sell order with equal probability. All orders have a one unit volume.

As usual, we denote  $p(t)$  the current price. In case the submitted order at time step  $t + 1$  is a limit ask (resp. bid) order, it is placed in the book at price  $p(t) + \Delta$  (resp.  $p(t) - \Delta$ ),  $\Delta$  being a random variable uniformly distributed in  $]0; \Delta^M = 4]$ . In case the submitted order at time step  $t + 1$  is a market order, one order at the opposite best quote is removed and the price  $p(t + 1)$  is recorded. In order to prevent the number of orders in the order book from large increase, two mechanisms are proposed by the author: either keeping a fixed maximum number of orders (by discarding new limit orders when this maximum is reached), or removing them after a fixed lifetime if they have not been executed.

Numerical simulations show that this model exhibits non-Gaussian heavy-tailed distributions of returns. On

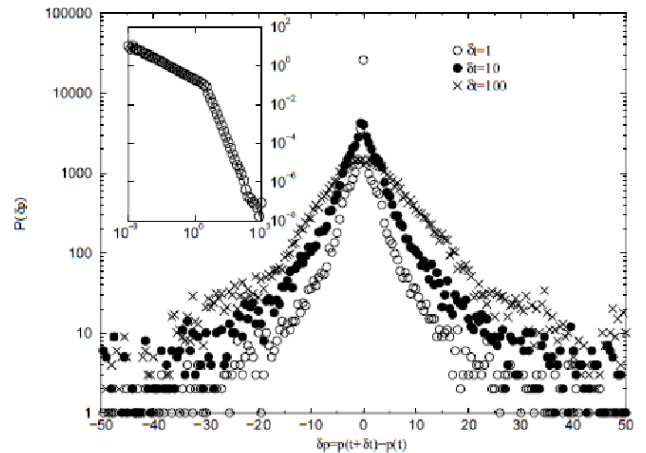


FIG. 3. Empirical probability density functions of the price increments in the Maslov model. In inset, log-log plot of the positive increments. Reproduced from Maslov (2000).

figure 3, the empirical probability density of the price increments for several time scales are plotted. For a time scale  $\delta t = 1$ , the author fit the tails distribution with a power law with exponent 3.0, i.e. reasonable compared to empirical value. However, the Hurst exponent of the price series is still  $H = 1/4$  with this model. It should also be noted that Slanina (2001) proposed an analytical study of the model using a mean-field approximation (See below section II E).

This model brings very interesting innovations in order book simulation: order book with (fixed) limit orders, market orders, necessity to cancel orders waiting too long in the order book. These features are of prime importance in any following order book model.

### 3. The order book as a deposition-evaporation process

Challet and Stinchcombe (2001) continue the work of Bak *et al.* (1997) and Maslov (2000), and develop the analogy between dynamics of an order book and an infinite one dimensional grid, where particles of two types (ask and bid) are subject to three types of events: *deposition* (limit orders), *annihilation* (market orders) and *evaporation* (cancellation). Note that annihilation occurs when a particle is deposited on a site occupied by a particle of another type. The analogy is summarized in table II. Hence, the model goes as follows: At each time step, a bid (resp. ask) order is deposited with probability  $\lambda$  at a price  $n(t)$  drawn according to a Gaussian distribution centred on the best ask  $a(t)$  (resp. best bid  $b(t)$ ) and with variance depending linearly on the spread  $s(t) = a(t) - b(t)$ :  $\sigma(t) = Ks(t) + C$ . If  $n(t) > a(t)$  (resp.  $n(t) < b(t)$ ), then it is a market order: annihilation takes place and the price is recorded. Otherwise, it is a limit order and it is stored in the book. Finally, each limit or-

TABLE II. Analogy between the deposition-evaporation process and the order book in Challet and Stinchcombe (2001).

| Physics          | Challet and Stinchcombe (2001) |
|------------------|--------------------------------|
| Particles        | Orders                         |
| Infinite lattice | Order book                     |
| Deposition       | Limit orders submission        |
| Evaporation      | Limit orders cancellation      |
| Annihilation     | Transaction                    |

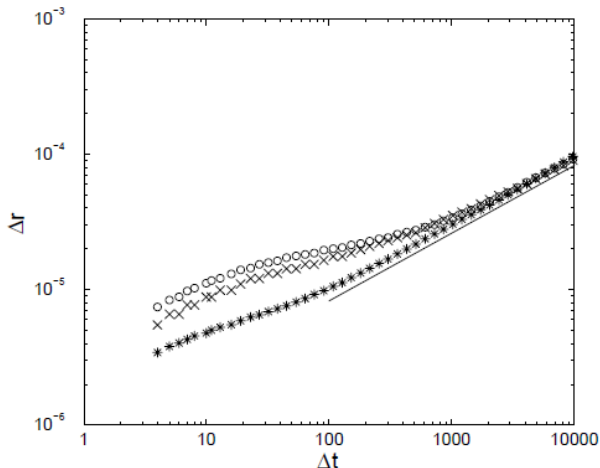


FIG. 4. Average return  $\langle r_{\Delta t} \rangle$  as a function of  $\Delta t$  for different sets of parameters and simultaneous depositions allowed in the Challet and Stinchcombe model. Reproduced from Challet and Stinchcombe (2001).

der stored in the book has a probability  $\delta$  to be cancelled (evaporation).

Figure 4 shows the average return as a function of the time scale. It appears that the series of price returns simulated with this model exhibit a Hurst exponent  $H = 1/4$  for short time scales, and that tends to  $H = 1/2$  for larger time scales. This behaviour might be the consequence of the random evaporation process (which was not modelled in Maslov (2000), where  $H = 1/4$  for large time scales). Although some modifications of the process (more than one order per time step) seem to shorten the sub-diffusive region, it is clear that no over-diffusive behaviour is observed.

#### D. Empirical zero-intelligence models

The three models presented in the previous section II C have successively isolated essential mechanisms that are to be used when simulating a “realistic” market: one order is the smallest entity of the model; the submission of one order is the time dimension (i.e. event time is used, not an exogenous time defined by market clearing and “tatonnement” on exogenous supply and demand func-

tions); submission of market orders (as such in Maslov (2000), as “crossing limit orders” in Challet and Stinchcombe (2001)) and cancellation of orders are taken into account. On the one hand, one may try to describe these mechanisms using a small number of parameters, using Poisson process with constant rates for order flows, constant volumes, etc. This might lead to some analytically tractable models, as will be described in section II E. On the other hand, one may try to fit more complex empirical distributions to market data without analytical concern.

This type of modelling is best represented by Mike and Farmer (2008). It is the first model that proposes an advanced calibration on the market data as for order placement and cancellation methods. As for volume and time of arrivals, assumptions of previous models still hold: all orders have the same volume, discrete event time is used for simulation, i.e. one order (limit or market) is submitted per time step. Following Challet and Stinchcombe (2001), there is no distinction between market and limit orders, i.e. market orders are limit orders that are submitted across the spread  $s(t)$ . More precisely, at each time step, one trading order is simulated: an ask (resp. bid) trading order is randomly placed at  $n(t) = a(t) + \delta a$  (resp.  $n(t) = b(t) + \delta b$ ) according to a Student distribution with scale and degrees of freedom calibrated on market data. If an ask (resp. bid) order satisfies  $\delta a < -s(t) = b(t) - a(t)$  (resp.  $\delta b > s(t) = a(t) - b(t)$ ), then it is a buy (resp. sell) market order and a transaction occurs at price  $a(t)$  (resp.  $b(t)$ ).

During a time step, several cancellations of orders may occur. The authors propose an empirical distribution for cancellation based on three components for a given order:

- the position in the order book, measured as the ratio  $y(t) = \frac{\Delta(t)}{\Delta(0)}$  where  $\Delta(t)$  is the distance of the order from the opposite best quote at time  $t$ ,
- the order book imbalance, measured by the indicator  $N_{imb}(t) = \frac{N_a(t)}{N_a(t) + N_b(t)}$  (resp.  $N_{imb}(t) = \frac{N_b(t)}{N_a(t) + N_b(t)}$ ) for ask (resp. bid) orders, where  $N_a(t)$  and  $N_b(t)$  are the number of orders at ask and bid in the book at time  $t$ ,
- the total number  $N(t) = N_a(t) + N_b(t)$  of orders in the book.

Their empirical study leads them to assume that the cancellation probability has an exponential dependance on  $y(t)$ , a linear one in  $N_{imb}$  and finally decreases approximately as  $1/N_t(t)$  as for the total number of orders. Thus, the probability  $P(C|y(t), N_{imb}(t), N_t(t))$  to cancel an ask order at time  $t$  is formally written :

$$P(C|y(t), N_{imb}(t), N_t(t)) = A(1 - e^{-y(t)})(N_{imb}(t) + B) \frac{1}{N_t(t)}, \quad (5)$$

where the constants  $A$  and  $B$  are to be fitted on market data. Figure 5 shows that this empirical formula provides a quite good fit on market data.

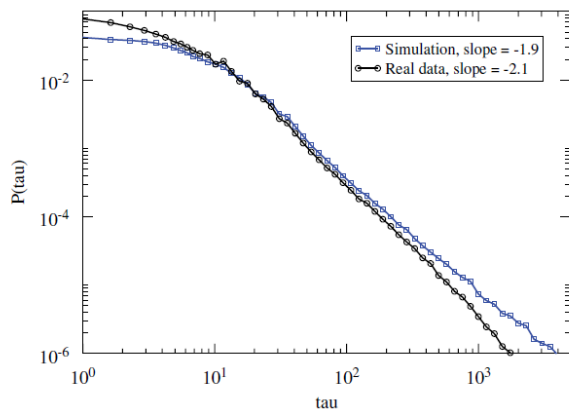


FIG. 5. Lifetime of orders for simulated data in the Mike and Farmer model, compared to the empirical data used for fitting. Reproduced from Mike and Farmer (2008).

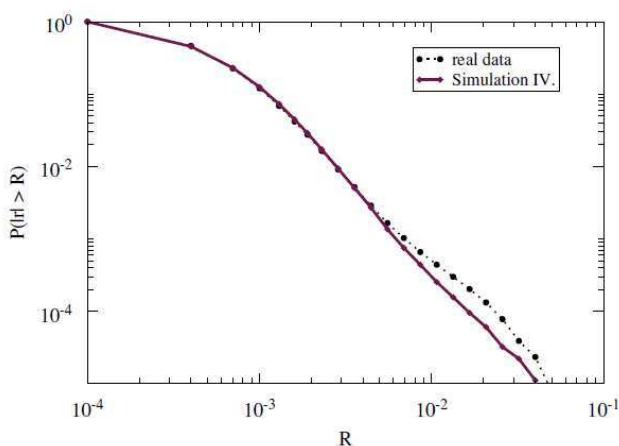


FIG. 6. Cumulative distribution of returns in the Mike and Farmer model, compared to the empirical data used for fitting. Reproduced from Mike and Farmer (2008).

Finally, the authors mimic the observed long memory of order signs by simulating a fractional Brownian motion. The auto-covariance function  $\Gamma(t)$  of the increments of such a process exhibits a slow decay :

$$\Gamma(k) \sim H(2H - 1)t^{2H-2} \quad (6)$$

and it is therefore easy to reproduce exponent  $\beta$  of the decay of the empirical autocorrelation function of order signs observed on the market with  $H = 1 - \beta/2$ .

The results of this empirical model are quite satisfying as for return and spread distribution. The distribution of returns exhibit fat tails which are in agreement with empirical data, as shown on figure 6. The spread distribution is also very well reproduced. As their empirical model has been built on the data of only one stock, the authors test their model on 24 other data sets of stocks on the same market and find for half of them a good agreement between empirical and simulated properties. However, the bad results of the other half suggest that

such a model is still far from being “universal”.

Despite these very nice results, some drawbacks have to be pointed out. The first one is the fact that the stability of the simulated order book is far from ensured. Simulations using empirical parameters in the simulations may bring situations where the order book is emptied by large consecutive market orders. Thus, the authors require that there is at least two orders in each side of the book. This exogenous trick might be important, since it is activated precisely in the case of rare events that influence the tails of the distributions. Also, the original model does not focus on volatility clustering. Gu and Zhou (2009) propose a variant that tackles this feature. Another important drawback of the model is the way order signs are simulated. As noted by the authors, using an exogenous fractional Brownian motion leads to correlated price returns, which is in contradiction with empirical stylized facts. We also find that at long time scales it leads to a dramatic increase of volatility. As we have seen in the first part of the review, the correlation of trade signs can be at least partly seen as an artefact of execution strategies. Therefore this element is one of the numerous that should be taken into account when “programming” the agents of the model. In order to do so, we have to leave the (quasi) “zero-intelligence” world and see how modelling based on heterogeneous agents might help to reproduce non-trivial behaviours. Prior to this development below in IIF, we briefly review some analytical works on the “zero-intelligence” models.

## E. Analytical treatments of zero-intelligence models

In this section we present some analytical results obtained on zero-intelligence models where processes are kept sufficiently simple so that a mean-field approximation may be derived (Slanina (2001)) or probabilities conditionally to the state of the order book may be computed (Cont *et al.* (2008)). The key assumptions here are such that the process describing the order book is stationary. This allows either to write a stable density equation, or to fit the model into a nice mathematical framework such as ergodic Markov chains.

### 1. Mean-field theory

Slanina (2001) proposes an analytical treatment of the model introduced by Maslov (2000) and reviewed above. Let us briefly described the formalism used. The main hypothesis is the following: on each side of the current price level, the density of limit orders is uniform and constant (and  $\rho_+$  on the ask side,  $\rho_-$  on the bid side). In that sense, this is a “mean-field” approximation since the individual position of a limit order is not taken into account. Assuming we are in a stable state, the arrival of a market order of size  $s$  on the ask (resp. bid) side will make the price change by  $x_+ = s/\rho_+$  (resp.  $x_- = s/\rho_-$ ).

It is then observed that the transformations of the vector  $X = (x_+, x_-)$  occurring at each event (new limit order, new buy market order, new sell market order) are linear transformation that can easily and explicitly be written. Therefore, an equation satisfied by the probability distribution  $P$  of the vector  $X$  of price changes can be obtained. Finally, assuming further simplifications (such as  $\rho_+ = \rho_-$ ), one can solve this equation for a tail exponent and find that the distribution behaves as  $P(x) \approx x^{-2}$  for large  $x$ . This analytical result is slightly different from the one obtained by simulation in Maslov (2000). However, the numerous approximations make the comparison difficult. The main point here is that some sort of mean-field approximation is natural if we assume the existence of a stationary state of the order book, and thus may help handling order book models.

Smith *et al.* (2003) also propose some sort of mean-field approximation for zero-intelligence models. In a similar model (but including a cancellation process), mean field theory and dimensional analysis produces interesting results. For example, it is easy to see that the book depth (i.e. number of orders)  $N_e(p)$  at a price  $p$  far away from the best quotes is given by  $N_e(p) = \lambda/\delta$ , where  $\lambda$  is the rate of arrival of limit orders per unit of time and per unit of price, and  $\delta$  the probability for an order to be cancelled per unit of time. Indeed, far from the best quotes no market orders occurs, so that if a steady-state exists, the number of limit orders per time step  $\lambda$  must be balanced by the number of cancellation  $\delta N_e(p)$  per unit of time, hence the result.

## 2. Explicit computation of probabilities conditionally on the state of the order book

Cont *et al.* (2008) is an original attempt at analytical treatments of limit order books. In their model, the price is constrained to be on a grid  $\{1, \dots, N\}$ . The state of the order book can then be described by a vector  $X(t) = (X_1(t), \dots, X_N(t))$  where  $|X_i(t)|$  is the quantity offered in the order book at price  $i$ . Conventionally,  $X_i(t), i = 1, \dots, N$  is positive on the ask side and negative on the bid side. As usual, limit orders arrive at level  $i$  at a constant rate  $\lambda_i$ , and market orders arrive at a constant rate  $\mu$ . Finally, at level  $i$ , each order can be cancelled at a rate  $\theta_i$ . Using this setting, Cont *et al.* (2008) show that each event (limit order, market order, cancellation) transforms the vector  $X$  in a simple linear way. Therefore, it is shown that under reasonable conditions,  $X$  is an ergodic Markov chain, and thus admits a stationary state. The original idea is then to use this formalism to compute conditional probabilities on the processes. More precisely, it is shown that using Laplace transform, one may explicitly compute the probability of an increase of the mid price conditionally on the current state of the order book.

This original contribution could allow explicit evaluation of strategies and open new perspectives in high-

frequency trading. However, it is based on a simple model that does not reproduce empirical observations such as volatility clustering. Complex models trying to include market interactions will not fit into these analytical frameworks. We review some of these models in the next section.

## F. Towards non-trivial behaviours: modelling market interactions

In all the models we have reviewed until now, flows of orders are treated as independent processes. Under some (strong) modelling constraints, we can see the order book as a Markov chain and look for analytical results (Cont *et al.* (2008)). In any case, even if the process is empirically detailed and not trivial (Mike and Farmer (2008)), we work with the assumption that orders are independent and identically distributed. This very strong (and false) hypothesis is similar to the “representative agent” hypothesis in Economics: orders being successively and independently submitted, we may not expect anything but regular behaviours. Following the work of economists such as Kirman (1992, 1993, 2002), one has to translate the heterogeneous property of the markets into the agent-based models. Agents are not identical, and not independent.

