Probability in the Engineering and Informational Sciences, **26**, 2012, 509–534.
doi:10.1017/S0269964812000162

AN URN MODEL FOR CASCADING FAILURES ON A LATTICE

PASQUALE CIRILLO AND JÜRG HÜSLER
Institute of Mathematical Statistics and Actuarial Sciences
University of Bern
Sidlerstrasse 5, Bern CH3008, Switzerland
E-mails: pasquale.cirillo@stat.unibe.ch; juerg.huesler@stat.unibe.ch

A cascading failure is a failure in a system of interconnected parts, in which the breakdown of one element can lead to the subsequent collapse of the others. The aim of this paper is to introduce a simple combinatorial model for the study of cascading failures. In particular, having in mind particle systems and Markov random fields, we take into consideration a network of interacting urns displaced over a lattice. Every urn is Pólya-like and its reinforcement matrix is not only a function of time (time contagion) but also of the behavior of the neighboring urns (spatial contagion), and of a random component, which can represent either simple fate or the impact of exogenous factors. In this way a non-trivial dependence structure among the urns is built, and it is used to study default avalanches over the lattice. Thanks to its flexibility and its interesting probabilistic properties, the given construction may be used to model different phenomena characterized by cascading failures such as power grids and financial networks.

1. INTRODUCTION

A cascading failure is a failure in a system of interconnected parts in which the breakdown of a part can lead to the default of successive parts. As pointed out in Lindley and Singpurwalla [41], the concept of cascading failure is not at all an abstract one, being conversely very present in the actual world.

Examples of cascading failures are available in different fields of science and engineering. Think for example of a power grid, where one of the elements collapses, shifting its load to the nearby elements in the system. These neighboring elements are then pushed beyond their capacity so that they get overloaded and shift their load onto other elements. This surge process can induce the already overloaded nodes into

failure, setting off more overloads and thus breaking down the whole system in a very short time.

In biology, a cascading failures process can manifest itself as a cancer growing over a cell tissue or as the release of toxins caused by an ischemic attack that destroys far more cells than the initial damage, resulting in more toxins being released.

In economics and finance, cascading failures are often referred to as systemic risk (Lorenz et al. [42]), the typical example being represented by the joint defaults of firms in an industrial district. Moreover, it is not difficult to see cascading failures in the current world economic crisis, started in 2008. The failure of one big financial institution may cause other financial institutions, that is, its counterparts, to go bankrupt, creating an avalanche effect throughout the system. Institutions that are believed to foster systemic risk, because of the great number of interconnections they have with the rest of the system, are called either "too big to fail" (TBTF) or "too interconnected to fail" (TICTF).

There are two macro categories of cascading models (see Lindley and Singpurwalla [41] and Swift [53]): those in which the failure of a component has a permanent effect on the probability of default of the other elements of the system, and those in which the negative effect is just temporary, so that after a certain time the system absorbs it. In other words, the first type of models does not allow recovery, while the second one does. Finally, a defaulted element may be substituted or not. Obviously, this increases or decreases the overall strength of the system. We refer to Lindley and Singpurwalla [41], Wang et al. [54] and Dobson et al. [24] for a more detailed taxonomy.

In the literature, there are several works that fall under the large family of cascading failure models. For what concerns epidemiological applications, without the assumption of being exhaustive, it is natural to cite Liggett [40]. In this book, the author describes several models, among which the well-known *contact process* has become the prototype for studying the spread of infections.

The contact process is based on an interacting particle system and, in some sense, it can be seen as a supercritical oriented percolation model. In a nutshell, it is a continuous time Markov process with state space $\{0,1\}^S$, where S is a finite or countable graph, typically $\mathcal{Z}^d$. If the state of the process at a given time is represented by ξ, then a site $s \in S$ is infected if $\xi(s) = 1$ and healthy if $\xi(s) = 0$. Infected sites become healthy at a constant rate, while healthy sites become infected at a rate proportional to the number of infected neighbors. The dynamics of the basic contact process is defined by the following transitions: $1 \to 0$ at rate 1, and $0 \to 1$ at rate $\rho \sum_{y:y\sim s} \xi(y)$, where the sum is over all the neighbors in S of s. This means that each site waits an exponential time with the corresponding rate, and then flips. For every graph S there exists a critical value ρ_c for the parameter ρ such that, if $\rho > \rho_c$, then the 1's survive with positive probability, while if $\rho < \rho_c$ the process dies out. In the last years, the contact process has been generalized and perfected several times in the literature, for example by setting the state space to $\{0, \dots, \kappa\}^S$, as in the so-called multi-type contact process. An interesting version is presented in Durrett [26].

