Physica A 390 (2011) 3120–3135

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Replicating financial market dynamics with a simple self-organized critical lattice model

B. Dupoyet [a], H.R. Fiebig [b,*], D.P. Musgrove [b]

[a] Department of Finance, Florida International University, Miami, FL 33199, USA
[b] Department of Physics, Florida International University, Miami, FL 33199, USA

ARTICLE INFO

Article history:
Received 28 October 2010
Received in revised form 4 February 2011
Available online 22 April 2011

Keywords:
Econophysics
Financial markets
Statistical field theory
Self-organized criticality

ABSTRACT

We explore a simple lattice field model intended to describe statistical properties of high-frequency financial markets. The model is relevant in the cross-disciplinary area of econophysics. Its signature feature is the emergence of a self-organized critical state. This implies scale invariance of the model, without tuning parameters. Prominent results of our simulation are time series of gains, prices, volatility, and gains frequency distributions, which all compare favorably to features of historical market data. Applying a standard GARCH(1,1) fit to the lattice model gives results that are almost indistinguishable from historical NASDAQ data.

Published by Elsevier B.V.

1. Introduction

From a reductionist perspective the statistical physics of a large number of dynamical systems in nature originates from nonlinear processes at the microscopic level. In many cases, this leads to phenomena characterized in the literature by the terms of chaos, complexity, fractal geometry, and criticality. These scenarios are quite ubiquitous, thus not limited to basic physical systems, where turbulence comes to mind for example, but also to applications in biology, geology, social networks, economic systems, and finance, to name a few [1]. The literature on the subject is prodigious.

Our current interest in the subject stems from a recent simulation of financial market dynamics [2]. At the root of that study is a microscopic model based on the principle of gauge invariance, assuming that one of the key mechanisms of trader behavior is independent of any scale (currency unit, for example) used in the market transactions [3]. In technical terms, the model is a quantum field theory based on the gauge group $G = \mathbb{R}^+$, the dilation group, which implies scale invariance of the market model with respect to ordinary multiplication of prices with positive real numbers. The quantum aspect of that model implements the empirical observation that arbitrage opportunities, i.e. realizing a profit via transactions in different markets, vanish quickly because of market dynamics.

At this stage, the model does not provide a mechanism for describing a complex system, as it should, given the empirical evidence. The distribution of market returns, if analyzed appropriately [2], exhibits fat tails (probabilities larger than Gaussian) the likes of which are observed in many high frequency financial markets. However, this is not a consequence of the intrinsic dynamics of the model. In order to remedy this situation, in the present article, we study an abridged model with local interactions that lead to a self-organized complex market model. Although our goal is to eventually combine the gauge model with features of the abridged model discussed here, the latter, despite its simplicity, produces salient characteristics of actual financial markets surprisingly well. Among those are the semblance of return time series, returns frequency distributions, and the nature of volatility. The market volatility, in particular, is a subject of intense research [4–6].

* Corresponding address: Department of Physics, Florida International University, 11200 SW 8th Street, Miami, FL 33199, USA. Tel.: +1 305 348 2605; fax: +1 305 348 6700.
E-mail address: fiebig@fiu.edu (H.R. Fiebig).

0378-4371/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.physa.2011.04.017

These features are promising enough to study this simple model in its own right, which is the subject of this work. Along a one-dimensional lattice representing discrete time, the field on the sites are interpreted as returns in a model market. An updating algorithm is then applied which is loosely fashioned after the well-known proposal by Bak, Tang and Wiesenfeld (BTW) [7]; see also Refs. [8,9,1,10]. The idea is to develop a market model that is driven by microscopic entities, say traders, such that their interaction leaves the lattice field in a self-organized critical (SOC) state. In a critical state, among other things, long-range correlations of suitable observables lead to power law behavior with respect to scaling transformations. Self-organization means that the system is driven to criticality without fine-tuning any external parameters, i.e. solely by its intrinsic dynamics.

It is generally realized that financial markets, being prime examples of social systems, exhibit SOC [11–15]. However, to the best of our knowledge, attempts to model those from a microscopic point of view are rare [16]. In Refs. [17,18] percolation clusters act as investors. Assigning random percolation probabilities, power law behavior is found for the usual observables derived from stock market prices [19]. Another example of such an attempt close to the BTW evolution model is the work of Bartolozzi et al. [20]. Though close to our work in spirit, significant differences exist in terms of the updating strategy and interpretation. Our implementation of our lattice field model will produce price time series, returns and their distributions, volatility time series and their clustering features that are all strikingly similar to historical market data.

Why would one like to have a market model in the first place? After all there are myriads of historical data being collected every day. A good reason is the fact that all of the collected data are merely instances of a random draw from some probability distribution, just like one throw of dice gives only one result, hiding the statistics behind it. A stochastic model on the other hand will enable us to study any number of market instances, and collect ensembles in the language of statistical physics. Observables, as averages endowed with errors, could be computed. Ultimately, a successful model could provide probability distributions for future prices, and thus be an invaluable tool for risk analysis, and the like.

2. Lattice model

We consider the simplest lattice market model conceivable, a one-dimensional chain of $n+1$ sites with labels $j = 0 \ldots n$, where j indicates discrete time $t = j\Delta$ in steps of some arbitrary unit Δ. The sites are populated with a real-valued field r with components $r_j \in \mathbb{R}$. As it turned out it is essential to interpret the field components r_j directly as investment returns. The returns are defined as

$$r_j = \log(\Phi_j/\Phi_{j-1}) \tag{1}$$

where $\Phi_j = P_j/C$ is the price of an investment instrument, such as a stock or index fund for example, and C is a unit (currency, shares, etc.). The continuum version of (1) can be surmised from taking the limit $\Delta \to 0$ in

$$\log(\Phi(t)/\Phi(t-\Delta)) = \Delta \frac{\mathrm{d}}{\mathrm{d}t} \log \Phi(t) + \mathcal{O}(\Delta^2), \tag{2}$$

where $\Phi(t) = \Phi_j$ at $t = j\Delta$.

If we understand the linear lattice as a stand-alone model, the interpretation of r_j as investment returns is supported, with hindsight, only by the outcome of the simulation. However, there is a rather revealing connection to the gauge field model mentioned in the Introduction [2,3] that lends additional support to this interpretation. The discussion of a single-asset gauge model in Section 5:4 of Ref. [3] is relevant to our case. Fixing the gauge such that cash and the asset are measured in the same unit turns out to be convenient. The curvature field, living on the dual lattice, inherits the arbitrage gains from the plaquettes of the gauge field. In Appendix A, we have adapted those deliberations within the context of Ref. [2], adding a few facets. Within that framework the returns r_j naturally correspond to the arbitrage gains of the gauge model.

In order to endow the field r with dynamics we find inspiration in the popular evolutionary model by Bak and Sneppen [9,1,10]. In that context, the field components are fitness values, say $f_j \in [0, 1]$, assigned to the sites of a lattice. The updating process consists in finding the site j_s with $f_{j_s} = \min\{f_j : j = 0 \ldots n\}$, i.e. the least adapted species. Then f_{j_s} and the values $f_{j_s \pm 1}$ of the two next neighbors are replaced with uniformly distributed random numbers from $[0, 1]$. This prescription, when iterated many times $s = 0, 1 \ldots \infty$, leads to a stationary state of the lattice field where a single perturbation can lead to a burst of activity, called an avalanche. The frequency distribution of avalanche sizes is found to follow a power law. A power law is a signature feature of a critical state. Since no tuning of a model parameter is needed, the phenomenon is known as self-organized criticality (SOC). The model is very robust in the sense that changing the updating prescription, within reasonable bounds, will still lead to SOC. A rigorous discussion, containing analytical results, may be found in Ref. [10].

In the context of the financial market model we adopt a modified version of Bak's updating prescription. We select periodic boundary conditions with period $n + 1$, such that $r_{n+1} = r_0$ and $r_{-1} = r_n$. In terms of the returns r_j, we define

$$v_j = r_j(r_{j+1} - r_{j-1}) \tag{3}$$

$$V_j = |v_j| \tag{4}$$

and call

$$V = \max\{V_j : j = 0 \ldots n\} \tag{5}$$

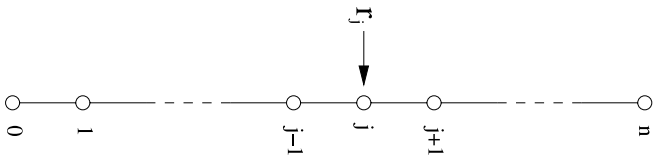

Fig. 1. Illustration of the geometry of the lattice model and the label scheme for the sites. Periodic boundary conditions $r_{j+n+1} = r_j$ are implemented. Updating is done on the field component with signal V, and its two next neighbors.

the signal of the field configuration r. The updating strategy then proceeds with finding a site j_s from[1]

$$j_s \in \{j = 0 \ldots n : V_j = V\}, \tag{6}$$

and then replacing the returns on sites j_s and its two neighbors according to

$$r_{j_s} \leftarrow x_0 \quad \text{and} \quad r_{j_s \pm 1} \leftarrow x_{\pm 1} \tag{7}$$

where x_0 and $x_{\pm 1}$ are three random numbers drawn from a normal distribution $p(x) \propto \exp(-x^2/2w)$ while enforcing the constraint $x_{-1} + x_0 + x_{+1} = 0$. The variance w is a parameter. Fig. 1 illustrates the situation.

