Eur. Phys. J. B **57**, 453–462 (2007)
DOI: 10.1140/epjb/e2007-00185-4

THE EUROPEAN PHYSICAL JOURNAL B

Interacting gaps model, dynamics of order book, and stock-market fluctuations

A. Svorenčík[1] and F. Slanina[2,a]

[1] Department of Theoretical Physics and Didactics of Physics, FMFI, Comenius University, Mlynska Dolina, 84248 Bratislava, Slovakia

[2] Institute of Physics, Academy of Sciences of the Czech Republic, Na Slovance 2, 18221 Praha, Czech Republic and Center for Theoretical Study, Jilská 1, Prague, Czech Republic

Received 13 December 2006 / Received in final form 15 May 2007
Published online 29 June 2007 –

Abstract. Inspired by order-book models of financial fluctuations, we investigate the Interacting gaps model, which is the schematic one-dimensional system mimicking the order-book dynamics. We find by simulations the power-law tail in return distribution, power-law decay of volatility autocorrelation with exponent 0.5 and Hurst exponent close to 1/2. Surprisingly, when we make a mean-field approximation, i.e. replace the one-dimensional system by effectively infinite-dimensional one, we obtain analytically the return exponent 5/2, in perfect accord with one-dimensional simulations.

PACS. 89.65.-s Social and economic systems – 05.40.-a Fluctuation phenomena, random processes, noise, and Brownian motion – 02.50.-r Probability theory, stochastic processes, and statistics

1 Introduction

Despite much effort, understanding the fluctuations observed in social phenomena, namely movements of prices of various commodities, remains a challenging problem [1–4]. In short, the system is generically characterised by a power-law tail in return distribution, with exponent $1 + \alpha \simeq 4$, power-law autocorrelation of volatility, with exponent ranging between 0.3 to 0.5, anomalous Hurst exponent $H \simeq 2/3$, multifractality, and other, more subtle features [5–13]. Econophysics is based on the belief that physical models, methods and procedures may be appropriate to explain these stylised facts.

Among the numerous models of stock-market fluctuations investigated by physicists in the last about ten years [14–26], the order-book modelling [27–46] is particularly appealing, for at least two reasons. First, the models of this kind typically use much less arbitrary and ill-justified assumptions than other approaches. Second, the similarity to widely-studied deposition and evaporation models [47] promises to shed some light on genuinely physical problems as well. Moreover, very detailed empirical studies are available, looking into the deep mechanisms in work within the order book [40,42,48–57].

Perhaps the first important step in this direction was marked by the model of Bak et al. (BPS) [29]. On a line representing the price axis two kinds of particles are placed. The first kind, denoted A (ask), corresponds to the sell orders, while the second, B (bid), corresponds to the buy orders. The position of the particle is the price at which the order is to be satisfied . A trade can occur only when two particles of opposite type meet. If that happens, the orders are satisfied and the particles are removed from the system. This can be described as annihilation reaction $A+B \to \emptyset$. It is evident that all B particles must lie on the left with respect to all A particles. The particles diffuse freely and in order to keep their concentration constant on average, new orders are inserted from the left (B type) and from the right (A type). The whole picture of this order-book model is therefore identical to the two-species diffusion-annihilation process. The changes in the price are mapped on the movement of the reaction front.

Many analytical results are known for this model. Most importantly, the Hurst exponent can be calculated exactly [30,31] and the result is $H = 1/4$. This value is well below the empirically established value $H \simeq 2/3$. To cure this discrepancy, the pure diffusion-annihilation process was modified so that the new orders are placed close to already existing ones, thus mimicking certain level of "copying" or "herding" mechanism, which is surely present in the real-world price dynamics. The diffusion constant of the orders can also be coupled to the past volatility, introducing a positive feedback effect. This way the Hurst exponent can be enhanced up to the level consistent with the empirical value.

The diffusion of orders contradicts reality. Indeed, orders can be placed into the order book, and later either cancelled or satisfied, but change in price is very uncommon. In the BPS model diffusion was a necessary ingredient to bring the orders together and let them annihilate. Maslov [32] introduced a different model, equally simple

[a] e-mail: slanina@fzu.cz

but closer to actual functioning of the order book. Besides the obvious classification of orders as A and B, they differ in the timing of their satisfaction. Limit orders are placed at certain price and their position is never changed, until they are satisfied. The market orders do not specify any price, but require to be executed immediately. Therefore, the best limit order (i.e. highest B or lowest A) annihilates with the currently introduced market order.

The Maslov model has several appealing features. Especially, the return distribution characterised by exponent $1+\alpha \simeq 3$ seems to be close to the empirically found power law. The scaling in return distribution is clearly seen as well as the volatility clustering manifested by power-law decay of the autocorrelation of absolute returns. However, the Hurst exponent is 1/4 as in the BPS model, which is bad news.

Sophisticated models with many ingredients making them as realistic as possible were introduced subsequently [37–39]. The most important feature which was absent in Maslov model is the cancellation (or evaporation) of orders [42]. In fact, in a typical stock market, the fraction of orders cancelled before execution is quite substantial. And it is the evaporation of orders that increases the value of the Hurst exponent closer to the empirical value, as demonstrated e.g. in [42]. Analytical results were obtained for two types of "mean-field" approximation [37]. The exclusion principle, forbidding more than one order at the same place, significantly facilitates the treatment. In the first type of approximation, the average density of orders was calculated using the master-equation approach. In the second and complementary one, the quantity of interest was the distribution of distances (gaps) between adjacent orders. It is the latter approach that we took inspiration from in this work.

Important progress was achieved in the articles [43,44] using the assumption that the orders are "hard-core" particles. The exclusion principle follows and the authors were able to exploit it to the maximum possible extent. By mapping the model on the well-studied asymmetric exclusion process, it was possible to obtain the exact value of the Hurst exponent $H = 2/3$. This is indeed very satisfactory result, because the empirical value seems to lie very near.

Despite of the optimistic view presented in the preceding paragraphs the true dynamics of the order book is far from being fully understood. On one side, the trading in the stock market is much more intricate than mere play of limit and market orders. There are many more types of them, sometimes rather complicated. Moreover, all the above mentioned models are appropriate only to markets without official market maker. In other markets the role of the market maker, which can also follow certain non-trivial rules of behaviour, cannot be passed over.

