Understanding the complex dynamics of stock markets through cellular automata

G. Qiu, D. Kandhai, and P. M. A. Sloot
Section Computational Science, Faculty of Science, University of Amsterdam, Kruislaan 403, 1098 SJ Amsterdam, The Netherlands

We present a cellular automaton (CA) model for simulating the complex dynamics of stock markets. Within this model, a stock market is represented by a two-dimensional lattice, of which each vertex stands for a trader. According to typical trading behavior in real stock markets, agents of only two types are adopted: fundamentalists and imitators. Our CA model is based on local interactions and adopts simple rules for representing the behavior of traders and a simple rule for price updating. This model can reproduce, in a simple and robust manner, the main characteristics observed in empirical financial time series. Heavy-tailed return distributions due to large price variations can be generated through the imitation behavior of agents. In contrast to other microscopic simulation (MS) models, our results suggest that it is not necessary to assume a certain network topology in which agents group together, e.g., a random graph or a percolation network. Long-range interactions can form from local interactions. Volatility clustering, which also leads to heavy tails, seems to be related to the combined effect of a fast process and a slow process: the evolution of the influence of news and the evolution of agents' activity respectively. In a general sense, these causes of heavy tails and volatility clustering appear to be common among some notable MS models that can confirm the main characteristics of financial markets.

PACS numbers: 89.65.Gh

I. INTRODUCTION

The complex dynamics of financial markets can be characterized by some "stylized facts", which are common across many financial instruments, markets, and time horizons. Most of them are counter-intuitive and contrary to the expectations of traditional financial theories. These stylized features have been observed or discussed in many independent studies [1–7]. On long time scales (typically a week or longer), empirical distributions of financial return [27] generally fit to the Gaussian distribution. However, most financial returns over short timescales are described well by a non-Gaussian (heavy-tailed or fat-tailed) distribution. A commonly used, although not rigorous, criterion for the normality of a distribution is its kurtosis (k): $k = 3$ corresponds to a Gaussian distribution, $k > 3$ indicates a so-called leptokurtic distribution with a sharp peak and heavy tails. The kurtosis of financial returns is far from that of a Gaussian distribution. Our estimate for the kurtosis of S&P 500 [28] daily returns over the period June 1950 to June 2005 is around 38. Figure 1(a) shows the distribution of these returns, together with a Gaussian probability density function (PDF) and a Lorentz PDF for comparison. Clearly, daily returns of S&P 500 follow a non-Gaussian (fat-tailed) distribution, implying a greater frequency of extreme events than would be expected if they followed a normal distribution. However, the variance of the distribution is finite, whereas that of a Lorentz distribution (or a stable Lévy distribution in general) is infinite. Furthermore, the autocorrelation function (ACF) of the daily returns quickly converges to the noise range, whereas the corresponding ACF of volatility [29] decays slowly (see Fig. 1(b)). The long-term autocorrelation of volatility is the reflection of the phenomenon termed "volatility clustering" — high (positive or negative) returns tend to group together. Figure 1(c) shows the time series of return over the period. In this figure, the effect of volatility clustering is clearly illustrated.

Traditional (analytical) approaches in finance and economics to aggregative phenomena either are purely macroscopic, or rely on top-down construction based on a number of unrealistic assumptions mainly for the sake of analytical tractability. Interactions between traders play no role in the explanation of the phenomena [8]. In fact, markets consist of a large number of agents. The interactions between them give rise to some macroeconomic regularity, which in turn influences the microscopic interactions. This dynamics is highly non-linear and difficult to describe analytically.

In recent years, researchers have used microscopic simulation to explore complex economic dynamics from bottom up. With MS, we study a complex system by directly modeling its individual agents and their interactions. The macroscopic behavior of the system will eventually emerge from the micro-dynamics. MS has shown great potential for more realistically modeling complex dynamical systems in economics and finance [9]. In addition, it facilitates the testing of existing economic or financial models and theories, and the development of new theories and models [8]. At present, most research of MS in finance focuses on understanding the characteristics of financial markets. To achieve this objective, many MS models of financial markets have been developed during the last decade.

However, as shown in Sec. II, researchers in this field have not yet reached an agreement on explaining the

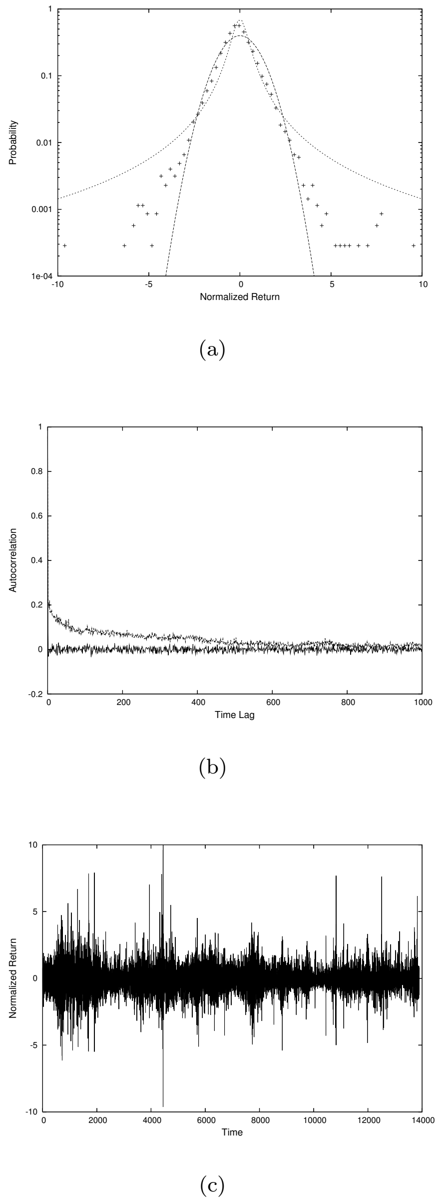


FIG. 1: (a) Distribution of the daily returns of S&P 500 over the period from June 1950 to June 2005 (the points), compared with a Gaussian PDF (the curve that decays faster) and a Lorentz PDF. A logarithmic scale is used for the vertical axis. (b) Autocorrelation function of the daily returns (the lower line) and the corresponding ACF of volatility. (c) Time series of the daily returns.

complex dynamics of financial markets characterized by the stylized facts. In addition, as recently pointed out by Cont [10], due to the complexity of the existing (agent-based) models, it is often not clear which aspects of the models are responsible for generating the stylized facts and whether all their ingredients are indeed required for explaining empirical observations.

In view of this fact, our general motivation is to develop a MS model with a simple structure that can reproduce the main stylized facts. More importantly, the causalities of the dynamics generated by the model can be clearly identified. In particular, we wish to confirm the main stylized features within a parsimonious CA framework by dealing with local agent interactions and adopting simple rules for representing agents' behavior and a simple rule for price updating. For this reason, we have taken a simple CA model developed by Bandini *et al.* [11], which is not able to reproduce the stylized facts, as our starting point and included simple modifications that have clear economic meanings.

In Sec. II, we first give an overview of the most notable MS models of financial markets that have been reported in literature and the CA model that is our starting point. We then give a detailed description of our CA model in Sec. III. We present it in the order of increasing sophistication, so that the cause(s) of certain stylized fact(s) can be identified at each level of sophistication. The simulation results of our model is shown in Sec. IV. Section V is devoted to a thorough investigation of the simulated dynamics through computational experiments and mathematical analysis. In the final section, we present our conclusions.