Within reason, we have experimented with numerous alternative definitions for the signal, defined through (3)–(5). Overall, it appears that Bak updating is very robust and SOC is easily achieved. However, the choice (3)–(5) proved to best match the stylized features of historical financial market data.

Although (3)–(5) were mostly selected on empirical grounds, in retrospect, more motivation may be provided: note that (3) is just a discretized version of

$$v(t) = r(t) 2\Delta \frac{\mathrm{d}r(t)}{\mathrm{d}t} = \Delta \frac{\mathrm{d}}{\mathrm{d}t} r(t)^2 \tag{8}$$

where $r(t)$, in view of (1) and (2), has been identified with

$$r(t) = \Delta \frac{\mathrm{d}}{\mathrm{d}t} \log \Phi(t). \tag{9}$$

In finance, an established approach is to treat the returns $r(t)$ as a stochastic process [21]. Returns are typically modeled by a (generalized) Wiener process, i.e. assuming normal distributed random variables with a time dependent (random) variance

$$W(t) = E[r(t)^2] - E[r(t)]^2, \tag{10}$$

where $E[\cdots]$ indicates the stochastic expected value. Discretized versions of the time derivatives of the two terms are

$$\frac{\mathrm{d}}{\mathrm{d}t} E[r(t)^2] = 2E[r(t)r'(t)] \simeq \langle r_j (r_{j+1} - r_{j-1}) \rangle \Delta^{-1} \tag{11}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} E[r(t)]^2 = 2E[r(t)]E[r'(t)] \simeq \langle r_j \rangle \langle r_{j+1} - r_{j-1} \rangle \Delta^{-1}. \tag{12}$$

On the right-hand sides, we have changed the notation for the expectation value from $E[\cdots]$ to $\langle \cdots \rangle$, the latter indicating the averages for lattice-generated returns. In the current simulation, standard stochastic financial modeling dictates that $\langle r_j \rangle$ be independent of time. This implies that the discretized part of Eq. (12) is exactly equal to zero.

Therefore, the dynamics of the lattice model is driven by eliminating extreme, sudden, changes of the variance on the returns deemed 'unfit' in the spirit of Ref. [9]. Combining (10) and (11), (12) the lattice version of the latter turns out to be

$$\frac{\mathrm{d}W(t)}{\mathrm{d}t} \simeq [\mathrm{cov}(r_j, r_{j+1}) - \mathrm{cov}(r_j, r_{j-1})] \Delta^{-1}, \tag{13}$$

where $\mathrm{cov}(r_\alpha, r_\beta) = \langle r_\alpha r_\beta \rangle - \langle r_\alpha \rangle \langle r_\beta \rangle$ is the covariance of the two random variables. Hence, extreme changes of the covariance of returns between adjacent time slices are discouraged as part of the dynamics of the market model.

As a side remark we comment on the use of the two-step time derivative to approximate $r'(t)$ in (11) and (12) as opposed to employing one-step forward $(r_{j+1} - r_j)\Delta^{-1}$, or backward $(r_j - r_{j-1})\Delta^{-1}$, discretizations. In both cases the interaction (3), which drives the lattice dynamics, would mutate from a proper next-nearest-neighbor coupling to a term dominated by self-interactions $\propto r_j^2$. This does not lead to a sensible physical system, and demonstrably gives absurd results in a numerical simulation. The corresponding approximations to the time derivative of the variance $\mathrm{d}W/\mathrm{d}t$ are $\simeq [\mathrm{cov}(r_{j+1}, r_j) - \mathrm{var}(r_j)] 2\Delta^{-1}$ and $\simeq [\mathrm{var}(r_j) - \mathrm{cov}(r_j, r_{j-1})] 2\Delta^{-1}$, respectively. Those are variance driven and, equally, not suitable to define market dynamics.

[1] There is at least one, and almost always only one element in this set.

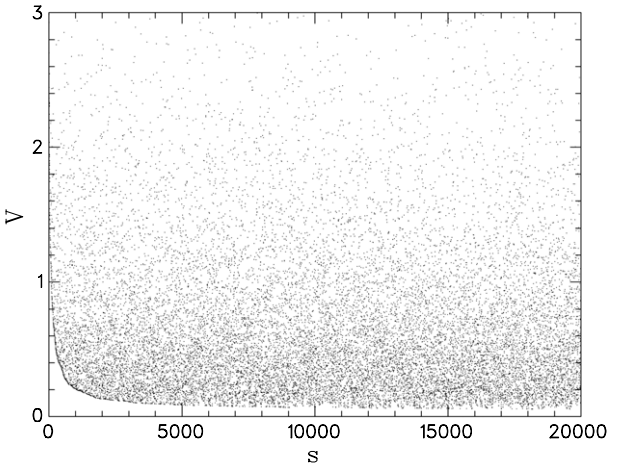

Fig. 2. Updating evolution of the signal V versus the simulation time s. The signal at $s = 0$ is due to a random start of the lattice field configuration. (Its value is =4.74.)

Additional insight may be gleaned from the commentary in Appendix B. There, an action-based classical point of view, extracted from the update rules, sheds some light on the flow dynamics of the returns. Perhaps most importantly, that discussion brings out the significance of the paradigm followed in this article: modeling a market far from equilibrium.

Finally, we should caution that the purpose of the above exposition is only to provide some motivation. In fact we should expect that the returns time series has fractal geometry. This makes perfect sense on a discretized time lattice, with the usual implications [22,23], but it certainly discourages using the concept of continuous time derivatives on everyday manifestations of market data. In this vein the lattice model, in combination with next-nearest-neighbor interactions, appears to be a rational approach.

3. Simulation

For good measure, we have chosen a lattice with $n = 780$ time steps. Starting with a random lattice field r we show in Fig. 2 an example of the updating evolution of the signal $V = \max\{|r_j(r_{j+1} - r_{j-1})| : j = 0 \ldots n\}$ versus the simulation 'time' s. Clearly visible is a significant drop of the lower envelope of $V(s)$ as s approaches $\approx$10,000. Beyond that the distribution of signal dots appears to have stabilized at around $\approx$20,000, at least this is the visual impression.

Following Ref. [10] we take a closer look at the envelope and define the 'gap' function

$$G(x) = \min\{V(s) : s \in \mathbb{N} \cup \{0\} \text{ and } s \leq x\} \quad \text{with } x \in \mathbb{R}^+ \cup \{0\}, \tag{14}$$

which is meant to trace the lower envelope of the signal. By construction, $G(x)$ is a decreasing piecewise constant function with discontinuities at certain discrete values x_k, $k \in \mathbb{N}$. Adding $x_0 = 0$ and assuming an ordered sequence the length of a plateau is $\Lambda_k = x_k - x_{k-1}$, and its height is $G(x_{k-1})$. By definition, we say that an avalanche of length Λ_k starts at x_{k-1} and ends at x_k. At $s = x_{k-1}$ all lattice sites have local signals $V_j \leq G(x_{k-1})$, see (5). As long as the avalanche lasts, there is at least one lattice site with a local signal larger than $G(x_{k-1})$ and thus the updating activity continues until $s = x_k$. Since the gap function is decreasing and bounded from below by zero it will eventually approach a constant $\lim_{x\to\infty} G(x) = G_C$. In this regime the avalanche size diverges and the system has reached the desired state of criticality [10]. An example of $G(x)$ from our simulation is shown in Fig. 3. It corresponds to the data of Fig. 2 but up to much larger simulation times. The evidence points to a critical value G_C with $0 \leq G_C \lesssim 0.01$. Note that $G_C = 0$ is a possibility, we do not know if it is realized. Unlike for the evolution model [10] no analytic results are available. Strictly speaking, the above narrative applies to the thermodynamic limit, implying a lattice with infinitely many sites. Thus, in principle, we expect finite-size effects to afflict our simulation. However, because subsequent results are very sensible we do not expect those to be an obstacle to practical application of this model.