On the other side, the more realistic the model is, the lower are the chances for analytic solution. Except for the BPS model seen as diffusion-annihilation process and the mapping on asymmetric exclusion process in [43], no exact results are available. Maslov model was solved within a sort of mean-field approximation [58], but the exponent of the return distribution calculated this way was $1+\alpha = 2$, in contradiction to the value $1+\alpha = 3$ observed in simulations. Several approximative schemes were applied to the models with evaporation [37] and the static properties, namely the order density and the response function, were successfully computed. The fluctuation properties, i.e. the return distribution and the Hurst exponent, turn to be harder problem. Therefore, it would be desirable to devise a schematic model, which, while preserving general features of the order-book dynamics, would promise deeper analytic insight. In order to fill, at least partially, the gap between more realistic and analytically solvable models, we investigate here the Interacting Gaps model (IGM).

2 Order book as one-dimensional IGM

2.1 Motivation

It is well established empirically [56] that the distribution of returns after one trade follows faithfully the distribution of the gap between lowest ask and highest bid (the spread). The idea is therefore to calculate the probability distribution of this gap, as it follows from the stochastic process used to model the order-book dynamics.

Indeed, the state of the order book can be described in terms of the positions of the orders c_i. Equally well the state is represented by the sequence of gaps, i.e. distances between orders $g_i = |c_{i+1} - c_i|$. Here and in all what follows we suppose that all orders have the same unit volume. We also simplify the situation assuming that the gaps can be only positive integer numbers. This means that we allow at most one order at each site and the same exclusion principle as in previous works [37,43,44] is also effective here.

When the orders are deposited, executed and evaporated, the lengths of the gaps change, but such change affects only at most two neighbouring gaps, as if the two gaps in contact underwent a reaction producing one or two new gaps at the place of the old ones. Indeed, if an order is executed or evaporated, two adjacent gaps collapse into one. When an order is placed, one gap is split into a pair of adjacent gaps. To mimic such dynamics, Solomon introduced the Interacting gaps model, which was first investigated in [59].

The idea of the model consists of three (admittedly brute) simplifications. First, we do not make any distinction between the spread and any other gap. All gaps are treated on equal basis. This also implies that there is no intrinsic difference between asks and bids. All orders can substitute both, depending on the context; more specifically, on the current position of the price. As all gaps are equivalent, the distribution of one-trade returns is assumed equal to the distribution of lengths of all gaps.

This first simplification certainly deviates most from the reality. Te spread, the second, the third etc. gaps are not equivalent. On the other hand, the first gap interacts strongly with the second one, so the distribution of the former reflects to large extent the distribution of the latter. Hence the temptation to consider the distribution of the first two gaps equal and by iteration we extend the hypothesis of equal distribution to each and every gap. But then, it is also irrelevant where the reaction of the limit and market orders takes place. We only need that the place is not far from the reaction site in the previous instant. It would be desirable to support the assumption of equal distribution of at least two or three first gaps by empirical data. Unfortunately, we are not aware of any published results in this direction — which may be only due to our ignorance, perhaps.

The second simplification consists in keeping the total number of orders constant. We also do not include the evaporation explicitly. So, the deposition of an order is immediately followed by execution of another order and vice versa. Therefore, instead of keeping the average number of orders constant, as in the models [32,37,42], we rather impose strict conservation law for the number of orders, like in [29]. Although the conservation or not may in principle change the universality class of the model, we do not believe it is the case here, because the numerical simulation of e.g. the model of reference [37] shows rather mild fluctuations of the number of orders around its average value.

The third simplification concerns the position of the price. A more realistic definition could be that the price is the position of the last executed order. However, since all our orders have unit volume and there may be at most one order at the same place, the price would be situated at an empty site. Such definition would bring an additional complication in the relation between the distribution of returns and distribution of gaps. Instead, we consider the position of the existing order next to the last executed order as the current price. We take the left and right neighbour with equal probability $1/2$. It can be understood as identifying the price with either the highest bid or the lowest ask. This way we make a systematic error in the absolute price position. Typical size of the error is half of the spread. However, we do not expect it to have significant impact on the qualitative features of the return distribution, namely on the power-law tail, if present. Note also that the commonly used definition of price as the exact middle between lowest ask and highest bid is burdened by the same type of systematic error, without considerable consequences for the empirical return distribution.

Now, the question is how to implement the model with such simplifications, while keeping as much fidelity to the real dynamics as possible. Of course, many realistic features will be irreparably lost, but still we believe that the gap kinetics we apply says something on the true order books. Thus, we introduce the following dynamical rules, illustrated also in Figure 1.

One of the existing orders marks the current price. Let us call it i. A new order is deposited next to it, at distance 1. Then, one order is executed immediately. We admit three possibilities. If the executed order is the newest one, no apparent change in the order book results. Or, the order i is executed. (E.g. if i is a bid, new order is an ask and then a market order to sell arrives.) In this case the sequence of deposition plus execution is equivalent to a shift of the order i to the left or right by distance 1. In other words, one of the reacting gaps is shortened by 1 and the other one gains 1 in its length. The third possibility is that the order j, next to the newest one, but on the other side than the order i, is executed. (E.g. if both i and the new orders are bids, j is an ask and a market order to buy arrives.) This is equivalent to the collapse of two adjacent gaps into two another gaps, one of them having length 1 and the second gaining the rest of the sum of lengths of the original gaps. We suppose that the collapse occurs with probability p and the mere shift by 1 with complementary probability $1-p$.

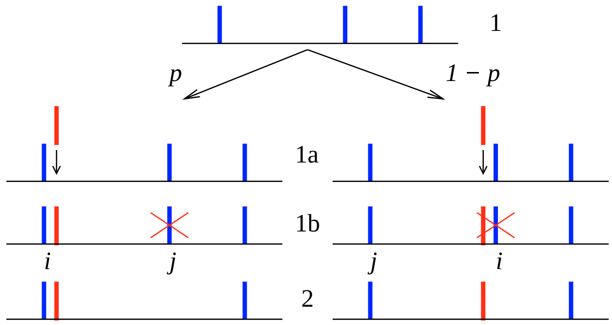

Fig. 1. Illustration of the Interacting Gaps model. From an original configuration at step 1, two ways can lead to a final configuration at step 2. In the intermediate step 1a a new order (red in colour version) is deposited at the distance 1 from an existing order (call it i for the moment). Then, at intermediate step 1b one of the previously existing orders is executed. This way the total number of orders is kept unchanged. The executed order is just the order i close to which the new order was placed (this happens with probability $1-p$) or it is the order j situated on the other side with respect to the new order (with probability p).

In the IGM there is no obvious constraint on the choice of the pair of interacting gaps, besides the fact that they must be adjacent. However, it is natural to assume that the pair to interact now lies next to the pair which resulted from the interaction in the last step. Were it gaps g_j and g_{j+1}, we let interact either the pair g_{j-1} and g_j or g_{j+1} and g_{j+2}, with equal probability. This way the index $j=k(t)$ determining the pair of gaps to interact at time t performs a simple random walk. In reality, of course, the process $k(t)$ results from complicated dynamics of deposition, execution, etc. of orders. Here we relax all bounds $k(t)$ must follow and keep only the elementary fact that in absence of evaporation (which is what we assume) the subsequent interaction sites must not lie far from each other. The simplest way how to implement it is to require that $k(t)-k(t+1)=\pm 1$ with equal probability for both signs.