A recent application of cascading failures models in engineering-related problems can be found in Dobson et al. [23], where the CASCADE model is introduced. The model, which is analytically tractable, considers n identical components with random initial loads. For each component the minimum initial load is L_{min} and the maximum initial load is L_{max}. For $j = 1, 2, \ldots, n$, component j has initial load L_j, which is a random variable uniformly distributed in $[L_{min}, L_{max}]$. $L_1, L_2, \ldots, L_n$ are assumed independent. Components fail when their load exceeds the threshold L_{fail}. When a component breaks down, a fixed amount P of its load is transferred to each of the other components. To start the cascade, the authors assume an initial disturbance that loads each component by an additional amount D. Other components may then fail depending on their initial loads L_j, and the failure of any of these components will distribute an additional load $P \geq 0$ that can cause further failures in an avalanche effect. An interesting feature of the CASCADE model is its ability of reproducing the empirical evidence that the frequency of big black-outs on power grids follows a power law. It will be clear in the next sections that the model presented in Dobson et al. [23] can be obtained as a special case of the construction we are going to show.

As far as the study of cascading failures in economics and finance is concerned, there are several good models in the literature, e.g. Blume and Durlauf [9], Frey and Backhaus [31], Giesecke and Weber [32], Dai Pra et al. [20], Bhargava and Mukherjee [8] and Delli Gatti et al. [22]. In Dai Pra et al. [20], for example, a particle system approach is used to investigate the propagation of financial distress in a network of firms facing credit risk. Through the development of a mean-field interaction model, the authors are able to compute the limiting distribution of the financial system, also determining the time evolution of some interesting credit quality indicators for the firms. For a complete description of the model we refer to the original work.

Regarding this paper, our aim is to introduce a special Pólya-like urn process for modeling cascading failures. In particular, our goal is to study the behavior of a network of urns, each representing a unit subject to default. The default of every system can occur for internal reasons, for example, natural obsolescence, but also for external causes, such as spatial dependence from defaulting neighbors and even the impact of fate on the network (some big calamity). Naturally, the default of every unit may start an avalanche effect through the grid, as expected in a cascading failure model.

2. URN PROCESSES

Introduced by Jakob Bernoulli [6], urn problems have played an important role in probability theory. They form a very large family of probabilistic models in which the probability of certain events is described in terms of sampling, replacing and adding balls in one or more urns or boxes.

Their success is essentially due to their ability of simplifying complex probabilistic ideas, making them intuitive and concrete, and yet guaranteeing a good level of abstraction, that allows for general results. Further details can be found in Johnson and Kotz [36], Balakrishnan [5] and Pemantle [51].

Urn models possess several interesting features, among which:

1. urns are efficient tools for studying chance experiments, especially when characterized by countable spaces; moreover, they represent an excellent way to describe the concept of "random choice";
2. simple urn schemes can be easily compounded into new ones in order to study more complex problems, the typical examples being urn chains and hyperurns;
3. urns have powerful and elegant combinatorial properties, that allow for general, complex results in a rather concise form;
4. there are many relationships and isomorphisms between urn models and other well-known mathematical objects, and this increases their flexibility;
5. for their natural connections with sampling schemes, urns are very useful tools in simulations.

The employment of urn schemes for modeling and representing infectious and epidemic phenomena, such as cascading failures, dates back to the pioneering works of Eggenberger and Pólya [27] at the beginning of the twentieth century.