A signature feature of a critical system is scale invariance, implying power law behavior of certain quantities. Again following Ref. [10], we display in Fig. 4 the frequency distribution of the avalanche sizes $\Delta N/\Delta\Lambda$ where $\Delta\Lambda$ is a binning interval for the avalanche sizes and ΔN is the count of avalanches within that interval. We have used 10,000 bins with a binning interval of $\Delta\Lambda = 1$. The data points come from an ensemble average over 2000 independent lattice simulations with 2×10^6 update steps each. This allows one to calculate statistical errors, also displayed in Fig. 4. A power law behavior is beyond doubt. A least-χ^2 fit including data points in the interval $10^1 \leq \Lambda \leq 10^2$ gives $\Delta N/\Delta\Lambda = 301\,\Lambda^{-1.39}$. Integrating this gives a total number of $N \approx 780$ avalanches out of a run with 2×10^6 updates. However, the integrated average N is clearly less than the actual avalanche count per simulation due to the fact that the averaging operation gives rise to counts less than unity, which is realistically unfeasible. An analysis over many simulations gives a more accurate count average

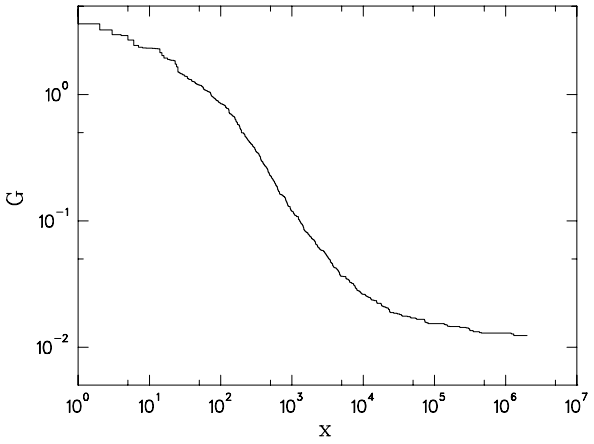

Fig. 3. Plot of the gap function (14) of Fig. 2, but up to $x = 2 \times 10^6$.

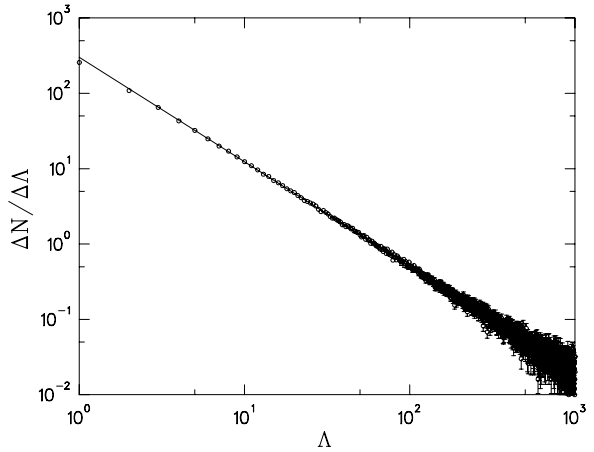

Fig. 4. Frequency distribution of avalanche sizes. The straight line corresponds to a power law fit.

where $780 \leq N \leq 1000$ for a single run with 2-million updates. Hence long avalanches clearly are key to the updating process.

In Fig. 5 we display the activity of the lattice sites during an updating sequence. By definition, H is the number of times a lattice site j has been visited by an updating hit since the start of the simulation. Initially, with random r_j assigned to the sites the signal V tends to visit each site with comparable probability, leading to a flat activity plot. The plot in Fig. 5 reflects the activity after 20,000 updates. A glance at Figs. 2 and 3, reveals that this corresponds to the onset of criticality. The peak-like structures in the activity plot give evidence of the emergence of avalanches with larger sizes.

A closely related plot shown in Fig. 6 shows a zoom window on the simulation time dependence of the updated sites around $j \approx 420$ and $s \approx 10^6$. Each dot in Fig. 6 indicates an updating event of site j at simulation time s. Disconnected sets of dots clearly show the presence of avalanches. By visual inspection the curve is fractal in nature, a signature feature of complexity. In the context of the evolution model [9] the term 'punctuated equilibrium' has been used to describe a similar observation.

Finally, we have used an entropy-like quantity to monitor the approach of the field toward a complex state. Using the exponentiated returns $R_j = \exp(r_j)$ define

$$S = \frac{1}{n} \sum_{j=1}^{n} R_j \log R_j. \tag{15}$$

Then Fig. 7, in which is displayed S versus the simulation time s, shows that the initially random system becomes organized, again at around $s = 20{,}000$. From then on the information content of the field has stabilized, as indicated by bounded fluctuations of S between 10^{-1} and 10^{-2}.

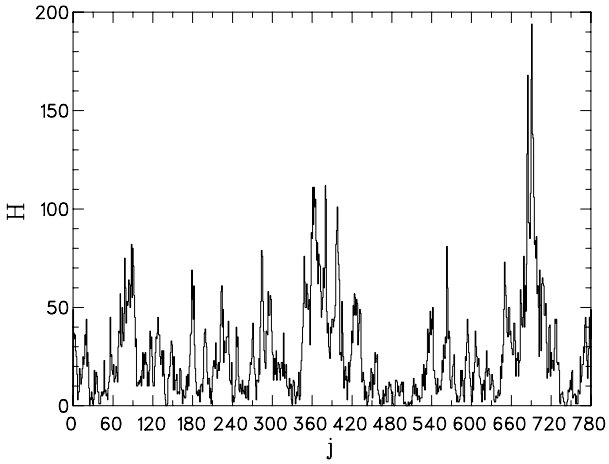

Fig. 5. Plot of the activity across the lattice sites after 20,000 update steps. The appearance of peaks signals the emergence of avalanche dynamics.

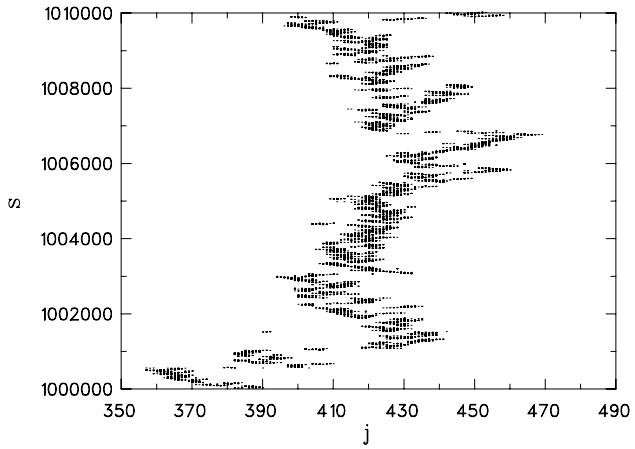

Fig. 6. A zoom window on the updating evolution. Each dot at site j and simulation time s indicates an updating hit.

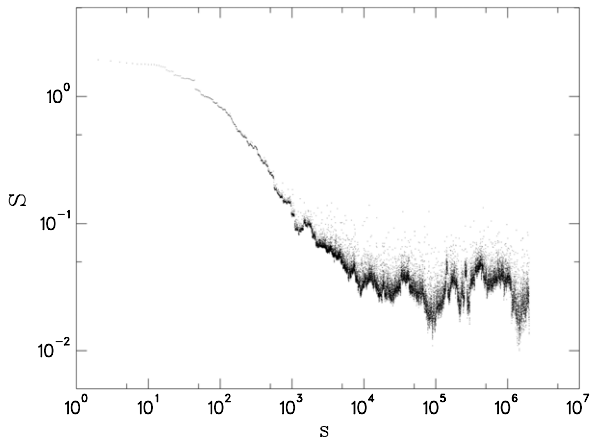

Fig. 7. Entropy S of the exponentiated returns $R_j = \exp(r_j)$ as a function of the simulation times s.

4. Results

As mentioned above, with our lattice geometry and size, it takes at least 20,000 updating hits to reach a critical state. The lattice has $n = 780$ time intervals, and we use $w = 1$. To obtain the results discussed in this section we have used 4×10^6 initial updates before collecting field configurations from independent simulations.

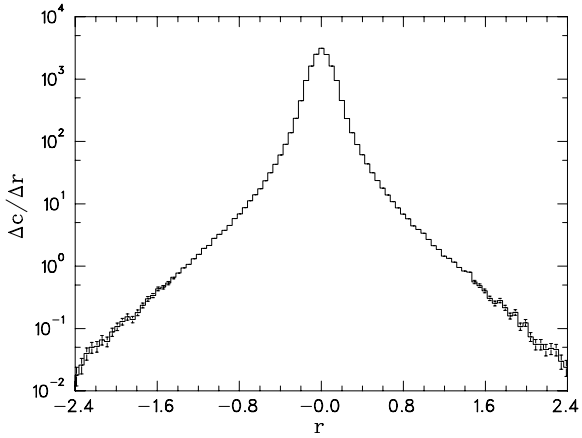

Fig. 8. Lattice returns (gains) distribution. $\Delta c/\Delta r$ is the number of returns of size r at a binning interval of $\Delta r = 0.05$.

