Emerging properties of financial time series in the “Game of Life”

Article (Published Version)

Citation:

This version is available from Sussex Research Online: http://sro.sussex.ac.uk/50244/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher’s version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.
Emerging properties of financial time series in the “Game of Life”

Departamento de Inteligencia Artificial, Facultad de Física e Inteligencia Artificial, Universidad Veracruzana, Sebastián Camacho 5, Xalapa Veracruz 91000, Mexico

M. Rodríguez-Achach
Departamento de Física, Facultad de Física e Inteligencia Artificial, Universidad Veracruzana, Lomas del Estadio S/N, Xalapa, Veracruz, Mexico

M. Politi
SSRI & Department of Economics and Business, International Christian University, 3-10-2 Osawa, Mitaka, Tokyo, 181-8585 Japan and Basque Center for Applied Mathematics, Bizkaia Technology Park, Building 500, ES-48160, Derio, Spain

E. Scalas
Dipartimento di Scienze e Tecnologie Avanzate, Università del Piemonte Orientale “Amedeo Avogadro,” Viale T. Michel 11, IT-15121 Alessandria, Italy and Basque Center for Applied Mathematics, Bizkaia Technology Park, Building 500, ES-48160, Derio, Spain

We explore the spatial complexity of Conway’s “Game of Life,” a prototypical cellular automaton by means of a geometrical procedure generating a two-dimensional random walk from a bidimensional lattice with periodical boundaries. The one-dimensional projection of this process is analyzed and it turns out that some of its statistical properties resemble the so-called stylized facts observed in financial time series. The scope and meaning of this result are discussed from the viewpoint of complex systems. In particular, we stress how the supposed peculiarities of financial time series are, often, overstated in their importance.

DOI: 10.1103/PhysRevE.84.066104 PACS number(s): 89.75.Fb, 05.40.—a, 89.65.Gh

I. INTRODUCTION

Advances in physics (especially computational statistical physics), applied mathematics (information and stochastic theories), and computer science have allowed us to successfully attack and study the problem of many interacting units. The affected disciplines range from pure physics to biology, sociology, and economics. The common tools used to study such problems are collectively known as complex-system science. Due to important implications and the accessibility of data, economic systems such as the economy of a country or stock markets are one of the main research subjects. Current research focuses on topics such as the study of the distributional properties of the price fluctuations in stock markets, network analysis of economical systems, financial crashes, and wealth distributions [1–3]. In all these mentioned phenomena, the presence of power-law distributions is ubiquitous and is often recognized as a sign of complexity. Those distributions, together with a full set of common peculiar statistical properties, are omnipresent in market data. They are known as ”stylized facts” and include absence of price-increment correlations, long-range correlation of their absolute values, volatility clustering, aggregational Gaussianity, etc. [4,5].

In order to study and create models of financial markets under a microscopic point of view, new techniques named microscopic simulation (MS) [6] are being intensively applied. They consist in studying a system by individually following each agent and its interactions with other agents, simulating the overall evolution. This line of conduct generated various models able to reproduce “stylized facts” [7,8].

In a closely related way, a great amount of work was devoted to constructing artificial stock markets by means of cellular automata (CA) models [9–11]. Cellular automata are space-time-like discrete deterministic dynamical systems, the behavior of which is defined completely in terms of local interactions. CA were introduced by von Neumann as a tool to understand the biological mechanisms of self-reproduction [12]. Because of their intrinsic mathematical interest and their success in modeling complex phenomena in physical, chemical, economical, and biological systems, design of parallel computing architectures, traffic models, programming environments, etc. [13], they are now much studied.

Cellular automata became very popular at the beginning of the 1970’s thanks to an article written by Gardner [14] about the cellular automaton called the “Game of Life” (GOL), or just “Life.” This automaton was proposed by Conway at the end of 1960’s and since then has displayed a very rich and interesting behavior; very soon, it became the favorite game of the community of computer fans in those times. In practice, by their simplicity, CA are probably the most simple type of abstract complex systems [15–21]. “Life” is a class IV (shows complex behavior), two-state, bidimensional, totalitarian cellular automaton [22–24]. In discrete time, the updating rules determining “Life” evolution are applied on a Moore neighborhood as follows: (a) a dead cell surrounded by exactly three living cells is born again, and (b) a live cell will die if either it has less than two or more than three living neighbors. These simple rules produce a very rich behavior, generating self-organized structures and also producing important and interesting emergent properties (formation of self-replicating structures, universal computation, etc.) [23].