Pólya urn, thanks to its intuitive reinforcement mechanism, has become the prototype for many probabilistic models for studying contagion and after effects. The behavior of this urn model is very simple yet ingenious. In its simplest version, we have an urn containing balls of two different colors. Every time we sample the urn we look at the color of the chosen ball and then put it back into the urn together with another ball of the same color. In this way, the more a given color has been sampled in the past, the more likely it will be sampled in the future. It is easy to understand how this naive replacement rule is able to reproduce a lot of self-reinforcing and contagious phenomena. For a detailed list, see once again Johnson and Kotz [36] or Mahmoud [43].

More recently other Pólya-like reinforced urns schemes have been used in epidemiology (Marshall and Olkin [44]) and more in general in the study of infectious phenomena, like the spreading of information in the society (Mahmoud [43]), innovation processes (Bottazzi and Secchi [10]) and firms' defaults (see e.g., Cirillo and Hüsler [17]). Nevertheless, all these interesting models often lack a spatial dimension, as they only deal with temporal contagion.

In the following section we will introduce a new urn model in which different urns are displaced and interact on a lattice. In particular, the reinforcement matrix of every urn will depend on the behavior of the neighboring urns, thus introducing a clear spatial relationship among them. Every urn represents an individual, a firm, a particle, or generally speaking, a system subject to shocks and infectious phenomena on the grid. Through spatial dependence, the default of every single urn can trigger off an avalanche of cascading failures. Applications can go from the analysis of the spreading of flu and electrical black-outs to the propagation of risk in the inter-banking system.

Since our model is meant to study interacting and joint defaults on a lattice, in some sense it can be considered as a spatial generalization of the urn-based generalized

extreme shock model of Cirillo and Hüsler [15], and it can thus be also seen as a spatiotemporal urn-based extreme shock model.

3. AN URN LATTICE MODEL FOR CASCADING FAILURES

We now introduce our model, based on the spatiotemporal interaction of different urns on a lattice. We call it *the urn lattice model*. The aim is to develop a framework for the analysis of spatiotemporal contagious events, such as cascading failures. Some inspiring works for the construction of our model are Cirillo [13], Marsili and Valleriani [45] and Muliere et al. [48]. In Cirillo [13] and Muliere et al. [48], some urn chains have been introduced to model self-reinforcing phenomena and knowledge updating, with some first applications [13] to credit risk and firms' defaults.

In particular, from Cirillo [13] we borrow some ideas related to urns' reinforcement. Self-reinforcement is definitely one of the main differences between our model and existing models such as Dai Pra et al. [20]. It is indeed known that reinforcement processes may lead to the rise of self-organized criticalities (see e.g., Bottazzi and Secchi [10] and Mahmoud [43], but also Marsili and Valleriani [45]) and power law behaviors, through preferential attachment and the so-called Matthew's principle. Since all these phenomena are widely present in nature, we do believe it is useful to enter reinforcement in cascading failures models.

In Marsili and Valleriani [45], the authors introduce a simple model that presents non-trivial self-organized critical properties. In particular, the construction describes a system of interacting units, modeled by simple urns, subject to perturbations and which occasionally break down. Each urn contains initially b black and 1 white balls. As in standard sandpile models, at each time step a site is randomly selected and the attempt to add a "grain of sand", that is, a white ball to the corresponding urn, is made. A ball is drawn from the selected urn: if the ball is white the attempt is successful and a new white ball is added to the urn. If it is black a "fatal accident" occurs, the urn becomes unstable and it collapses. At this point the urn is reset to 1 white ball and b black ones and, for each white ball previously present in the urn, a similar attempt to add a white ball is made on a randomly chosen nearest neighbor urn. In this way, white balls released by an unstable urn can provoke some fatal accident in nearby urns. The process stops when all balls are redistributed and no further break-down is observed.

The main differences between the model in Marsili and Valleriani [45] and our construction are: (1) our use of generalized randomly reinforced Pólya-like urns instead of the simple urns of Marsili and Valleriani [45], thus introducing a more flexible construction that allows for the modeling of systemic risk and fate; (2) the introduction of spatial dependence through the reinforcement mechanism of every urn rather than through the use of the "fatal accident" tool; (3) the possibility of allowing recovery from defaults (as in Marsili and Valleriani [45]) or not; and (4) the different stopping rules. Anyway, it is important to stress that both models produce interesting results related to self-organized criticalities and power laws.