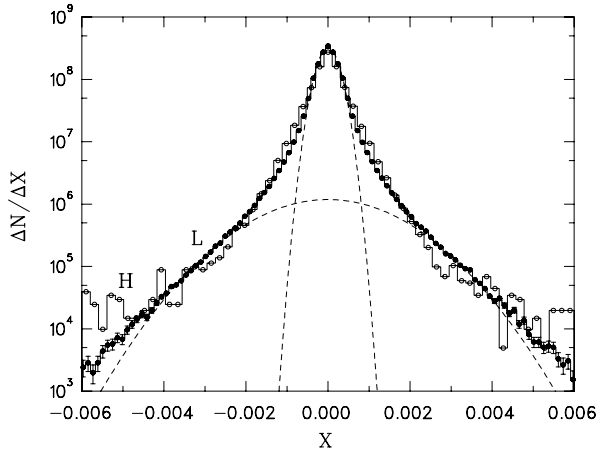

Fig. 9. Historical NASDAQ returns (gains) distribution (open circles, H) compared to the results of the lattice simulation (filled circles, L). The dashed lines are Gaussian fits to the center and the tails, respectively, of the lattice data.

4.1. Gains distribution

The advantage of working with a stochastic model is that observables can be estimated from ensembles, i.e. multiple realizations of market time series. Thus we can take a look at the gains distribution, defined as the frequency plot of returns against the size of the returns. The lattice data displayed in Fig. 8 are averages from 10,000 simulations, taking one sample after 4×10^6 updates each. The binning interval for the returns is $\Delta r = 0.05$ and the number of counts per interval $\Delta c/\Delta r$ is normalized such that the total number of counts is $n = 780$. Statistical errors are visible at the tails of the distribution. A well-known feature of gains distributions is that they exhibit 'fat tails', meaning that for extreme values of r the distribution is considerably enhanced over a normal (Gaussian) distribution. In Ref. [2], where we studied the effects of arbitrage using a local gauge field, this feature was only obtained after analyzing the returns performing a weighted time average of past returns. In the present model the fat tails distribution emerges naturally. Since the returns field develops into a critical state the resulting scale invariance implies loss of memory of past holding patterns.

The same gains distribution is displayed in Fig. 9 along with historical market data from the NASDAQ index compiled from minute data between 2005-Aug-26 and 2008-Aug-25 [24]. Scale factors as in $\Delta X = 2.4 \times 10^{-3} \Delta r$ and $\Delta N/\Delta X = 1.1 \times 10^5 \Delta c/\Delta r$ have been applied to match the historical data. The dashed lines correspond to Gaussian distributions fit to the lattice data in the center (7 points) and the tails (2×38 points), respectively. The Gaussian distribution is clearly ruled out. We observe that the lattice model accounts remarkably well for the empirical gains distribution over many orders of magnitude.

4.2. Time series

In Figs. 10–13 we show four sets of sample time series from the lattice and selected historical data [24] from the NASDAQ index. The latter are included to demonstrate that the lattice model has the 'distinctive air' of a real market. It goes without

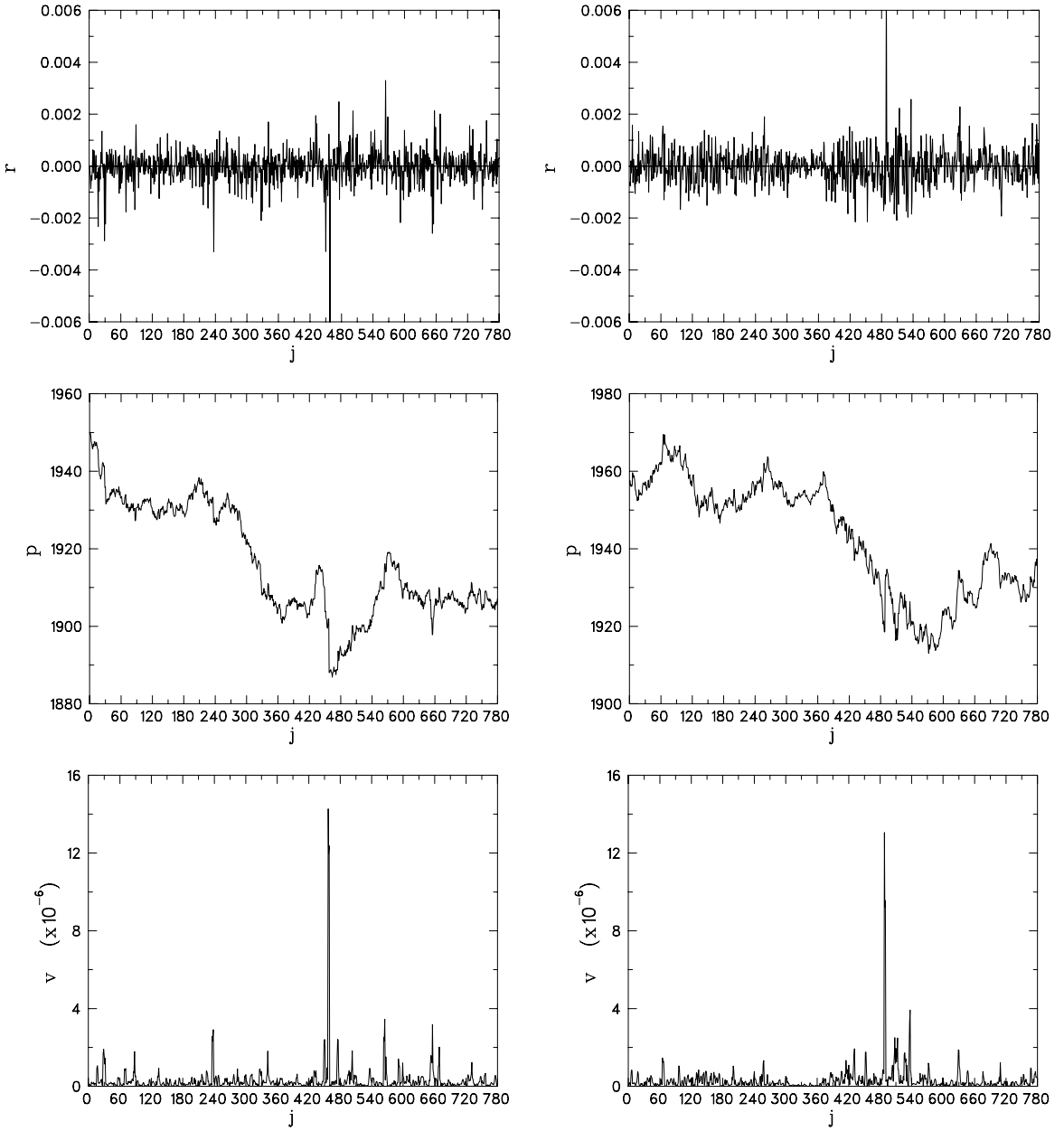

Fig. 10. Set 1: Lattice-generated times series of returns r, prices p, and volatility v are shown in the left column. The three panels on the right show selected historical r, p, v times series from the NASDAQ index.

saying that a statistical model is at a loss predicting time series, nor can one expect that a single model will describe subtle details of stochastic features of every market. What we are looking for has been very eloquently laid out in Ref. [22] by Mandelbrot. In the Introduction we read: "It is worth noting that fully fleshed out and detailed pictures … put a heavy premium on the ability of the eye to recognize patterns that existing analytic techniques were not designed to identify or assess". It is in this spirit that Figs. 10–13 are presented.

The NASDAQ data from Ref. [24] are at minute intervals, which naively translates to $\Delta = 60$ s for the parameter introduced at the beginning of Section 2. There are discontinuities by end-of-the-day and over-the-weekend interruptions in the time series. Ignoring this, we have picked random time series of length n from the historical data. However, the situation is more complicated. Assuming that the time series has fractal geometry, and thus is devoid of a scale, the choice of any Δ would be equally valid. We are not in a position to pursue this issue here, but will rather live with the above naive choice, adopting the scaling argument.