In this section we present some toy models implementing mechanisms that aim at bringing heterogeneity: herding behaviour on markets in Cont and Bouchaud (2000), trend following behaviour in Lux and Marchesi (2000) or in Preis *et al.* (2007), threshold behaviour Cont (2007). Most of the models reviewed in this section are not order book models, since a persistent order book is not kept during the simulations. They are rather price models, where the price changes are determined by the aggregation of excess supply and demand. However, they identify essential mechanisms that may clearly explain some empirical data. Incorporating these mechanisms in an order book model is not yet achieved but is certainly a future prospective.

### 1. Herding behaviour

The model presented in Cont and Bouchaud (2000) considers a market with  $N$  agents trading a given stock with price  $p(t)$ . At each time step, agents choose to buy or sell one unit of stock, i.e. their demand is  $\phi_i(t) = \pm 1, i = 1, \dots, N$  with probability  $a$  or are idle with probability  $1-2a$ . The price change is assumed to be linearly linked with the excess demand  $D(t) = \sum_{i=1}^N \phi_i(t)$  with a factor  $\lambda$  measuring the liquidity of the market :

$$p(t+1) = p(t) + \frac{1}{\lambda} \sum_{i=1}^N \phi_i(t). \quad (7)$$



$\lambda$  can also be interpreted as a market depth, i.e. the excess demand needed to move the price by one unit. In order to evaluate the distribution of stock returns from Eq.(7), we need to know the joint distribution of the individual demands  $(\phi_i(t))_{1 \leq i \leq N}$ . As pointed out by the authors, if the distribution of the demand  $\phi_i$  is independent and identically distributed with finite variance, then the Central Limit Theorem stands and the distribution of the price variation  $\Delta p(t) = p(t+1) - p(t)$  will converge to a Gaussian distribution as  $N$  goes to infinity.

The idea here is to model the diffusion of the information among traders by randomly linking their demand through clusters. At each time step, agents  $i$  and  $j$  can be linked with probability  $p_{ij} = p = \frac{c}{N}$ ,  $c$  being a parameter measuring the degree of clustering among agents. Therefore, an agent is linked to an average number of  $(N-1)p$  other traders. Once clusters are determined, the demand are forced to be identical among all members of a given cluster. Denoting  $n_c(t)$  the number of cluster at a given time step  $t$ ,  $W_k$  the size of the  $k$ -th cluster,  $k = 1, \dots, n_c(t)$  and  $\phi_k = \pm 1$  its investment decision, the price variation is then straightforwardly written :

$$\Delta p(t) = \frac{1}{\lambda} \sum_{k=1}^{n_c(t)} W_k \phi_k. \quad (8)$$

This modelling is a direct application to the field of finance of the random graph framework as studied in Erdos and Renyi (1960). Kirman (1983) previously suggested it in economics. Using these previous theoretical works, and assuming that the size of a cluster  $W_k$  and the decision taken by its members  $\phi_k(t)$  are independent, the author are able to show that the distribution of the price variation at time  $t$  is the sum of  $n_c(t)$  independent identically distributed random variables with heavy-tailed distributions :

$$\Delta p(t) = \frac{1}{\lambda} \sum_{k=1}^{n_c(t)} X_k, \quad (9)$$

where the density  $f(x)$  of  $X_k = W_k \phi_k$  is decaying as :

$$f(x) \sim_{|x| \rightarrow \infty} \frac{A}{|x|^{5/2}} e^{\frac{-(c-1)|x|}{W_0}}. \quad (10)$$

Thus, this simple toy model exhibits fat tails in the distribution of prices variations, with a decay reasonably close to empirical data. Therefore, Cont and Bouchaud (2000) show that taking into account a naive mechanism of communication between agents (herding behaviour) is able to drive the model out of the Gaussian convergence and produce non-trivial shapes of distributions of price returns.

## 2. Fundamentalists and trend followers

Lux and Marchesi (2000) proposed a model very much in line with agent-based models in behavioural finance,

but where trading rules are kept simple enough so that they can be identified with a presumably realistic behaviour of agents. This model considers a market with  $N$  agents that can be part of two distinct groups of traders:  $n_f$  traders are “fundamentalists”, who share an exogenous idea  $p_f$  of the value of the current price  $p$ ; and  $n_c$  traders are “chartists” (or trend followers), who make assumptions on the price evolution based on the observed trend (mobile average). The total number of agents is constant, so that  $n_f + n_c = N$  at any time. At each time step, the price can be moved up or down with a fixed jump size of  $\pm 0.01$  (a tick). The probability to go up or down is directly linked to the excess demand  $ED$  through a coefficient  $\beta$ . The demand of each group of agents is determined as follows :

- Each fundamentalist trades a volume  $V_f$  proportional, with a coefficient  $\gamma$ , to the deviation of the current price  $p$  from the perceived fundamental value  $p_f$ :  $V_f = \gamma(p_f - p)$ .
- Each chartist trades a constant volume  $V_c$ . Denoting  $n_+$  the number of optimistic (buyer) chartists and  $n_-$  the number of pessimistic (seller) chartists, the excess demand by the whole group of chartists is written  $(n_+ - n_-)V_c$ .

Therefore, assuming that there exists some noise traders on the market with random demand  $\mu$ , the global excess demand is written :

$$ED = (n_+ - n_-)V_c + n_f \gamma (p_f - p) + \mu. \quad (11)$$

The probability that the price goes up (resp. down) is then defined to be the positive (resp. negative) part of  $\beta ED$ .

As observed in Wyart and Bouchaud (2007), fundamentalists are expected to stabilize the market, while chartists should destabilize it. In addition, following Cont and Bouchaud (2000), the authors expect non-trivial features of the price series to results from herding behaviour and transitions between groups of traders. Referring to Kirman’s work as well, a mimicking behaviour among chartists is thus proposed. The  $n_c$  chartists can change their view on the market (optimistic, pessimistic), their decision being based on a clustering process modelled by an opinion index  $x = \frac{n_+ - n_-}{n_c}$  representing the weight of the majority. The probabilities  $\pi_+$  and  $\pi_-$  to switch from one group to another are formally written :

$$\pi_{\pm} = v \frac{n_c}{N} e^{\pm U}, \quad U = \alpha_1 x + \alpha_2 p/v, \quad (12)$$

where  $v$  is a constant, and  $\alpha_1$  and  $\alpha_2$  reflect respectively the weight of the majority’s opinion and the weight of the observed price in the chartists’ decision. Transitions between fundamentalists and chartists are also allowed, decided by comparison of expected returns (see Lux and Marchesi (2000) for details).

The authors show that the distribution of returns generated by their model have excess kurtosis. Using a

Hill estimator, they fit a power law to the fat tails of the distribution and observe exponents grossly ranging from 1.9 to 4.6. They also check hints for volatility clustering: absolute returns and squared returns exhibit a slow decay of autocorrelation, while raw returns do not. It thus appears that such a model can grossly fit some “stylized facts”. However, the number of parameters involved, as well as the complicated rules of transition between agents, make clear identification of sources of phenomena and calibration to market data difficult and intractable.

Alfi *et al.* (2009a,b) provide a somewhat simplifying view on the Lux-Marchesi model. They clearly identify the fundamentalist behaviour, the chartist behaviour, the herding effect and the observation of the price by the agents as four essential effects of an agent-based financial model. They show that the number of agents plays a crucial role in a Lux-Marchesi-type model: more precisely, the stylized facts are reproduced only with a finite number of agents, not when the number of agents grows asymptotically, in which case the model stays in a fundamentalist regime. There is a finite-size effect that may prove important for further studies.

The role of the trend following mechanism in producing non-trivial features in price time series is also studied in Preis *et al.* (2007). The starting point is an order book model similar to Challet and Stinchcombe (2001) and Smith *et al.* (2003): at each time step, liquidity providers submit limit orders at rate  $\lambda$  and liquidity takers submit market orders at rate  $\mu$ . As expected, this zero-intelligence framework does not produce fat tails in the distribution of (log-)returns nor an over-diffusive Hurst exponent. Then, a stochastic link between order placement and market trend is added: it is assumed that liquidity providers observing a trend in the market will act consequently and submit limit orders at a wider depth in the order book. Although the assumption behind such a mechanism may not be empirically confirmed (a questionable symmetry in order placement is assumed) and should be further discussed, it is interesting enough that it directly provides fat tails in the log-return distributions and an over-diffusive Hurst exponent  $H \approx 0.6 - 0.7$  for medium time-scales, as shown in figure 7.

### 3. Threshold behaviour

We finally review a model focusing primarily on reproducing the stylized fact of volatility clustering, while most of the previous models we have reviewed were mostly focused on fat tails of log returns. Cont (2007) proposes a model with a rather simple mechanism to create volatility clustering. The idea is that volatility clustering characterizes several regimes of volatility (quite periods vs bursts of activity). Instead of implementing an exogenous change of regime, the author defines the following trading rules.

At each period, an agent  $i \in \{1, \dots, N\}$  can issue a buy

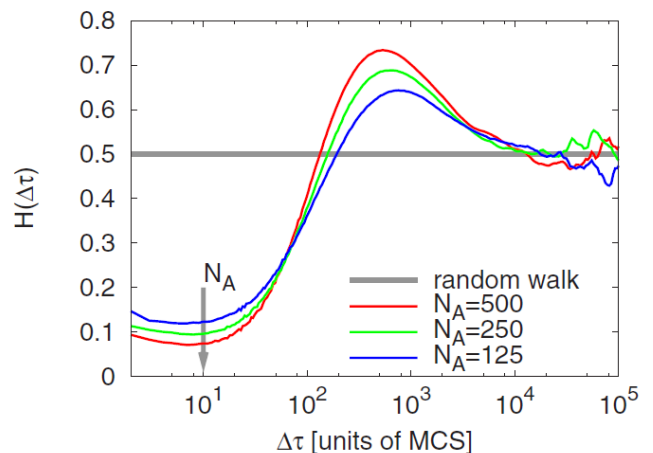


FIG. 7. Hurst exponent found in the Preis model for different number of agents when including random demand perturbation and dynamic limit order placement depth. Reproduced from Preis *et al.* (2007).

or a sell order:  $\phi_i(t) = \pm 1$ . Information is represented by a series of i.i.d Gaussian random variables.  $(\epsilon_t)$ . This public information  $\epsilon_t$  is a forecast for the value  $r_{t+1}$  of the return of the stock. Each agent  $i \in \{1, \dots, N\}$  decides whether to follow this information according to a threshold  $\theta_i > 0$  representing its sensibility to the public information:

$$\phi_i(t) = \begin{cases} 1 & \text{if } \epsilon_i(t) > \theta_i(t) \\ 0 & \text{if } |\epsilon_i(t)| < \theta_i(t) \\ -1 & \text{if } \epsilon_i(t) < -\theta_i(t) \end{cases} \quad (13)$$

Then, once every choice is made, the price evolves according to the excess demand  $D(t) = \sum_{i=1}^N \phi_i(t)$ , in a way similar to Cont and Bouchaud (2000). At the end of each time step  $t$ , threshold are asynchronously updated. Each agent has a probability  $s$  to update its threshold  $\theta_i(t)$ . In such a case, the new threshold  $\theta_i(t+1)$  is defined to be the absolute value  $|r_t|$  of the return just observed. In short:

$$\theta_i(t+1) = \mathbf{1}_{\{u_i(t) < s\}} |r_t| + \mathbf{1}_{\{u_i(t) > s\}} \theta_i(t). \quad (14)$$

The author shows that the time series simulated with such a model do exhibit some realistic facts on volatility. In particular, long range correlations of absolute returns is observed. The strength of this model is that it directly links the state of the market with the decision of the trader. Such a feedback mechanism is essential in order to obtain non trivial characteristics. Of course, the model presented in Cont (2007) is too simple to be fully calibrated on empirical data, but its mechanism could be used in a more elaborate agent-based model in order to reproduce the empirical evidence of volatility clustering.

## G. Remarks

Let us attempt to make some concluding remarks on these developments of agent-based models for order books. In table III, we summarize some key features of some of the order book models reviewed in this section. Among important elements for future modelling, we may mention the cancellation of orders, which is the less realistic mechanism implemented in existing models ; the order book stability, which is always exogenously enforced (see our review of Mike and Farmer (2008) above) ; and the dependence between order flows (see e.g. Muni Toke (2010) and reference therein). Empirical estimation of these mechanisms is still challenging.

Emphasis has been put in this section on order book modelling, a field that is at the crossroad of many larger disciplines (market microstructure, behavioural finance and physics). Market microstructure is essential since it defines in many ways the goal of the modelling. We pointed out that it is not a coincidence if the work by Garman (1976) was published when computerization of exchanges was about to make the electronic order book the key of all trading. Regulatory issues that pushed early studies are still very important today. Realistic order book models could be a invaluable tool in testing and evaluating the effects of regulations such as the 2005 Regulation NMS<sup>3</sup> in the USA, or the 2007 MiFID<sup>4</sup> in Europe.

## III. AGENT-BASED MODELLING FOR WEALTH DISTRIBUTIONS: KINETIC THEORY MODELS

The distributions of money, wealth or income, i.e., how such quantities are shared among the population of a given country and among different countries, is a topic which has been studied by economists for a long time. The relevance of the topic to us is twofold: From the point of view of the science of Complex Systems, wealth distributions represent a unique example of a quantitative outcome of a collective behavior which can be directly compared with the predictions of theoretical models and numerical experiments. Also, there is a basic interest in wealth distributions from the social point of view, in particular in their degree of (in)equality. To this aim, the Gini coefficient (or the Gini index, if expressed as a percentage), developed by the Italian statistician Corrado Gini, represents a concept commonly employed to measure inequality of wealth distributions or, more in general, how uneven a given distribution is. For a cumulative distribution function  $F(y)$ , that is piecewise differentiable, has a finite mean  $\mu$ , and is zero for  $y < 0$ ,

the Gini coefficient is defined as

$$\begin{aligned} G &= 1 - \frac{1}{\mu} \int_0^\infty dy (1 - F(y))^2 \\ &= \frac{1}{\mu} \int_0^\infty dy F(y)(1 - F(y)). \end{aligned} \quad (15)$$

It can also be interpreted statistically as half the relative mean difference. Thus the Gini coefficient is a number between 0 and 1, where 0 corresponds with perfect equality (where everyone has the same income) and 1 corresponds with perfect inequality (where one person has all the income, and everyone else has zero income). Some values of  $G$  for some countries are listed in Table IV.

Let us start by considering the basic economic quantities: money, wealth and income.

### A. Money, wealth and income

A common definition of *money* suggests that money is the “[c]ommodity accepted by general consent as medium of economics exchange”<sup>5</sup>. In fact, money circulates from one economic agent (which can represent an individual, firm, country, etc.) to another, thus facilitating trade. It is “something which all other goods or services are traded for” (for details see Shostak (2000)). Throughout history various commodities have been used as money, for these cases termed as “commodity money”, which include for example rare seashells or beads, and cattle (such as cow in India). Recently, “commodity money” has been replaced by other forms referred to as “fiat money”, which have gradually become the most common ones, such as metal coins and paper notes. Nowadays, other forms of money, such as electronic money, have become the most frequent form used to carry out transactions. In any case the most relevant points about money employed are its basic functions, which according to standard economic theory are

- to serve as a medium of exchange, which is universally accepted in trade for goods and services;
- to act as a measure of value, making possible the determination of the prices and the calculation of costs, or profit and loss;
- to serve as a standard of deferred payments, i.e., a tool for the payment of debt or the unit in which loans are made and future transactions are fixed;
- to serve as a means of storing wealth not immediately required for use.

<sup>3</sup> National Market System

<sup>4</sup> Markets in Financial Instruments Directive

<sup>5</sup> In Encyclopædia Britannica. Retrieved June 17, 2010, from Encyclopædia Britannica Online

| Model               | Stigler (1961)   | Garman (1976)   | Bak, Paczuski and Shubik (1997)                                    | Maslov (2000)  | Challet and Stinchcombe (2001)  | Mike and Farmer (2008)   |
|---------------------|--|---|--|--|---|--|
| Price range         | Finite grid  | Finite grid   | Finite grid  | Unconstrained  | Unconstrained   | Unconstrained  |
| Clock               | Trade time   | Physical Time   | Aggregated time  | Event time   | Aggregated time   | Aggregated time  |
| Flows / Agents      | One zero-intelligence agent / One flow                             | One zero-intelligence agent / Two flows (buy/sell)                                  | N agents owning each one limit order                               | One zero-intelligence flow (limit order with fixed probability, else market order) | One zero-intelligence agent / One flow  | One zero-intelligence agent / One flow   |
| Limit orders        | Uniform distribution on the price grid                             | Two Poisson processes for buy and sell orders                                       | Moving at each time step by one tick                               | Uniformly distributed in a finite interval around last price                       | Normally distributed around best quote  | Student-distributed around best quote  |
| Market orders       | Defined as crossing limit orders                                   | Defined as crossing limit orders  | Defined as crossing limit orders                                   | Submitted as such  | Defined as crossing limit orders  | Defined as crossing limit orders   |
| Cancellation orders | Pending orders are cancelled after a fixed number of time steps    | None  | None (constant number of pending orders)                           | Pending orders are cancelled after a fixed number of time steps                    | Pending orders can be cancelled with fixed probability at each time step        | Pending orders can be cancelled with 3-parameter conditional probability at each time step |
| Volume              | Unit   | Unit  | Unit   | Unit   | Unit  | Unit   |
| Order signs         | Independent  | Independent   | Independent  | Independent  | Independent   | Correlated with a fractional Brownian motion   |
| Claimed results     | Return distribution is power-law 0.3 / Cut-off because finite grid | Microstructure is responsible for negative correlation of consecutive price changes | No fat tails for returns / Hurst exponent 1/4 for price increments | Fat tails for distributions of returns / Hurst exponent 1/4                        | Hurst exponent 1/4 for short time scales, tending to 1/2 for larger time scales | Fat tails distributions of returns / Realistic spread distribution / Unstable order book   |

TABLE III. Summary of the characteristics of the reviewed limit order book models.