Currently, 40 years after its birth, “Life” is still a very actively researched CA; it is being studied in an interdisciplinary way by physicists, mathematicians, and computer scientists,
The purpose of this paper is twofold. First, we believe “Life” can be a laboratory in which we can try to understand some of the underlying statistical mechanisms behind stylized facts. Second, we would like to highlight how, sometimes, researchers’ focus is pointed on statistical properties too common to justify all the emphasis put among them. The paper is organized as follows. In Sec. II, we present a mapping procedure from “Life” to a one-dimensional time series showing statistical properties very similar to those of financial processes, as already noticed by the fathers of statistical physics in the second half of the 19th century [3,44]. More specifically, to obtain the one-dimensional observable analyzed in this paper [45], we construct the vector \( R_{c.m.}(i) \) for each time step \( i = 1, 2, 3, \ldots \), as follows:

\[
R_{c.m.}(i) := (X_{c.m.}(i), Y_{c.m.}(i)) = \frac{1}{N} \sum_{x=1}^{N} \sum_{y=1}^{N} C_{xy}(i),
\]

where \( C_{xy}(i) \) denotes the state (1 or 0) of the cell in the coordinates \((x, y)\) at time step \( i \). The subscript c.m. stands for center of mass; indeed, \( R_{c.m.}(i) \) is the center of mass of the living particles at step \( i \) (with unitary mass). Figure 1 (right panel) shows 10 000 time steps evolution of the vector \( R_{c.m.}(i) \). Our observable \( r_i \) will be the length of \( R_{c.m.}(i) \), i.e., the distance of c.m. from the origin: \( r_i := \sqrt{X_{c.m.}(i)^2 + Y_{c.m.}(i)^2}, \)

\[ i = 1, 2, 3, \ldots, M. \]

In order to analyze this time series, we construct the returns or logarithmic differences \( S_i \) for each realization of the experiment:

\[
S_i := \log r_{i+1} - \log r_i, \quad i = 1, 2, 3, \ldots, M. \tag{1}
\]

We employed a lattice of size \( N \times N = 3000 \times 3000 \) and the simulation has been initialized configuring randomly the 20\%, 40\%, 60\%, or the 80\% of cells as alive. This means that we have randomly chosen as alive exactly 1 800 000, 3 600 000, 5 400 000, or 7 200 000 of the 9 000 000 total cells. For each one of these densities, we generated 20 random walks of 20 000 steps. Considering certain characteristics such as the finite size of the lattice, or a particular initial configuration, generated fluctuations tend to die out (completely dead lattice) or become periodic after an unknown number of time steps. To overcome this problem, we consider only the first 5000 returns of each one of the 80 original samples. The final four
Emerging Proper Ties of Financial Time Series

PHYSICAL REVIEW E 84, 066104 (2011)

Time series are obtained by concatenating the similar series, eliminating the 19 boundary returns.

III. NUMERICAL RESULTS

Figure 2(a) shows 10,000 time steps of our observable \( r_i \) as a function of time, whereas Fig. 2(b) shows the plot of the corresponding log returns. Both figures were obtained from one of the studied time series. Already, a visual inspection of the return time series displays an intermittent behavior typical of volatility clustering in finance.

A. Return distribution

The empirical densities of standardized returns \( S_i \) are reported in Fig. 3. Left and right tails are shown in Fig. 4. The same figure displays the power-law exponents obtained by a fit using an optimal cutoff parameter together with the Hill estimator as explained in Ref. [46]. Fit parameters are shown in Table I and they are consistent with a power-law behavior.

TABLE I. Fit parameters from cumulative distribution function tails, for right and left tails for each of our four standardized return samples. The second column reports estimated power-law exponents, the third column Anderson-Darling (AD) statistics, the fourth column number of observations fitted in the tail, and the last column the chosen cutoff value (see Ref. [46]). Only the values in italics are larger than the critical value at the 5% significance level.