II. OVERVIEW OF SOME MS MODELS

P. Bak *et al.* have developed a MS model [12] of which the stock market contains fundamental value traders and noise traders. The former set their prices based on a utility function. The behavior of the latter is characterized by drifting their prices to spot prices and copying other traders' prices (imitation). The strength of their drifting is proportional to recent price variation. This mechanism is called "volatility feedback". The model can generate non-Gaussian distributions of return and volatility clustering. "Volatility feedback" of this model is similar to the mechanism of activity adjustment within our CA model, whereas other features are different.

Within the model of T. Lux and M. Marchesi [13], the market consists of two groups of traders: fundamentalists and noise traders. Fundamentalists buy (sell) the asset when its market price is below (above) its fundamental value. Noise traders are further differentiated into optimists and pessimists. The former believe in a rising market and buy the asset, whereas the latter believe in a declining market and sell. Agents move to the other group or subgroup when they think that the traders there are more successful (imitation or herding). Volatility clustering and fat-tailed distributions of return can be produced by this model. While they are different in other aspects, this model and our model are virtually identical with regard to the behavior of fundamentalists and price updating.

The model of R. Cont and J.-P. Bouchaud [14] deals with homogeneous traders who group together in clusters through binary links between them. Such a structure is known as a random graph. In a certain condition, the cluster sizes follow a power law distribution. The members in each cluster coordinate their individual demands to decide whether to buy, sell, or not to trade (herding). This model gives rise to probability distributions of price change with heavy tails. However, it cannot generate volatility clustering. In principle, the price updating rule of this model is identical to the corresponding rule within our model, although these two models are very different in other features.

Cellular automata have been widely applied to study complex phenomena in different fields such as physics, chemistry, biology, social and economic sciences, etc. [15]. Recently, some researchers have used cellular automata to model financial markets for studying their complex dynamics.

G. Iori has developed an CA-type model [16] that represents a market as a two-dimensional lattice, of which each node is an agent connecting with four neighbors. Each agent's decision making is driven by his own signal and the signals of his neighbors (imitation). When the aggregate signal exceeds his activation threshold, he will transact. Agents' thresholds are adjusted over time in response to price changes. This model generates fat-tailed distributions of return and volatility clustering. Differing from each other in other features, both this model and our model adopt local interactions.

In [17], M. Bartolozzi and A. W. Thomas present a stochastic CA model of stock markets. In simulations using this model, clusters of active traders form and evolve over time through percolation dynamics (herding). This process produces a power law distribution of cluster size. Similar to the process of a random Ising model, traders within each cluster exchange information and update their states. The model can produce heavy tails and volatility clustering. While this model and our model are distinct in other aspects, they are similar in price updating.

The model of S. Bandini *et al.* [11] adopts agents of two types: fundamentalists and imitators. The former trade in quantities proportional to the differences between their perceived fundamental values of the asset and spot prices. The latter follow the actions of their neighbors (imitation). Ref. [11] does not report any stylized facts. This model and our CA model are similar in concept with regard to the basic behavior of agents and price updating, whereas other features are different.

We can see that, although most of these MS models can confirm the stylized facts, they are different in agent types, description of agents' behavior and ultimately market dynamics. However, it should be noticed that all of them involve a typical kind of behavior in financial markets, namely imitation. Some of these models have been studied in Ref. [18].

III. DESCRIPTION OF OUR CA MODEL

A. Modeling stock markets with cellular automata

We represent a stock market as a two-dimensional $L \times L$ lattice, every vertex of which contains an agent. Each agent has a radius one (Moore) neighborhood. Within our model, speculative traders of only two types are adopted: fundamentalists and imitators. All the agents trade in a single stock.

Fundamentalists are those traders who are informed of the nature of the stock being traded and in agreement with regard to its fundamental value. They believe that the price of the stock may temporarily deviate from, but will eventually return to the fundamental value. They therefore buy (sell) the asset whenever its price is lower (higher) than its fundamental value perceived by them. In stock markets, there are also some traders who do not know or do not care about fundamental values. A trader of this type will follow his acquaintances and adopt the majority trading opinion in the light of the current price. An agent of this type is referred to as an imitator.

News influences both fundamentalists and imitators. However, the effects of news on them are distinct in many aspects. For example, fundamentalists pay comparatively more attention to news about the specific company that has issued the stock, while imitators respond comparatively more frequently to news related to the whole stock market.

We can adopt other types of agents to model stock markets more realistically. However, we think that the two kinds of behavior discussed here are the most typical. The behavior of other speculative traders has no obvious characteristics. For example, we cannot find a general trait of chartists, because even using the same data they may come to different conclusions due to distinctions among their techniques. We can treat these agents as noise traders who randomly influence the price to different extents. However, because their adoption within our model does not fundamentally influence the dynamics characterized by the stylized facts, we ignore them.

The real fundamental value of a stock is related to the current and prospective states of the company that has issued the stock, among many others. The modeling of its variations is beyond the scope of this work. Instead, we are more interested in the reason(s) for excess volatility, i.e., the extra factor(s) causing the price of a stock to be more volatile than its real fundamental value. For this reason, within our model we assume that the real fundamental value of the asset (F) is a constant. (We can alternatively add a drift to F to denote the time value of money without influencing the characteristics of the returns.)

B. Level I model

We consider each agent's trading opinion in the light of the current price level. For the moment, we assume that it is the only factor determining his transaction quantity.

Fundamentalists

Empirically, the larger the difference between the price of a stock and its fundamental value perceived by a fundamentalist, the more likely he will trade it (or statistically trade it more). We assume for the moment that the fundamentalists perceive the real fundamental value accurately. We can then adopt Eq. (1) to express the trading opinion at time $t+1$ of a fundamentalist located at i, $V_{i,fu}^{t+1}$, and his transaction quantity at the same time, $q_{i,fu}^{t+1}$:

$$\begin{aligned} q_{i,fu}^{t+1} &= V_{i,fu}^{t+1} \\ &= F - P^t, \end{aligned} \tag{1}$$

where F is the real fundamental value, P^t is the price at time t.

Imitators

We take the average trading opinion of an imitator's neighbors at the previous time step as his current trading opinion, i.e., $V_{i,im}^{t+1} = \langle V_{i,nb}^t \rangle$. We can then use Eq. (2) to express the transaction quantity of an imitator located at i at time $t+1$, $q_{i,im}^{t+1}$:

$$\begin{aligned} q_{i,im}^{t+1} &= V_{i,im}^{t+1} \\ &= \langle V_{i,nb}^t \rangle. \end{aligned} \tag{2}$$

C. Level II model

Within this model, we further consider the influence of continuously arriving news on traders' trading opinions.

Fundamentalists

News influences fundamentalists' perceptions of fundamental values. Positive (negative) news can cause them to overestimate (undervalue) values of assets. Within our model, we assume that at each time step, all the fundamentalists perceive the fundamental value identically. (We can alternatively assume that their perceived values at each time step are normally distributed, without fundamentally influencing the dynamics.)