A related feature relevant for the present investigation is that money is the medium in which prices or values of all commodities as well as costs, profits, and transactions can be determined or expressed. *Wealth* is usually understood as things that have economic utility (monetary value or value of exchange), or material goods or property; it also represents the abundance of objects of value (or riches) and the state of having accumulated these objects; for our purpose, it is important to bear in mind that wealth can be measured in terms of money. Also *income*, defined in Case and Fair (2008) as “the sum of all the wages, salaries, profits, interests payments, rents and other forms of earnings received... in a given period of time”, is a quantity which can be measured in terms of money (per unit time).

## B. Modelling wealth distributions

It was first observed by Pareto (1897) that in an economy the higher end of the distribution of income  $f(x)$  follows a power-law,

$$f(x) \sim x^{-1-\alpha}, \quad (16)$$

with  $\alpha$ , now known as the Pareto exponent, estimated by him to be  $\alpha \approx 3/2$ . For the last hundred years the value

of  $\alpha \sim 3/2$  seems to have changed little in time and across the various capitalist economies (see Yakovenko and Rosser (2009) and references therein).

Gibrat (1931) clarified that Pareto’s law is valid only for the high income range, whereas for the middle income range he suggested that the income distribution is described by a log-normal probability density

$$f(x) \sim \frac{1}{x\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{\log^2(x/x_0)}{2\sigma^2} \right\}, \quad (17)$$

where  $\log(x_0) = \langle \log(x) \rangle$  is the mean value of the logarithmic variable and  $\sigma^2 = \langle [\log(x) - \log(x_0)]^2 \rangle$  the corresponding variance. The factor  $\beta = 1/\sqrt{2\sigma^2}$ , also known as Gibrat index, measures the equality of the distribution.

More recent empirical studies on income distribution have been carried out by physicists, e.g. those by Dragulescu and Yakovenko (2001b,a) for UK and US, by Fujiwara *et al.* (2003) for Japan, and by Nirei and Souma (2007) for US and Japan. For an overview see Yakovenko and Rosser (2009). The distributions obtained have been shown to follow either the log-normal (Gamma like) or power-law types, depending on the range of wealth, as shown in Fig. 8.

One of the current challenges is to write down the “microscopic equation” which governs the dynamics of the



TABLE IV. Gini indices (in percent) of some countries (from *Human Development Indicators of the United Nations Human Development Report 2004*, pp.50-53, available at <http://hdr.undp.org/en/reports/global/hdr2004>. More recent data are also available from their website.)

|              |      |
|--------------|------|
| Denmark      | 24.7 |
| Japan        | 24.9 |
| Sweden       | 25.0 |
| Norway       | 25.8 |
| Germany      | 28.3 |
| India        | 32.5 |
| France       | 32.7 |
| Australia    | 35.2 |
| UK           | 36.0 |
| USA          | 40.8 |
| Hong Kong    | 43.4 |
| China        | 44.7 |
| Russia       | 45.6 |
| Mexico       | 54.6 |
| Chile        | 57.1 |
| Brazil       | 59.1 |
| South Africa | 59.3 |
| Botswana     | 63.0 |
| Namibia      | 70.7 |

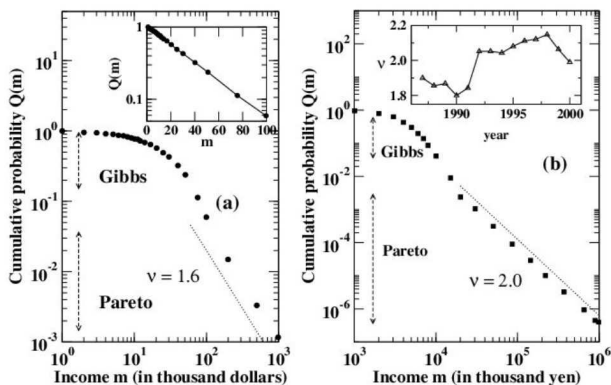


FIG. 8. Income distributions in the US (left) and Japan (right). Reproduced and adapted from Chakrabarti and Chatterjee (2003), available at [arXiv:cond-mat/0302147](https://arxiv.org/abs/cond-mat/0302147).

evolution of wealth distributions, possibly predicting the observed shape of wealth distributions, including the exponential law at intermediate values of wealth as well as the century-old Pareto law. To this aim, several studies have been made to investigate the characteristics of the real income distribution and provide theoretical models or explanations (see e.g. reviews by Lux (2005), Chatterjee and Chakrabarti (2007), Yakovenko and Rosser (2009)).

The model of Gibrat (1931) and other models for-

mulated in terms of a Langevin equation for a single wealth variable, subjected to multiplicative noise (Mandelbrot (1960); Levy and Solomon (1996); Sornette (1998); Burda *et al.* (2003)), can lead to equilibrium wealth distributions with a power law tail, since they converge toward a log-normal distribution. However, the fit of real wealth distributions does not turn out to be as good as that obtained using e.g. a  $\Gamma$ - or a  $\beta$ -distribution, in particular due to too large asymptotic variances (Angle (1986)). Other models use a different approach and describe the wealth dynamics as a wealth flow due to exchanges between (pairs of) basic units. In this respect, such models are basically different from the class of models formulated in terms of a Langevin equation for a single wealth variable. For example, Solomon and Levy (1996) studied the generalized Lotka-Volterra equations in relation to power-law wealth distribution. Ispolatov *et al.* (1998) studied random exchange models of wealth distributions. Other models describing wealth exchange have been formulated using matrix theory (Gupta (2006)), the master equation (Bouchaud and Mezard (2000); Dragulescu and Yakovenko (2000); Ferrero (2004)), the Boltzmann equation approach (Dragulescu and Yakovenko (2000); Slanina (2004); Repetowicz *et al.* (2005); Cordier *et al.* (2005); Matthes and Toscani (2007); Düring and Toscani (2007); Düring *et al.* (2008)), or Markov chains (Scalas *et al.* (2006, 2007); Garibaldi *et al.* (2007)). It should be mentioned that one of the earliest modelling efforts were made by Champernowne (1953). Since then many economists, Gabaix (1999) and Benhabib and Bisin (2009) amongst others, have also studied mechanisms for power laws, and distributions of wealth.

In the two following sections we consider in greater detail a class of models usually referred to as *kinetic wealth exchange models* (KWEM), formulated through finite time difference stochastic equations (Angle (1986, 2002, 2006); Chakrabarti and Chakrabarti (2000); Dragulescu and Yakovenko (2000); Chakrabarti (2002); Hayes (2002); Chatterjee *et al.* (2003); Das and Yarlagadda (2003); Scafetta *et al.* (2004); Iglesias *et al.* (2003, 2004); Ausloos and Pekalski (2007)). From the studies carried out using wealth-exchange models, it emerges that it is possible to use them to generate power law distributions.

### C. Homogeneous kinetic wealth exchange models

Here and in the next section we consider KWEMs, which are statistical models of closed economy. Their goal, rather than describing the market dynamics in terms of intelligent agents, is to predict the time evolution of the distribution of some main quantity, such as wealth, by studying the corresponding flow process among individuals. The underlying idea is that however complicated the detailed rules of wealth exchanges can be, their average behaviour can be described in a relatively more simple way and will share some universal properties with other transport processes, due to general

conservation constraints and the effect of the fluctuations due to the environment or associated to the individual behaviour. In this, there is a clear analogy with the general theory of transport phenomena (e.g. of energy).

In these models the states of agents are defined in terms of the wealth variables  $\{x_n\}$ ,  $n = 1, 2, \dots, N$ . The evolution of the system is carried out according to a trading rule between agents which, for obtaining the final equilibrium distribution, can be interpreted as the actual time evolution of the agent states as well as a Monte Carlo optimization. The algorithm is based on a simple update rule performed at each time step  $t$ , when two agents  $i$  and  $j$  are extracted randomly and an amount of wealth  $\Delta x$  is exchanged,

$$\begin{aligned} x'_i &= x_i - \Delta x, \\ x'_j &= x_j + \Delta x. \end{aligned} \quad (18)$$

Notice that the quantity  $x$  is conserved during single transactions,  $x'_i + x'_j = x_i + x_j$ , where  $x_i = x_i(t)$  and  $x_j = x_j(t)$  are the agent wealth before, whereas  $x'_i = x_i(t+1)$  and  $x'_j = x_j(t+1)$  are the final ones after the transaction. Several rules have been studied for the model defined by Eqs. (18). It is noteworthy, that though this theory has been originally derived from the entropy maximization principle of statistical mechanics, it has recently been shown that the same could be derived from the utility maximization principle as well, following a standard exchange-model with Cobb-Douglas utility function (as explained later), which bridge physics and economics together.

### 1. Exchange models without saving

In a simple version of KWEM considered in the works by Bennati (1988a,b, 1993) and also studied by Dragulescu and Yakovenko (2000) the money difference  $\Delta x$  in Eqs. (18) is assumed to have a constant value,  $\Delta x = \Delta x_0$ . Together with the constraint that transactions can take place only if  $x'_i > 0$  and  $x'_j > 0$ , this leads to an equilibrium exponential distribution, see the curve for  $\lambda = 0$  in Fig. 9.

Various other trading rules were studied by Dragulescu and Yakovenko (2000), choosing  $\Delta x$  as a random fraction of the average money between the two agents,  $\Delta x = \epsilon(x_i + x_j)/2$ , corresponding to a  $\Delta x = (1-\epsilon)x_i - \epsilon x_j$  in (18), or of the average money of the whole system,  $\Delta x = \epsilon \langle x \rangle$ .

The models mentioned, as well as more complicated ones (Dragulescu and Yakovenko (2000)), lead to an equilibrium wealth distribution with an exponential tail

$$f(x) \sim \beta \exp(-\beta x), \quad (19)$$

with the effective temperature  $1/\beta$  of the order of the average wealth,  $\beta^{-1} = \langle x \rangle$ . This result is largely independent of the details of the models, e.g. the multi-agent nature of the interaction, the initial conditions, and the

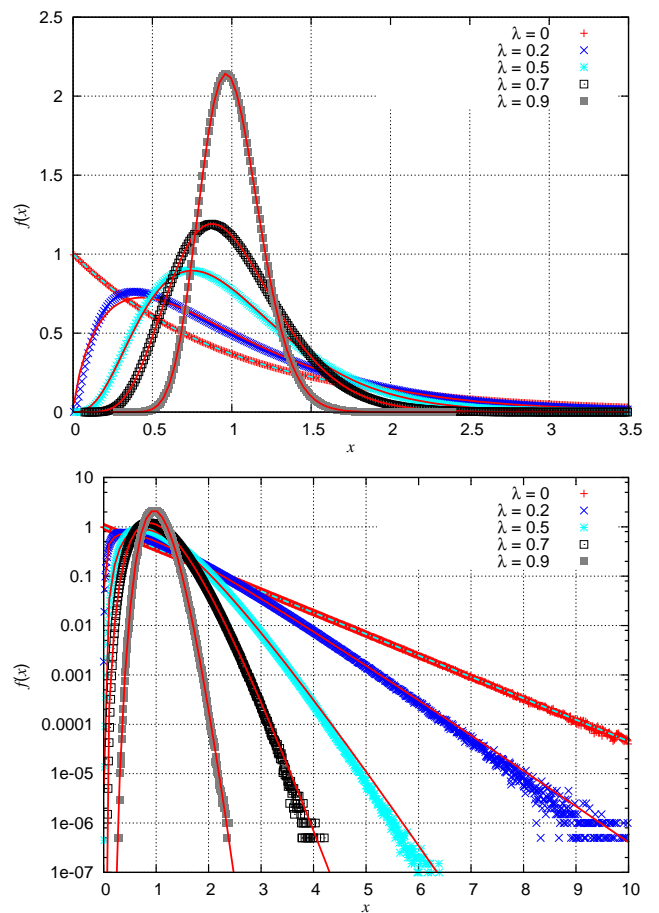


FIG. 9. Probability density for wealth  $x$ . The curve for  $\lambda = 0$  is the Boltzmann function  $f(x) = \langle x \rangle^{-1} \exp(-x/\langle x \rangle)$  for the basic model of Sec. IIIC 1. The other curves correspond to a global saving propensity  $\lambda > 0$ , see Sec. IIIC 2.

random or consecutive order of extraction of the interacting agents. The Boltzmann distribution is characterized by a majority of poor agents and a few rich agents (due to the exponential tail), and has a Gini coefficient of 0.5.

### 2. Exchange models with saving

As a generalization and more realistic version of the basic exchange models, a saving criterion can be introduced. Angle (1983), motivated by the surplus theory, introduced a unidirectional model of wealth exchange, in which only a fraction of wealth smaller than one can pass from one agent to the other, with a  $\Delta x = \epsilon x_i$  or  $(-\omega x_j)$ , where the direction of the flow is determined by the agent wealth (Angle (1983, 1986)). Later Angle introduced the One-Parameter Inequality Process (OPIP) where a constant fraction  $1-\omega$  is saved before the transaction (Angle (2002)) by the agent whose wealth decreases, defined by an exchanged wealth amount  $\Delta x = \omega x_i$  or  $-\omega x_j$ , again

with the direction of the transaction determined by the relative difference between the agent wealth.

A “saving parameter”  $0 < \lambda < 1$  representing the fraction of wealth *saved*, was introduced in the model by Chakraborti and Chakrabarti (2000). In this model (CC) wealth flows simultaneously toward and from each agent during a single transaction, the dynamics being defined by the equations

$$\begin{aligned} x'_i &= \lambda x_i + \epsilon(1 - \lambda)(x_i + x_j), \\ x'_j &= \lambda x_j + (1 - \epsilon)(1 - \lambda)(x_i + x_j), \end{aligned} \quad (20)$$

or, equivalently, by a  $\Delta x$  in (18) given by

$$\Delta x = (1 - \lambda)[(1 - \epsilon)x_i - \epsilon x_j]. \quad (21)$$

These models, apart from the OPIP model of Angle which has the remarkable property of leading to a power law in a suitable range of  $\omega$ , can be well fitted by a  $\Gamma$ -distribution. The  $\Gamma$ -distribution is characterized by a mode  $x_m > 0$ , in agreement with real data of wealth and income distributions (Dragulescu and Yakovenko (2001a); Ferrero (2004); Silva and Yakovenko (2005); Sala-i Martin and Mohapatra (2002); Sala-i Martin (2002); Aoyama *et al.* (2003)). Furthermore, the limit for small  $x$  is zero, i.e.  $P(x \rightarrow 0) \rightarrow 0$ , see the example in Fig. 9. In the particular case of the model by Chakraborti and Chakrabarti (2000), the explicit distribution is well fitted by

$$\begin{aligned} f(x) &= n\langle x \rangle^{-1} \gamma_n(nx/\langle x \rangle) \\ &= \frac{1}{\Gamma(n)} \frac{n}{\langle x \rangle} \left( \frac{nx}{\langle x \rangle} \right)^{n-1} \exp\left(-\frac{nx}{\langle x \rangle}\right), \end{aligned} \quad (22)$$

$$n(\lambda) \equiv \frac{D_\lambda}{2} = 1 + \frac{3\lambda}{1 - \lambda}. \quad (23)$$

where  $\gamma_n(\xi)$  is the standard  $\Gamma$ -distribution. This particular functional form has been conjectured on the base of the excellent fitting provided to numerical data (Angle (1983, 1986); Patriarca *et al.* (2004b,a); Patriarca and Heinsalu (2009)). For more information and a comparison of similar fittings for different models see Patriarca *et al.* (2010). Very recently, Lallouache *et al.* (2010) have shown using the distributional form of the equation and moment calculations that strictly speaking the Gamma distribution is not the solution of Eq. (20), confirming the earlier results of Repetowicz *et al.* (2005). However, the Gamma distribution is a very very good approximation.

The ubiquitous presence of  $\Gamma$ -functions in the solutions of kinetic models (see also below heterogeneous models) suggests a close analogy with kinetic theory of gases. In fact, interpreting  $D_\lambda = 2n$  as an effective dimension, the variable  $x$  as kinetic energy, and introducing the effective temperature  $\beta^{-1} \equiv T_\lambda = \langle x \rangle / 2D_\lambda$  according to the equipartition theorem, Eqs. (22) and (23) define the canonical distribution  $\beta\gamma_n(\beta x)$  for the kinetic energy of a gas in  $D_\lambda = 2n$  dimensions, see Patriarca *et al.* (2004a) for details. The analogy is illustrated in Table V and the dependences of  $D_\lambda = 2n$  and of  $\beta^{-1} = T_\lambda$  on the saving parameter  $\lambda$  are shown in Fig. 10.