We should also note that similar models have been studied in physics since long ago, starting with the work by Smoluchowski [60] on coalescence in colloidal suspensions. Indeed, the gaps can be considered as sizes of particles, which may merge or split due to aggregation and dissociation. Such processes were studied in open [61–63] as well as closed [64,65] circumstances. Especially the model with aggregation and chipping [66–70] seems to be close to

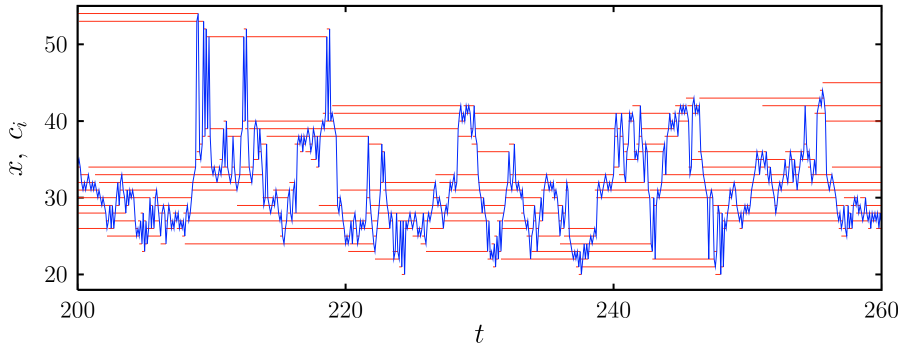

Fig. 2. Example of the evolution of the order-book configuration. Horizontal lines represent the positions of the orders, the rugged line describes the movement of the price. The parameters are $N = 10$, $\overline{g} = 5$, $p = 0.3$.

our Interacting Gaps model. We shall spend a few words of comparison in the conclusions.

2.2 Definition

Let us repeat the rules just explained in a more formal way. We have $N+1$ orders of unit volume placed at integer positions, $c_i \in \mathbb{Z}$, $i = 0, 1, \ldots, N$. Alternatively, the state is uniquely described by the position c_1 and the widths of the N gaps $g_i = c_i - c_{i-1}$, $i = 1, 2, \ldots, N$. The pair of gaps to interact at time t is determined by the order $k(t)$ which lies at the border between the two gaps. Thus, the interacting gaps are g_j and g_{j+1}, where $j = k(t)$. Next time the interaction takes place at the neighbouring order, so $k(t+1) = k(t) \pm 1$ with equal probability, and with the only (obvious) condition that $0 \leq k(t) \leq N$ for all times t. We shall explain later what happens at the boundaries, when $k(t) = N$ or 0. Let us assume for a while that $0 < k(t) < N$.

After the interaction the two gaps affected change in the following way. In all cases we first generate a random number $\sigma = \pm 1$ (the direction along the price axis) where both signs come with equal probability 1/2. Then, with probability p a "collapse" occurs, and we perform the update

$$(g_j, g_{j+1})(t+1) = \begin{cases} (1, g_j(t) + g_{j+1}(t) - 1) & \text{for} \quad \sigma = +1 \\ (g_j(t) + g_{j+1}(t) - 1, 1) & \text{for} \quad \sigma = -1. \end{cases} \tag{1}$$

Conversely, with probability $1 - p$ a "shift" manifests itself as

$$(g_j, g_{j+1})(t+1) = (g_j(t) + \sigma, g_{j+1}(t) - \sigma) \tag{2}$$

on condition that both $g_j(t) + \sigma \geq 1$ and $g_{j+1}(t) - \sigma \geq 1$; otherwise the gaps do not change.

For the orders at extremal positions, $k(t) = N$ or 0, the rules must be appropriately adapted, because there is only one gap to change, instead of two. In principle, there are several options how to define the dynamics. We found the following modification reasonable. Again, the collapse and shift occur with probabilities p and $1-p$, respectively and $\sigma = \pm 1$ is chosen with equal probabilities 1/2. Then, the single affected gap is g_j, where $j = 1$ if $k(t) = 0$ and $j = N$ if $k(t) = N$. The collapse implies that the gap shrinks

$$g_j(t+1) = 1 \tag{3}$$

but this happens only if $\sigma = +1$ and $k(t) = N$ or $\sigma = -1$ and $k(t) = 0$, i.e. the rightmost and leftmost gaps can collapse only towards the bulk of the other orders. Otherwise the gap does not change. On the other hand, the shift will be free, i.e.

$$g_j(t+1) = g_j(t) + \sigma \tag{4}$$

if only $g_j(t) + \sigma \geq 1$.

The gap dynamics implies unambiguously the dynamics of the positions of the orders. Interaction of gaps g_{j+1} and g_j affects only the position of order j. Its old position $c_j(t) = c_{j-1}(t) + g_j(t)$ is changed to $c_j(t+1) = c_{j-1}(t) + g_j(t+1)$. (To be precise, for $j = 0$ we must use $c_j(t+1) = c_{j+1}(t) - g_{j+1}(t+1)$ instead.) All other orders keep their positions, $c_i(t+1) = c_i(t)$ for all $i \neq j$.

The remaining piece to be specified is the movement of price. The most realistic prescription for the location of the current price $x(t)$ is the position of the order separating the two gaps after their interaction, i.e.

$$x(t+1) = c_{k(t)}(t+1) \tag{5}$$

therefore the return after one step of the dynamics is

$$r(t) \equiv x(t+1) - x(t) = g_{k'}(t+1) \tag{6}$$

where we denoted $k' = \max(k(t), k(t-1))$. (The case $k' = k(t)$ applies when the price went up, the other case corresponds to a downward movement.) We can see that the return equals one of the gaps exactly, which is just the feature we built our model upon.

2.3 Simulation: dynamics

Let us see how the configuration of gaps evolves in computer simulations. We found convenient to use rescaled time, depending on the number of orders. So, in one update the time is actually advanced by $1/N$. The initial condition will be always the uniform one, will all gaps

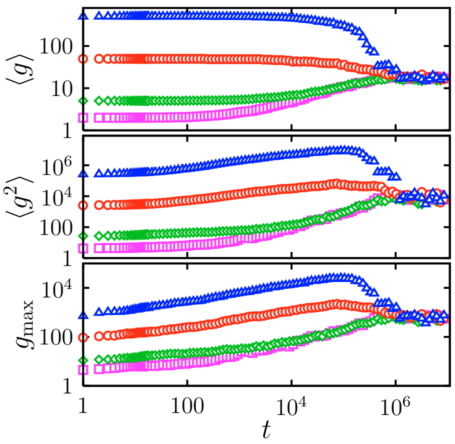

Fig. 3. Evolution of the first two moments of the gap size distribution and the largest gap. The parameters are $N = 100$; $\overline{g} = 2$ (□), 5 (◇), 50 (○), and 500 (△); $p = 0.01$.