Our model is also indebted to the huge physical literature related to particle fields, cellular automata and Ising models. In fact, in some sense, our construction can be considered an urn-based ferromagnetic Ising model, where the standard particles are substituted by urns.

The ferromagnetic Ising Model is one of the pillars of statistical mechanics: each site on a lattice can have two values (red/white, $1/0, +/-, \ldots$), and neighboring sites have an energetic preference to be the same value. As a system of $+/-$ spins, it represents a model for magnetism. Thought of as sites either occupied or vacant (1/0) on a lattice, it is a model for the liquid–gas transition: dense regions of occupied liquid are surrounded by dilute regions of mostly gas.

In a few words, the Ising model tries to reproduce the situations in which individual elements (e.g. atoms, animals, cells, human beings) modify their behavior in order to conform to the behavior of other individuals in their vicinity. More than 12,000 papers have been published between 1969 and 1997 using the Ising model (and the number is still growing). Given the large amount of papers in the field, we simply refer to Young [55] for more details.

3.1. The construction

We start by defining the basic ingredients of our urn construction. For simplicity, we use two-color urns, but all the results can be easily generalized using a countable set of colors. Let us consider:

1. A countable state space V, whose elements $v \in V$ are displaced to form a two-dimensional lattice.[1] We can also consider the points $v \in V$ as the vertices of an undirected graph $G = (V, E)$, where E is the set of edges connecting the points in V.
2. For every $v \in V$, at time $t = 0$, there exists an urn $U_0(v)$ containing $b_0(v) \geq 0$ black and $w_0(v) \geq 0$ white balls, such that $n_0(v) = b_0(v) + w_0(v) > 0$. The urns will be used to model the behavior of the different elements forming the network.
3. For every $v = (v_i, v_j) \in V$ define the closed neighborhood[2] set $S(v)$.
   For example, on a square lattice, rolled up to form an imaginary discrete torus, so that there is no ending point nor border problem, we could define the Minkowski distance of order 1 (1-norm distance) $d(u, v)$ and set $S(v) = \{u : d(u, v) \leq 1\} = \{(v_i, v_j), (v_{i+1}, v_j), (v_{i-1}, v_j), (v_i, v_{j+1}), (v_i, v_{j-1})\}$. Similarly we can define $S'(v) = \{u : d(u, v) = 1\}$ to be the open neighborhood. Anyway, it is worth underlining that, in our construction, $S(v)$ can be different for every v.

[1] Please note that v is a two-dimensional vector. We do not use bold characters in order not to perplex the notation.

[2] Heuristically, "closed" means that the neighborhood of v contains v itself. If v is not part of its neighborhood, then we speak of "open" neighborhood.

4. For every urn $v \in V$ we define, for $t = 0, 1, \ldots$, the quantities

$$Z_t^S(v) = \frac{\sum_{u \in S(v)} b_t(u)}{\sum_{u \in S(v)} n_t(u)}, \tag{1}$$

as the proportion of black balls in the complete neighborhood of state v, and

$$Z_t(v) = \frac{b_t(v)}{n_t(v)}, \tag{2}$$

as the proportion of black balls in urn $U(v)$ only. Then we define, for every $v \in V$, the random variable $X_t(v)$ distributed as a $Bernoulli(Z_t^S(v))$. $X_t(v)$ is thus the indicator of the random event "extraction of a black ball in the neighborhood of v".

5. Urn $U_t(v)$ defaults in t if $Z_t(v) \geq \xi$. Moreover, set $D_t(v) = 1_{Z_t(v) \geq \xi}$. We can define a stopping rule for which the entire lattice collapses when

$$\sum_{v \in V} D_t(v) \geq \gamma. \tag{3}$$

6. The evolution of every urn is Pólya-like, that is, for $t = 1, 2, \ldots$

$$b_t(v) = b_{t-1}(v) + X_{t-1}(v) Y_t(v)(1 - D_{t-1}(v)); \tag{4}$$

$$w_t(v) = w_{t-1}(v) + (1 - X_{t-1}(v)) Y_t(v)(1 - D_{t-1}(v)). \tag{5}$$

where $Y_t(v)$'s are random variables with marginal distributions $F_t(v)$. The specification of diverse $F_t(v)$ may produce several different behaviors.