<table>
<thead>
<tr>
<th>Sample</th>
<th>( \alpha )</th>
<th>AD Statistic</th>
<th>( N_o )</th>
<th>Cutoff</th>
</tr>
</thead>
<tbody>
<tr>
<td>20% positive tail</td>
<td>2.79</td>
<td>0.84</td>
<td>441</td>
<td>4.20</td>
</tr>
<tr>
<td>40% positive tail</td>
<td>3.17</td>
<td>1.38</td>
<td>233</td>
<td>5.80</td>
</tr>
<tr>
<td>60% positive tail</td>
<td>3.58</td>
<td>0.39</td>
<td>216</td>
<td>5.70</td>
</tr>
<tr>
<td>80% positive tail</td>
<td>3.28</td>
<td>0.58</td>
<td>235</td>
<td>5.50</td>
</tr>
<tr>
<td>20% negative tail</td>
<td>3.02</td>
<td>0.49</td>
<td>379</td>
<td>4.50</td>
</tr>
<tr>
<td>40% negative tail</td>
<td>2.68</td>
<td>1.75</td>
<td>446</td>
<td>4.30</td>
</tr>
<tr>
<td>60% negative tail</td>
<td>3.20</td>
<td>0.94</td>
<td>222</td>
<td>5.50</td>
</tr>
<tr>
<td>80% negative tail</td>
<td>2.29</td>
<td>0.29</td>
<td>378</td>
<td>4.20</td>
</tr>
</tbody>
</table>

B. Aggregation properties

Although its formation mechanism is not well understood, it is well known that financial return distributions converge to the normal distribution extremely slowly when the time scale increases [5,47]. This statistical property is called aggregational Gaussianity (or normality), and we want to show its presence in the data.

The analysis is performed summing the simple log returns of Eq. (1), i.e., considering the following definition:

\[
S^\Delta_i := \log r_{i+(\Delta)} - \log r_i, \quad i = 1, 2, 3, \ldots, M - \Delta
\]

where \( \Delta \) stands for the time scale used to aggregate the data. Here, we underline that the plots in Fig. 5 are obtained using overlapping time windows, a method that is not reliable, but can be used for illustrative purposes. One can see that, as the time scale increases, the empirical probability density functions of all four samples converge to a normal probability density slowly. Tables II and III contain the estimated excess kurtosis and skewness of empirical return distributions for every sample and for different time scales \( \Delta \).

TABLE II. Return kurtosis for our four samples and increasing time scales \( \Delta \) used in Fig. 5.

<table>
<thead>
<tr>
<th>( \Delta )</th>
<th>( \kappa_{20} )</th>
<th>( \kappa_{40} )</th>
<th>( \kappa_{60} )</th>
<th>( \kappa_{80} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>36.1</td>
<td>41.0</td>
<td>31.9</td>
<td>141.7</td>
</tr>
<tr>
<td>10</td>
<td>15.4</td>
<td>18.9</td>
<td>10.6</td>
<td>63.7</td>
</tr>
<tr>
<td>100</td>
<td>4.5</td>
<td>5.2</td>
<td>3.4</td>
<td>18.5</td>
</tr>
<tr>
<td>1000</td>
<td>1.3</td>
<td>1.6</td>
<td>1.1</td>
<td>3.3</td>
</tr>
<tr>
<td>10000</td>
<td>0.97</td>
<td>1.3</td>
<td>0.87</td>
<td>0.57</td>
</tr>
</tbody>
</table>
From the inspection of Tables II and III, it turns out that both skewness and excess kurtosis of return distributions vanish slowly as the time scale \( \Delta \) increases. These empirical results are compatible with the convergence of return distributions to the normal distribution. This analysis is not sufficient to infer that our data samples satisfy the property of aggregational Gaussianity, but the moment behavior is definitely in agreement with the phenomenon.

### C. Autocorrelation properties

The upper and lower panels of Fig. 6 show the estimate for the autocorrelation functions (ACF) of returns and of absolute returns, respectively. It can be seen that the ACF for returns shows no memory, immediately decaying to the level of noise; indeed, Fig. 6 looks similar to the ACF of daily financial returns. On the other hand, the ACF of absolute returns decays slowly, showing long-range memory. Both these results are in agreement with the stylized facts found in financial data. Figure 7 displays the average squared ACFs over the 20 realizations of each initial coverage. Power-law fits have been performed on the estimated squared returns ACF. Fit parameters can be found in Table IV. From this table, it is possible to see that fits are excellent for samples 20%, 40%, and 80%, and good for the 60% sample (see Fig. 8). Based on

<table>
<thead>
<tr>
<th>Sample</th>
<th>AD statistic</th>
<th>( \beta )</th>
<th>Remaining Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>20%</td>
<td>0.536</td>
<td>2.13</td>
<td>297</td>
</tr>
<tr>
<td>40%</td>
<td>0.351</td>
<td>2.46</td>
<td>123</td>
</tr>
<tr>
<td>60%</td>
<td>1.408</td>
<td>2.02</td>
<td>126</td>
</tr>
<tr>
<td>80%</td>
<td>0.299</td>
<td>2.16</td>
<td>135</td>
</tr>
</tbody>
</table>