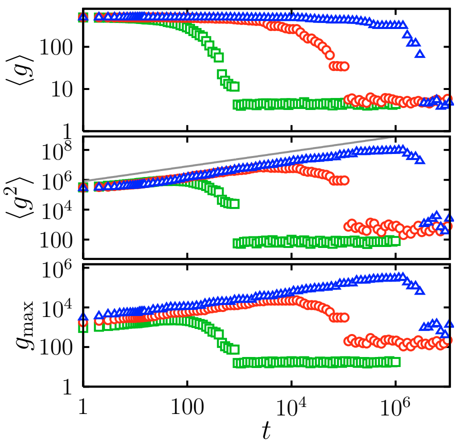

Fig. 4. Evolution of the first two moments of the gap size distribution and the largest gap. The parameters are $N = 10$ (□), 100 (○), and 1000 (△); $\overline{g} = 500$; $p = 0.1$. The line in the middle panel is the dependence $\propto t^{1/2}$.

equal, $g_i(0) = \overline{g}$. In Figure 2 we show a typical record of the gap dynamics in relatively small order book of 10 orders. We can clearly see the orders "waiting" unchanged until the diffusing price hits them. The widths of the gaps are far from uniform and quite large gaps occur regularly, although most of the gaps are small.

Deeper insight into the dynamics of gaps distribution can be gained through the evolution of the average and the second moment of the gap distribution

$$\langle g^n \rangle \equiv \frac{1}{N} \sum_{i=1}^{N} g_i^n(t) \tag{7}$$

for $n = 1,\ 2$. An interesting quantity will be also the width of the largest gap $g_{\max}$. Obviously, the initial condition implies $\langle g \rangle(0) = g_{\max}(0) = \overline{g}$ and $\langle g^2 \rangle(0) = \overline{g}^2$.

We can see in Figures 3 and 4 how these three quantities evolve in the course of the simulation.

The average gap $\langle g \rangle$ keeps close to its initial value for quite a long time and then quickly settles to a stationary value. The latter is independent of the number of gaps N and of the initial gap size $\overline{g}$. However, it depends on the probability of collapse p. For larger p the stationary $\langle g \rangle$ is smaller, which can be easily understood intuitively, as collapses produce short gaps (of size 1). The independence of N is also to be expected, while convergence to a common value irrespectively of $\overline{g}$ implies that after long enough time the initial condition embodied in the starting placement of orders is forgotten. In other words, the dynamics is nicely ergodic, only the characteristic time to reach the stationary state grows relatively fast (faster than linearly) with the number of gaps N.

These conclusions are supported by the same qualitative behaviour of the second moment and the maximum gap. Interestingly, the second moment grows slowly, $\langle g^2 \rangle \sim t^{1/2}$ in the transient regime, but then suddenly settles at its stationary value. The behaviour of $g_{\max}$ is nearly identical, suggesting that the transient values of $\langle g^2 \rangle$ are dominated by the largest gap. The slight increase in the stationary value of $g_{\max}$ with N is due to the fact that the largest gap can be estimated as $\int_{g_{\max}}^{\infty} P_{\text{stat}}(g) \simeq 1/N$ where $P_{\text{stat}}(g)$ is the $N \to \infty$ limit of the stationary distribution of gap sizes.

To conclude, the study of the dynamics suggest that there exists a well-defined stationary state, whose structure depends only on a single parameter p. The next paragraph will be devoted to the investigation of its properties.

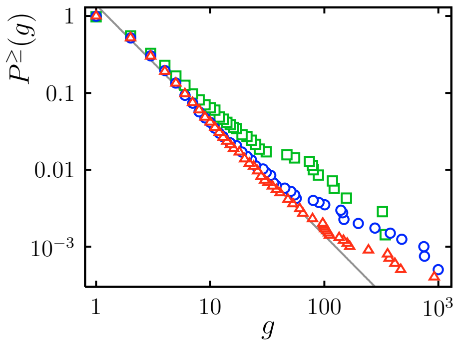

Fig. 5. Cumulative distribution of gap sizes. The parameters are $N = 100$, (□), 1000 (○), and 10000 (△); $\overline{g} = 500$; $p = 0.1$. Line denotes the power law $\propto g^{-1.5}$.

2.4 Simulation: stationary state

Let us look first at the distribution of gap sizes. It is shown in Figure 5 and the data strongly suggest that for $N \to \infty$ the tail if the distribution becomes a power law

$$P^{\geq}(g) \sim g^{-3/2}. \tag{8}$$

The deviation from this dependence is due to finite number of gaps.

We should not forget that the motivation for the model is mimicking the price fluctuations. So, let us see how the one-step returns look like. The distribution of returns is shown in Figure 6. Again, we can clearly see the power-law

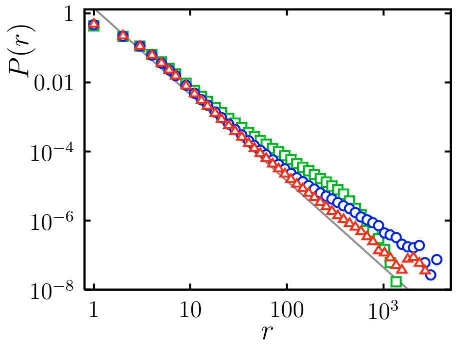

Fig. 6. Histogram of one-step returns. The parameters are $N = 100$, (□), 1000 (○), and 10000 (△); $\overline{g} = 500$; $p = 0.1$. Line denotes the power law $\propto r^{-2.5}$.

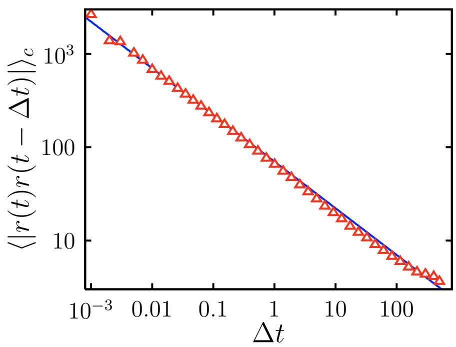

Fig. 7. Autocorrelation of absolute returns. Parameters are $N = 1000$, $\overline{g} = 50$, $p = 0.1$. The line marks the dependence $\propto (\Delta t)^{-1/2}$.