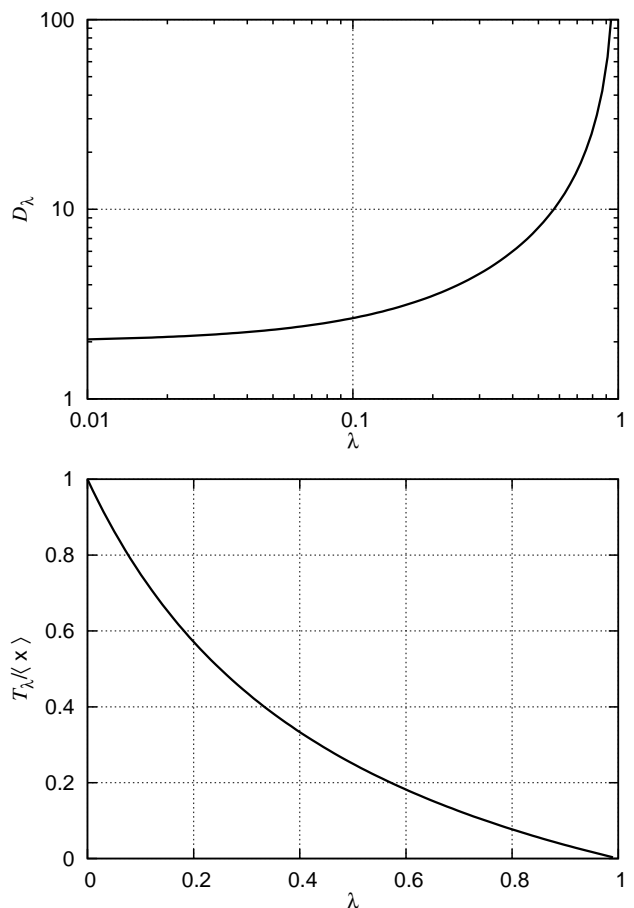


FIG. 10. Effective dimension  $D_\lambda$  and temperature  $T$  as a function of the saving parameter  $\lambda$ .

TABLE V. Analogy between kinetic the theory of gases and the kinetic exchange model of wealth

|                          | Kinetic model                    | Economy model                                |
|--------------------------|----------------------------------|--|
| variable                 | $K$ (kinetic energy)             | $x$ (wealth)                                 |
| units                    | $N$ particles                    | $N$ agents                                   |
| interaction              | collisions                       | trades                                       |
| dimension                | integer $D$                      | real number $D_\lambda$                      |
| temperature definition   | $k_B T = 2\langle K \rangle / D$ | $T_\lambda = 2\langle x \rangle / D_\lambda$ |
| reduced variable         | $\xi = K / k_B T$                | $\xi = x / T_\lambda$                        |
| equilibrium distribution | $f(\xi) = \gamma_{D/2}(\xi)$     | $f(\xi) = \gamma_{D_\lambda/2}(\xi)$         |

The exponential distribution is recovered as a special case, for  $n = 1$ . In the limit  $\lambda \rightarrow 1$ , i.e. for  $n \rightarrow \infty$ , the distribution  $f(x)$  above tends to a Dirac  $\delta$ -function, as shown in Patriarca *et al.* (2004a) and qualitatively illustrated by the curves in Fig. 9. This shows that a large saving criterion leads to a final state in which economic agents tend to have similar amounts of money and, in the limit of  $\lambda \rightarrow 1$ , exactly the same amount  $\langle x \rangle$ .

The equivalence between a kinetic wealth-exchange

model with saving propensity  $\lambda \geq 0$  and an  $N$ -particle system in a space with dimension  $D_\lambda \geq 2$  is suggested by simple considerations about the kinetics of collision processes between two molecules. In one dimension, particles undergo head-on collisions in which the whole amount of kinetic energy can be exchanged. In a larger number of dimensions the two particles will not travel in general exactly along the same line, in opposite verses, and only a fraction of the energy can be exchanged. It can be shown that during a binary elastic collision in  $D$  dimensions only a fraction  $1/D$  of the total kinetic energy is exchanged on average for kinematic reasons, see Chakraborti and Patriarca (2008) for details. The same  $1/D$  dependence is in fact obtained inverting Eq. (23), which provides for the fraction of exchanged wealth  $1 - \lambda = 6/(D_\lambda + 4)$ .

Not all homogeneous models lead to distributions with an exponential tail. For instance, in the model studied in Chakraborti (2002) an agent  $i$  can lose all his wealth, thus becoming unable to trade again: after a sufficient number of transactions, only one trader survives in the market and owns the entire wealth. The equilibrium distribution has a very different shape, as explained below:

In the toy model it is assumed that both the economic agents  $i$  and  $j$  invest the same amount  $x_{min}$ , which is taken as the minimum wealth between the two agents,  $x_{min} = \min\{x_i, x_j\}$ . The wealth after the trade are  $x'_i = x_i + \Delta x$  and  $x'_j = x_j - \Delta x$ , where  $\Delta x = (2\epsilon - 1)x_{min}$ . We note that once an agent has lost all his wealth, he is unable to trade because  $x_{min}$  has become zero. Thus, a trader is effectively driven out of the market once he loses all his wealth. In this way, after a sufficient number of transactions only one trader survives in the market with the entire amount of wealth, whereas the rest of the traders have zero wealth. In this toy model, only one agent has the entire money of the market and the rest of the traders have zero money, which corresponds to a distribution with Gini coefficient equal to unity.

Now, a situation is said to be Pareto-optimal “if by reallocation you cannot make someone better off without making someone else worse off”. In Pareto’s own words:

“We will say that the members of a collectivity enjoy maximum opheimity in a certain position when it is impossible to find a way of moving from that position very slightly in such a manner that the opheimity enjoyed by each of the individuals of that collectivity increases or decreases. That is to say, any small displacement in departing from that position necessarily has the effect of increasing the opheimity which certain individuals enjoy, and decreasing that which others enjoy, of being agreeable to some, and disagreeable to others.”

— Vilfredo Pareto, Manual of Political Economy (1906), p.261.

However, as Sen (1971) notes, *an economy can be Pareto-optimal, yet still “perfectly disgusting” by any ethi-*

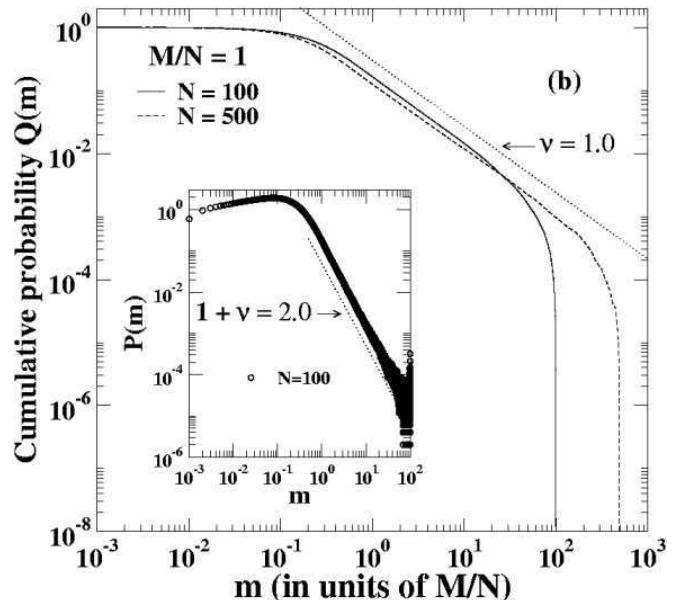


FIG. 11. Results for randomly assigned saving parameters. Reproduced and adapted from Chakraborti and Chatterjee (2003), available at [arXiv:cond-mat/0302147](https://arxiv.org/abs/cond-mat/0302147).

cal standards . It is important to note that Pareto-optimality, is merely a descriptive term, a property of an “allocation”, and there are no ethical propositions about the desirability of such allocations inherent within that notion. Thus, in other words there is nothing inherent in Pareto-optimality that implies the maximization of social welfare.

This simple toy model thus also produces a Pareto-optimal state (it will be impossible to raise the well-being of anyone except the *winner*, i.e., the agent with all the money, and vice versa ) but the situation is economically undesirable as far as social welfare is concerned!

Note also, as mentioned above, the OPIP model of Angle (2006, 2002), for example, depending on the model parameters, can also produce a power law tail. Another general way to produce a power law tail in the equilibrium distribution seems to diversify the agents, i.e. to consider heterogeneous models, discussed below.

#### D. Heterogeneous kinetic wealth exchange models

##### 1. Random saving propensities

The models considered above assume the all agents have the same statistical properties. The corresponding equilibrium wealth distribution has in most of the cases an exponential tail, a form which well interpolates real data at small and intermediate values of wealth. However, it is possible to conceive generalized models which lead to even more realistic equilibrium wealth distributions. This is the case when agents are diversified by



assigning different values of the saving parameter. For instance, Angle (2002) studied a model with a trading rule where diversified parameters  $\{\omega_i\}$  occur,

$$\Delta x = \omega_i \epsilon x_i \quad \text{or} \quad -\omega_j \epsilon x_j, \quad (24)$$

with the direction of wealth flow determined by the wealth of agents  $i$  and  $j$ . Diversified saving parameters were independently introduced by Chatterjee *et al.* (2003, 2004) by generalizing the model introduced in Chakraborti and Chakrabarti (2000):

$$\begin{aligned} x'_i &= \lambda_i x_i + \epsilon[(1 - \lambda_i)x_i + (1 - \lambda_j)x_j], \\ x'_j &= \lambda x_j + (1 - \epsilon)[(1 - \lambda_i)x_i + (1 - \lambda_j)x_j], \end{aligned} \quad (25)$$

corresponding to a

$$\Delta x = (1 - \epsilon)(1 - \lambda_i)x_i - \epsilon(1 - \lambda_j)x_j. \quad (26)$$

The surprising result is that if the parameters  $\{\lambda_i\}$  are suitably diversified, a power law appears in the equilibrium wealth distribution, see Fig. 11. In particular if the  $\lambda_i$  are uniformly distributed in  $(0, 1)$  the wealth distribution exhibits a robust power-law tail,

$$f(x) \propto x^{-\alpha-1}, \quad (27)$$

with the Pareto exponent  $\alpha = 1$  largely independent of the details of the  $\lambda$ -distribution. It may be noted that the exponent value unity is strictly for the tail end of the distribution and not for small values of the income or wealth (where the distribution remains exponential). Also, for finite number  $N$  of agents, there is always an exponential (in  $N$ ) cut off at the tail end of the distribution. This result is supported by independent theoretical considerations based on different approaches, such as a mean field theory approach (Mohanty (2006), see below for further details) or the Boltzmann equation (Das and Yarlagadda (2003, 2005); Repetowicz *et al.* (2005); Chatterjee *et al.* (2005a)). For derivation of the Pareto law from variational principles, using the KWEM context, see Chakraborti and Patriarca (2009).

## 2. Power-law distribution as an overlap of Gamma distributions

A remarkable feature of the equilibrium wealth distribution obtained from heterogeneous models, noticed in Chatterjee *et al.* (2004), is that the individual wealth distribution  $f_i(x)$  of the generic  $i$ -th agent with saving parameter  $\lambda_i$  has a well defined mode and exponential tail, in spite of the resulting power-law tail of the marginal distribution  $f(x) = \sum_i f_i(x)$ . In fact, Patriarca *et al.* (2005) found by numerical simulation that the marginal distribution  $f(x)$  can be resolved as an overlap of individual Gamma distributions with  $\lambda$ -dependent parameters; furthermore, the mode and the average value of the distributions  $f_i(x)$  both diverge for  $\lambda \rightarrow 1$  as  $\langle x(\lambda) \rangle \sim 1/(1-\lambda)$  (Chatterjee *et al.* (2004); Patriarca *et al.* (2005)). This

fact was justified theoretically by Mohanty (2006). Consider the evolution equations (25). In the mean field approximation one can consider that each agent  $i$  has an (average) wealth  $\langle x_i \rangle = y_i$  and replace the random number  $\epsilon$  with its average value  $\langle \epsilon \rangle = 1/2$ . Indicating with  $y_{ij}$  the new wealth of agent  $i$ , due to the interaction with agent  $j$ , from Eqs. (25) one obtains

$$y_{ij} = (1/2)(1 + \lambda_i)y_i + (1/2)(1 - \lambda_j)y_j. \quad (28)$$

At equilibrium, for consistency, average over all the interaction must give back  $y_i$ ,

$$y_i = \sum_j y_{ij}/N. \quad (29)$$

Then summing Eq. (28) over  $j$  and dividing by the number of agents  $N$ , one has

$$(1 - \lambda_i)y_i = \langle (1 - \lambda)y \rangle, \quad (30)$$

where  $\langle (1 - \lambda)y \rangle = \sum_j (1 - \lambda_j)y_j/N$ . Since the right hand side is independent of  $i$  and this relation holds for arbitrary distributions of  $\lambda_i$ , the solution is

$$y_i = \frac{C}{1 - \lambda_i}, \quad (31)$$

where  $C$  is a constant. Besides proving the dependence of  $y_i = \langle x_i \rangle$  on  $\lambda_i$ , this relation also demonstrates the existence of a power law tail in the equilibrium distribution. If, in the continuous limit,  $\lambda$  is distributed in  $(0, 1)$  with a density  $\phi(\lambda)$ , ( $0 \leq \lambda < 1$ ), then using (31) the (average) wealth distribution is given

$$f(y) = \phi(\lambda) \frac{d\lambda}{dy} = \phi(1 - C/x) \frac{C}{y^2}. \quad (32)$$

Figure 12 illustrates the phenomenon for a system of  $N = 1000$  agents with random saving propensities uniformly distributed between 0 and 1. The figure confirms the importance of agents with  $\lambda$  close to 1 for producing a power-law probability distribution (Chatterjee *et al.* (2004); Patriarca and Heinsalu (2009)).

However, when considering values of  $\lambda$  close enough to 1, the power law can break down at least for two reasons. The first one, illustrated in Fig. 12-bottom right, is that the power-law can be resolved into almost disjoint contributions representing the wealth distributions of single agents. This follows from the finite number of agents used and the fact that the distance between the average values of the distributions corresponding to two consecutive values of  $\lambda$  grows faster than the corresponding widths (Patriarca *et al.* (2005); Chatterjee *et al.* (2005b)). The second reason is due to the finite cutoff  $\lambda_M$ , always present in a numerical simulation. However, to study this effect, one has to consider a system with a number of agents large enough that it is not possible to resolve the wealth distributions of single agents for the sub-intervals of  $\lambda$  considered. This was done in by Patriarca *et al.* (2006) using a system with  $N = 10^5$  agents with saving parameters distributed uniformly between 0 and  $\lambda_M$ .

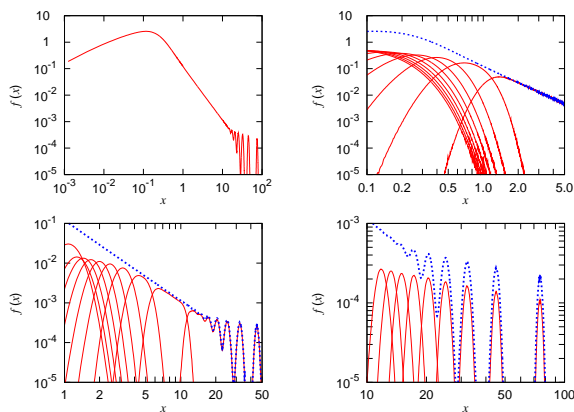


FIG. 12. Wealth distribution in a system of 1000 agents with saving propensities uniformly distributed in the interval  $0 < \lambda < 1$ . Top left: marginal distribution. Top right: marginal distribution (dotted line) and distributions of wealth of agents with  $\lambda \in (j\Delta\lambda, (j+1)\Delta\lambda)$ ,  $\Delta\lambda = 0.1$ ,  $j = 0, \dots, 9$  (continuous lines). Bottom-left: the distribution of wealth of agents with  $\lambda \in (0.9, 1)$  has been further resolved into contributions from subintervals  $\lambda \in (0.9 + j\Delta\lambda, 0.9 + (j+1)\Delta\lambda)$ ,  $\Delta\lambda = 0.01$ . Bottom-right: the partial distribution of wealth of agents with  $\lambda \in (0.99, 1)$  has been further resolved into those from subintervals  $\lambda \in (0.99 + j\Delta\lambda, 0.99 + (j+1)\Delta\lambda)$ ,  $\Delta\lambda = 0.001$ .

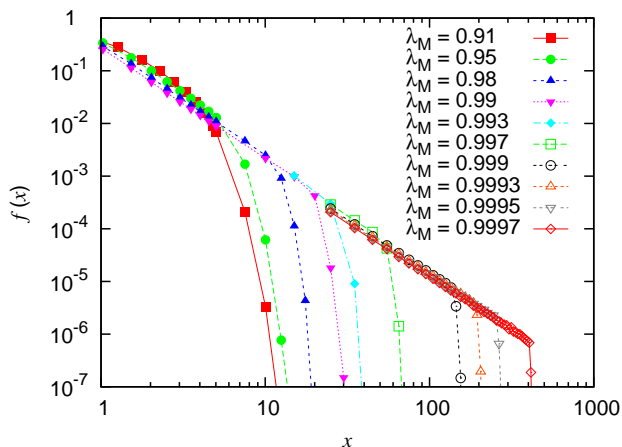


FIG. 13. Wealth distribution obtained for the uniform saving propensity distributions of  $10^5$  agents in the interval  $(0, \lambda_M)$ .