### TABLE III. Return skewness for our four samples and increasing time scales \( \Delta \).

<table>
<thead>
<tr>
<th>( \Delta )</th>
<th>skew(_{20})</th>
<th>skew(_{40})</th>
<th>skew(_{60})</th>
<th>skew(_{80})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.08</td>
<td>0.015</td>
<td>-0.02</td>
<td>-0.29</td>
</tr>
<tr>
<td>10</td>
<td>-0.32</td>
<td>-0.55</td>
<td>-0.03</td>
<td>-1.15</td>
</tr>
<tr>
<td>100</td>
<td>-0.06</td>
<td>-0.12</td>
<td>-0.01</td>
<td>-1.09</td>
</tr>
<tr>
<td>1000</td>
<td>0.11</td>
<td>0.05</td>
<td>0.08</td>
<td>-0.202</td>
</tr>
<tr>
<td>10000</td>
<td>-0.001</td>
<td>0.005</td>
<td>-0.09</td>
<td>-0.203</td>
</tr>
</tbody>
</table>
FIG. 5. (Color online) Aggregation of the empirical return distributions. When the time lag increases, the empirical probability densities converge to a normal density. The dashed line is a normal density.

the above considerations, we can safely conclude that returns are not correlated, whereas squared, or equivalently, absolute returns ACFs decay very slowly, following a power law. All of this is in agreement with financial stylized facts.

D. Leverage effect

Empirical studies of volatility for financial data have shown that volatility estimates and returns are negatively correlated for positive time lag [5,48]. Following Ref. [48], we investigate this effect by estimating the leverage correlation function as

$$L(\tau) = \frac{\langle S^2(i + \tau)S(i) \rangle}{\text{var}[S(i)]^2}.$$ (3)

The estimates for our four samples are shown in Fig. 9. Weak negative correlations are observed for 20% to 60% initial coverage and no correlations for the 80% sample. The same figure shows that the leverage correlation function can be fitted by an exponential function. We found no correlation between past volatility and future price changes and a weak but clear negative correlation with an exponential time decay between

FIG. 6. Upper panel: return ACF (linear scale). Lower Panel: absolute return ACF (logarithmic scale). Both estimates are based on 15 realizations of our experiment.

FIG. 7. (Color online) Absolute returns ACF for all the four samples.

FIG. 8. (Color online) Squared returns ACF for our four samples in a log-log plot. Power-law fit exponents are displayed. Straight lines are used to guide the eye. Fit parameters can be found in Table IV.
future volatility and past returns changes:

\[ L(\tau) = \begin{cases} -Ce^{-a\tau}, \\ 0. \end{cases} \]

Fitted parameters for the exponent \( a \) for each one of our four samples are shown in Table V. Therefore, from the analysis of this section, we can conclude that at least three data sets generated with the lowest initial coverage of living states display a weak leverage effect.

E. Volatility analysis

Volatility \( v(t) \) is calculated \([49]\) by averaging the absolute returns over a time window \( T = n\Delta t \) as follows:

\[ v(t) := \frac{1}{n} \sum_{t'=t}^{t+n-1} |S(t')|. \]  

(4)

<table>
<thead>
<tr>
<th>Sample</th>
<th>( a )</th>
</tr>
</thead>
<tbody>
<tr>
<td>20%</td>
<td>-0.487 ± 0.048</td>
</tr>
<tr>
<td>40%</td>
<td>1.098 ± 0.164</td>
</tr>
<tr>
<td>60%</td>
<td>-0.958 ± 0.109</td>
</tr>
<tr>
<td>80%</td>
<td>No leverage effect</td>
</tr>
</tbody>
</table>

Here, we have set up \( \Delta t = 1 \) time lag and a window of 50 time steps. Figure 10 displays the volatility for the first 3000 time steps of our observable for one of our generated random walks. Volatility distributions for the four samples are shown in Fig. 11. The data were analyzed in order to fit a suitable distribution based on the full sample values of the volatility. A three-parameter log-normal distribution closely describes

FIG. 9. (Color online) Leverage effect. It is not there for the sample with initial density of 80% living cells.