As for Figs. 10–13 the three panels on the left column display, top to bottom, for $j = 1 \ldots n$ the returns r_j, the prices p_j and the volatility v_j of returns derived from the lattice model. While the returns come directly from the simulation, the price

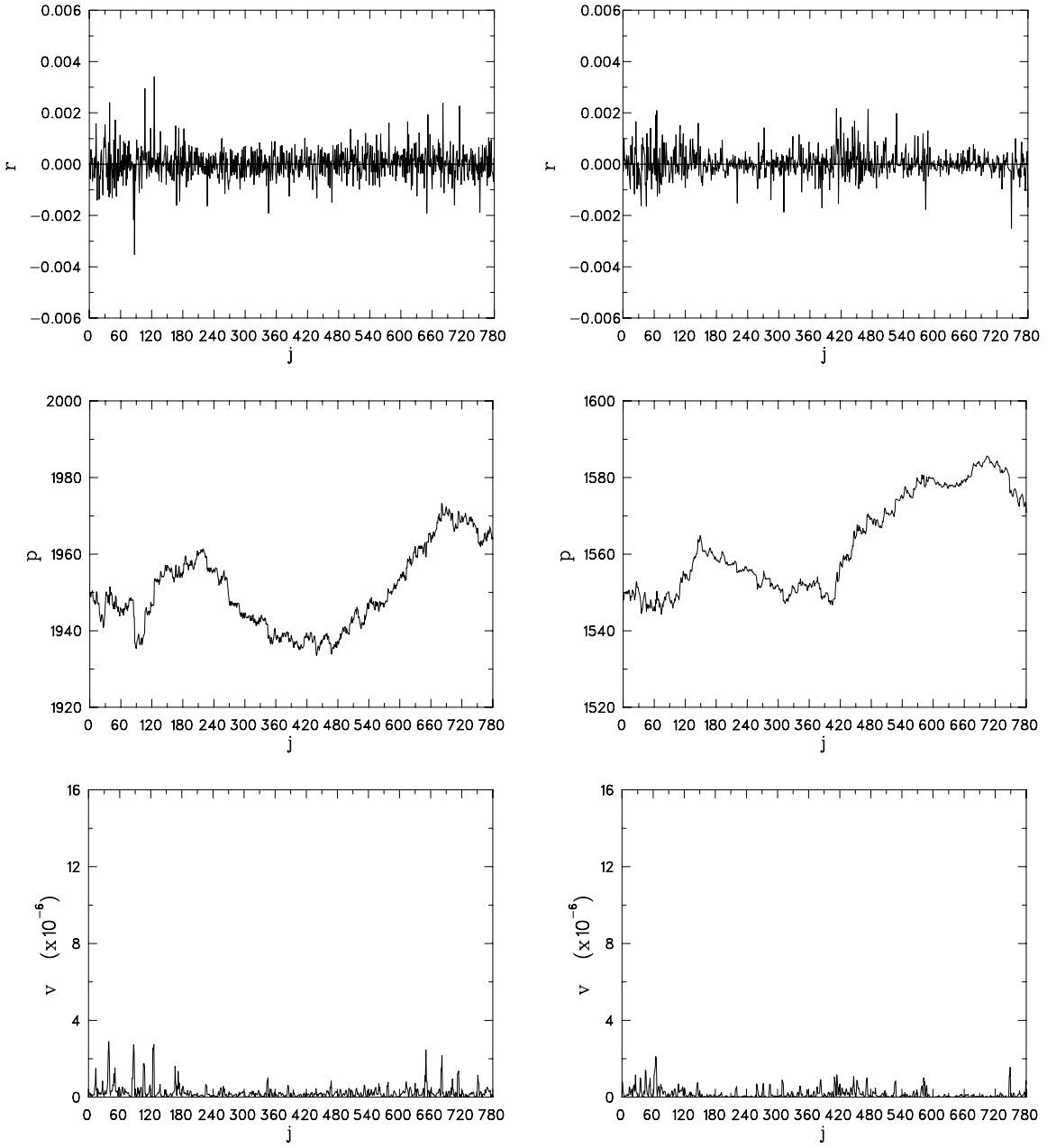

Fig. 11. Set 2: See caption of Fig. 10.

time series is obtained by integrating (9)

$$\Phi(t) = \Phi(t_0) \exp\left[\frac{1}{\Delta}\int_{t_0}^{t} \mathrm{d}t' r(t')\right]. \tag{16}$$

The discrete version of (16) is equivalent to rewriting (1) as a recursion

$$p_j = p_{j-1} \exp(r_j) \tag{17}$$

with $p_j = \Phi_j C$ and initial condition p_0. Finally, the time series of the variance, or volatility in financial terms, is computed from just three time slices through

$$v_j = \frac{1}{3}\sum_{j'=j-1}^{j+1} (r_{j'} - \bar{r})^2 \quad \text{with } \bar{r} = \frac{1}{3}\sum_{j'=j-1}^{j+1} r_{j'}. \tag{18}$$

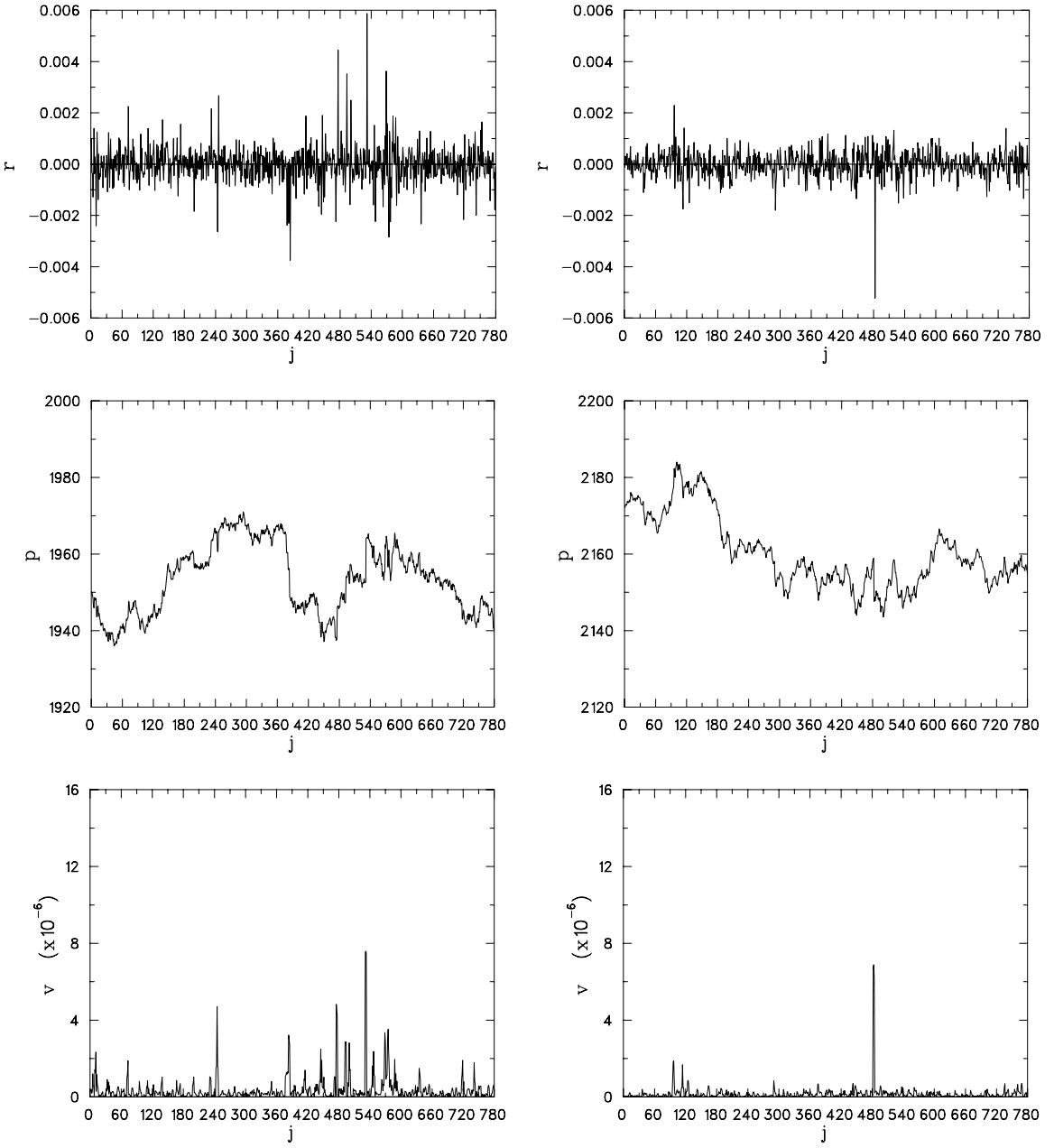

Fig. 12. Set 3: See caption of Fig. 10.

Note that Δ, p_0 and C are all trivial parameters (units, etc.) in the sense that they have no effect on the simulation. The same is true for the variance w used to draw random numbers, from $p(x) \propto \exp(-x^2/2w)$, while updating the lattice, see (7). Since the signal (5) of the field configuration only involves a comparison (max) of returns the effect of changing w will be a rescaling of the field. Specifically

$$w \to \lambda w \quad \text{then } r_j \to \sqrt{\lambda} r_j. \tag{19}$$

Therefore, doing an entire simulation at only one fixed, arbitrary, choice for w is sufficient. In this sense the model has no adjustable parameters. Observables then scale accordingly, for example $v_j \to \lambda v_j$. The price time series (17) behaves in a less straightforward, though well defined, manner because the scaling factor appears with the argument of an exponential function. The plots in Figs. 10–13 were consistently produced with $w = 1, \lambda = 2 \times 10^{-5}$, and $p_0 = 1950$. The resulting scales are similar to those of the historical NASDAQ data. Aside from lateral shifts of the price viewing windows, comparable scales in all panels of Figs. 10–13 are identical.