Results are shown in Fig. 13, in which curves from left to right correspond to increasing values of the cutoff  $\lambda_M$  from 0.9 to 0.9997. The transition from an exponential to a power-law tail takes place continuously as the cut-off  $\lambda_M$  is increased beyond a critical value  $\lambda_M \approx 0.9$  toward  $\lambda_M = 1$ , through the enlargement of the  $x$ -interval in which the power-law is observed.

### 3. Relaxation process

Relaxation in systems with constant  $\lambda$  had already been studied by Chakraborti and Chakrabarti (2000), where a systematic increase of the relaxation time with  $\lambda$ , and eventually a divergence for  $\lambda \rightarrow 1$ , was found. In fact, for  $\lambda = 1$  no exchanges occurs and the system is frozen. The relaxation time scale of a heterogeneous system had been studied by Patriarca *et al.* (2007). The system is observed to relax toward the same equilibrium wealth distribution from any given arbitrary initial distribution of wealth. If time is measured by the number of transactions  $n_t$ , the time scale is proportional to the number of agents  $N$ , i.e. defining time  $t$  as the ratio  $t = n_t/N$  between the number of trades and the total number of agents  $N$  (corresponding to one Monte Carlo cycle or one sweep in molecular dynamics simulations) the dynamics and the relaxation process become independent of  $N$ . The existence of a natural time scale independent of the system size provides a foundation for using simulations of systems with finite  $N$  in order to infer properties of systems with continuous saving propensity distributions and  $N \rightarrow \infty$ .

In a system with uniformly distributed  $\lambda$ , the wealth distributions of each agent  $i$  with saving parameter  $\lambda_i$  relaxes toward different states with characteristic shapes  $f_i(x)$  (Patriarca *et al.* (2005); Chatterjee *et al.* (2005b); Patriarca *et al.* (2006)) with different relaxation times  $\tau_i$  (Patriarca *et al.* (2007)). The differences in the relaxation process can be related to the different relative wealth exchange rates, that by direct inspection of the evolution equations appear to be proportional to  $1 - \lambda_i$ . Thus, in general, higher saving propensities are expected to be associated to slower relaxation processes with a relaxation time  $\propto 1/(1 - \lambda)$ .

It is also possible to obtain the relaxation time distribution. If the saving parameters are distributed in  $(0, 1)$  with a density  $\phi(\lambda)$ , it follows from probability conservation that  $\tilde{f}(\bar{x})d\bar{x} = \phi(\lambda)d\lambda$ , where  $\bar{x} \equiv \langle x \rangle_\lambda$  and  $\tilde{f}(\bar{x})$  the corresponding density of average wealth values. In the case of uniformly distributed saving propensities, one obtains

$$\tilde{f}(\bar{x}) = \phi(\lambda) \frac{d\lambda(\bar{x})}{d\bar{x}} = \phi\left(1 - \frac{k}{\bar{x}}\right) \frac{k}{\bar{x}^2}, \quad (33)$$

showing that a uniform saving propensity distribution leads to a power law  $\tilde{f}(\bar{x}) \sim 1/\bar{x}^2$  in the (average) wealth distribution. In a similar way it is possible to obtain the associated distribution of relaxation times  $\psi(\tau)$  for the global relaxation process from the relation  $\tau_i \propto 1/(1 - \lambda_i)$ ,

$$\psi(\tau) = \phi(\lambda) \frac{d\lambda(\tau)}{d\tau} \propto \phi\left(1 - \frac{\tau'}{\tau}\right) \frac{\tau'}{\tau^2}, \quad (34)$$

where  $\tau'$  is a proportionality factor. Therefore  $\psi(\tau)$  and  $\tilde{f}(\bar{x})$  are characterized by power law tails in  $\tau$  and  $\bar{x}$  respectively with the same Pareto exponent.

In conclusion, the role of the  $\lambda$ -cut-off is also related to the relaxation process. This means that the slowest convergence rate is determined by the cut-off and is  $\propto 1 - \lambda_M$ . In numerical simulations of heterogeneous KWEMs, as well as in real wealth distributions, the cut-off is necessarily finite, so that the convergence is fast (Gupta (2008)). On the other hand, if considering a hypothetical wealth distribution with a power law extending to infinite values of  $x$ , one cannot find a fast relaxation, due to the infinite time scale of the system, due to the agents with  $\lambda = 1$ .

### E. Microeconomic formulation of Kinetic theory models

Very recently, Chakrabarti and Chakrabarti (2009) have studied the framework based on microeconomic theory from which the kinetic theory market models could be addressed. They derived the moments of the model by Chakraborti and Chakrabarti (2000) and reproduced the exchange equations used in the model (with fixed savings parameter). In the framework considered, the utility function deals with the behaviour of the agents in an exchange economy.

They start by considering two exchange economy, where each agent produces a single perishable commodity. Each of these goods is different and money exists in the economy to simply facilitate transactions. Each of these agents are endowed with an initial amount of money  $M_1 = m_1(t)$  and  $M_2 = m_2(t)$ . Let agent 1 produce  $Q_1$  amount of commodity 1 only, and agent 2 produce  $Q_2$  amount of commodity 2 only. At each time step  $t$ , two agents meet randomly to carry out transactions according to their utility maximization principle.

The utility functions as defined as follows: For agent 1,  $U_1(x_1, x_2, m_1) = x_1^{\alpha_1} x_2^{\alpha_2} m_1^{\alpha_m}$  and for agent 2,  $U_2(y_1, y_2, m_2) = y_1^{\alpha_1} y_2^{\alpha_2} m_2^{\alpha_m}$  where the arguments in both of the utility functions are consumption of the first (i.e.  $x_1$  and  $y_1$ ) and second good (i.e.  $x_2$  and  $y_2$ ) and amount of money in their possession respectively. For simplicity, they assume that the utility functions are of the above Cobb-Douglas form with the sum of the powers normalized to 1 i.e.  $\alpha_1 + \alpha_2 + \alpha_m = 1$ .

Let the commodity prices to be determined in the market be denoted by  $p_1$  and  $p_2$ . Now, the budget constraints are as follows: For agent 1 the budget constraint is  $p_1 x_1 + p_2 x_2 + m_1 \leq M_1 + p_1 Q_1$  and similarly, for agent 2 the constraint is  $p_1 y_1 + p_2 y_2 + m_2 \leq M_2 + p_2 Q_2$ , which mean that the amount that agent 1 can spend for consuming  $x_1$  and  $x_2$  added to the amount of money that he holds after trading at time  $t + 1$  (i.e.  $m_1$ ) cannot exceed the amount of money that he has at time  $t$  (i.e.  $M_1$ ) added to what he earns by selling the good he produces (i.e.  $Q_1$ ), and the same is true for agent 2.

Then the basic idea is that both of the agents try to maximize their respective utility subject to their respective budget constraints and the *invisible hand* of the market that is the price mechanism works to clear the market for both goods (i.e. total demand equals total supply for

both goods at the equilibrium prices), which means that agent 1's problem is to maximize his utility subject to his budget constraint i.e. maximize  $U_1(x_1, x_2, m_1)$  subject to  $p_1 x_1 + p_2 x_2 + m_1 = M_1 + p_1 Q_1$ . Similarly for agent 2, the problem is to maximize  $U_2(y_1, y_2, m_2)$  subject to  $p_1 y_1 + p_2 y_2 + m_2 = M_2 + p_2 Q_2$ . Solving those two maximization exercises by Lagrange multiplier and applying the condition that the market remains in equilibrium, the competitive price vector  $(\hat{p}_1, \hat{p}_2)$  as  $\hat{p}_i = (\alpha_i / \alpha_m)(M_1 + M_2) / Q_i$  for  $i = 1, 2$  is found (Chakrabarti and Chakrabarti (2009)).

The outcomes of such a trading process are then:

1. At optimal prices  $(\hat{p}_1, \hat{p}_2)$ ,  $m_1(t) + m_2(t) = m_1(t + 1) + m_2(t + 1)$ , i.e., demand matches supply in all market at the market-determined price in equilibrium. Since money is also treated as a commodity in this framework, its demand (i.e. the total amount of money held by the two persons after trade) must be equal to what was supplied (i.e. the total amount of money held by them before trade).
2. If a restrictive assumption is made such that  $\alpha_1$  in the utility function can vary randomly over time with  $\alpha_m$  remaining constant. It readily follows that  $\alpha_2$  also varies randomly over time with the restriction that the sum of  $\alpha_1$  and  $\alpha_2$  is a constant  $(1 - \alpha_m)$ . Then in the money demand equations derived, if we suppose  $\alpha_m$  is  $\lambda$  and  $\alpha_1 / (\alpha_1 + \alpha_2)$  is  $\epsilon$ , it is found that money evolution equations become

$$m_1(t + 1) = \lambda m_1(t) + \epsilon(1 - \lambda)(m_1(t) + m_2(t))$$

$$m_2(t + 1) = \lambda m_2(t) + (1 - \epsilon)(1 - \lambda)(m_1(t) + m_2(t)).$$

For a fixed value of  $\lambda$ , if  $\alpha_1$  (or  $\alpha_2$ ) is a random variable with uniform distribution over the domain  $[0, 1 - \lambda]$ , then  $\epsilon$  is also uniformly distributed over the domain  $[0, 1]$ . This limit corresponds to the Chakraborti and Chakrabarti (2000) model, discussed earlier.

3. For the limiting value of  $\alpha_m$  in the utility function (i.e.  $\alpha_m \rightarrow 0$  which implies  $\lambda \rightarrow 0$ ), the money transfer equation describing the random sharing of money without saving is obtained, which was studied by Dragulescu and Yakovenko (2000) mentioned earlier.

This actually demonstrates the equivalence of the two maximizations principles of entropy (in physics) and utility (in economics), and is certainly noteworthy.

## IV. AGENT-BASED MODELLING BASED ON GAMES

### A. Minority Game models

#### 1. El Farol Bar Problem

Arthur (1994) introduced the ‘El Farol Bar’ problem as a paradigm of complex economic systems. In this problem, a population of agents have to decide whether to go to the bar opposite Santa Fe, every Thursday night. Due to a limited number of seats, the bar cannot entertain more than  $X\%$  of the population. If less than  $X\%$  of the population go to the bar, the time spent in the bar is considered to be satisfying and it is better to attend the bar rather than staying at home. But if more than  $X\%$  of the population go to the bar, then it is too crowded and people in the bar have an unsatisfying time. In this second case, staying at home is considered to be better choice than attending the bar. So, in order to optimise its own utility, each agent has to predict what everybody else will do.

In particular Arthur was also interested in agents who have bounds on “rationality”, i.e. agents who:

- do not have perfect information about their environment, in general they will only acquire information through interaction with the dynamically changing environment;
- do not have a perfect model of their environment;
- have limited computational power, so they can’t work out all the logical consequences of their knowledge;
- have other resource limitations (e.g. memory).

In order to take these limitations into account, each agent is randomly given a fixed menu of models potentially suitable to predict the number of people who will go the bar given past data (e.g. the same as two weeks ago, the average of the past few weeks, etc.). Each week, each agent evaluates these models against the past data. He chooses the one that was the best predictor on this data and then uses it to predict the number of people who will go to the bar this time. If this prediction is less than  $X$ , then the agent decides to go to the bar as well. If its prediction is more than  $X$ , the agent stays home. Thus, in order to make decisions on whether to attend the bar, all the individuals are equipped with certain number of “strategies”, which provide them the predictions of the attendance in the bar next week, based on the attendance in the past few weeks. As a result the number who go to the bar oscillates in an apparently random manner around the critical  $X\%$  mark.

This was one of the first models that led a way different from traditional economics.

### 2. Basic Minority game

The Minority Games (abbreviated MGs) (Challet *et al.* (2004)) refer to the multi-agent models of financial markets with the original formulation introduced by Challet and Zhang (1997), and all other variants (Coolen (2005); Lamper *et al.* (2002)), most of which share the principal features that the models are repeated games and agents are inductive in nature. The original formulation of the Minority Game by Challet and Zhang (1997) is sometimes referred as the “Original Minority Game” or the “Basic Minority Game”.

The basic minority game consists of  $N$  (odd natural number) agents, who choose between one of the two decisions at each round of the game, using their own simple inductive strategies. The two decisions could be, for example, “buying” or “selling” commodities/assets, denoted by 0 or 1, at a given time  $t$ . An agent wins the game if it is one of the members of the minority group, and thus at each round, the minority group of agents win the game and rewards are given to those strategies that predict the winning side. All the agents have access to finite amount of public information, which is a common bit-string “memory” of the  $M$  most recent outcomes, composed of the winning sides in the past few rounds. Thus the agents with finite memory are said to exhibit “bounded rationality” (Arthur (1994)).

Consider for example, memory  $M = 2$ ; then there are  $P = 2^M = 4$  possible “history” bit strings: 00, 01, 10 and 11. A “strategy” consists of a response, i.e., 0 or 1, to each possible history bit strings; therefore, there are  $G = 2^P = 2^{2^M} = 16$  possible strategies which constitute the “strategy space”. At the beginning of the game, each agent randomly picks  $k$  strategies, and after the game, assigns one “virtual” point to a strategy which would have predicted the correct outcome. The actual performance  $r$  of the player is measured by the number of times the player wins, and the strategy, using which the player wins, gets a “real” point. A record of the number of agents who have chosen a particular action, say, “selling” denoted by 1,  $A_1(t)$  as a function of time is kept (see Fig. 14). The fluctuations in the behaviour of  $A_1(t)$  actually indicate the system’s total utility. For example, we can have a situation where only one player is in the minority and all the other players lose. The other extreme case is when  $(N - 1)/2$  players are in the minority and  $(N + 1)/2$  players lose. The total utility of the system is obviously greater for the latter case and from this perspective, the latter situation is more desirable. Therefore, the system is more efficient when there are smaller fluctuations around the mean than when the fluctuations are larger.

As in the El Farol bar problem, unlike most traditional economics models which assume agents are “deductive” in nature, here too a “trial-and-error” *inductive* thinking approach is implicitly implemented in process of decision-making when agents make their choices in the games.



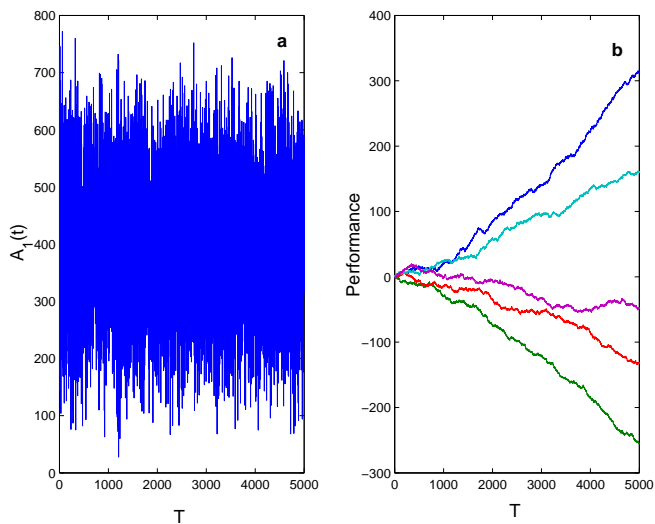


FIG. 14. Attendance fluctuation and performances of players in Basic Minority Game. Plots of (a) attendance and (b) performance of the players (five curves are: the best, the worst and three randomly chosen) for the basic minority game with  $N = 801$ ;  $M = 6$ ;  $k = 10$  and  $T = 5000$ . Reproduced from Sysi-Aho *et al.* (2003b).

### 3. Evolutionary minority games

Challet generalized the basic minority game (see Challet and Zhang (1997, 1998)) mentioned above to include the Darwinian selection: the worst player is replaced by a new one after some time steps, the new player is a “clone” of the best player, i.e. it inherits all the strategies but with corresponding virtual capitals reset to zero (analogous to a new born baby, though having all the predispositions from the parents, does not inherit their knowledge). To keep a certain diversity they introduced a mutation possibility in cloning. They allowed one of the strategies of the best player to be replaced by a new one. Since strategies are not just recycled among the players any more, the whole strategy phase space is available for selection. They expected this population to be capable of “learning” since bad players are weeded out with time, and fighting is among the so-to-speak the “best” players. Indeed in Fig. 15, they observed that the learning emerged in time. Fluctuations are reduced and saturated, this implies the average gain for everybody is improved but never reaches the ideal limit.

Li *et al.* (2000a,b) also studied the minority game in the presence of “evolution”. In particular, they examined the behaviour in games in which the dimension of the strategy space,  $m$ , is the same for all agents and fixed for all time. They found that for all values of  $m$ , not too large, evolution results in a substantial improvement in overall system performance. They also showed that after evolution, results obeyed a scaling relation among games played with different values of  $m$  and different numbers of

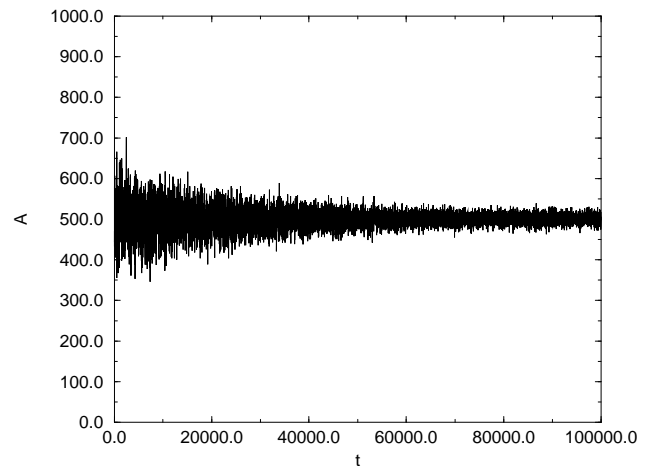


FIG. 15. Temporal attendance of  $A$  for the genetic approach showing a learning process. Reproduced from Challet and Zhang (1997)

agents, analogous to that found in the non-evolutionary, adaptive games (see remarks on section IV A 5). Best system performance still occurred, for a given number of agents, at  $m_c$ , the same value of the dimension of the strategy space as in the non-evolutionary case, but system performance was nearly an order of magnitude better than the non-evolutionary result. For  $m < m_c$ , the system evolved to states in which average agent wealth was better than in the random choice game. As  $m$  became large, overall systems performance approached that of the random choice game.