FIG. 10. (Color online) Volatility for a typical generated random walk with a time window of 50 time steps.
the behavior of the set of volatility values. The cumulative distribution function (CDF) is

\[ F(v) = \Phi \left( \frac{\ln(v - \lambda) - \mu}{\sigma} \right) \quad \text{for} \ v > \lambda, \quad (5) \]

where \( \Phi \) denotes the Laplace integral (or the CDF of a standard normal random variable) and \( \mu, \sigma, \) and \( \lambda \) denote the location, scale, and threshold parameters, respectively. In performing the fit, the values of the parameters were found using the maximum likelihood estimates. The results are shown in Table VI for 20\%, 40\%, 60\%, and 80\% return samples, together with values of the Kolmogorov-Smirnov (KS) statistic measuring the maximum distance between the empirical and the fitted cumulative distribution functions. In general, the fits appear to be fairly good; from Fig. 12, it can be seen that the empirical distribution function (EDF) (solid) and the fitted cumulative distribution function (CDF) (dashed) overlap. It is necessary to remark that no statistical goodness-of-fit test can be carried out given that, by construction, the values of the volatility are not independent as required by the tests. In this case, the KS statistic is presented as a descriptive measure to assess the fit quality. In this context, the three-parameter log-normal distribution was used as a model to describe the general behavior of the data.

### IV. DISCUSSION

Since GOL is a class IV cellular automaton and exhibits the property of criticality, it is not a big surprise to detect power-law distributions emerging from different observables, such as spatial and temporal duration of “activity avalanches,” density of living cells and its fluctuations, etc. [33,37,38,50]. In this paper, we chose the summary statistics \( |R_{c.m.}(i)| \) physically suited to study spatial complexity. In doing so, we assumed \textit{a priori} that somehow the process of mapping the cell distribution from each GOL’s microstate \((3000 \times 3000\) dichotomous state variables) onto the “center of mass” (a single continuous variable) would conserve the complexity properties of GOL. This assumption turns out to be correct, and we obtain the expected and ubiquitous power-law distribution of \( |R_{c.m.}(i)| \) variations as well as other unexpected statistical emergent properties. Once more, these facts display the high complexity of GOL and its importance for complexity sciences.

To rephrase, in a very simple way, we have shown how “Life’s” dynamics is able to generate time series displaying most of the statistical properties called stylized facts, usually connected with financial market data and recognized as signs of their complexity. By means of a very simple geometrical mapping, and using a two-dimensional cellular automaton, we generated synthetic data having increments with fat-tailed distributions, clustered volatility, no autocorrelations, long memory in the autocorrelation of absolute returns, aggregational Gaussianity, and leverage effect.

We believe the use of our scheme can be of help in understanding the underlying mechanisms governing the formation of the stylized facts. Of course, “Life” is not the direct explanation for the emergence of these statistical properties, but, given its simplicity, we are confident it can help to build some convincing analogies in order to shed new light on this difficult and important problem. Moreover, we think researchers in both complex systems and quantitative finance should be aware of how they can be easily fooled by all these phenomena. Are power laws really distinct signs of “anomalies”? Is volatility clustering a pregnant concept? Is leverage correlation so physically

---

Table VI. Fit parameter and KS values for the log-normal fits.

<table>
<thead>
<tr>
<th>Sample</th>
<th>( \lambda )</th>
<th>( \mu )</th>
<th>( \sigma )</th>
<th>KS statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>20%</td>
<td>0.08711</td>
<td>-1.1167</td>
<td>0.8712</td>
<td>0.01362</td>
</tr>
<tr>
<td>40%</td>
<td>0.08308</td>
<td>-1.2306</td>
<td>0.9496</td>
<td>0.01446</td>
</tr>
<tr>
<td>60%</td>
<td>0.03108</td>
<td>-0.9190</td>
<td>0.7666</td>
<td>0.02615</td>
</tr>
<tr>
<td>80%</td>
<td>0.03361</td>
<td>-1.3572</td>
<td>0.9918</td>
<td>0.04072</td>
</tr>
</tbody>
</table>

FIG. 11. (Color online) Volatility frequency histogram for all the samples.

FIG. 12. (Color online) Complementary cumulative probability distribution function (1-CDF) for the four samples. Solid lines: volatility empirical probability distribution function. Dashed lines: fitted curve. Both lines almost overlap for all fits. We plot and fit the empirical complementary cumulative distribution function for easier comparison.
we need to think about all these facts as pillars of our models or is it more likely they are just “natural” results of nonstationarity and nonlinear dependencies? At this stage, we do not further speculate on the reported findings that are far from giving an explanation and an interpretation of the phenomena. However, this manuscript can be seen as an example of how misleading the direct study of these phenomena can be; phenomena that are complex not because necessarily connected with something complicated, but because we do not yet have any consistent tool to address them.

[45] The software needed to generate the data sample can be obtained upon request sending an email to manurodriguez@uv.mx.