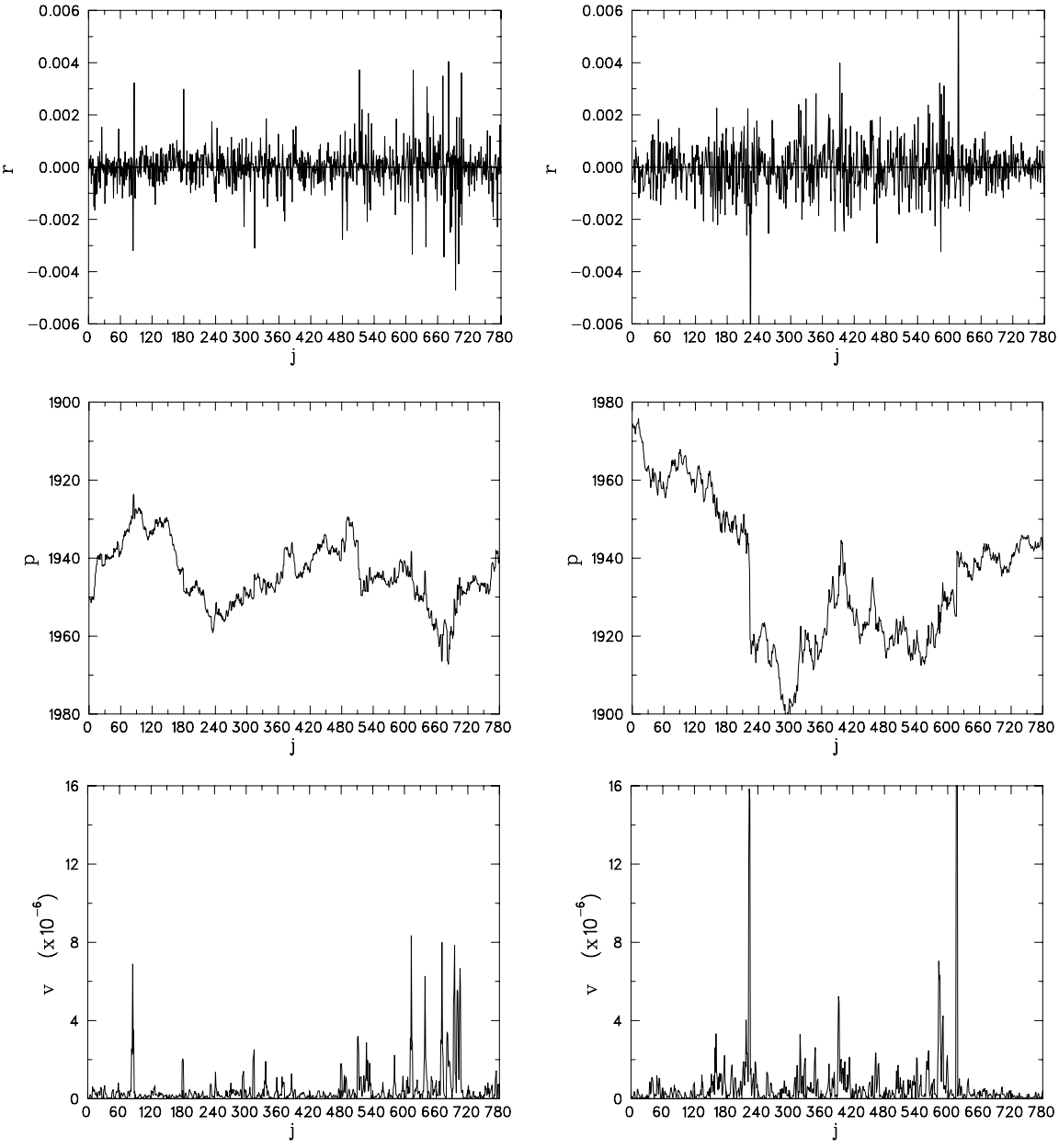

Fig. 13. Set 4: See caption of Fig. 10.

The four data sets in Figs. 10–13 were chosen somewhat randomly, except that they were paired to point out common features appearing in both the lattice and historical sets. For example, in Fig. 10 there is a clearly visible time interval around $j \approx 330$ where the historical returns data exhibits a region of quenched volatility. In financial market data the phenomenon of volatility clustering is well known and even exploited as an industry standard modeling technique [21]. Periods of small/large volatility are often separated by price shocks that manifest themselves in large spikes of the volatility. This is clearly a common feature of both the historical and the lattice data sets of Fig. 10, best visible in the v panels. The price evolutions are similar, but this is accidental. The lattice model generates up or down markets with equal probability. (Although this can easily be changed by modifying the updating algorithm.)

The remaining sets exhibit the same characteristics. In Set 2, Fig. 11, we have paired lattice and historical data that share a small overall volatility. This demonstrates that the lattice model is capable of producing quiescent market periods as well. The price times series look strikingly similar, which is fallacious, the statistical model is not capable of making predictions. The subsequent Sets 3 and 4 provide additional evidence that the lattice model is able to emulate stylized featured of real markets. Again, volatility clusters are clearly present separated by spikes of various sizes. Similarly, Set 4 exhibits fairly active volatility patterns for both the lattice and the historical data.

4.3. Financial market dynamics

We now turn to the ability of our model to capture some of the well-known dynamics observed in financial markets. One of the most important characteristics of a financial time series is its volatility, and more importantly, how the volatility evolves over time. Most financial time series exhibit time-varying volatility clustering, which means that periods of large swings tend to be followed by periods of large swings, while periods of calm tend to be followed by periods of calm. These dynamics can be modeled by the Auto Regressive Conditional Heteroskedasticity (ARCH) model of Engle [21] and by its generalized version, the Generalized Auto Regressive Conditional Heteroskedasticity (GARCH) model of Bollerslev [25].

The fact that these specifications imply that a large shock on average tends to be followed by another large shock means that the resulting distribution of returns will exhibit 'fat tails' or higher-than-Gaussian probability masses in the extreme regions. Similarly, the fact that these specifications also imply that a small shock on average tends to be followed by another small shock means that the resulting distribution of returns will exhibit a higher-than-Gaussian probability mass around the origin, see Fig. 9. ARCH and GARCH models have grown into incredibly popular tools as they are able to replicate these salient features of financial returns distributions, namely more probability mass in the tails and around the center of the distribution than in the benchmark Gaussian case. Capturing these features of financial markets is crucial for many purposes including but not limited to: derivatives pricing, hedging, forecasting volatility, portfolio management, regulatory issues, value-at-risk, and so on.

The traditional ARCH specification expresses the current volatility level as a function of past squared shocks (about a mean or average return) at various lags. This implies that the volatility is not constant over time anymore but that it depends on how large recent deviations (positive or negative) from the mean have been. In the more general GARCH specification, the current volatility level is still a function of past squared shocks to the returns, but it is also a function of past lagged levels of itself, thus making the model even more flexible. The GARCH(p,q) specification for the volatility σ_t^2 at time t can be described by the following equation:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \tag{20}$$

where q and p are the maximum lags allowed by the model for past shocks and volatility levels respectively, and ϵ_{t-i} is the return shock (about the mean) at lag i.

One can obviously allow q and p to be as large as one wants, but a GARCH(1,1) model in practice turns out to be surprisingly flexible. Moreover, overfitting is often the recipe for poor out-of-sample performance, and parsimonious models often end up defeating more complex ones when tested outside of the in-sample period. Hansen and Lunde [26] compare 330 (G)ARCH-type models in terms of their ability to describe the conditional volatility of exchange rate and IBM return data, and find no evidence that a GARCH(1,1) is outperformed by more sophisticated models in the analysis. Therefore we choose to focus on the parsimonious, yet very apt, GARCH(1,1) model for purposes of comparing the dynamics of time-varying volatility in our lattice-generated returns and in NASDAQ historical returns. Our model can thus be written as

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2. \tag{21}$$

The goal here is twofold. First, it is to investigate whether our lattice model is able to produce returns volatility dynamics displaying some form, if any, of ARCH/GARCH effects. Second, it is to examine whether our lattice model is able to produce returns volatility dynamics that are rather consistent with those of NASDAQ historical returns. In Figs. 10 through 13, for each set L and N (lattice and NASDAQ returns), we fit a GARCH(1,1) model onto the returns data through a Levenberg–Marquardt optimization algorithm [27,28] and report the results in Tables 1 through 4.

Before getting to the tables, one may first notice that Figs. 10 through 13 display similar-looking sets of charts when one compares the lattice-generated graphs with the NASDAQ historical ones. The top portion represents returns (gains or losses) over time, and one can readily see that there are clusters of volatility in each, indicating that the model seems to be capable of capturing them. The middle portion simply represents the evolution of the asset value over time. Visually, the price dynamics generated by the lattice model appear 'plausible' as those of a financial market such as the NASDAQ. Finally, the bottom plot represents the evolution of the volatility over time, for both the lattice and the NASDAQ data. Here again, the volatility patterns generated by the lattice model seem credible as those of financial markets, displaying a variety of spikes of various sizes and frequencies.