Li *et al.* (2000a,b) continued the study of evolution in minority games by examining games in which agents with poorly performing strategies can trade in their strategies for new ones from a different strategy space, which meant allowing for strategies that use information from different numbers of time lags,  $m$ . They found, in all the games, that after evolution, wealth per agent is high for agents with strategies drawn from small strategy spaces (small  $m$ ), and low for agents with strategies drawn from large strategy spaces (large  $m$ ). In the game played with  $N$  agents, wealth per agent as a function of  $m$  was very nearly a step function. The transition was found to be at  $m = m_t$ , where  $m_t \simeq m_c - 1$ , and  $m_c$  is the critical value of  $m$  at which  $N$  agents playing the game with a fixed strategy space (fixed  $m$ ) have the best emergent coordination and the best utilization of resources. They also found that overall system-wide utilization of resources is independent of  $N$ . Furthermore, although overall system-wide utilization of resources after evolution varied somewhat depending on some other aspects of the evolutionary dynamics, in the best cases, utilization of resources was on the order of the best results achieved in evolutionary games with fixed strategy spaces.

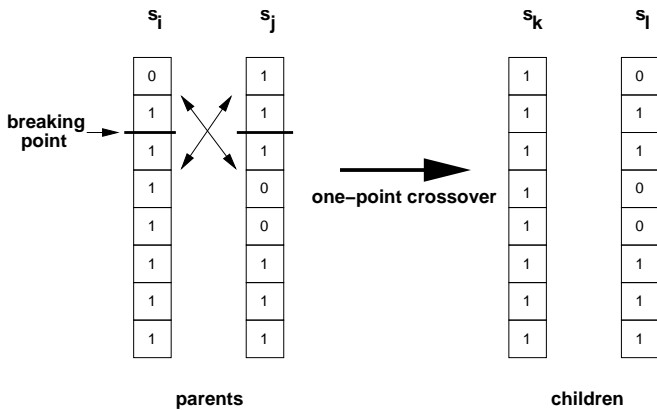


FIG. 16. Schematic diagram to illustrate the mechanism of one-point genetic crossover for producing new strategies. The strategies  $s_i$  and  $s_j$  are the parents. We choose the breaking point randomly and through this one-point genetic crossover, the children  $s_k$  and  $s_l$  are produced and substitute the parents. Reproduced from Sysi-Aho *et al.* (2003b).

#### 4. Adaptive minority games

Sysi-Aho *et al.* (2003a,c,b, 2004) presented a simple modification of the basic minority game where the players modify their strategies periodically after every time interval  $\tau$ , depending on their performances: if a player finds that he is among the fraction  $n$  (where  $0 < n < 1$ ) who are the worst performing players, he adapts himself and modifies his strategies. They proposed that the agents use hybridized one-point genetic crossover mechanism (as shown in Fig. 16), inspired by genetic evolution in biology, to modify the strategies and replace the bad strategies. They studied the performances of the agents under different conditions and investigate how they adapt themselves in order to survive or be the best, by finding new strategies using the highly effective mechanism. They also studied the measure of total utility of the system  $U(x_t)$ , which is the number of players in the minority group; the total utility of the system is maximum  $U_{max}$  as the highest number of players win is equal to  $(N-1)/2$ . The system is more efficient when the deviations from the maximum total utility  $U_{max}$  are smaller, or in other words, the fluctuations in  $A_1(t)$  around the mean become smaller.

Interestingly, the fluctuations disappear totally and the system stabilizes to a state where the total utility of the system is at maximum, since at each time step the highest number of players win the game (see Fig. 17). As expected, the behaviour depends on the parameter values for the system (see Sysi-Aho *et al.* (2003b, 2004)). They used the utility function to study the efficiency and dynamics of the game as shown in Fig. 18. If the parents are chosen randomly from the pool of strategies then the mechanism represents a “one-point genetic crossover” and if the parents are the best strategies then the mech-

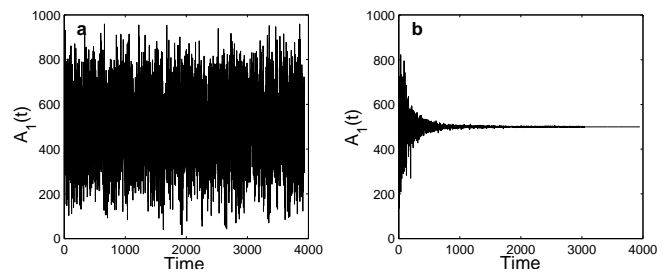


FIG. 17. Plot to show the time variations of the number of players  $A_1$  who choose action 1, with the parameters  $N = 1001$ ,  $m = 5$ ,  $s = 10$  and  $t = 4000$  for (a) basic minority game and (b) adaptive game, where  $\tau = 25$  and  $n = 0.6$ . Reproduced from Sysi-Aho *et al.* (2003b).

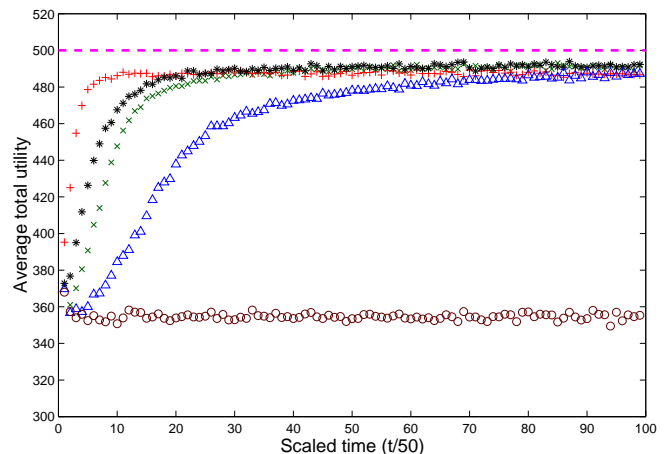


FIG. 18. Plot to show the variation of total utility of the system with time for the basic minority game for  $N = 1001$ ,  $m = 5$ ,  $s = 10$ ,  $t = 5000$ , and adaptive game, for the same parameters but different values of  $\tau$  and  $n$ . Each point represents a time average of the total utility for separate bins of size 50 time-steps of the game. The maximum total utility ( $= (N-1)/2$ ) is shown as a dashed line. The data for the basic minority game is shown in circles. The plus signs are for  $\tau = 10$  and  $n = 0.6$ ; the asterisk marks are for  $\tau = 50$  and  $n = 0.6$ ; the cross marks for  $\tau = 10$  and  $n = 0.2$  and triangles for  $\tau = 50$  and  $n = 0.2$ . The ensemble average over 70 different samples was taken in each case. Reproduced from Sysi-Aho *et al.* (2003b).

anism represents a “hybridized genetic crossover”. The children may replace parents or two worst strategies and accordingly four different interesting cases arise: (a) one-point genetic crossover with parents “killed”, i.e. parents are replaced by the children, (b) one-point genetic crossover with parents “saved”, i.e. the two worst strategies are replaced by the children but the parents are retained, (c) hybridized genetic crossover with parents “killed” and (d) hybridized genetic crossover with parents “saved”.

In order to determine which mechanism is the most

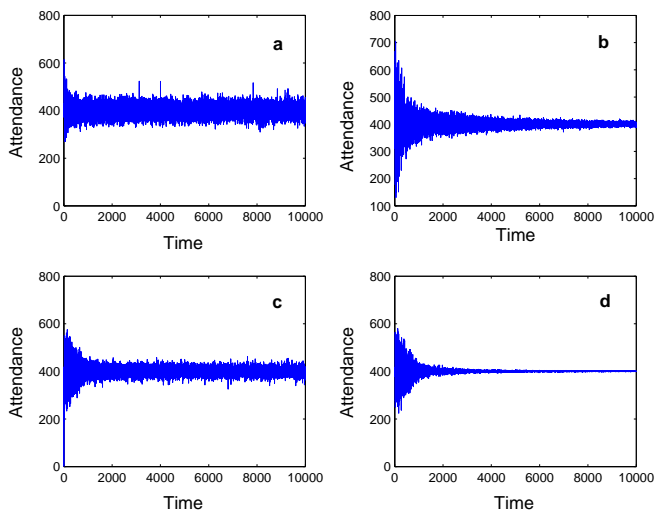


FIG. 19. Plots of the attendances by choosing parents randomly (a) and (b), and using the best parents in a player's pool (c) and (d). In (a) and (c) case parents are replaced by children and in (b) and (d) case children replace the two worst strategies. Simulations have been done with  $N = 801$ ,  $M = 6$ ,  $k = 16$ ,  $t = 40$ ,  $n = 0.4$  and  $T = 10000$ .

efficient, we have made a comparative study of the four cases, mentioned above. We plot the attendance as a function of time for the different mechanisms in Fig. 19. In Fig. 20 we show the total utility of the system in each of the cases (a)-(d), where we have plotted results of the average over 100 runs and each point in the utility curve represents a time average taken over a bin of length 50 time-steps. The simulation time is doubled from those in Fig. 19, in order to expose the asymptotic behaviour better. On the basis of Figs. 19 and 20, we find that the case (d) is the most efficient. In order to investigate what happens in the level of an individual agent, we created a competitive surrounding- "test" situation where after  $T = 3120$  time-steps, six players begin to adapt and modify their strategies such that three are using hybridized genetic crossover mechanism and the other three one point genetic crossover, where children replace the parents. The rest of the players play the basic minority game. In this case it turns out that in the end the best players are those who use the hybridized mechanism, second best are those using the one-point mechanism, and the bad players those who do not adapt at all. In addition it turns out that the competition amongst the players who adapt using the hybridized genetic crossover mechanism is severe.

It should be noted that the mechanism of evolution of strategies is considerably different from earlier attempts such as Challet and Zhang (1997) or Li *et al.* (2000a,b). This is because in this mechanism the strategies are changed by the agents themselves and even though the strategy space evolves continuously, its size and dimensionality remain the same.

Due to the simplicity of these models (Sysi-Aho *et al.*

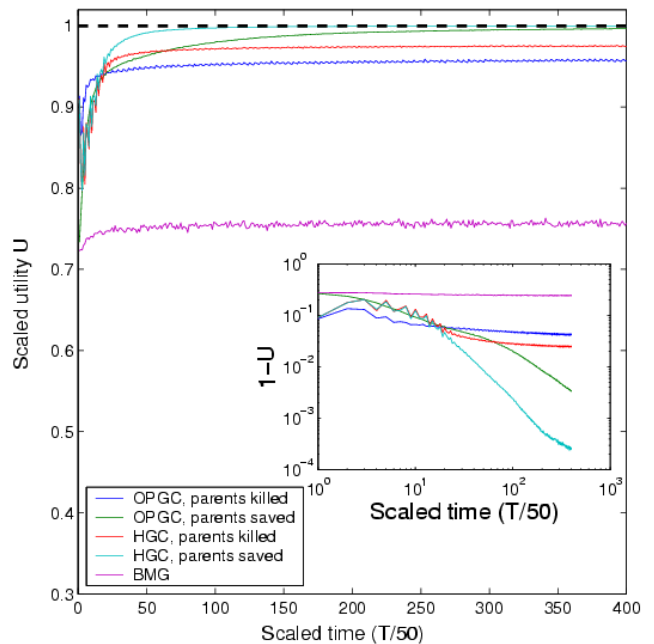


FIG. 20. Plots of the scaled utilities of the four different mechanisms in comparison with that of the basic minority game. Each curve represents an ensemble average over 100 runs and each point in a curve is a time average over a bin of length 50 time-steps. In the inset, the quantity  $(1 - U)$  is plotted against scaled time in the double logarithmic scale. Simulations are done with  $N = 801$ ,  $M = 6$ ,  $k = 16$ ,  $t = 40$ ,  $n = 0.4$  and  $T = 20000$ . Reproduced from Sysi-Aho *et al.* (2003b).

(2003a,c,b, 2004)), a lot of freedom is found in modifying the models to make the situations more realistic and applicable to many real dynamical systems, and not only financial markets. Many details in the model can be fine-tuned to imitate the real markets or behaviour of other complex systems. Many other sophisticated models based on these games can be setup and implemented, which show a great potential over the commonly adopted statistical techniques in analyses of financial markets.

## 5. Remarks

For modelling purposes, the minority game models were meant to serve as a class of simple models which could produce some macroscopic features observed in the real financial markets, which included the fat-tail price return distribution and volatility clustering (Challet *et al.* (2004); Coolen (2005)). Despite the hectic activity (Challet and Zhang (1998); Challet *et al.* (2000)) they have failed to capture or reproduce most important stylized facts of the real markets. However, in the physicists' community, they have become an interesting and established class of models where the physics of disordered systems (Cavagna *et al.* (1999); Challet *et al.* (2000)), lending a large amount of physical insights (Savit

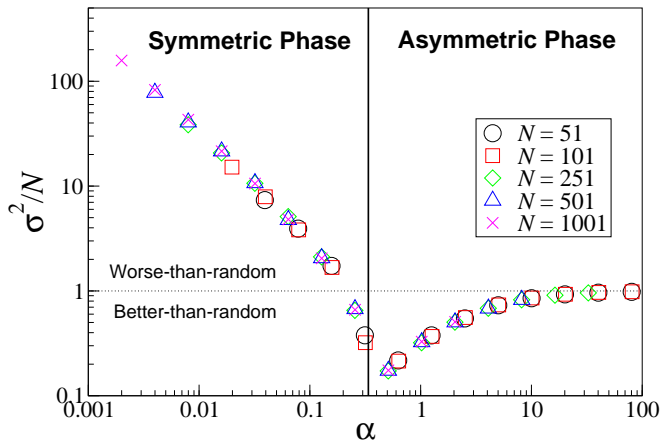


FIG. 21. The simulation results of the variance in attendance  $\sigma^2/N$  as a function of the control parameter  $\alpha = 2^M/N$  for games with  $k = 2$  strategies for each agent, ensemble averaged over 100 sample runs. Dotted line shows the value of volatility in random choice limit. Solid line shows the critical value of  $\alpha = \alpha_c \approx 0.3374$ . Reproduced from Yeung and Zhang [arxiv:0811.1479](https://arxiv.org/abs/0811.1479).

*et al.* (1999); Martino *et al.* (2004)). Since in the BMG model a Hamiltonian function could be defined and analytic solutions could be developed in some regimes of the model, the model was viewed with a more physical picture. In fact, it is characterized by a clear two-phase structure with very different collective behaviours in the two phases, as in many known conventional physical systems (Savit *et al.* (1999); Cavagna *et al.* (1999)).

Savit *et al.* (1999) first found that the macroscopic behaviour of the system does not depend independently on the parameters  $N$  and  $M$ , but instead depends on the ratio

$$\alpha \equiv \frac{2^M}{N} = \frac{P}{N} \quad (35)$$

which serves as the most important control parameter in the game. The variance in the attendance (see also Sysi-Aho *et al.* (2003c)) or volatility  $\sigma^2/N$ , for different values of  $N$  and  $M$  depend only on the ratio  $\alpha$ . Fig. 21 shows a plot of  $\sigma^2/N$  against the control parameter  $\alpha$ , where the data collapse of  $\sigma^2/N$  for different values of  $N$  and  $M$  is clearly evident. The dotted line in Fig. 21 corresponds to the “coin-toss” limit (random choice or pure chance limit), in which agents play by simply making random decisions (by coin-tossing) at every rounds of the game. This value of  $\sigma^2/N$  in coin-toss limit can be obtained by simply assuming a binomial distribution of the agents’ binary actions, with probability 0.5, such that  $\sigma^2/N = 0.5(1 - 0.5) \cdot 4 = 1$ . When  $\alpha$  is small, the value of  $\sigma^2/N$  of the game is larger than the coin-toss limit which implies the collective behaviours of agents are worse than the random choices. In the early literature, it was popularly called as the *worse-than-random* regime. When  $\alpha$  increases, the value of  $\sigma^2/N$  decreases and enters a region where agents are performing better than

the random choices, which was popularly called as the *better-than-random* regime. The value of  $\sigma^2/N$  reaches a minimum value which is substantially smaller than the coin-toss limit. When  $\alpha$  further increases, the value of  $\sigma^2/N$  increases again and approaches the coin-toss limit. This allowed one to identify two phases in the Minority Game, as separated by the minimum value of  $\sigma^2/N$  in the graph. The value of  $\alpha$  where the rescaled volatility attended its minimum was denoted by  $\alpha_c$ , which represented the phase transition point;  $\alpha_c$  has been shown to have a value of  $0.3374 \dots$  (for  $k = 2$ ) by analytical calculations Challet *et al.* (2000).