In other words, the sampling process of every urn does not only depend on the urn itself, but also on the composition of the neighboring urns through $X_t(v)$ (that depends on $Z_t^S(v)$), as if every neighborhood were a hyperurn[3] $U^S(v)$ given by the aggregation of all the smaller urns forming the neighborhood itself.[4] In this way, for all $v \in V$, the reinforcement process of $U(v)$ is not only a function of the history of the urn itself (temporal contagion), as in the classical Pólya case, but it also depends on the neighboring states $S(v)$ (spatial dependence), and on a random quantity $Y_t(v)$, that represents the strength of reinforcement and which can both include simple fate, but also the impact of exogenous factors. In fact, the distribution (and hence the behavior) of $Y_t(v)$ may depend on several systemic factors that, for example, may modify its support. In general, the support of $F_t(v)$ can be any subset of $(-\infty, +\infty)$, so that balls can be both added and removed from the urns. Nevertheless, whether $Y_t(v)$ assumes

[3] We use the term hyperurn as in Mahmoud [43] indicating an urn that can be seen as combination of simpler urns.

[4] Note that if v and v' are neighbors, $v' \in S(v)$ and $v \in S(v')$. v and v' affect each other with 1 time period delay. In fact, the ball sampled in v at t depends on the composition of its neighborhood, which includes v', at time $t - 1$. The same is true for v'.

negative values, it makes sense to modify Eqs. (4) and (5) by substituting the term $1 - D_{t-1}(v)$ with $(1 - D_{t-1}(v))(1 - E_{t-1}(v))$, where $E_{t-1}(v)$ is equal to 1 if all the balls of a given color have been depleted at time $t - 1$ and 0 otherwise. In this way we avoid a negative number of balls in the urns, assuming that the running up of one of the colors corresponds to a default.

It is also important to stress that $Y_t(v)$ can be split into a $Y_t^w(v)$ for white balls and a $Y_t^b(v)$ for the black ones, or even be Y_t for the entire grid. All the following results still hold.

Thanks to the variability of $Y_t(v)$, the default of urn $U(v)$ can happen after the extraction of several black balls (we associate black balls to negative events), for a cumulative effect, or even after the extraction of one single black ball whose reinforcement is very strong because of $Y_{t-1}(v)$. Moreover, the probability of picking and adding black balls to a particular urn increases as the number of black balls increases in the neighboring urns. This reinforces the cascading effect.

For simplicity, we assume that recovery is not possible and, once a system is defaulted, its site is not occupied by another system. Moreover, note that, once an urn has collapsed, its composition is no longer modified. Anyway, most of the following results can be easily adapted to the case of recovery.

3.2. Basic properties

The urn lattice process possesses many interesting properties. Some are specific to the process itself and are given here below; some others can be extended to a more general class of models and are discussed in the next section.

Let us start by considering the behavior of the hyperurn $U^S(t)$, which we can call the neighborhood urn. Through Eqs. (4) and (5), it is clear that the evolution of $U^S(t)$ strictly depends on $Y_t(v)$, and hence on its distribution $F_t(v)$.

PROPOSITION 3.1: *Set $F_t(v) = F$, for all v and t.*

If F is degenerate at $\alpha \in \mathbb{N}^+$, then for every $v \in V$ the hyperurn $U^S(v)$, aggregating all the urns in the neighborhood of v, behaves as a two-color randomly reinforced Pólya urn, as introduced in May et al. [46].

For a general F with nonnegative support, hyperurn $U^S(v)$ is a randomly reinforced urn, as the one studied in Muliere et al. [47].