We express the perceived fundamental value at time t as $F\eta_{fu}^t$, in which η_{fu}^t denotes the influence of the news at that time. We assume that $\eta_{fu}^t = 1 + c_{fu}\phi_{fu}^t$, where $\phi_{fu}^{(\cdot)}$ is an independent Gaussian random variable with mean 0 and standard deviation 1, c_{fu} is a parameter indicating the fundamentalists' sensitivity to news. At this point, we have a modified expression for the transaction quantity of a fundamentalist,

$$\begin{aligned} q_{i,fu}^{t+1} &= V_{i,fu}^{t+1} \\ &= F\eta_{fu}^{t+1} - P^t. \end{aligned} \tag{3}$$

Imitators

We also assume that news influences all the imitators identically. (We can alternatively assume that the effects of news on them at each time step are normally distributed, without fundamentally influencing the dynamics.) Significant (unimportant) news can make an imitator trade more (less) than his neighbors, and vice versa. We reformulate the transaction quantity of an imitator as

$$\begin{aligned} q_{i,im}^{t+1} &= V_{i,im}^{t+1} \\ &= \langle V_{i,nb}^t \rangle \eta_{im}^{t+1}, \end{aligned} \tag{4}$$

in which η_{im}^{t+1} indicates the influence of the news at time $t+1$ and is equal to $1 + c_{im}\phi_{im}^{t+1}$, where $\phi_{im}^{(\cdot)}$ is an independent Gaussian random variable with mean 0 and standard deviation 1, c_{im} is a parameter indicating the imitators' sensitivity to news.

D. Level III model

A common strategy used by traders is to buy low and sell high (BLASH). It aims for capital gains by taking advantage of changes in prices. Price fluctuations are therefore indispensable to this strategy.

Based on BLASH, capitals of traders move among different assets pursuing larger profits at lower risks. When the price fluctuation level of a stock is at the two extremities, i.e., very low and very high, the asset is the least desirable to this strategy: If it is very low, traders who hold the asset will not be able to find an opportunity to sell it profitably and will not even be able to cover their opportunity costs [30]. If it is very high, traders will consider the investment in the asset too risky. In the range between the two extremities, as the price fluctuation level rises, the desirability of the asset will first increase and then decrease.

When a stock is more desirable than other alternatives, traders will trade it more, and vice versa. However, BLASH is a risky approach itself, because there is no way to predict price changes accurately. Frequently, traders just end up going in the opposite direction, i.e., selling at a loss. In order to reduce this risk, traders typically consider more lagged price changes of a stock to figure out its desirability. They then gradually adjust their trading activity in accordance with its desirability. For the sake of simplicity, we assume that the desirability of a stock is identical for every agent and that, at each time step, the trading activity of each agent is equal to the desirability of the stock.

We represent the price fluctuation level of a stock at time t as

$$L^t = \frac{1}{k}\sum_{i=t-k}^{t-1} |P^i - \bar{P}|/\bar{P}, \tag{5}$$

where k is the length of a period before t, P^i is the price of the asset at time i in the period, $\bar{P}$ is the average price over the period. (We can alternatively assume that agents take different values of k that are normally distributed, without fundamentally influencing the dynamics.)

For the sake of simplicity, we adopt a straightforward function for the desirability of a stock (and hence agents' trading activity of it), M^t. It starts from zero as a monotonically increasing straight line then, at a certain point, becomes a monotonically decreasing straight line:

$$M^t = \begin{cases} c_l L^t, & L^t \leq L_m \\ c_l(-L^t + 2L_m), & L^t > L_m \end{cases} \tag{6}$$

where L_m is the value of L^t at the turning point of M^t, c_l is a parameter. (Simulations show that adopting other concave functions [31] with a same domain and a same range leads to qualitatively similar results.)

Within the level III model we consider that the transaction quantity of an agent is the product of his trading opinion in the light of the current price level and his current trading activity. The quantity of a fundamentalist is therefore

$$\begin{aligned} q_{i,fu}^{t+1} &= V_{i,fu}^{t+1} M^{t+1} \\ &= (F\eta_{fu}^{t+1} - P^t) M^{t+1}, \end{aligned} \tag{7}$$

whereas the quantity of an imitator is

$$\begin{aligned} q_{i,im}^{t+1} &= V_{i,im}^{t+1} M^{t+1} \\ &= \langle V_{i,nb}^{t} \rangle \eta_{im}^{t+1} M^{t+1}. \end{aligned} \tag{8}$$

Considering the fact that agents always have a number of exceptional reasons to transact, we adopt a lower bound for $M^{(\cdot)}$.

E. Rule of price updating

The price is updated according to the following rule:

$$P^{t+1} = P^t + \frac{c_p Q^t}{N}, \tag{9}$$

where Q^t is the total transaction quantity (or the excess demand) for the asset at time t, N is number of traders. Since Q^t is proportional to N, we rescale it with N. We adopt c_p to indicate the sensitivity of the price to the excess demand. Due to the fact that stock prices cannot be negative, the lower bound of $P^{(\cdot)}$ is 0.

Equation (9) can be explained as the action of a market maker [32] to balance the supply and the demand of the stock. In principle, however, it is merely the translation of the classic theory of supply and demand stating that price will move toward the point that equalizes quantities supplied and demanded [33].

IV. SIMULATION RESULTS

A. Simulation results of the level I model

In simulations using the level I model, we set an initial price ($P^0 = 105$) that deviates from the fundamental value ($F = 100$). The number of agents is 1×10^4. Figure 2 displays the price trajectories corresponding to two fractions of imitators: $\alpha_{im} = 20\%$ and 80% respectively.

The parameter c_p has an important impact on the price: When its value is increased up to 1 (other parameters are kept constant), the price process may start to switch from a convergent process to a divergent one, depending on the value of c_p itself and the value of α_{im}. We provide a theoretical analysis of this issue in Sec. V. To model a stable market, we adopt only those values of c_p smaller than 1.

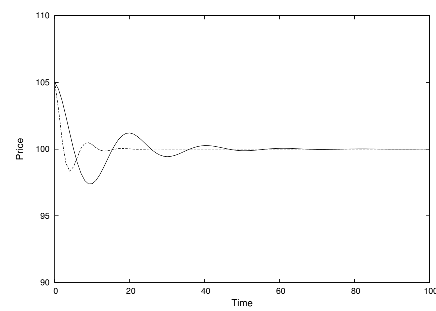


FIG. 2: Price trajectories obtained through the simulation using the level I model when $\alpha_{im} = 20\%$ and 80% respectively. The curve that decays faster is of the first instance. Parameter setting used: $N = 1 \times 10^4$, $c_p = 0.5$, $F = 100$, and $P^0 = 105$.

As shown in Fig. 2, although the level I model is not completely identical to the model of Bandini *et al.* [11], it generates price trajectories similar to those produced by the latter. Starting from an initial deviation from the fundamental value, the price either directly converges to it, or fluctuates around it for some time and eventually overlaps with it. Obviously, both models cannot produce sustained price movement. Besides, as the price quickly dies out, we cannot obtain any stylized fact by using the level I model.

B. Simulation results of the level II model

Within the level II model, we have added random factors η^t_{fu} and η^t_{im}, so that it can produce sustained price fluctuations. When the fraction of imitators is set to 70%, we obtain simulation results shown in Fig. 3. In our simulations, return is represented by the difference between two successive natural logarithms of price, i.e., log-return.

We see that the level II model can generate a non-Gaussian (fat-tailed) distribution of return, but is not able to confirm another important stylized fact, namely volatility clustering. It therefore has the same problem as the Cont-Bouchaud model [14] has. Nevertheless, through simulations using this model, we can further study how the fraction of imitators influences the distribution of return. Figure 4 shows the results for different instances: $\alpha_{im} = 20\%$, 50%, and 80% respectively. If the fraction is small, returns will follow a Gaussian distribution. Increasing it enlarges the tails of the return distribution.