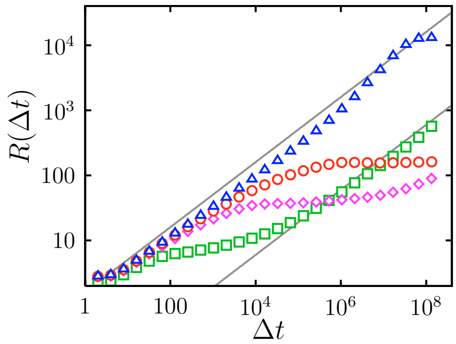

Fig. 8. Hurst plot for the systems of size $N = 10$ (□), $N = 100$ (◇), $N = 1000$ (○), and $N = 10000$ (△). Other parameters are $\overline{g} = 500$, $p = 0.1$. The straight lines are the powers $\propto (\Delta t)^{1/2}$.

behaviour in the tail

$$P(r) \sim r^{-5/2} \tag{9}$$

which is perfectly compatible with the gap-size distribution (8). The exponent observed in the simulations is very well compatible with the value $1 + \alpha = 5/2$, which is smaller than in the Maslov model, but it is still closer to the reality than the mean-field calculation of reference [58].

We have not observed any dependence of the exponent in the distributions on the parameter p. The only change concerns the small values of g, while the tails remain unaffected. Therefore the behaviour expressed by (8) and (9) appears to be universal.

The volatility clustering is manifested in the autocorrelation function of absolute returns $\langle |r(t)r(t-\Delta t)|\rangle_c = \langle |r(t)r(t-\Delta t)|\rangle - \langle |r(t)|\rangle\langle |r(t-\Delta t)|\rangle$. We show the simulation result for this quantity in Figure 7. We can clearly observe the slow power-law decay

$$\langle |r(t)r(t-\Delta t)|\rangle_c \sim (\Delta t)^{-\eta} \tag{10}$$

with exponent consistent with the value $\eta \simeq 0.5$, which is somewhat larger than in most empirical studies, but still agrees with at least some of the real data [13]. Therefore, the long-time correlations observed in the real time series of returns are quite well reproduced within the Interacting gaps model.

Finally, we looked at the Hurst plot, which is the dependence of the typical extent of fluctuations during a time interval Δt on the length of this interval. The convenient measure is provided by the quantity

$$R(\Delta t) = \left\langle \frac{\max_{t',t''} |x(t') - x(t'')|}{\sqrt{\langle r^2(t')\rangle_{t'} - \langle r(t')\rangle_{t'}^2}} \right\rangle_t \tag{11}$$

where t' and t'' go from t to $t + \Delta t$. The definition looks awkward, but in the numerator there is just the extent of the fluctuation from t to $t+\Delta t$ and in the denominator we have standard deviation of the fluctuations in this interval, or square root of the volatility. Everything is then averaged over the starting time t. The denominator becomes substantial when the volatility has long-time correlations, which is just the case here.

If the dependence on the time span goes like $R(\Delta t) \sim (\Delta t)^H$, then H is the Hurst exponent. We can see from the simulation results shown in Figure 8 that in IGM the situation is not so straightforward. We observe three regimes. For shorter times, the data indicate power-law behaviour with Hurst exponent $H \simeq 0.5$. Then, there is a crossover regime for medium times, where $R(\Delta t)$ is essentially constant, and for the largest times there is again power-law dependence with the same exponent $H \simeq 0.5$. All three regimes grow longer when the number of intervals increases. The behaviour is at first sight similar to those of reference [42], but the mechanism of the crossover to the asymptotic value $H = 1/2$ is different. In [42] it is related to the evaporation of orders, which is absent here.

It is not too difficult to interpret these results. We have seen that in the stationary state the orders concentrate into more or less coherent bunch. However, such bunch is free to move along the price axis. The regime of short times (which may last many decades, though, when N is large enough) is dominated by the movements of the price within the bunch. The rule determining which pair of gaps will interact, imposes a true random walk in terms of the indices of the gaps. Provided we can establish a typical gap size (and because the gap distribution has exponent $\simeq 5/2 > 2$ we can, indeed), this implies that also the price follows a process, which is essentially a random walk. Hence the value $1/2$ for the Hurst exponent, as long as the movements are limited to the interior of the bunch of orders. The bunch as a whole also moves, but at much larger timescale. In the intermediate times the price seems to be trapped at a fixed position, just where the bunch resides. This means that $R(\Delta t)$ remains constant, proportional to the size of the bunch. In this regime $H = 0$. However, at still larger scales of time the movement of the bunch determines also the movement of the price and because the bunch also moves like a random walk, the value $H = 1/2$ is restored.

3 Mean-field approximation

In the one-dimensional version only adjacent intervals can interact. Allowing interactions between any pair of intervals we make a mean-field approximation, in the same spirit as in reference [71] (and Refs. [66–68]).

The state is described by the set of gap lengths $\{g_i\}$, $i = 1, 2, \ldots, N$. The dynamics is the same as in the one-dimensional IGM, with some simplifications. First, in each step we choose randomly a pair of interacting gaps g_j and g_l. As j and l are a priori equivalent indices (more precisely, the dynamics is invariant with respect to any permutation of all indices, including the exchange $j \leftrightarrow l$ as a special case), the random choice of $\sigma = \pm 1$ is now superfluous. Second, there is no need for separate treatment of the first and last gap, so that boundary conditions have no effect. Thus, the collapse occurring with probability p means

$$(g_j, g_l)(t+1) = (1, g_j(t) + g_l(t) - 1). \tag{12}$$

The shift with probability $1 - p$ amounts to the update

$$(g_j, g_l)(t+1) = (g_j(t) + 1, g_l(t) - 1) \tag{13}$$

if $g_l(t) > 1$; otherwise there is no change.

We shall be interested in the probability distribution on the space of all configurations, $P(\{g_i\})$. However, the restricted information contained in one-gap distribution function $P_j(g) = \sum_{\{g_i\}} \delta(g_j - g) P(\{g_i\})$ will be enough for our purposes. Occasionally, we shall employ the two-gap distribution $P_{jk}(g, g') = \sum_{\{g_i\}} \delta(g_j - g)\delta(g_k - g')P(\{g_i\})$. The permutation invariance guarantees that $P_j(g) = P_1(g)$ and $P_{jk}(g, g') = P_{12}(g, g')$ for any indices j and k ($k \neq j$).