The non-negativity of the support of F is necessary in order to avoid the case of semi-sacrificial urns (see Flajolet et al. [30]), where sooner or later one of the colors disappears (i.e. all the corresponding balls are depleted), creating an absorbing state.

PROOF: If F is degenerate at $\alpha \in \mathbb{N}^+$, we have that $P(Y_t(v) = \alpha) = 1$, for all $v \in V$ and for all t.

Let us fix v and consider hyperurn $U^S(v)$ that, for simplicity and without any loss of generality, we assume to be based on a first-order neighborhood. For notational convenience set $W_t^S(v)$ and $B_t^S(v)$ to be the random numbers of white and black

balls respectively sampled from the urns forming the neighborhood $S(v)$ at time t. By construction both $W_t^S(v)$ and $B_t^S(v)$ are discretely distributed with support $[0, 5]$. Hence at every time step we add a total of 5α balls in $U^S(v)$: $\alpha B_t^S(v)$ are black and $\alpha W_t^S(v) = \alpha(5 - B_t^S(v))$ are white.

At time t, the probability of sampling (in the future) in a given sequence of k black balls out of a sample of m samples (i.e. bbwb…wwbwb) is

$$\frac{\sum_{u\in S(v)} b_t(u)}{\sum_{u\in S(v)} n_t(u)} \times \frac{\sum_{u\in S(v)} b_t(u) + \alpha B_{t+1}^S(v)}{\sum_{u\in S(v)} n_t(u) + 5\alpha} \times \frac{\sum_{u\in S(v)} w_{t+2}(u)}{\sum_{u\in S(v)} n_t(u) + 10\alpha} \cdots \frac{\sum_{u\in S(v)} b_{t+m-2}(u) + \alpha B_{t+m-1}^S(v)}{\sum_{u\in S(v)} n_t(u) + 5(m-1)\alpha}. \tag{6}$$

This corresponds to the behavior of the two-color randomly reinforced Pólya urn of May et al. [46].

In the case of general F with positive support, we have that $U^S(v)$ is a Pólya-like urn with random reinforcement. Eq. (6) is still valid, but now $B_t^S(v)$ can be any real number. Since the urn of May et al. [46] is nothing but a special case of the construction of Muliere et al. [47], the second statement of the proposition follows immediately. ■

In a very similar way, we can prove that also $U(v)$ behaves like the randomly reinforced urn of Muliere et al. [47], to which we refer for all the properties.

Since, in every case, $U^S(v)$ is Pólya-like, that is, it is positively reinforced, we have the following result.

PROPOSITION 3.2: *$\{Z_t^S(v)\}$ is a martingale with respect to the filtration $\mathcal{F}_t(v) = \sigma(X_1(v), Y_1(v), \ldots, X_{t-1}(v), Y_{t-1}(v))$ with values in $[0, 1]$. As a consequence of this $\{Z_t^S\}$ is a martingale w.r.t. $\mathcal{F}_t = \sigma(X_1, Y_1, \ldots, X_{t-1}, Y_{t-1})$ with values in $[0, 1]^V$, where $Z_t^S = \{Z_t^S(v) : v \in V\}$, $X_t = \{X_t(v) : v \in V\}$ and $Y_t = \{Y_t(v) : v \in V\}$.*

PROOF: Let us fix $v \in V$. Since $U^S(v)$ is a generalized Pólya urn, we automatically have that $Z_t^S(v)$ is a martingale, see May et al. [46] or Muliere et al. [47].

Hence, for every v there exists $Z_\infty^S(v)$ such that $P(\lim_{t\to\infty} Z_t^S(v) = Z_\infty^S(v)) = 1$. At this point, since V is countable, we have $\{Z_t^S\} \to_{a.s.} Z_\infty^S$. ■

PROPOSITION 3.3: *In the urn lattice model $P(\lim_{t\to\infty} t^{-1} \sum_{l=0}^{t-1} X_t = Z_\infty^S) = 1$.*

PROOF: For the proof we refer to the proof of Proposition 3.5 in the next section. The procedure is just the same. ■

In the case of $F_t(v) = F$ degenerate at $\alpha \in \mathbb{N}$, for all v and t, we have that Z_∞^S can be well approximated by a Beta distribution. This fact derives from the fact that $U^S(v)$ is a two-color randomly reinforced Pólya urn. Since this results has already been proven in May et al. [46], we refer to that paper for more details.