C. Simulation results of the level III model

Within the level III model we have further added a mechanism through which agents' activity is adjusted over time. Fixing α_{im} to 70%, we obtain simulation results shown in Fig. 5, in which

1. Fig. 5(a) records the price process. Some large "flights" can be observed, which correspond to large (positive or negative) returns.
2. Fig. 5(b) illustrates the time series of return. The effect of volatility clustering is clear.
3. Fig. 5(c) shows the probability distribution of return, together with a Gaussian PDF and a Lorentz PDF for comparison. The tails of the distribution are clearly heavier than those of a Gaussian PDF.
4. Fig. 5(d) displays the ACF of return (the lower curve) and that of volatility. The former converges quickly to the noise range, whereas the latter decays much more slowly.
5. Fig. 5(e) shows the time evolution of transaction volume [34].
6. Fig. 5(f) illustrates the time evolution of trading activity. It is a slow process in comparison with the fast evolution of the influence of news [35].

These simulation results indicate that our CA model (level III) is able to reproduce the main stylized facts. In addition, as shown below, this model is robust with regard to the stylized facts for wide ranges of the parameters.

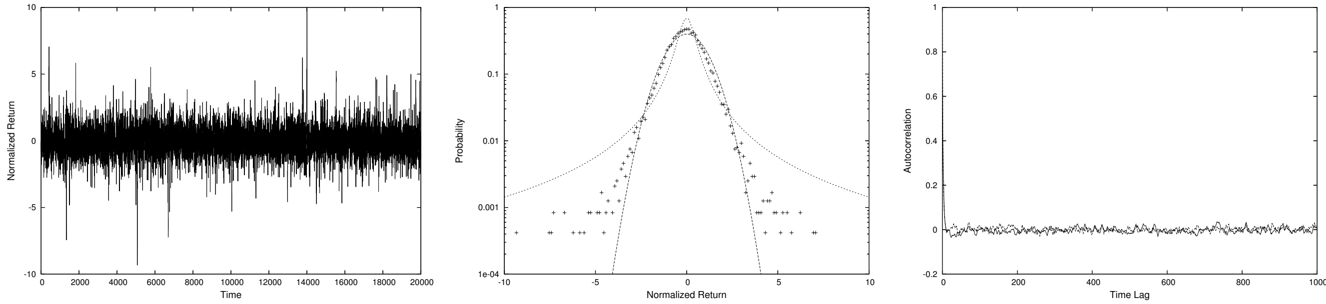
Table I compiles the kurtosis values of the return distributions for different values of α_{im}, c_{im} and c_{fu} respectively. We see that imitators have a strong influence on kurtosis, while the relation between fundamentalists and kurtosis is not explicit. Specifically, the fraction of imitators (α_{im}) and the sensitivity of imitators to news (c_{im}) are positively correlated with kurtosis. When α_{im} or c_{im} increase to a certain level, kurtosis suddenly becomes very large, implying that the system becomes unstable. For example, when $c_{im} = 0.9$, we obtain a price pattern with frequent dramatic 'flights' and a time series of return with many striking strokes. These are shown in Fig. 6.

TABLE I: Kurtosis values of the return distributions produced by the level III model for increasing values of α_{im}, c_{im} and c_{fu} respectively. Parameter setting used: $N = 1 \times 10^4$, $c_p = 0.05$, $F = 100$, $c_{fu} = 0.2$, $c_{im} = 0.7$, $P^0 = 100$, $k = 400$, $c_l = 20$ and $L_m = 0.01$. The lower bound of $M^{(\cdot)}$ is 0.05.

α_{im}	Kurtosis	c_{im}	Kurtosis	c_{fu}	Kurtosis
0%	4.44	0.1	4.36	0.1	15.53
10%	4.36	0.2	5.37	0.2	17.22
20%	4.57	0.3	6.37	0.3	42.94
30%	5.33	0.4	7.44	0.4	37.28
40%	6.07	0.5	8.60	0.5	35.39
50%	6.64	0.6	10.55	0.6	46.71
60%	8.51	0.7	17.22	0.7	43.98
70%	17.22	0.8	32.57	0.8	42.11
80%	69.33	0.9	119.14	0.9	35.79
90%	190.46	1.0	422.77	1.0	30.72

Keeping other parameters constant and adopting different values of k, we obtain autocorrelation functions of volatility shown in Fig.7(a). When k is smaller than 50, ACFs of volatility quickly drop to the noise range and correspondingly, the effects of volatility clustering are negligible. Volatility clustering becomes significant when k is increased to around 100. The importance of k to volatility clustering is further discussed in Sec. V D.

Choosing three values for the number of agents (lattice sizes) and keeping other parameters constant, simulations using our model give ACFs of volatility shown in Fig. 7(b). All these ACFs are qualitatively similar to that of S&P 500 shown in Fig. 1(b), indicating that the model can reproduce the stylized facts not only for markets with small numbers of agents, but also for very large markets. At this point, it differs from some MS models that behave realistically only for limited numbers of traders but not for large markets [22].



(a) (b) (c)

FIG. 3: Simulation results of the level II model when $\alpha_{im} = 70\%$. (a) Normalized return. (b) Distribution of return (the scale of the vertical axis is logarithmic). (c) Autocorrelation function of return and that of volatility. Parameter setting used: $N = 1 \times 10^4$, $c_p = 0.005$, $F = 100$, $c_{fu} = 0.2$, $c_{im} = 0.7$, and $P^0 = 100$.

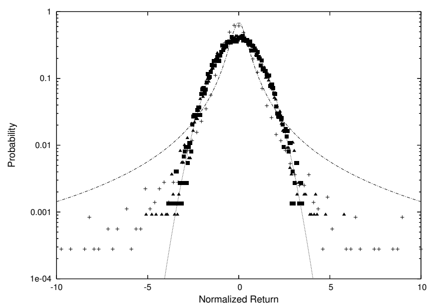


FIG. 4: Return distributions of the level II model for different fractions of imitators. The ■ points, the ▲ points and the + points refer to $\alpha_{im} = 20\%$, 50%, and 80% respectively. The scale of the vertical axis is logarithmic. Parameter setting used: $N = 1 \times 10^4$, $c_p = 0.005$, $F = 100$, $c_{fu} = 0.2$, $c_{im} = 0.7$, and $P^0 = 100$.

V. DISCUSSION: THE MARKET DYNAMICS REVEALED BY OUR CA MODEL

A. Long-range interactions can form from local interactions

In this section, for the sake of simplicity, we take a one-dimensional version of our CA model to derive analytical expressions. An agent located at i then has two neighbors at $i-1$ and $i+1$ respectively. Statistically, the total quantity at time $t+1$ can be expressed as

$$Q^{t+1} = \sum_{i=1}^{N} q_{i,(\cdot)}^{t+1} = \sum_{i=1}^{N} [u_i q_{i,fu}^{t+1} + (1-u_i) q_{i,im}^{t+1}], \quad (10)$$

where u_i is determined in the following way: We sample a variable γ ($0 \le \gamma \le 1$) that is uniformly distributed. If $0 \le \gamma \le \alpha_{fu}$, $u_i = 1$, else $u_i = 0$. Here, α_{fu} is the fraction of fundamentalists.