Tables 1 through 4 display the estimated parameters α_0, α_1 and β_1, followed by their standard errors in parentheses and their t-statistics, in square brackets, for both the lattice and the NASDAQ returns data. Except for the NASDAQ parameter α_0 in Table 4, every parameter is statistically significant at the 5% significance level, and most parameters are even statistically significant at the 1% significance level. This indicates that both the NASDAQ returns and our lattice-generated returns display ARCH/GARCH behavior in volatility. Moreover, and more remarkably, the estimated parameters coming from our lattice-generated returns are extremely close, especially in magnitude, to the estimated parameters coming from the NASDAQ historical returns. For instance, in Table 2, α_0, α_1 and β_1 are estimated to be 4.14×10^{-9}, 0.022450 and 0.966177 respectively for the lattice-generated model, while they are 4.20×10^{-9}, 0.031982 and 0.951740 respectively for the NASDAQ historical returns. Although this does not indicate forecasting abilities on the part of the lattice model, it does show that it is able to reproduce, with precision, important financial markets volatility dynamics and thus has the potential to provide future insights on the inner mechanics of such markets.

Table 1
GARCH fit parameters of Set 1, Fig. 10, for the lattice L and historical N data.

	Set 1L	Set 1N
α_0	3.23E−08(1.01E−08)[3.214841]	1.11E−08(5.35E−09)[2.079218]
α_1	0.043332(0.008898)[4.870008]	0.130702(0.019696)[6.636112]
β_1	0.891148(0.028338)[31.44721]	0.859163(0.026981)[31.84381]

Table 2
GARCH fit parameters of Set 2, Fig. 11, for the lattice L and historical N data.

	Set 2L	Set 2N
α_0	4.14E−09(1.61E−09)[2.564338]	4.20E−09(1.07E−09)[3.916649]
α_1	0.022450(0.005070)[4.428308]	0.031982(0.007263)[4.403283]
β_1	0.966177(0.007889)[122.4713]	0.951740(0.010025)[94.93278]

Table 3
GARCH fit parameters of Set 3, Fig. 12, for the lattice L and historical N data.

	Set 3L	Set 3N
α_0	6.38E−09(1.46E−09)[4.360360]	5.87E−09(2.02E−09)[2.907091]
α_1	0.015515(0.003301)[4.700237]	0.022312(0.008688)[2.568121]
β_1	0.972397(0.004485)[216.8057]	0.956839(0.014609)[65.49589]

Table 4
GARCH fit parameters of Set 4, Fig. 13, for the lattice L and historical N data.

	Set 4L	Set 4N
α_0	2.76E−08(5.42E−09)[5.092728]	3.73E−09(2.70E−09)[1.378933]
α_1	0.045674(0.007692)[5.937874]	0.069064(0.005237)[13.18884]
β_1	0.911153(0.013877)[65.65998]	0.936347(0.003615)[259.0275]

5. Summary and conclusion

By their very nature, historical market data constitute only instances of some stochastic process. It is thus desirable to have available a stochastic model of a financial market. As a result, such a model grants complete access to the market's stochastic features through the act of drawing multiple instances, or ensembles. This allows the possibility of a detailed investigation of market dynamics and the features that define them.

We have studied the properties of a stochastic lattice model which by design derives its features from next-nearest-neighbor interactions of microscopic entities that live on a linear chain in discrete time. The model is realized as a lattice field theory with field components being interpreted as market returns, and is subject to computer simulation with an updating algorithm inspired by the evolution model by Bak et al. [7] that drives the lattice field into a self-organized critical state. We present evidence, including power law behavior, that the critical state is indeed achieved. Its presence is the salient feature of the model.

We compute time series of market returns, prices, volatilities, and returns frequency distributions, all of which are remarkably consistent with historical market data for the NASDAQ stock index. In particular the 'fat tails' feature of the returns distribution comes out effortlessly. It is worth noting that, aside from (trivial) units, the lattice model has no adjustable parameters. However, if so desired, it is straightforward to modify the updating rules by introducing external parameters. For example, asymmetry between up and down markets is easily produced. Perhaps most remarkable is the observation that standard financial industry analysis tools, in our case the GARCH(1,1) model [25], produce fits that make the lattice model time series almost indistinguishable from real financial market data.

It seems that the most important conclusion derived from our study is that self-organized criticality, being a fundamental driving force for the model, is also key to characterizing real world financial markets. To say the least, self-organized criticality should be an important component of intrinsic market dynamics, allowing us to model financial instruments using lattice methods which may hopefully be competitive with current industry practice.

Of course, self-organization is surely not the only mechanism that drives financial market dynamics. Among other things, trading occurs within a background of arbitrage opportunities [3]. We are looking forward to combining past studies focused on arbitrage [2] with the present model in order to develop an even more realistic stochastic market model.

Appendix A. The one-asset lattice gauge model

Using the notation of Ref. [2] we here consider the lattice gauge model with only one asset ($m = 1$). The illustration in Fig. A.14 depicts the details. The vertical direction represents time, discretized by $j \in \mathbb{N}$. The horizontal (space) direction has

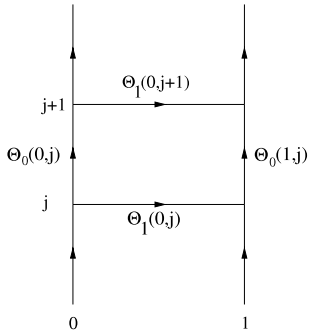

Fig. A.14. Elementary plaquette of the one-asset gauge model.

locations $i = 0$ indicating an interest bearing account (cash), and $i = 1$ being interpreted as an asset (stock). The gauge field $\Theta_\mu(x) \in \mathbb{R}^+$ has components on vertical links, of which $\Theta_0(0, j)$ are interpreted as interest rate factors and $\Theta_0(1, j)$ as the dynamical change of the asset values during one time increment $j \to j + 1$. The horizontal link variables $\Theta_1(0, j)$ are the conversion factors between the units used in the account (units of cash) and those used for the asset (units of stock), at time slices j. With this interpretation, which largely follows Ref. [3], the elementary plaquette

$$P_{01}(0, j) = \Theta_1(0, j)\Theta_0(1, j)\Theta_1^{-1}(0, j + 1)\Theta_0^{-1}(0, j) \tag{A.1}$$

describes the relative gain (profit or loss) realized via an arbitrage transaction between the markets $i = 0$ and $i = 1$ (cash and stocks) involving slices j and $j + 1$. The plaquette $P_{\mu\nu}(x)$ is invariant under a gauge transformation $\Theta_\mu(x) \to g(x)\Theta_\mu(x)g^{-1}(x + e_\mu)$ where $g(x) \in \mathbb{R}^+$, but otherwise arbitrary.

In Ref. [3] this model is elucidated by choosing a gauge where $\Theta_1(0, j) = 1$, meaning that the same unit is used for the cash and stock accounts. We now illustrate how this particular gauge maps the ladder geometry to the simple linear geometry (Fig. 1) used in the current work.

For a given field configuration Θ define a gauge transformation g through

$$g(1, j) = g(0, j)\Theta_1(0, j), \tag{A.2}$$

where $g(0, j) \in \mathbb{R}^+$ is arbitrary, for the time being. This leads to

$$\Theta_1(0, j) \to g(0, j)\Theta_1(0, j)g^{-1}(1, j) = 1, \tag{A.3}$$

thus arriving at the discussion in Ref. [3]. Going a step further, we fix $g(0, j)$ recursively with respect to time

$$g(0, j + 1) = g(0, j)\Theta_0(0, j)R^{-1}, \tag{A.4}$$

where $g(0, 0) \in \mathbb{R}^+$ is an arbitrary initial condition, hence

$$\Theta_0(0, j) \to g(0, j)\Theta_0(0, j)g^{-1}(0, j + 1) = R. \tag{A.5}$$

The number R is a choice, $R = 1$ means that no interest is payed in the cash account.[2] For given R and $g(0, 0)$ the gauge transformation is now completely determined by (A.4) and (A.2). The elementary plaquette is gauge invariant, it now reads

$$P_{01}(0, j) \to P_{01}(0, j) = g(1, j)\Theta_0(1, j)g^{-1}(1, j + 1)R^{-1}. \tag{A.6}$$

In this form the arbitrage gains live on a linear geometry along the asset time axis, $i = 1$, and can be mapped to the geometry Fig. 1. In the light of (1) we thus identify

$$\log P_{01}(0, j) \longleftrightarrow r_{j+1} \tag{A.7}$$

with the returns of the simple model.

The message of (A.7) is that it adds credence to our interpretation of the field components r_j directly as returns. Within the bounds of the gauge model r_j are returns realized via arbitrage transactions.

Appendix B. Classical lattice action

In this appendix we try to help with finding intuition for the updating prescription laid out in Section 2. Technically, none of this is used in the simulation.