From the elementary moves (12) and (13) we derive the following exact equation for the evolution of the one-gap distribution

$$\begin{aligned} N\Big[P_1(g; t+1) - P_1(g; t)\Big] &= -2P_1(g, t) \\ &+ (1-p)\Big[\delta(g-1)P_1(1; t) \\ &+ P_1(g+1; t) + P_1(g-1; t) \\ &+ P_{12}(g, 1; t) - P_{12}(g-1, 1; t)\Big] \\ &+ p\Big[\delta(g-1) + \sum_{u=1}^{g} P_{12}(u, g+1-u; t)\Big]. \end{aligned} \tag{14}$$

For $N \to \infty$ the two-gap distribution factorises, so that

$$P_{12}(g, g'; t) = P_1(g; t)P_1(g'; t) \tag{15}$$

and we obtain exact and closed equation for one-gap distribution function. However, when taking the thermodynamic limit $N \to \infty$ we miss certain essential features of the distribution. We shall discuss later how to bring these features back, for otherwise we fall into annoying inconsistencies.

In the stationary state the LHS of the master equation (14) vanishes. and applying the factorisation (15) we get a closed equation for the stationary distribution $P_1(g) = \lim_{t\to\infty} P_1(g; t)$. The discrete Laplace transform $\hat{f}(z) = \sum_{g=1}^{\infty} z^g f(g)$ converts it into a quadratic equation, thus

$$\begin{aligned} -2\widehat{P}_1(z) + p\Big[z + \frac{1}{z}\big(\widehat{P}_1(z)\big)^2\Big] \\ + (1-p)\Big[\Big(z + \frac{1}{z}\Big)\widehat{P}_1(z) \\ - (1-z)\big(1 - \widehat{P}_1(z)\big)P_1(1)\Big] = 0. \end{aligned} \tag{16}$$

Selecting the only physically admissible root we get the solution

$$\begin{aligned} \widehat{P}_1(z) =& 1 + \Big(1 + \frac{1-p}{2p}P_1(1)\Big)(z-1) \\ & - \frac{1-p}{2p}\Big(1 - P_1(1)\Big)(z-1)^2 \\ & + \sqrt{\Big\{\Big[1 + \Big(1 + \frac{1-p}{2p}P_1(1)\Big)(z-1) - \frac{1-p}{2p}\Big(1 - P_1(1)\Big)(z-1)^2\Big]^2 - \Big[1 + \Big(2 + \frac{1-p}{p}P_1(1)\Big)(z-1) + \Big(1 + \frac{1-p}{p}P_1(1)\Big)(z-1)^2\Big]\Big\}}. \end{aligned} \tag{17}$$

Now, the still unknown quantity $P_1(1)$, i.e. the probability that a given gap has the minimum size 1, should be established using the initial condition, where all gaps have the same size $\overline{g}$. Indeed, unlike the 1D dynamics, in the mean-field version the updates (12) and (13) preserve the average gap size. Therefore, $\langle g \rangle = \lim_{z\to 1^-} \frac{\mathrm{d}}{\mathrm{d}z} P_1(z; t) = \overline{g}$ holds at all times, so it must be satisfied also in the stationary state, on condition that the latter exists. From (17) we deduce that

$$\begin{aligned} \langle g \rangle = 1 + \frac{1-p}{p}P_1(1) \\ - \sqrt{\Big[\frac{1-p}{p}P_1(1)\Big]^2 - \frac{1-p}{p}\big(1 - P_1(1)\big)}. \end{aligned} \tag{18}$$

As a function of $P_1(1)$ this expression attains maximum at $P_1(1) = P_1^c(1) \equiv \frac{p}{1-p}\big(1/\sqrt{p} - 1\big)$. The maximum value of the average gap size is then

$$\langle g \rangle_{\max} = \frac{1}{\sqrt{p}}. \tag{19}$$

We must distinguish three regimes. For $\overline{g} < p^{-1/2}$ the one-gap distribution has exponential tail, because $\widehat{P}_1(z)$ has all derivatives at $z \to 1^-$. For example, in the limit $p \to 0$ we can compute the distribution explicitly

$$P_1(g) = (2 - \overline{g}^{-1})\big(\overline{g}^{-1} - 1\big)^{g-1} \tag{20}$$

for any value of $\overline{g}$.

The second regime pertains to the critical value $\overline{g} = p^{-1/2}$. If we insert the value $P_1^c(1)$ into the solution (17), we obtain

$$\widehat{P}_1^c(z) = 1 - \frac{1}{\sqrt{p}}(1-z) - \frac{(1-\sqrt{p})^2}{2p}(1-z)^2 + \frac{(1-\sqrt{p})^3}{p^{9/4}}|1-z|^{3/2}\sqrt{1 + \frac{(1-\sqrt{p})^2}{4\sqrt{p}}(1-z)}\,. \tag{21}$$

The $(1-z)^{3/2}$ singularity at $z \to 1$ is the fingerprint of the power-law tail in the distribution of gap sizes, with exponent $5/2$, so

$$P_1^c(g) \sim g^{-5/2}, \text{ for } g \to \infty. \tag{22}$$

The third regime with $\overline{g} > p^{-1/2}$ is the most subtle one. In infinite system, there is no stationary state. However, it is easy to understand what type of non-stationarity we face. As the time goes on, the gap distribution splits into two substantially different contributions. First, there is a portion proportional to the critical gap distribution $P_1^c(g)$. Second, there is a δ-function part which shifts to larger and larger values but its weight shrinks. Schematically we can write

$$P_1(g;t) \simeq \big(1 - \epsilon(t)\big)P_1^c(g) + \epsilon(t)\delta\big(g - g_{\max}(t)\big) \tag{23}$$

where $g_{\max}(t)$ should be interpreted as the size of the largest gap at time t. The functions $\epsilon(t)$ and $g_{\max}(t)$ are related by the requirement that the average gap size is conserved, so $\epsilon(t) = \big(\overline{g} - p^{-1/2}\big)/\big(g_{\max}(t) - p^{-1/2}\big)$.

In finite-size system the evolution does not proceed indefinitely, because $g_{\max}(t)$ cannot exceed the total sum of gaps $N\overline{g}$, as well as $\epsilon(t)$ cannot drop below the value $1/N$. A stationary state is reached after all. For large but still finite N we can easily deduce the stationary size of the single largest gap

$$g_{\max} = \big(\overline{g} - p^{-1/2}\big)N + p^{-1/2}. \tag{24}$$

Thus, the overall picture in the supercritical regime $\overline{g} > p^{-1/2}$ is as follows. We have in mind the situation for large but finite N. Most of the gaps are distributed according to the critical distribution $P_1^c(g)$ characterised by the power-law tail $\sim g^{-5/2}$. This part of the distribution does not depend neither on the probability p nor on the number of the gaps N and this is the relevant part to be taken into account when we ask about the distribution of one-step returns. So, we conclude that the return distribution has a power-law tail

$$P(r) \sim r^{-5/2}, \text{ for } r \to \infty, \tag{25}$$

for any value $p \geq 1/\overline{g}^2$. In this range, the return exponent $\alpha + 1 = 5/2$ is universal.