The term $q_{i,fu}^{t+1}$ and $q_{i,im}^{t+1}$ in Eq. (10) are determined by Eq. (7) and Eq. (8) respectively. However, because M^t changes much more slowly than Q^t, we can consider the former as a constant to study the basic dynamics of the latter. We set $M^{(\cdot)} = 1$, then $q_{i,fu}^{t+1}$ and $q_{i,im}^{t+1}$ are respectively determined by Eq. (3) and Eq. (4). Therefore,

$$\begin{aligned} q_{i,im}^{t+1} &= \eta_{im}^{t+1}(\frac{1}{2})[V_{i-1,(\cdot)}^{t} + V_{i+1,(\cdot)}^{t}] \\ &= \eta_{im}^{t+1}(\frac{1}{2})[[u_{i-1}V_{i-1,fu}^{t} + (1-u_{i-1})V_{i-1,im}^{t}] \\ &\quad +[u_{i+1}V_{i+1,fu}^{t} + (1-u_{i+1})V_{i+1,im}^{t}]]. \end{aligned} \quad (11)$$

Similarly, the terms $V_{i-1,im}^{t}$ and $V_{i+1,im}^{t}$ in Eq. (11) can be respectively expressed as

$$\begin{aligned} V_{i-1,im}^{t} &= \eta_{im}^{t}(\frac{1}{2})[[u_{i-2}V_{i-2,fu}^{t-1} + (1-u_{i-2})V_{i-2,im}^{t-1}] \\ &\quad +[u_{i}V_{i,fu}^{t-1} + (1-u_{i})V_{i,im}^{t-1}]], \end{aligned} \quad (12)$$

and

$$\begin{aligned} V_{i+1,im}^{t} &= \eta_{im}^{t}(\frac{1}{2})[[u_{i}V_{i,fu}^{t-1} + (1-u_{i})V_{i,im}^{t-1}] \\ &\quad +[u_{i+2}V_{i+2,fu}^{t-1} + (1-u_{i+2})V_{i+2,im}^{t-1}]]. \end{aligned} \quad (13)$$

Following the same scheme, we can further express the terms $V_{i-2,im}^{t-1}$, $V_{i,im}^{t-1}$ and $V_{i+2,im}^{t-1}$ in Eqs. (12) and (13) in terms of the trading opinions at time step $t-2$ of the neighbors of the agents located at $i-2$, i and $i+2$ respectively, and so on. Basically, in this way, we can replace the trading opinion of every imitator at each time step with the trading opinions of fundamentalists at the preceding time steps, noting that $V_{i,fu}^{t}$ is independent of

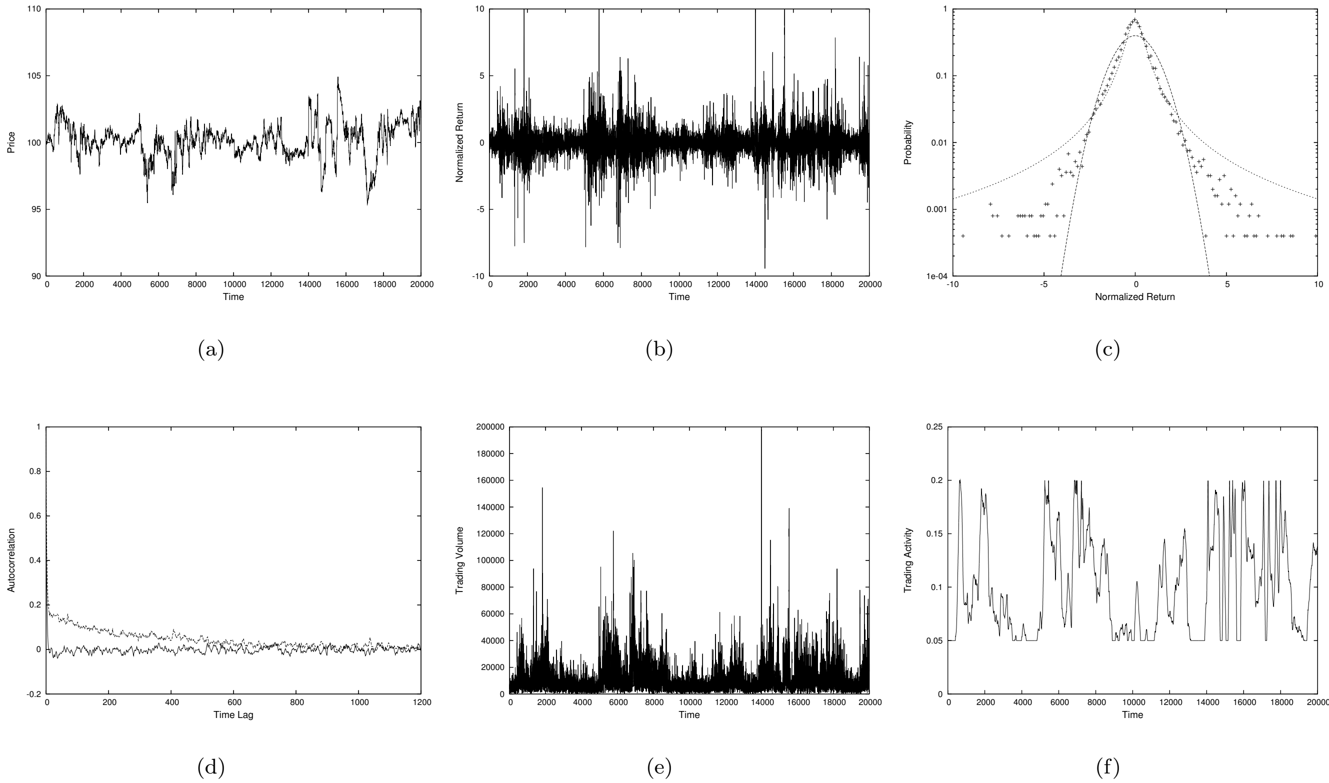


FIG. 5: Simulation results of the level III model when 70% of the agents are imitators. (a) Price. (b) Normalized return. (c) Distribution of return (the scale of the vertical axis is logarithmic). (d) Autocorrelation function of return (the lower curve) and that of volatility. (e) Trading volume. (f) Trading activity. Parameter setting used: $N = 1 \times 10^4$, $c_p = 0.05$, $F = 100$, $c_{fu} = 0.2$, $c_{im} = 0.7$, $P^0 = 100$, $k = 400$, $c_l = 20$ and $L_m = 0.01$. The lower bound of $M^{(\cdot)}$ is 0.05.

i. After substitutions, we have

$$\begin{aligned} Q^{t+1} &= \sum_{i=1}^{N}[A_i^{t+1}V_{i,fu}^{t+1} + A_i^t(\eta_{im}^{t+1})V_{i,fu}^t \\ &+A_i^{t-1}(\eta_{im}^{t+1}\eta_{im}^t)V_{i,fu}^{t-1} \\ &+A_i^{t-2}(\eta_{im}^{t+1}\eta_{im}^t\eta_{im}^{t-1})V_{i,fu}^{t-2} + \ldots \\ &+A_i^{t-\tau}(\eta_{im}^{t+1}\eta_{im}^t\eta_{im}^{t-1}\ldots\eta_{im}^{t-\tau+1})V_{i,fu}^{t-\tau} + \ldots], \end{aligned} \tag{14}$$

where $\tau = -1, 0, 1, 2, \ldots$. The first few instances of $A_i^{t-\tau}$ are

$$A_i^{t+1} = u_i,$$

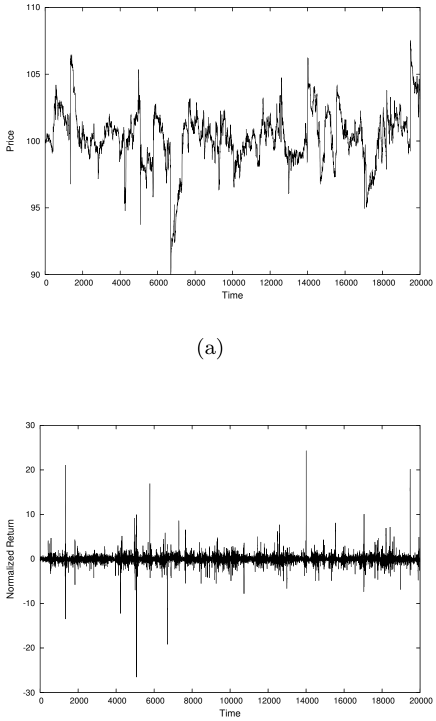
$$A_i^t = \frac{1}{2}((1-u_i)u_{i-1} + (1-u_i)u_{i+1}),$$