Besides these collective behaviours, physicists became also interested in the dynamics of the games such as crowd vs anti-crowd movement of agents, periodic attractors, etc. (Johnson *et al.* (1999a,b); Hart *et al.* (2001)). In this way, the Minority Games serve as a useful tool and provide a new direction for physicists in viewing and analysing the underlying dynamics of complex evolving systems such as the financial markets.

## B. The Kolkata Paise Restaurant (KPR) problem

The KPR problem (Chakrabarti *et al.* (2009); Ghosh and Chakrabarti (2009); Ghosh *et al.* (2010a,b)) is a repeated game, played between a large number  $N$  of agents having no interaction amongst themselves. In KPR problem, prospective customers (agents) choose from  $N$  restaurants each evening simultaneously (in parallel decision mode);  $N$  is fixed. Each restaurant has the same price for a meal but a different rank (agreed upon by all customers) and can serve only one customer any evening. Information regarding the customer distributions for earlier evenings is available to everyone. Each customer’s objective is to go to the restaurant with the highest possible rank while avoiding the crowd so as to be able to get dinner there. If more than one customer arrives at any restaurant on any evening, one of them is randomly chosen (each of them are anonymously treated) and is served. The rest do not get dinner that evening.

In Kolkata, there were very cheap and fixed rate “Paise Restaurants” that were popular among the daily labourers in the city. During lunch hours, the labourers used to walk (to save the transport costs) to one of these restaurants and would miss lunch if they got to a restaurant where there were too many customers. Walking down to the next restaurant would mean failing to report back to work on time! Paise is the smallest Indian coin and there were indeed some well-known rankings of these restaurants, as some of them would offer tastier items compared to the others. A more general example of such a problem would be when the society provides hospitals (and beds) in every locality but the local patients go to hospitals of better rank (commonly perceived) elsewhere, thereby competing with the local patients of those hospitals. Unavailability of treatment in time may be considered as lack of the service for those people and consequently as



(social) wastage of service by those unattended hospitals.

A dictator's solution to the KPR problem is the following: the dictator asks everyone to form a queue and then assigns each one a restaurant with rank matching the sequence of the person in the queue on the first evening. Then each person is told to go to the next ranked restaurant in the following evening (for the person in the last ranked restaurant this means going to the first ranked restaurant). This shift proceeds then continuously for successive evenings. This is clearly one of the most efficient solution (with utilization fraction  $\bar{f}$  of the services by the restaurants equal to unity) and the system arrives at this this solution immediately (from the first evening itself). However, in reality this cannot be the true solution of the KPR problem, where each agent decides on his own (in parallel or democratically) every evening, based on complete information about past events. In this game, the customers try to evolve a learning strategy to eventually get dinners at the best possible ranked restaurant, avoiding the crowd. It is seen, the evolution these strategies take considerable time to converge and even then the eventual utilization fraction  $\bar{f}$  is far below unity.

Let the symmetric stochastic strategy chosen by each agent be such that at any time  $t$ , the probability  $p_k(t)$  to arrive at the  $k$ -th ranked restaurant is given by

$$p_k(t) = \frac{1}{z} \left[ k^\alpha \exp \left( -\frac{n_k(t-1)}{T} \right) \right],$$

$$z = \sum_{k=1}^N \left[ k^\alpha \exp \left( -\frac{n_k(t-1)}{T} \right) \right], \quad (36)$$

where  $n_k(t)$  denotes the number of agents arriving at the  $k$ -th ranked restaurant in period  $t$ ,  $T > 0$  is a scaling factor and  $\alpha \geq 0$  is an exponent.

For any natural number  $\alpha$  and  $T \rightarrow \infty$ , an agent goes to the  $k$ -th ranked restaurant with probability  $p_k(t) = k^\alpha / \sum k^\alpha$ ; which means in the limit  $T \rightarrow \infty$  in (36) gives  $p_k(t) = k^\alpha / \sum k^\alpha$ .

If an agent selects any restaurant with equal probability  $p$  then probability that a single restaurant is chosen by  $m$  agents is given by

$$\Delta(m) = \binom{N}{m} p^m (1-p)^{N-m}. \quad (37)$$

Therefore, the probability that a restaurant with rank  $k$  is not chosen by any of the agents will be given by

$$\Delta_k(m=0) = \binom{N}{0} (1-p_k)^N; \quad p_k = \frac{k^\alpha}{\sum k^\alpha}$$

$$\simeq \exp \left( -\frac{k^\alpha N}{\tilde{N}} \right) \quad \text{as } N \rightarrow \infty, \quad (38)$$

where  $\tilde{N} = \sum_{k=1}^N k^\alpha \simeq \int_0^N k^\alpha dk = \frac{N^{\alpha+1}}{(\alpha+1)}$ . Hence

$$\Delta_k(m=0) = \exp \left( -\frac{k^\alpha (\alpha+1)}{N^\alpha} \right). \quad (39)$$

Therefore the average fraction of agents getting dinner in the  $k$ -th ranked restaurant is given by

$$\bar{f}_k = 1 - \Delta_k(m=0). \quad (40)$$

Naturally for  $\alpha = 0$ , the problem corresponding to random choice  $\bar{f}_k = 1 - e^{-1}$ , giving  $\bar{f} = \sum \bar{f}_k / N \simeq 0.63$  and for  $\alpha = 1$ ,  $\bar{f}_k = 1 - e^{-2k/N}$  giving  $\bar{f} = \sum \bar{f}_k / N \simeq 0.58$ .

In summary, in the KPR problem where the decision made by each agent in each evening  $t$  is independent and is based on the information about the rank  $k$  of the restaurants and their occupancy given by the numbers  $n_k(t-1) \dots n_k(0)$ . For several stochastic strategies, only  $n_k(t-1)$  is utilized and each agent chooses the  $k$ -th ranked restaurant with probability  $p_k(t)$  given by Eq. (36). The utilization fraction  $f_k$  of the  $k$ -th ranked restaurants on every evening is studied and their average (over  $k$ ) distributions  $D(f)$  are studied numerically, as well as analytically, and one finds (Chakrabarti *et al.* (2009); Ghosh and Chakrabarti (2009); Ghosh *et al.* (2010a)) their distributions to be Gaussian with the most probable utilization fraction  $\bar{f} \simeq 0.63, 0.58$  and  $0.46$  for the cases with  $\alpha = 0, T \rightarrow \infty$ ;  $\alpha = 1, T \rightarrow \infty$ ; and  $\alpha = 0, T \rightarrow 0$  respectively. For the stochastic crowd-avoiding strategy discussed in Ghosh *et al.* (2010b), where  $p_k(t+1) = \frac{1}{n_k(t)}$  for  $k = k_0$  the restaurant visited by the agent last evening, and  $= 1/(N-1)$  for all other restaurants ( $k \neq k_0$ ), one gets the best utilization fraction  $\bar{f} \simeq 0.8$ , and the analytical estimates for  $\bar{f}$  in these limits agree very well with the numerical observations. Also, the time required to converge to the above value of  $\bar{f}$  is independent of  $N$ .

The KPR problem has similarity with the Minority Game Problem (Arthur (1994); Challet *et al.* (2004)) as in both the games, herding behaviour is punished and diversity's encouraged. Also, both involves learning of the agents from the past successes etc. Of course, KPR has some simple exact solution limits, a few of which are discussed here. The real challenge is, of course, to design algorithms of learning mixed strategies (e.g., from the pool discussed here) by the agents so that the fair social norm emerges eventually (in  $N^0$  or  $\ln N$  order time) even when every one decides on the basis of their own information independently. As we have seen, some naive strategies give better values of  $\bar{f}$  compared to most of the "smarter" strategies like strict crowd-avoiding strategies, etc. This observation in fact compares well with earlier observation in minority games (see e.g. Satinover and Sornette (2007)).

It may be noted that all the stochastic strategies, being parallel in computational mode, have the advantage that they converge to solution at smaller time steps ( $\sim N^0$  or  $\ln N$ ) while for deterministic strategies the convergence time is typically of order of  $N$ , which renders such strategies useless in the truly macroscopic ( $N \rightarrow \infty$ ) limits. However, deterministic strategies are useful when  $N$  is small and rational agents can design appropriate punishment schemes for the deviators (see Kandori (2008)).

The study of the KPR problem shows that while a dictateded solution leads to one of the best possible solution to the problem, with each agent getting his dinner at the best ranked restaurant with a period of  $N$  evenings, and with best possible value of  $\bar{f}$  ( $= 1$ ) starting from the first evening itself. The parallel decision strategies (employing evolving algorithms by the agents, and past informations, e.g., of  $n(t)$ ), which are necessarily parallel among the agents and stochastic (as in democracy), are less efficient ( $\bar{f} \ll 1$ ; the best one discussed in Ghosh *et al.* (2010b), giving  $\bar{f} \simeq 0.8$  only). Note here that the time required is not dependent on  $N$ . We also note that most of the “smarter” strategies lead to much lower efficiency.

## V. CONCLUSIONS AND OUTLOOK

Agent-based models of order books are a good example of interactions between ideas and methods that are usually linked either to Economics and Finance (microstructure of markets, agent interaction) or to Physics (reaction-diffusion processes, deposition-evaporation process, kinetic theory of gases). As of today, existing models exhibit a trade-off between “realism” and calibration in its mechanisms and processes (empirical models such as Mike and Farmer (2008)), and explanatory power of simple observed behaviours (Cont and Bouchaud (2000); Cont (2007) for example). In the first case, some of the “stylized facts” may be reproduced, but using empirical processes that may not be linked to any behaviour observed on the market. In the second case, these are only toy models that cannot be calibrated on data. The mixing of many features, as in Lux and Marchesi (2000) and as is usually the case in behavioural finance, leads to poorly tractable models where the sensitivity to one parameter is hardly understandable. Therefore, no empirical model can tackle properly empirical facts such as volatility clustering. Importing toy model features explaining volatility clustering or market interactions in order book models is yet to be done. Finally, let us also note that to our knowledge, no agent-based model of order books deals with the multidimensional case. Implementing agents trading simultaneously several assets in a way that reproduces empirical observations on correlation and dependence remains an open challenge.

We believe this type of modelling is crucial for future developments in finance. The financial crisis that occurred in 2007-2008 is expected to create a shock in classic modelling in Economics and Finance. Many scientists have expressed their views on this subject (e.g. Bouchaud (2008); Lux and Westerhoff (2009); Farmer and Foley (2009)) and we believe as well that agent-based models we have presented here will be at the core of future modelling. As illustrations, let us mention Iori *et al.* (2006), which models the interbank market and investigates systemic risk, Thurner *et al.* (2009), which investigates the effects of use of leverage and margin calls on the stability of a market and Yakovenko and Rosser (2009), which

provides a brief overview of the study of wealth distributions and inequalities. No doubt these will be followed by many other contributions.

## ACKNOWLEDGEMENTS

The authors would like to thank their collaborators and two anonymous reviewers whose comments greatly helped improving the review. AC is grateful to B.K. Chakrabarti, K. Kaski, J. Kertesz, T. Lux, M. Marsili, D. Stauffer and V. Yakovenko for invaluable suggestions and criticisms.

- Alfi, V., Cristelli, M., Pietronero, L. and Zaccaria, A., Minimal agent based model for financial markets I. *The European Physical Journal B - Condensed Matter and Complex Systems*, 2009a, **67**, 385–397.
- Alfi, V., Cristelli, M., Pietronero, L. and Zaccaria, A., Minimal agent based model for financial markets II. *The European Physical Journal B - Condensed Matter and Complex Systems*, 2009b, **67**, 399–417.
- Angle, J., The surplus theory of social stratification and the size distribution of personal wealth. In *Proceedings of the Proceedings of the American Social Statistical Association, Social Statistics Section*, p. 395, 1983 (Alexandria, VA).
- Angle, J., The surplus theory of social stratification and the size distribution of personal wealth. *Social Forces*, 1986, **65**, 293.
- Angle, J., The Statistical signature of pervasive competition on wage and salary incomes. *J. Math. Sociol.*, 2002, **26**, 217.
- Angle, J., The Inequality Process as a wealth maximizing process. *Physica A*, 2006, **367**, 388.
- Aoyama, H., Souma, W. and Fujiwara, Y., Growth and fluctuations of personal and company’s income. *Physica A*, 2003, **324**, 352.
- Arthur, W.B., Inductive Reasoning and Bounded Rationality. *The American Economic Review*, 1994, **84**, 406–411.
- Ausloos, M. and Pekalski, A., Model of wealth and goods dynamics in a closed market. *Physica A*, 2007, **373**, 560.
- Bak, P., Paczuski, M. and Shubik, M., Price variations in a stock market with many agents. *Physica A: Statistical and Theoretical Physics*, 1997, **246**, 430 – 453.
- Benhabib, J. and Bisin, A., The distribution of wealth and fiscal policy in economies with finitely lived agents. *National Bureau of Economic Research Working Paper Series*, 2009, **No. 14730**.
- Bennati, E., *La simulazione statistica nell’analisi della distribuzione del reddito: modelli realistici e metodo di Monte Carlo* 1988a, ETS Editrice, Pisa.
- Bennati, E., Un metodo di simulazione statistica nell’analisi della distribuzione del reddito. *Rivista Internazionale di Scienze Economiche e Commerciali*, 1988b, **35**, 735–756.
- Bennati, E., Il metodo Monte Carlo nell’analisi economica. In *Proceedings of the Rassegna di lavori dell’ISCO X*, Vol. 31, 1993.
- Biais, B., Hillion, P. and Spatt, C., An empirical analysis of the limit order book and the order flow in the Paris Bourse. *Journal of Finance*, 1995, pp. 1655–1689.
- Bouchaud, J.P. and Mezard, M., Wealth condensation in a simple model of economy. *Physica A*, 2000, **282**, 536.
- Bouchaud, J.P., Mézard, M. and Potters, M., Statistical properties of stock order books: empirical results and models. *Quantitative Finance*, 2002, **2**, 251.
- Bouchaud, J., Farmer, J.D., Lillo, F., Hens, T. and Schenk-Hopp, K.R., How Markets Slowly Digest Changes in Supply and Demand. In *Handbook of Financial Markets: Dynamics and Evolution*, pp. 57–160, 2009 (North-Holland: San Diego).
- Bouchaud, J., Economics needs a scientific revolution. *Nature*, 2008, **455**, 1181.
- Burda, Z., Jurkiewicz, J. and Nowak, M.A., Is Econophysics a solid science?. *Acta Physica Polonica B*, 2003, **34**, 87.