The remaining part of the gap distribution corresponds to the condensation of the excess total length of the gaps into the single largest gap. For finite system, its stationary length scales linearly with N, while in the infinite system it keeps growing forever. For the distribution of returns this is irrelevant, though. Indeed, the largest gap is to be interpreted as if the price axis was compactified by periodic boundary conditions, joining the highest price with the lowest one. The orders are then placed on a ring instead of a line or a segment. The largest gap is then the distance from the highest to the lowest order measured around the whole ring. Its size depends largely on the measure of the ring, i.e. on the way the boundary conditions were introduced, and does not bear any physically relevant information. To sum up, it is an artifact of the mean-field approximation.

So, for small p the distribution is exponential, while for large enough p it is power-law. One may ask a seemingly academic question, what happens if $p = 1$ strictly. Then, the stationary state is trivial, all gaps having size 1 except a single gap collecting all the remaining length. However, the transient state is of interest. We can easily find that at each time the distribution of gap lengths is exponential, $P_1(g) \propto a^{-g}$, but a depends on time. Interestingly, the result coincides exactly with Smoluchowski's solution of the coagulation kinetics [60].

4 Conclusions

We investigated a schematic model of stock market dynamics, based on placing and executing orders in the order book. We reduced the distinction between buy and sell orders to a minimum, which leads to formulation of one-dimensional model of pairwise interacting intervals. The only relevant parameter is the probability p that the two gaps collapse.

The simulations show that for any initial conditions, the dynamics settles down in a stationary state, characterised by the power-law distribution of interval sizes, with exponent $\simeq 5/2$. The same exponent $1 + \alpha \simeq 5/2$ is found in the tail of the distribution of returns. An important fact is that the exponent does nor depend on the parameter p and therefore it seems to be universal. On the other hand, the empirical value is larger, about $\alpha \simeq 3$ and also in the "classical" order book model by Maslov the value $\alpha = 2$ exceeds our result. This means that the Interacting Gaps model exaggerates large price changes. This may be due to the fact that by definition it does not take fully into account that the new orders are placed close to the current price. Therefore, in reality the price is trapped and squeezed in an area of high density of orders, bids on one side and asks on the other. This effect is suppressed in the IGM.

It seems rather evident that our model belongs to the same universality class as the aggregation-chipping models [67,68], which have the same critical exponent 5/2 in mean-field. However, the choice of interacting pairs is different. Ours is much more correlated and our model should differ significantly, when fluctuations play a role. We expect much higher level of spatial correlations in IGM than in aggregation-chipping models. Why these correlations are not reflected in the super-universal value 5/2 of the exponent is not clear to us; we shall return to this point several paragraphs below.

We also observed volatility clustering. We measured the autocorrelation function of absolute returns, which decays as a power law with exponent $\eta \simeq 0.5$, agreeing qualitatively with empirical data. Unfortunately, we were not able to analytical results in this direction. The temporal correlations of volatility reflect the spatial correlations of the gap lengths. Within the mean-field approximation the correlation of lengths of neighbouring gaps is absent and therefore also the temporal correlations in volatility disappear. We consider likely that such temporal correlations are absent also in one-dimensional aggregation models [67,68], but we have not checked it.

The measurement of the Hurst exponent shows certain ambiguity. As a short statement we can conclude that we found that $H \simeq 1/2$. This is exactly the random-walk value, lower than the empirical one. On the other hand we are much closer to reality than the results of e.g. the Maslov model and not farther than the improved models, which take into account evaporation of orders. Generically, we observe that in all "zero-intelligence" models, including IGM, the Hurst exponent is 1/2 at most. This also suggests that the over-diffusive empirical value $H \simeq 2/3$ might be due to some, maybe utterly simple, strategic behaviour. Speculations about various "zero-plus-epsilon-intelligence" models are open.

The striking result appears when we compare the simulations with analytically solvable mean-field version of the model. Despite of the fact that we have one-dimensional system on one side and infinite-dimensional variant on the other, the exactly found mean-field return exponent $1 + \alpha = 5/2$ is the same as the value obtained numerically in one dimension. In fact, in the aggregation-chipping model, which apparently belongs to the same universality class, it was proved [70] that the phase diagram does not depend on spatial dimension, so the mean-filed and one-dimensional results coincide. The same work also presents a strong argument in favour of super-universality of the exponent 5/2. We believe the same argument holds also in our model. Analogous behaviour was also observed in one dimensional rice-pile model [72], but general explanation of the effect eludes us.

Finally, let us mention the possible modification including a "zero-intelligence" market maker. In the order-book models mentioned so far the update was sequential, as the orders were executed one by one. The role of the market maker can be mimicked by partially parallel update of the order book. The orders are processed in chunks, containing larger or smaller number of individual orders. This way we can incorporate here some of the features of the Minority Game [19], one of the prominent abstract models of a stock market. The possible fruits of such fusion are subject of our further study.

This work was supported by the MŠMT of the Czech Republic, grant No. 1P04OCP10.001, and by the Research Program CTS MSM 0021620845. We wish to thank to the anonymous referee, who pointed us to references [66–70] we had not known before.