$$\begin{aligned} A_i^{t-1} &= \frac{1}{2^2}[(1-u_i)(1-u_{i-1})u_{i-2} \\ &+(1-u_i)(1-u_{i-1})u_i \\ &+(1-u_i)(1-u_{i+1})u_i \\ &+(1-u_i)(1-u_{i+1})u_{i+2}], \end{aligned}$$

$$\begin{aligned} A_i^{t-2} &= \frac{1}{2^3}[(1-u_i)(1-u_{i-1})(1-u_{i-2})u_{i-3} \\ &+(1-u_i)(1-u_{i-1})(1-u_{i-2})u_{i-1} \\ &+(1-u_i)(1-u_{i-1})(1-u_i)u_{i-1} \\ &+(1-u_i)(1-u_{i-1})(1-u_i)u_{i+1} \\ &+(1-u_i)(1-u_{i+1})(1-u_i)u_{i-1} \\ &+(1-u_i)(1-u_{i+1})(1-u_i)u_{i+1} \\ &+(1-u_i)(1-u_{i+1})(1-u_{i+2})u_{i+1} \\ &+(1-u_i)(1-u_{i+1})(1-u_{i+2})u_{i+3}]. \end{aligned}$$

In each product within $A_i^{t-\tau}$, the sequence of $1 - u_i$ terms indicates the propagation of imitation over time (backwards) and space (agents). However, if at least one of the terms is equal to zero, which corresponds to a fundamentalist, the whole product will be zero. As the fraction of imitators (fundamentalists) increases (decreases), some $A_i^{t-\tau}$ terms with larger τ values are greater than zero.

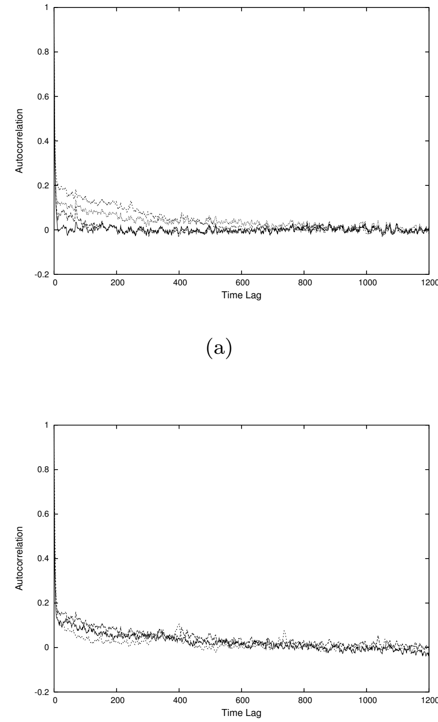
The imitation chains show that long-range interactions can form from local imitations. In the resultant networks, each agent is influenced, directly or indirectly, by some other near or remote agents. The strengths and time lags of influence differ. In this respect, our CA model



(a)

(b)

FIG. 6: Simulation results of the level III model when $\alpha_{im} = 70\%$ and $c_{im} = 0.9$. (a) Price. (b) Normalized return. Parameter setting used: $N = 1 \times 10^4$, $c_p = 0.05$, $F = 100$, $c_{fu} = 0.2$, $P^0 = 100$, $k = 400$, $c_l = 20$ and $L_m = 0.01$. The lower bound of $M^{(\cdot)}$ is 0.05.



(a)

(b)

FIG. 7: Autocorrelation functions of volatility produced by the level III model when $\alpha_{im} = 70\%$. (a) ACFs when $k = 10$ (the lowest curve), $k = 100$ (the second lowest), $k = 300$ (the highest) and $k = 500$ (the second highest) respectively. (b) ACFs when $N = 10 \times 10$ (the middle curve), 100×100 (the upper one) and 1000×1000 (the lower one) respectively. Parameter setting used: $N = 1 \times 10^4$, $c_p = 0.05$, $F = 100$, $c_{fu} = 0.2$, $c_{im} = 0.7$, $P^0 = 100$, $k = 400$, $c_l = 20$ and $L_m = 0.01$. The lower bound of $M^{(\cdot)}$ is 0.05.

is different from the Cont-Bouchaud model, where any two agents can be directly linked, and agents in a group behave identically. It is also distinct from the model of Bartolozzi *et al.*, within which agents in a cluster influence each other with an equivalent strength.

B. Price and volatility are mean-reverting

Fundamentalists behave according to price while imitators follow other agents but do not directly respond to price. We therefore argue that it is fundamentalists' behavior that determines price trend. This argument can be confirmed by our simulations: If $\alpha_{fu} = 0$ (all the agents are imitators), price fluctuations die out; in other cases we obtain price trajectories similar in shape but distinct only in amplitude.

Therefore, for the sake of simplicity, we can take the special instance that all the agents are fundamentalists to study the basic dynamics of the price. In such an instance, $\alpha_{fu} = 1$, hence $u_{(\cdot)} = 1$. Then, Eq. (14) gives

$$\begin{aligned} Q^t &= NV^t_{i,fu} \\ &= N(F\eta^t_{fu} - P^{t-1}). \end{aligned} \tag{15}$$

The noise term η^t_{fu} is indispensable for a sustained price process, but is not responsible for any regularity in price trends. We therefore set $\eta^{(\cdot)}_{fu} = 1$ for the sake of simplicity. Then, Eq. (15) becomes

$$Q^t = N(F - P^{t-1}). \tag{16}$$

Equation (9) gives,

$$Q^t = (\frac{N}{c_p})(P^{t+1} - P^t). \tag{17}$$

Substituting Eq. (17) into Eq. (16), we obtain

$$P^{t+1} - P^t + c_p P^{t-1} = c_p F. \tag{18}$$

Equation (18) is a second-order difference equation. Depending on the value of c_p, the price can follow a monotonically decaying process ($c_p < 0.25$), or a damped fluctuating process ($0.25 < c_p < 1$), or an explosive fluctuating process ($c_p > 1$). Figure 8 shows the three typical price trajectories when the initial price is 105. In all these instances, the price is mean-reverting. Some researchers have studied the mean-reverting nature of price processes for different behavioral types, as well as different stabilizing-destabilizing endogenous mechanisms of financial markets [19–21].

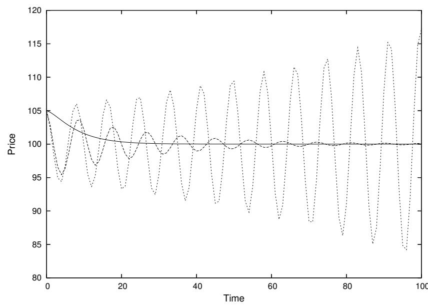


FIG. 8: Price trajectories given by Eq. (18) for different values of c_p. The monotonously decaying curve, the convergent fluctuating curve and the divergent fluctuating curve correspond to $c_p = 0.1$, 0.9 and 1.02 respectively. Parameter setting used: $N = 1 \times 10^4$, $F = 100$, $P^0 = 105$.