- Case, K.E. and Fair, R.C., *Principles of Economics*, International edition 2008, Pearson Education Limited.
- Cavagna, A., Garrahan, J.P., Giardinà, I. and Sherrington, D., Thermal Model for Adaptive Competition in a Market. *Phys. Rev. Lett.*, 1999, **21**, 4429.
- Chakrabarti, A.S. and Chakrabarti, B.K., Microeconomics of the ideal gas like market models. *Physica A: Statistical Mechanics and its Applications*, 2009, **388**, 4151 – 4158.
- Chakrabarti, A.S., Chakrabarti, B.K., Chatterjee, A. and Mitra, M., The Kolkata Paise Restaurant problem and resource utilization. *Physica A: Statistical Mechanics and its Applications*, 2009, **388**, 2420–2426.
- Chakrabarti, B.K. and Chatterjee, A., Ideal Gas-Like Distributions in Economics: Effects of Saving Propensity. In *Proceedings of the Application of Econophysics*, edited by H. Takayasu, 2003 (Springer-Verlag: Tokyo).
- Chakraborti, A., Distributions of money in model markets of economy. *International Journal of Modern Physics C-Physics and Computer*, 2002, **13**, 1315–1322.
- Chakraborti, A. and Chakrabarti, B., Statistical mechanics of money: how saving propensity affects its distribution. *The European Physical Journal B - Condensed Matter and Complex Systems*, 2000, **17**, 167–170.
- Chakraborti, A. and Patriarca, M., Gamma-distribution and Wealth inequality. *Pramana J. Phys.*, 2008, **71**, 233.
- Chakraborti, A. and Patriarca, M., Variational Principle for the Pareto Power Law. *Physical Review Letters*, 2009, **103**, 228701.
- Challet, D., Marsili, M. and Zecchina, R., Statistical Mechanics of Systems with Heterogeneous Agents: Minority Games. *Phys. Rev. Lett.*, 2000, **84**, 1824.
- Challet, D. and Stinchcombe, R., Analyzing and modeling 1+ 1d markets. *Physica A: Statistical Mechanics and its Applications*, 2001, **300**, 285–299.
- Challet, D. and Zhang, Y.C., Emergence of cooperation and organization in an evolutionary game. *Physica A: Statistical and Theoretical Physics*, 1997, **246**, 407–418.
- Challet, D., Marsili, M. and Zhang, Y., *Minority Games* 2004, Oxford University Press.
- Challet, D. and Zhang, Y.C., On the minority game: Analytical and numerical studies. *Physica A*, 1998, **256**, 514.
- Champernowne, D.G., A Model of Income Distribution. *The Economic Journal*, 1953, **63**, 318–351.
- Chatterjee, A. and Chakrabarti, B.K., Kinetic exchange models for income and wealth distributions. *Eur. Phys. J. B*, 2007, **60**, 135.
- Chatterjee, A., Chakrabarti, B.K. and Stinchcombe, R.B., Master equation for a kinetic model of trading market and its analytic solution. *Phys. Rev. E*, 2005a, **72**, 026126.
- Chatterjee, A., Yarlagadda, S. and Chakrabarti, B.K. (Eds) *Econophysics of Wealth Distributions* 2005b (Springer-Verlag: Milan).
- Chatterjee, A., Chakrabarti, B.K. and Manna, S.S., Money in Gas-Like Markets: Gibbs and Pareto Laws. *Physica Scripta*, 2003, **T106**, 36–38.
- Chatterjee, A., Chakrabarti, B.K. and Manna, S.S., Pareto law in a kinetic model of market with random saving propensity. *Physica A: Statistical Mechanics and its Applications*, 2004, **335**, 155–163.
- Chiarella, C. and Iori, G., A simulation analysis of the microstructure of double auction markets. *Quantitative Finance*, 2002, **2**, 346–353.
- Cliff, D. and Bruten, J., Zero is not enough: On the lower limit of agent intelligence for continuous double auction markets. 1997, Technical report HPL-97-141, Hewlett-Packard Laboratories, Bristol, UK.
- Cohen, M., Report of the Special Study of the Securities Markets of the Securities and Exchanges Commission. 1963a, Technical report U.S. Congress, 88th Cong., 1st sess., H. Document 95.
- Cohen, M.H., Reflections on the Special Study of the Securities Markets. 1963b, Technical report Speech at the Practising Law Institute, New York.
- Cont, R. and Bouchaud, J.P., Herd behavior and aggregate fluctuations in financial markets. *Macroeconomic dynamics*, 2000, **4**, 170–196.
- Cont, R., Volatility Clustering in Financial Markets: Empirical Facts and Agent-Based Models. In *Long Memory in Economics*, pp. 289–309, 2007.
- Cont, R., Stoikov, S. and Talreja, R., A Stochastic Model for Order Book Dynamics. *SSRN eLibrary*, 2008.
- Coolen, A., *The Mathematical Theory Of Minority Games: Statistical Mechanics Of Interacting Agents* 2005, Oxford University Press.
- Cordier, S., Pareschi, L. and Toscani, G., On a Kinetic Model for a Simple Market Economy. *J. Stat. Phys.*, 2005, **120**, 253.
- Das, A. and Yarlagadda, S., A distribution function analysis of wealth distribution. [arxiv.org/cond-mat/0310343](http://arxiv.org/cond-mat/0310343), 2003.
- Das, A. and Yarlagadda, S., An analytic treatment of the Gibbs–Pareto behavior in wealth distribution. *Physica A*, 2005, **353**, 529.
- Dragulescu, A. and Yakovenko, V., Statistical mechanics of money. *The European Physical Journal B*, 2000, **17**, 7 pages.
- Dragulescu, A. and Yakovenko, V., Evidence for the exponential distribution of income in the USA. *The European Physical Journal B*, 2001a, **20**, 5 pages.
- Dragulescu, A. and Yakovenko, V.M., Exponential and power-law probability distributions of wealth and income in the United Kingdom and the United States. *Physica A: Statistical Mechanics and its Applications*, 2001b, **299**, 213–221.
- Düring, B., Matthes, D. and Toscani, G., Kinetic equations modelling wealth redistribution: A comparison of approaches. *Phys. Rev. E*, 2008, **78**, 056103.
- Düring, B. and Toscani, G., Hydrodynamics from kinetic models of conservative economies. *Physica A*, 2007, **384**, 493.
- Erdos, P. and Renyi, A., On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, 1960, **5**, 17–61.
- Farmer, J.D. and Foley, D., The economy needs agent-based modelling. *Nature*, 2009, **460**, 685–686.
- Ferrero, J.C., The statistical distribution of money and the rate of money transference. *Physica A*, 2004, **341**, 575.
- Fujiwara, Y., Souma, W., Aoyama, H., Kaizoji, T. and Aoki, M., Growth and fluctuations of personal income. *Physica A: Statistical Mechanics and its Applications*, 2003, **321**, 598–604.
- Gabaix, X., Zipf’s Law For Cities: An Explanation\*. *Quarterly Journal of Economics*, 1999, **114**, 739–767.
- Garibaldi, U., Scalas, E. and Viarengo, P., Statistical equilibrium in simple exchange games II. The redistribution game. *Eur. Phys. J. B*, 2007, **60**, 241.
- Garman, M.B., Market microstructure. *Journal of Financial Economics*, 1976, **3**, 257 – 275.
- Ghosh, A., Chakrabarti, A.S. and Chakrabarti, B.K., Kolkata Paise Restaurant problem in some uniform learning strategy limits. In *Econophysics & Economics of Games, Social Choices & Quantitative Techniques*, pp. 3–9, 2010a (Springer: Milan).
- Ghosh, A. and Chakrabarti, B.K., Mathematica Demonstration of the Kolkata Paise Restaurant (KPR) problem. <http://demonstrations.wolfram.com/> 2009.
- Ghosh, A., Chatterjee, A., Mitra, M. and Chakrabarti, B.K., Statistics of the Kolkata Paise Restaurant Problem. *New J. Phys.*, 2010b (in press).
- Gibrat, R., *Les Inégalités Economiques* 1931 (Sirey: Paris).
- Gode, D.K. and Sunder, S., Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality. *The Journal of Political Economy*, 1993, **101**, 119–137.
- Gu, G. and Zhou, W., Emergence of long memory in stock volatility from a modified Mike-Farmer model. *EPL (Europhysics Letters)*, 2009, **86**, 48002.
- Gupta, A.K., Relaxation in the wealth exchange models. *Physica A*, 2008, **387**, 6819.
- Gupta, A.K., Money exchange model and a general outlook. *Physica A*, 2006, **359**, 634.



- Hakansson, N., Beja, A. and Kale, J., On the feasibility of automated market making by a programmed specialist. *Journal of Finance*, 1985, pp. 1–20.
- Hart, M., Jefferies, P., Hui, P. and Johnson, N., Crowd-anticrowd theory of multi-agent market games. *The European Physical Journal B*, 2001, **20**, 4.
- Hasbrouck, J., *Empirical Market Microstructure: The Institutions, Economics, and Econometrics of Securities Trading* 2007, Oxford University Press, USA.
- Hayes, B., Follow the Money. *Am. Sci.*, 2002, **90**, 400.
- Iglesias, J.R., Goncalves, S., Abramson, G. and Vega, J.L., Correlation between risk aversion and wealth distribution. *Physica A*, 2004, **342**, 186.
- Iglesias, J.R., Goncalves, S., Pianegonda, S., Vega, J.L. and Abramson, G., Wealth redistribution in our small world. *Physica A*, 2003, **327**, 12.
- Iori, G., Jafarey, S. and Padilla, F.G., Systemic risk on the inter-bank market. *Journal of Economic Behavior & Organization*, 2006, **61**, 525–542.
- Ispolatov, S., Krapivsky, P.L. and Redner, S., Wealth distributions in asset exchange models. *Eur. Phys. J. B*, 1998, **2**, 267.
- Johnson, N.F., Hui, P.M., Jonson, R. and Lo, T.S., Self-Organized Segregation within an Evolving Population. *Physical Review Letters*, 1999a, **82**, 3360.
- Johnson, N.F., Hart, M. and Hui, P.M., Crowd effects and volatility in markets with competing agents. *Physica A: Statistical Mechanics and its Applications*, 1999b, **269**, 1–8.
- Kandori, M., Repeated Games. In *The New Palgrave Dictionary of Economics*, 2008, Palgrave Macmillan.
- Kirman, A., In *Individual Forecasting and Aggregate Outcomes: Rational Expectations Examined*, edited by R. Frydman and E. Phelps, chap. On mistaken beliefs and resultant equilibria, 1983, Cambridge University Press.
- Kirman, A., Ants, rationality, and recruitment. *The Quarterly Journal of Economics*, 1993, pp. 137–156.
- Kirman, A., Reflections on interaction and markets. *Quantitative Finance*, 2002, **2**, 322–326.
- Kirman, A., Whom or what does the representative individual represent?. *The Journal of Economic Perspectives*, 1992, pp. 117–136.
- Lallouache, M., Jedidi, A. and Chakraborti, A., Wealth distribution: To be or not to be a Gamma?. *Science and Culture (Kolkata, India)*, special issue on "Econophysics", 2010.
- Lamper, D., Howison, S.D. and Johnson, N.F., Predictability of Large Future Changes in a Competitive Evolving Population. *Phys. Rev. Lett.*, 2002, **88**, 017902.
- LeBaron, B., Agent-based computational finance. *Handbook of computational economics*, 2006a, **2**, 1187–1233.
- LeBaron, B., Agent-Based Financial Markets: Matching Stylized Facts with Style. In *Post Walrasian macroeconomics*, edited by D.C. Colander, p. 439, 2006b, Cambridge University Press.
- Levy, M. and Solomon, S., Power laws are logarithmic Boltzmann laws. *Int. J. Mod. Phys. C*, 1996, **7**, 595.
- Li, Y., Riolo, R. and Savit, R., Evolution in minority games. (I). Games with a fixed strategy space. *Physica A: Statistical Mechanics and its Applications*, 2000a, **276**, 234–264.
- Li, Y., Riolo, R. and Savit, R., Evolution in minority games. (II). Games with variable strategy spaces. *Physica A: Statistical Mechanics and its Applications*, 2000b, **276**, 265–283.
- Lux, T. and Marchesi, M., Volatility clustering in financial markets. *Int. J. Theo. Appl. Finance*, 2000, **3**, 675–702.
- Lux, T. and Westerhoff, F., Economics crisis. *Nature Physics*, 2009, **5**, 2–3.
- Lux, T., Emergent Statistical Wealth Distributions in Simple Monetary Exchange Models: A Critical Review. In *Proceedings of the Econophysics of Wealth Distributions*, edited by A. Chatterjee, S. Yarlagadda and B.K. Chakrabarti, p. 51, 2005, Springer.
- Mandelbrot, B., The Pareto-Levy law and the distribution of income. *International Economic Review*, 1960, pp. 79–106.
- Mandelbrot, B., The variation of certain speculative prices. *Journal of business*, 1963, **36**, 394.
- Martino, A.D., Giardina, I., Tedeschi, A. and Marsili, M., Generalized minority games with adaptive trend-followers and contrarians. *Physical Review E*, 2004, **70**, 025104.
- Maslov, S., Simple model of a limit order-driven market. *Physica A: Statistical Mechanics and its Applications*, 2000, **278**, 571–578.
- Matthes, D. and Toscani, G., On Steady Distributions of Kinetic Models of Conservative Economies. *J. Stat. Phys.*, 2007, **130**, 1087.
- Mike, S. and Farmer, J.D., An empirical behavioral model of liquidity and volatility. *Journal of Economic Dynamics and Control*, 2008, **32**, 200–234.
- Mohanty, P.K., Generic features of the wealth distribution in ideal-gas-like markets. *Phys. Rev. E*, 2006, **74**, 011117.
- Muni Toke, I., "Market making" in an order book model and its impact on the bid-ask spread. In *Econophysics of Order-Driven Markets*, 2010 (Springer: Milan).
- Nirei, M. and Souma, W., A Two Factor Model of Income Distribution Dynamics. *The Review of Income and Wealth*, 2007, **53**, 440–459.
- O'Hara, M., *Market Microstructure Theory* 1995, Blackwell Publishers.
- Pareto, V., *Cours d'economie politique* 1897, Rouge, Lausanne.
- Patriarca, M., Chakraborti, A. and Germano, G., Influence of saving propensity on the power law tail of wealth distribution. *Physica A*, 2006, **369**, 723.
- Patriarca, M., Chakraborti, A. and Kaski, K., Statistical model with a standard gamma distribution. *Phys. Rev. E*, 2004a, **70**, 016104.
- Patriarca, M., Chakraborti, A., Kaski, K. and Germano, G., Kinetic theory models for the distribution of wealth: Power law from overlap of exponentials. In *Proceedings of the Econophysics of Wealth Distributions*, edited by A. Chatterjee, S. Yarlagadda and B.K. Chakrabarti, p. 93, 2005, Springer.
- Patriarca, M., Chakraborti, A., Heinsalu, E. and Germano, G., Relaxation in statistical many-agent economy models. *Eur. J. Phys. B*, 2007, **57**, 219.
- Patriarca, M., Chakraborti, A. and Kaski, K., Gibbs versus non-Gibbs distributions in money dynamics. *Physica A: Statistical Mechanics and its Applications*, 2004b, **340**, 334–339.
- Patriarca, M. and Heinsalu, E., Stochastic resonance in a general bistable system. *Eur. Phys. J. B*, submitted, 2009.
- Patriarca, M., Heinsalu, E. and Chakraborti, A., Basic kinetic wealth-exchange models: common features and open problems. *Eur. J. Phys. B* 73, 2010, p. 145.
- Potters, M. and Bouchaud, J.P., More statistical properties of order books and price impact. *Physica A: Statistical Mechanics and its Applications*, 2003, **324**, 133–140.
- Preis, T., Golke, S., Paul, W. and Schneider, J.J., Statistical analysis of financial returns for a multiagent order book model of asset trading. *Physical Review E*, 2007, **76**.
- Raberto, M., Cincotti, S., Focardi, S. and Marchesi, M., Agent-based simulation of a financial market. *Physica A: Statistical Mechanics and its Applications*, 2001, **299**, 319–327.
- Repetowicz, P., Hutzler, S. and Richmond, P., Dynamics of Money and Income Distributions. *Physica A*, 2005, **356**, 641.
- Sala-i Martin, I., The World Distribution of Income. *NBER Working Paper Series*, 2002.
- Sala-i Martin, X. and Mohapatra, S., Poverty, Inequality and the Distribution of Income in the G20. *Columbia University, Department of Economics, Discussion Paper Series*, 2002.
- Samanidou, E., Zschischang, E., Stauffer, D. and Lux, T., Agent-based models of financial markets. *Reports on Progress in Physics*, 2007, **70**, 409–450.
- Satinover, J.B. and Sornette, D., Illusion of control in Time-Horizon Minority and Parrondo Games. *The European Physical Journal B - Condensed Matter and Complex Systems*, 2007, **60**, 369–384.



- Savit, R., Manuca, R. and Riolo, R., Adaptive Competition, Market Efficiency, and Phase Transitions. *Phys. Rev. Lett.*, 1999, **82**, 2203.
- Scafetta, N., Picozzi, S. and West, B.J., An out-of-equilibrium model of the distributions of wealth. 2004.
- Scalas, E., Garibaldi, U. and Donadio, S., Statistical equilibrium in simple exchange games I - Methods of solution and application to the Benaï-Dragulescu-Yakovenko (BDY) game. *Eur. Phys. J. B*, 2006, **53**, 267.
- Scalas, E., Garibaldi, U. and Donadio, S., *Erratum*. Statistical equilibrium in simple exchange games I. *Eur. Phys. J. B*, 2007, **60**, 271.
- Sen, A.K., *Collective Choice and Social Welfare* 1971, Oliver & Boyd.
- Shostak, F., The Mystery of The Money Supply Definition. *Quart. J. Aust. Economics*, 2000, **3**, 69.
- Silva, A.C. and Yakovenko, V.M., Temporal evolution of the ‘thermal’ and ‘superthermal’ income classes in the USA during 1983-2001. *Europhysics Letters*, 2005, **69**, 304.
- Slanina, F., Inelastically scattering particles and wealth distribution in an open economy. *Phys. Rev. E*, 2004, **69**, 046102.
- Slanina, F., Critical comparison of several order-book models for stock-market fluctuations. *The European Physical Journal B-Condensed Matter and Complex Systems*, 2008, **61**, 225-240.
- Slanina, F.c.v., Mean-field approximation for a limit order driven market model. *Phys. Rev. E*, 2001, **64**, 056136.
- Smith, E., Farmer, J.D., Gillemot, L. and Krishnamurthy, S., Statistical theory of the continuous double auction. *Quantitative Finance*, 2003, **3**, 481-514.
- Solomon, S. and Levy, M., Spontaneous Scaling Emergence in Generic Stochastic Systems. , 1996.
- Sornette, D., Multiplicative processes and power laws. *Phys. Rev. E*, 1998, **57**, 4811.
- Stigler, G., Public regulation of the securities markets. *Journal of Business*, 1964, pp. 117-142.
- Sysi-Aho, M., Chakraborti, A. and Kaski, K., Adaptation using hybridized genetic crossover strategies. *Physica A*, 2003a, **322**, 701.
- Sysi-Aho, M., Chakraborti, A. and Kaski, K., Biology Helps You to Win a Game. *Physica Scripta T*, 2003b, **106**, 32.
- Sysi-Aho, M., Chakraborti, A. and Kaski, K., Intelligent minority game with genetic crossover strategies. *Eur. Phys. J. B*, 2003c, **34**, 373.
- Sysi-Aho, M., Chakraborti, A. and Kaski, K., Searching for good strategies in adaptive minority games. *Phys. Rev. E*, 2004, **69**, 036125.
- Turner, S., Farmer, J.D. and Geanakoplos, J., Leverage Causes Fat Tails and Clustered Volatility. *Preprint*, 2009.
- Wyart, M. and Bouchaud, J.P., Self-referential behaviour, overreaction and conventions in financial markets. *Journal of Economic Behavior & Organization*, 2007, **63**, 1-24.
- Yakovenko, V.M. and Rosser, J.B., Colloquium: Statistical mechanics of money, wealth, and income. *Reviews of Modern Physics*, 2009, **81**, 1703.