Another interesting property related to the behavior of $U^S(v)$ is that, for every $v \in V$, the sequence $\{X_t(v)\}_{t>0}$ is at least partially exchangeable, according to the definition of Aldous [2]. In very special cases of deterministic reinforcement, partial exchangeability is substituted by exchangeability. We better clarify these concepts in the next section.

3.3. Generalizing the urn lattice model

The simple urn lattice model can be easily generalized, in order to increase its flexibility. This more general construction can also be seen as an extension of the model of Paganoni and Secchi [50].

Let us consider:

1. A countable state space V, whose points $v \in V$ are displaced on a lattice. We consider the points $v \in V$ as the vertices of an undirected graph $G = (V, E)$, where E is the set of edges connecting the points in V.
2. For every $v \in V$, at time $t = 0$, there exists an urn $U_0(v)$ containing $b_0(v) \geq 0$ black and $w_0(v) \geq 0$ white balls, such that $n_0(v) = b_0(v) + w_0(v) > 0$.
3. For every urn $U_0(v)$ we define $Z_0(v) = b_0(v)/n_0(v)$ as the proportion of black balls in the urn. For $t = 0, 1, \ldots$ we also set $X_t(v) \sim Bernoulli(Z_t(v))$.
4. The evolution of every urn is Pólya-like, that is, for $t = 1, 2, \ldots$

$$b_t(v) = b_{t-1}(v) + X_{t-1}(v)\rho(v, S(v), Y_t(v)), \tag{7}$$

$$w_t(v) = w_{t-1}(v) + (1 - X_{t-1}(v))\rho(v, S(v), Y_t(v)), \tag{8}$$

$$Z_t(v) = \frac{b_t(v)}{n_t(v)}, \tag{9}$$

   where $\rho(\cdot)$ represents the reinforcement rule of the urn. This rule is a function of the state v itself, of a r.v. $Y_t(v)$ with distribution $F_t(v)$, and of $S(v)$, that represents the set of neighbors[5] of state v. The set $S(v)$ can be defined in several ways and it is related to the topology of the graph G. See Pérez [52] for more details.
5. Urn $U_t(v)$ defaults in t if $Z_t(v) \geq \xi$.
6. Let $D_t(v) = 1_{Z_t(v) \geq \xi}$. We can now define a stopping rule for which the entire lattice collapses when

$$\sum_{v \in V} D_t(v) \geq \gamma. \tag{10}$$

The generalized urn lattice model shows several interesting properties that hold notwithstanding the specification of parameters.

[5] Note that in this general formulation, $S(v)$ needs not to be a closed neighborhood, that is, point v can be excluded.

PROPOSITION 3.4: *$\{Z_t(v)\}$ is a martingale with respect to the filtration $\mathcal{F}_t(v) = \sigma(X_1(v), Y_1(v), \ldots, X_{t-1}(v), Y_{t-1}(v))$ with values in $[0,1]$. As a consequence of this $\{Z_t\}$ is a martingale w.r.t. $\mathcal{F}_t = \sigma(X_1, Y_1, \ldots, X_{t-1}, Y_{t-1})$ with values in $[0,1]^V$, where $Z_t = \{Z_t(v) : v \in V\}$, $X_t = \{X_t(v) : v \in V\}$ and $Y_t = \{Y_t(v) : v \in V\}$.*

PROOF: Let us fix $v \in V$. Then

$$E[Z_{t+1}(v)|\mathcal{F}_t(v)] = E\left[\frac{b_t(v)}{w_t(v)+b_t(v)}\frac{b_t(v)+\rho(v,S(v),Y_t(v))}{w_t(v)+b_t(v)+\rho(v,S(v),Y_t(v))}\right.$$
$$\left. + \frac{w_t(v)}{w_t(v)+b_t(v)}\frac{b_t(v)}{w_t(v)+b_t(v)+\rho(v,S(v),Y_t(v))}|\mathcal{F}_t(v)\right]$$
$$= E\left[\frac{b_t(v)}{w_t(v)+b_t(v)}|\mathcal{F}_t(v)\right] = Z_t(v). \quad (11)$$