In some sense, the updating rules (5)–(7) may be interpreted as an attempt to drive the absolute value of $\mathrm{d}W(t)/\mathrm{d}t$ to zero, see the discussion around (13). This is a hint that $W(t)$ may be viewed as giving rise to an action S, with $\delta S = 0$ leading

[2] Of course $R = R(j)$ could be chosen such that it is time dependent.

to the classical (Euler–Lagrange) equations of motion. Specifically, given the lattice discretization, and ignoring boundary conditions, the action is

$$S = \sum_j W_j. \tag{B.1}$$

For a given field configuration, now approximate the variance $W_j = \langle r_j^2 \rangle - \langle r_j \rangle^2$ from the lattice field on just three time slices

$$W_j \simeq \frac{1}{3}(r_{j-1}^2 + r_j^2 + r_{j+1}^2) - \frac{1}{9}(r_{j-1} + r_j + r_{j+1})^2 \tag{B.2}$$

$$= \frac{2}{9}(r_{j-1}^2 + r_j^2 + r_{j+1}^2 - r_{j-1}r_j - r_j r_{j+1} - r_{j+1}r_{j-1}). \tag{B.3}$$

The classical equation of motion comes from requiring that S is an extremum, thus calculate

$$\frac{\partial S}{\partial r_k} = \sum_j \frac{\partial W_j}{\partial r_k} = \frac{2}{9}(6r_k - 2r_{k-1} - 2r_{k+1} - r_{k-2} - r_{k+2}). \tag{B.4}$$

This result can be expressed in terms of discrete derivatives. For $r(t)$ define

$$\dot{r}(t) = \frac{r(t+\Delta) - r(t-\Delta)}{2\Delta} + \mathcal{O}(\Delta^2) \tag{B.5}$$

$$\ddot{r}(t) = \frac{r(t+\Delta) + r(t+\Delta) - 2r(t)}{\Delta^2} + \mathcal{O}(\Delta^2), \tag{B.6}$$

where the discretization error consistently is of order $\mathcal{O}(\Delta^2)$. With this notation, rewrite (B.4) as

$$\frac{\partial S}{\partial r_k} = -\frac{4}{9}[\Delta^2 \ddot{r}_k + \Delta(\dot{r}_{k+1} - \dot{r}_{k-1})] = -\frac{4}{9}[\Delta^2 \ddot{r}_k + 2\Delta^2 \ddot{r}_k] = -\frac{4\Delta^2}{3}\ddot{r}_k. \tag{B.7}$$

Hence

$$\delta S = \sum_k \frac{\partial S}{\partial r_k}\delta r_k = 0 \Longrightarrow \ddot{r}_k = 0. \tag{B.8}$$

As in Appendix A we may interpret r_k as the returns realized from arbitrage transactions. Then, (B.8) informs us that the flow dynamics of the returns are free of acceleration. In this simple classical sketch there are no forces on the returns.

We emphasize again that the above deliberations are only meant to help with intuition. The dynamics flowing out of the above considerations are *not* used in the simulation. In fact, those are quite contrary to the spirit of this article, because they would lead to an equilibrium system. In contrast, the rules of Section 2 result in a critical system far from equilibrium, a quite different paradigm.

References

[1] P. Bak, How Nature Works: The Science of Self-Organized Criticality, Copernicus, New York, 1996.
[2] B. Dupoyet, H.R. Fiebig, D.P. Musgrove, Gauge invariant lattice quantum field theory: Implications for statistical properties of high frequency financial markets, Physica A: Statistical Mechanics and its Applications 389 (1) (2010) 107–116. doi:10.1016/j.physa.2009.09.002.
[3] K. Ilinski, Physics of Finance—Gauge Modelling in Non-equilibrium Pricing, John Wiley & Sons, New York, 2001.
[4] S.-H. Poon, C.W.J. Granger, Forecasting volatility in financial markets: A review, Journal of Economic Literature 41 (2) (2003) 478–539. URL: http://ideas.repec.org/a/aea/jeclit/v41y2003i2p478-539.html.
[5] T. Rydén, T. Teräsvirta, S. Åsbrink, Stylized facts of daily return series and the hidden markov model, Journal of Applied Econometrics 13 (3) (1998) 217–244. URL: http://www.jstor.org/stable/223228.
[6] C.W.J. Granger, Z. Ding, Some properties of absolute return: an alternative measure of risk, Annales d'conomie et de Statistique 40 (1995) 67–91. URL: http://www.jstor.org/stable/20076016.
[7] P. Bak, C. Tang, K. Wiesenfeld, Self-organized criticality: an explanation of the 1/f noise, Physical Review Letters 59 (4) (1987) 381–384. doi:10.1103/PhysRevLett.59.381.
[8] P. Bak, C. Tang, K. Wiesenfeld, Self-organized criticality, Physical Review A 38 (1) (1988) 364–374. doi:10.1103/PhysRevA.38.364.
[9] P. Bak, K. Sneppen, Punctuated equilibrium and criticality in a simple model of evolution, Physical Review Letters 71 (24) (1993) 4083–4086. doi:10.1103/PhysRevLett.71.4083.
[10] M. Paczuski, S. Maslov, P. Bak, Avalanche dynamics in evolution, growth, and depinning models, Physical Review E 53 (1) (1996) 414–443. doi:10.1103/PhysRevE.53.414.
[11] R.N. Mantegna, H.E. Stanley, An Introduction to Econophysics: Correlations and Complexity in Finance, Cambridge University Press, New York, 2000.
[12] H.J. Jensen, Self-Organized Criticality: Emergent Complex Behavior in Physical and Biological Systems, Cambridge University Press, 1998.
[13] P. Bak, M. Paczuski, M. Shubik, Price variations in a stock market with many agents, Physica A: Statistical and Theoretical Physics 246 (1997) 430–453. doi:10.1016/S0378-4371(97)00401-9.
[14] D.L. Turcotte, Self-organized criticality, Reports on Progress in Physics 62 (10) (1999) 1377. URL: http://stacks.iop.org/0034-4885/62/i=10/a=201.
[15] J.A. Feigenbaum, Financial physics, Reports on Progress in Physics 66 (10) (2003) 1611–1649. doi:10.1088/0034-4885/66/10/R02.
[16] H. Levy, S. Solomon, M. Levy, Microscopic Simulation of Financial Markets: From Investor Behavior to Market Phenomena, Academic Press, Inc., Orlando, FL, USA, 2000.

[17] D. Stauffer, D. Sornette, Self-organized percolation model for stock market fluctuations, Physica A: Statistical Mechanics and its Applications 271 (1999) 496–506. arXiv:cond-mat/9906434 doi:10.1016/S0378-4371(99)00290-3.
[18] R. Cont, J.-P. Bouchaud, Herd behavior and aggregate fluctuations in financial markets, Science & Finance (CFM) working paper archive 500028, Science & Finance, Capital Fund Management (Dec. 1997). URL: http://ideas.repec.org/p/sfi/sfiwpa/500028.html.
[19] A.P. Marcel Ausloos, Paulette Clippe, Evolution of economic entities under heterogeneous political/environmental conditions within a bak-sneppen-like dynamics, Physica A: Statistical Mechanics and its Applications 332 (2004) 394–402. arXiv:physics/0309007.
[20] M. Bartolozzi, D.B. Leinweber, A.W. Thomas, Symbiosis in the bak-sneppen model for biological evolution with economic applications, Physica A: Statistical Mechanics and its Applications 365 (2) (2006) 499–508.
[21] R.F. Engle, Autoregressive conditional heteroscedasticity with estimates of variance of united kingdom inflation, Econometrica 50 (1982) 987–1008.
[22] B.B. Mandelbrot, Fractals and Scaling in Finance: Discontinuity, Concentration, Risk, Springer Verlag, New York, Berlin, Heidelberg, 1997.
[23] D. Sornette, Critical Phenomena in Natural Sciences: Chaos, Fractals, Selforganization and Disorder: Concepts and Tools, in: Series in Synergetics, Springer Verlag, Berlin, Heidelberg, New York, 2006.
[24] Finam Investment Company [link]. URL: http://www.fin-rus.com/analysis/export/.
[25] T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics 31 (3) (1986) 307–327. URL: http://econpapers.repec.org/RePEc:eee:econom:v:31:y:1986:i:3:p:307-327.
[26] A. Lunde, P.R. Hansen, A forecast comparison of volatility models: does anything beat a garch(1,1)?, Journal of Applied Econometrics 20 (7) (2005) 873–889. URL: http://ideas.repec.org/a/jae/japmet/v20y2005i7p873-889.html.
[27] K. Levenberg, A method for the solution of certain non-linear problems in least squares, The Quarterly of Applied Mathematics 2 (1944) 164–168.
[28] D.W. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, SIAM Journal on Applied Mathematics 11 (2) (1963) 431–441. doi:10.1137/0111030. URL: http://link.aip.org/link/?SMM/11/431/1.