To analyze the process of volatility generated by our model, we need to consider the mechanism as well as the noise. First of all, we have assumed that the noise, which causes the volatility, follows an independent Gaussian random process. Within this process, those values more close to the mean have higher probabilities. Second, by examining Eqs. (5) through (9), we can recognize that when L^t is smaller (greater) than L_m, a positive (negative) feedback loop will form between L^t and M^t. Namely, small (large) values of L^t tend to be enlarged (lessened). Therefore, the nature of the noise, the trading behavior of the agents and the rule of price updating ensure that the volatility also follows a mean-reverting process.

C. Heavy tails due to large price variations are caused by imitations

In this section, for the sake of simplicity, we adopt price change as return, i.e., $R^{t+1} = P^{t+1} - P^t$. According to Eq. (9), we can then examine Eq. (14) in order to investigate the cause of the resultant non-Gaussian return distributions.

In the simulations demonstrated in Sec. IV B, if $\alpha_{fu} = 1$, we cannot generate fat tails. In this case, the total quantity is described by Eq. (15), a special instance of Eq. (14) when all the terms with a product of η_{im}^t terms are equal to zero. Heavy tails are generated when $\alpha_{fu} < 1$ and some of these terms are present in Eq. (14). We therefore suppose that it is the multiplication of the various ϕ_{im}^t terms in different η_{im}^t terms that is responsible for the non-Gaussian distributions, although all these terms themselves follow a Gaussian distribution.

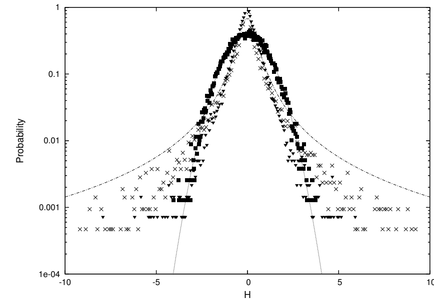


FIG. 9: Probability distributions of H^t defined by Eq. (19). Different values of ϵ are taken for comparison. The ■ points, the ▼ points and the × points refer to the instances of $\epsilon = 1$, $\epsilon = 0.5$ and $\epsilon = 0$ respectively. The scale of the vertical axis is logarithmic.

To confirm this supposition, we define a simple reference model:

$$H^t = \epsilon \phi^t + (1 - \epsilon) \phi^t \phi^{t-1} \phi^{t-2}, \tag{19}$$

where ϕ^t is an independent Gaussian random variable with mean 0 and standard deviation 1, ϵ is a parameter. Recall that, in Eq. (14), $\eta_{im}^t = 1 + c_{im} \phi_{im}^t$. Since Eq. (19), as with Eq. (14), deals with the sum of products of Gaussian terms, it represents the basic structure of the latter.

Figure 9 presents the experimental PDFs of H^t. We take different values of ϵ for comparison: 1, 0.5 and 0. In this figure we see that when ϵ decreases, the distribution of H^t gradually changes from being pure Gaussian to being very fat-tailed non-Gaussian. Thus, the more the product of ϕ^t terms is weighted, the heavier the tails of the consequent distribution. From the discussion in Sec. V A we know that, if α_{fu} is small, the products of more η_{im}^t factors in Eq. (14) will have more weight. This experiment therefore explains the regularity discussed in Sec. IV B and Sec. IV C: Larger fractions of imitators correspond to return distributions with heavier tails. In addition, products of η_{im}^t terms give rise to continued products of c_{im}. The multiplication of c_{im} explains the exponential growth of kurtosis following the increase of c_{im}, as shown in Sec. IV C.

D. Volatility clustering is related to the evolution of trading activity

According to Eq. (9) and the definition of return adopted in this section,

$$R^{t+1} \propto Q^t \tag{20}$$

Since M^t changes much more slowly than Q^t, we have $M^t \simeq M^{t-1} \simeq \ldots \simeq M^{t-\tau}$ for small values of τ. (Note that the analysis here is by no means rigorous.) Then, for a small value of τ, according to Eqs. (7), (8) and (10) as well as the scheme conveyed by Eqs. (11) through (13), Eq. (20) gives

$$R^{t+1} \propto M^t U^t \qquad (21)$$

where

$$\begin{aligned} U^t \; = \; & \sum_{i=1}^{N} [A_i^t (F\eta_{fu}^t - P^{t-1}) \\ & + A_i^{t-1} (\eta_{im}^t)(F\eta_{fu}^{t-1} - P^{t-2}) \\ & + \ldots \\ & + A_i^{t-\tau-1} (\eta_{im}^t \eta_{im}^{t-1} \ldots \eta_{im}^{t-\tau})(F\eta_{fu}^{t-\tau-1} - P^{t-\tau-2})] \end{aligned}$$

In Eq. (21), M^t is a factor that emerges from the agents' trading and in turn reinforces it. Because it changes more slowly than U^t, successive values of $|R^{t+1}|$ are positively correlated with each other. However, consecutive values of R^{t+1} are only weakly correlated due to the fast variation in its sign, which is caused by the fast variation in the sign of U^t due to news and the mean-reverting nature of the price. These explain the stylized facts: long-term autocorrelation of volatility and short-term autocorrelation of return.

E. The regularity can be identified within other MS models

At a higher level, the basic reason for large price variations and that of volatility clustering are common to those MS models discussed in Sec. I that can confirm these phenomena.

Although explaining imitation from different angles, all these MS models and our CA model show that the fraction of abnormally large price variations is much larger when agents imitate each other than when they are mutually independent. In the latter instance and in the limit of a large number of agents, returns follow a Gaussian distribution.

Within these MS models and our CA model, we can ultimately attribute volatility clustering to the evolution of agents' activity, although the corresponding processes of the models that indicate activity are quite distinct. (Notice that all these processes are positively correlated with the evolution of trading volume.) These processes are, respectively, the evolution of volatility (the model of Bak *et al.* [12]), the development of the fraction of noise traders (Lux *et al.* [13]), the evolution of agents' activation thresholds (Iori [16]), the percolation process (Bartolozzi *et al.* [17]), and the progression of the desirability of an asset (our CA model). In addition, these processes are slower than their corresponding "source" processes. Therefore, volatility clustering generated by these MS models and our CA model is the combined effect of two processes on different time scales.

In literature, on the one hand, there is not yet a common agreement on the origins of the stylized facts [10]. On the other hand, various analytical models for describing the phenomena, e.g., GARCH models [36], stochastic volatility models [37], and a recently published Itô-Langevin model described in [25], do not provide explicit economic explanations for the underlying dynamics. The regularity discussed here can help us achieve a more accurate understanding of the complex dynamics of stock markets.

VI. CONCLUSION

In this paper, a cellular automaton model for simulating the complex dynamics of stock markets has been described. The model can confirm the main stylized facts observed in empirical financial time series. Our simulations and analytical analysis suggest that price and volatility are mean-reverting. Long-range agent interactions, which are responsible for large price variations, can form from local interactions. Volatility clustering is associated with the variation in agents' trading activity, a slower process compared with the variation in the influence of news. Heavy-tailed return distributions are related to both large price variations and volatility clustering. After all, these non-Gaussian distributions are produced by agents' behavior in response to the arrival of news, even though the influence of news on agents' perceived fundamental value of a stock is assumed to follow a Gaussian distribution.

Acknowledgments

The authors wish to thank Breanndán Ó Nualláin, Zhengdao Wang, Joost Geerdink, Ketan Maheshwari, Dhiraj Hegde, Thomas Miedema and Stanley Jaddoe for their valuable suggestions and discussions about microscopic simulations in finance. We are grateful to our anonymous referees for their constructive comments and suggestions. This work was carried out in the context of VLe (Virtual Laboratory for eScience), a project supported by a BSIK grant from the Dutch Ministry of Education, Culture and Science.