References

1. *The Economy as an Evolving Complex System*, edited by P.W. Anderson, K.J. Arrow, D. Pines (Addison Wesley, Reading, 1988)
2. R.N. Mantegna, H.E. Stanley, *Introduction to Econophysics: Correlations Complexity in Finance* (Cambridge University Press, Cambridge, 1999)
3. J.-P. Bouchaud, M. Potters, *Theory of Financial Risks* (Cambridge University Press, Cambridge, 2000)
4. *Modeling Complexity in Economic and Social Systems*, edited by F. Schweitzer (World Scientific, Singapore, 2002)
5. R.N. Mantegna, H.E. Stanley, Nature **376**, 46 (1995)
6. Y. Liu, P. Cizeau, M. Meyer, C.-K. Peng, H.E. Stanley, Physica A **245**, 437 (1997)
7. P. Gopikrishnan, M. Meyer, L.A.N. Amaral, H.E. Stanley, Eur. Phys. J. B **3**, 139 (1998)
8. V. Plerou, P. Gopikrishnan, L.A.N. Amaral, M. Meyer, H.E. Stanley, Phys. Rev. E **60**, 6519 (1999)
9. N. Vandewalle, M. Ausloos, Eur. Phys. J. B **4**, 257 (1998)
10. J.F. Muzy, J. Delour, E. Bacry, Eur. Phys. J. B **17**, 537 (2000)
11. L. Kullmann, J. Töyli, J. Kertesz, A. Kanto, K. Kaski, Physica A **269**, 98 (1999)
12. J.-P. Bouchaud, A. Matacz, M. Potters, Phys. Rev. Lett. **87**, 228701 (2001)
13. F. Lillo, S. Mike, J.D. Farmer, Phys. Rev. E **71**, 066122 (2005)
14. H. Takayasu, H. Miura, T. Hirabayashi, K. Hamada, Physica A **184**, 127 (1992)
15. P. Bak, K. Chen, J. Scheinkman, M. Woodford, Ricerche Economiche **47**, 3 (1993)
16. M. Levy, H. Levy, S. Solomon, J. Phys. I France **5**, 1087 (1995)
17. D. Sornette, R. Cont, J. Phys. I France **7**, 431 (1997)
18. G. Caldarelli, M. Marsili, Y.-C. Zhang, Europhys. Lett. **40**, 479 (1997)
19. D. Challet, Y.-C. Zhang, Physica A **246**, 407 (1997)
20. A. Arnéodo, J.-F. Muzy, D. Sornette, Eur. Phys. J. **2**, 277 (1998)
21. N.F. Johnson, M. Hart, P.M. Hui, D. Zheng, Int. J. Theor. Appl. Finance **3**, 443 (2000)
22. A. Cavagna, J.P. Garrahan, I. Giardina, D. Sherrington, Phys. Rev. Lett. **83**, 4429 (1999)
23. G. Iori, J. Economic Behavior and Organization **49**, 269 (2002)
24. T. Lux, M. Marchesi, Nature **397**, 498 (1999)
25. R. Cont, J.-P. Bouchaud, Macroeconomic Dynamics **4**, 170 (2000)
26. V.M. Eguíluz, M.G. Zimmermann, Phys. Rev. Lett. **85**, 5659 (2000)
27. M.F.M. Osborne, Econometrica **33**, 88 (1965)

28. D.K. Gode, S. Sunder, The J. Political Economy **101**, 119 (1993)
29. P. Bak, M. Paczuski, M. Shubik, Physica A **246**, 430 (1997)
30. L.-H. Tang, G.-S. Tian, Physica A **264**, 543 (1999)
31. D. Eliezer, I.I. Kogan, e-print `arXiv:cond-mat/9808240`
32. S. Maslov, Physica A **278**, 571 (2000)
33. L. Kullmann, J. Kertész, Physica A **299**, 234 (2001)
34. F. Franci, R. Marschinski, L. Matasini, Physica A **294**, 213 (2001)
35. L. Matassini, F. Franci, Physica A **289**, 526 (2001)
36. J.D. Farmer, S. Joshi, J. Economic Behavior and Organization **49**, 149 (2002)
37. E. Smith, J.D. Farmer, L. Gillemot, S. Krishnamurthy, Quant. Finance **3**, 481 (2003)
38. M.G. Daniels, J.D. Farmer, L. Gillemot, G. Iori, E. Smith, Phys. Rev. Lett. **90**, 108102 (2003)
39. G. Iori, M.G. Daniels, J.D. Farmer, L. Gillemot, S. Krishnamurthy, E. Smith, Physica A **324**, 146 (2003)
40. J.-P. Bouchaud, M. Mézard, M. Potters, Quant. Finance **2**, 251 (2002)
41. J.-P. Bouchaud, Y. Gefen, M. Potters, M. Wyart, Quant. Finance **4**, 176 (2004)
42. D. Challet, R. Stinchcombe, Physica A **300**, 285 (2001)
43. R.D. Willmann, G.M. Schütz, D. Challet, Physica A **316**, 430 (2002)
44. D. Challet, R. Stinchcombe, e-print `arXiv:cond-mat/0208025`
45. D. Challet, R. Stinchcombe, Physica A **324**, 141 (2003)
46. D. Challet, e-print `arXiv:physics/0608013` (2006)
47. T. Halpin-Healy, Y.-C. Zhang, Phys. Rep. **254**, 215 (1995)
48. S. Maslov, M. Mills, Physica A **299**, 234 (2001)
49. X. Gabaix, P. Gopikrishnan, V. Plerou, H.E. Stanley, Nature **423**, 267 (2003)
50. V. Plerou, P. Gopikrishnan, H.E. Stanley, Phys. Rev. E **71**, 046131 (2005)
51. P. Weber, B. Rosenow, Quant. Finance **5**, 357 (2005)
52. P. Weber, B. Rosenow, Quant. Finance **6**, 7 (2006)
53. M. Potters, J.-P. Bouchaud, Physica A **324**, 133 (2003)
54. I. Zovko, J.D. Farmer, Quant. Finance **2**, 387 (2002)
55. F. Lillo, J.D. Farmer, R.N. Mantegna, Nature **421**, 129 (2003)
56. J.D. Farmer, L. Gillemot, F. Lillo, S. Mike, A. Sen, e-print `arXiv:cond-mat/0312703`
57. J.D. Farmer, P. Patelli, I.I. Zovko, Proc. Natl. Acad. Sci. USA **102**, 2254 (2005)
58. F. Slanina, Phys. Rev. E **64**, 0561136 (2001)
59. L. Muchnik, F. Slanina, S. Solomon, Physica A **330**, 232 (2003)
60. M.v. Smoluchowski, Phys. Z. **17**, 557 (1916)
61. H. Takayasu, I. Nishikawa, H. Tasaki, Phys. Rev. A **37**, 3110 (1988)
62. H. Takayasu, Phys. Rev. Lett. **63**, 2563 (1989)
63. S.N. Majumdar, C. Sire, Phys. Rev. Lett. **71**, 3729 (1993)
64. D. ben-Avraham, M.A. Burschka, C.R. Doering, J. Stat. Phys. **60**, 695 (1990)
65. S.N. Majumdar, D.A. Huse, Phys. Rev. E **52**, 270 (1995)
66. R.D. Vigil, R.M. Ziff, B. Lu, Phys. Rev. B **38**, 942 (1988)
67. P.L. Krapivsky, S. Redner, Phys. Rev. E **54**, 3553 (1996)
68. S.N. Majumdar, S. Krishnamurthy, M. Barma, J. Stat. Phys. **99**, 1 (2000)
69. S.N. Majumdar, S. Krishnamurthy, M. Barma, Phys. Rev. Lett. **81**, 3691 (1998)
70. R. Rajesh, S.N. Majumdar, Phys. Rev. E **63**, 036114 (2001)
71. F. Slanina, Phys. Rev. E **69**, 046102 (2004)
72. F. Slanina, Eur. Phys. J. B **25**, 209 (2002)