[1] B. Mandelbrot, J. Bus. **36**, 394 (1963).

[2] A. Pagan, J. Empirical Finance **3**, 15 (1996).

[3] Z. Ding, C. W. J. Granger, and R. F. Engle, J. Empirical Finance **1**, 83 (1993).
[4] R. Cont, Quant. Finance **1**, 223 (2001).
[5] D. M. Guillaume, M. M. Dacorogna, R. R. Davé, U. A. Müller, R. B. Olsen, and O. V. Pictet, Financ. Stochast. **1**, 95 (1997).
[6] R. N. Mantegna and H. E. Stanley, *An introduction to econophysics* (Cambridge University Press, 2000).
[7] J. Voit, *The statistical mechanics of financial markets* (Springer-Verlag Berlin Heidelberg New York, 2003).
[8] L. Tesfatsion, Artif. Life **8**, 55 (2002).
[9] M. Levy, H. Levy, and S. Solomon, *Microscopic simulation of financial markets* (Academic Press, 2000).
[10] R. Cont, in *Long Memory in Economics*, edited by A. Kirman and G. Teyssiere (Springer, 2005), p. 289.
[11] S. Bandini, S. Manzoni, A. Naimzada, and G. Pavesi, in *Proceedings of the 6th International Conference on Cellular Automata for Research and Industry, ACRI 2004, Amsterdam*, edited by P. M. A. Sloot, B. Chopard, and A. G. Hoekstra (Springer-Verlag Berlin Heidelberg New York, 2004), p. 591.
[12] P. Bak, M. Paczuski, and M. Shubik, Physica A **246**, 430 (1997).
[13] T. Lux and M. Marchesi, Nature **397**, 498 (1999).
[14] R. Cont and J.-P. Bouchaud, Macroecon. Dyn. **4**, 170 (2000).
[15] D. Talia and P. M. A. Sloot, Future Gener. Comp. Sy. **16**, v (1999).
[16] G. Iori, J. Econ. Behav. Organ. **49**, 269 (2002).
[17] M. Bartolozzi and A. W. Thomas, Phys. Rev. E **69**, 046112 (2004).
[18] Z. Wang, *Microsimulation of financial markets* (MSc thesis, University of Amsterdam, 2005).
[19] W. J. Baumol, Rev. Econ. Stat. **39**, 263 (1957).
[20] A. Beja and M. B. Goldman, J. Financ. **35**, 235 (1980).
[21] R. H. Day and W. Huang, J. Econ. Behav. Organ. **14**, 299 (1990).
[22] E. Egenter, T. Lux, and D. Stauffer, Physica A **268**, 250 (1999).
[23] T. Bollerslev, J. Econometrics **31**, 307 (1986).
[24] Z. Ding and C. W. J. Granger, J. Econometrics **73**, 185 (1996).
[25] C. Anteneodo and R. Riera, Phys. Rev. E **72**, 026106 (2005).
[26] J. Downes and J. E. Goodman, *Dictionary of finance and investment terms* (Barron's Educational Series, Inc. 1998).
[27] Generally, return is defined as $R_1^{t+1} = \ln P^{t+1} - \ln P^t$, where R_1^{t+1} is the return at time $t+1$, P^t is the price at time t, and so on. The basic relation is $R_2^{t+1} = (P^{t+1} - P^t)/P^t$. Sometimes, return is defined as price change, $R_3^{t+1} = P^{t+1} - P^t$. For high-frequency data, $|R_3^{t+1}| \ll P^t$. Hence, $R_1^{t+1} = \ln[1 + R_3^{t+1}/P^t] \simeq R_3^{t+1}/P^t = R_2^{t+1}$. Since R_3^t is a fast variable and P^t is a slow variable, $R_2^t \simeq CR_3^t$, where the time dependence of C is negligible. (See p. 35-39 of Ref. [6]) In our simulations, $|P^{t+1} - P^t| \ll P^t$, so $R_1^t \simeq R_2^t \simeq CR_3^t$. Within our CA model, these three indicators are alternatives to each other in analyzing the regularity in return distributions.
[28] An index is a sample list of stocks that is representative of a whole stock market. It is used by investors to track the performance of the stock market. Different methods are being used for calculating the price of an index. For example, the Dow Jones Industrial Average (DJIA), which contains 30 of the most influential companies in the U.S., is the price-based weighted average of the prices of the included stocks. The Standard and Poor's 500 Index (S&P 500), which includes 500 large publicly held companies that trade on major US stock exchanges, weights companies by market capitalization (the overall value of a company's stock on the market).
[29] In the finance literature, volatility refers to the spread of asset returns measured as the standard deviation of a sample of returns over a period of time, i.e., $\sigma = \sqrt{(1/T)\sum_{t=1}^{T}(R^t - \overline{R})^2}$, where T is the length of the period, R^t is the return at time t, and $\overline{R}$ is the average return over the period. Substituting $T = 1$ and $\overline{R} = 0$ into this equation, we obtain the absolute value of the return over a period of one time unit, $|r|$, which is the most commonly used proxy for volatility in practice. The other commonly used proxy is r^2. We adopt $|r|$ as volatility.
[30] Opportunity cost, or cost of capital, is the rate of return that a business could earn if it chose another investment with equivalent risk [26].
[31] A concave function is a continuous function whose graph at each point in its domain runs below its tangent at the point.
[32] To ensure liquidity, many organized exchanges use market makers, individuals who maintain inventories of their chosen securities and stand ready to buy or sell whenever the public wishes to sell or buy.
[33] In 1890, Alfred Marshall published *Principles of Economics*, in which he discussed how both supply and demand interact to determine price. His supply-demand model has become one of the fundamental concepts of economics. According to the model, if all other factors remain equal, the higher the price, the lower the quantity demanded and the higher the quantity supplied, vice versa. In a price(ordinate) - quantity(abscissa) chart the curve of demand is a downward slope, the supply relationship shows an upward slope. Equilibrium occurs at the intersection point of the two curves. In the chart, if straight lines are drawn instead of the more general curves (the shapes of the curves do not change the general relationships), we immediately obtain Eq. (9).
[34] Volume is defined as the sum of absolute aggregate demand and absolute aggregate supply.
[35] Here, we define the rate of time evolution of a variable X as $(|\Delta X|/|X|)/\Delta t$, where ΔX is the change of X within time increment Δt.
[36] GARCH is the abbreviation for "Generalized Autoregressive Conditional Heteroskedasticity", here "heteroskedasticity" means a situation in which the variance of a variable varies. It is a technique for modeling economic time series with time-varying volatility. A GARCH(p,q) process is defined as $\epsilon(t) = \sigma(t)e(t)$, $e(t) =$ i.i.d. $D(0,1)$, $\sigma^2(t) = \alpha_0 + \sum_{i=1}^{q}\alpha_i\epsilon^2(t-i) + \sum_{j=1}^{p}\beta_j\sigma^2(t-j)$, where $\alpha_{(\cdot)}$ and $\beta_{(\cdot)}$ are parameters [7, 23, 24].
[37] A class of stochastic volatility models considers volatility to be independent of return. Price is then assumed to follow a geometric Brownian motion with a time-dependent volatility: $dS(t) = \mu S(t)dt + \sigma(t)S(t)dz_1$. Here dz_1 describes a Wiener process. With $v(t) = \sigma^2(t)$, the time-dependent variance follows a different stochastic process $dv(t) = m(v(t))dt + s(v(t))dz_2$, where dz_2 is another Wiener process. Different forms of $m(v(t))$ and $s(v(t))$ correspond to some popular models of this type [7].