In other words, for every $v \in V$, $\{Z_t(v)\}$ is a martingale with respect to $\mathcal{F}_t(v)$. Hence, for every v there exists $Z_\infty(v)$ such that $P(\lim_{t\to\infty} Z_t(v) = Z_\infty(v)) = 1$. Since V is countable, $\{Z_t\} \to_{a.s.} Z_\infty$. ■

PROPOSITION 3.5: *In the generalized urn lattice model, the averages $\left[t^{-1}\sum_{l=0}^{t} X_t\right]$ converge almost surely to Z_∞, that is, a law of large numbers applies.*

PROOF: We want to show that $P(\lim_{t\to\infty} t^{-1}\sum_{l=0}^{t} X_t = Z_\infty) = 1$. To do that, we make use of a common technique in martingale theory [35], that is, we construct an *ad hoc* martingale to prove the proposition.

We start by fixing $v \in V$ and we set $M_0(v) = 0$. Then for every $t \geq 1$ we define $M_t(v) = tZ_t(v) - \sum_{l=0}^{t} X_l(v)$.

$\{M_t(v)\}$ is clearly a martingale w.r.t. $\mathcal{F}_t(v)$:

$$E[M_t(v)|\mathcal{F}_{t-1}(v)] = E\left[tZ_t(v) - \sum_{l=0}^{t} X_l(v)|\mathcal{F}_{t-1}(v)\right]$$
$$= tZ_{t-1}(v) - \sum_{l=0}^{t-1} X_l(v) - E[X_t(v)|\mathcal{F}_{t-1}(v)]$$
$$= tZ_{t-1}(v) - \sum_{l=0}^{t-1} X_l(v) - Z_{t-1}(v)$$
$$= (t-1)Z_{t-1}(v) - \sum_{l=0}^{t-1} X_l(v)$$
$$= M_{t-1}(v).$$

Now, let us define $G_t(v) = M_t(v) - M_{t-1}(v)$. We aim to show that $\sum_{t=1}^{\infty} E[(G_t(v))^2]/t^2 < +\infty$; this is essential to complete the proof.

First note that for $n \geq 1$, $G_t(v) = t(Z_t(v) - Z_{t-1}(v)) + Z_{t-1} + X_t(v)$ and, thanks to basic inequalities,

$$\frac{E[(G_t(v))^2]}{t^2} \leq 3\left[E[(Z_t(v) - Z_{t-1}(v))^2] + \frac{1}{t^2}E[(Z_{t-1}(v))^2] + \frac{1}{t^2}E[(X_t(v))^2]\right].$$

Since $Z_t(v)$ is a (L^2-bounded) martingale, we know that $\sum_{t=1}^{\infty} E[(Z_t(v) - Z_{t-1}(v))^2] < +\infty$. Moreover,

$$\sum_{t=1}^{\infty} \frac{1}{t^2}E[(Z_{t-1}(v))^2] < +\infty \quad \text{and} \quad \sum_{t=1}^{\infty} \frac{1}{t^2}E[(X_t(v))^2] < +\infty,$$

because $E[(Z_{t-1}(v))^2] < +\infty$ and $E[(X_t(v))^2] < +\infty$.

Hence, $\sum_{t=1}^{\infty} E[(G_t(v))^2]/t^2 < +\infty$ and, by Burkholder's inequality [35],

$$\lim_{t\to\infty} \frac{M_t(v)}{t} = \lim_{t\to\infty} \left(Z_t(v) - \frac{\sum_{l=0}^{t-1} X_l(v)}{t}\right) = 0.$$

As usual, the countability of V allows us to complete the proof. ■

Let us now change for a moment our point of view and look at our generalized urn lattice model as a process on the indirected graph $G = (V, E)$. Then, let us introduce a useful definition.

DEFINITION 3.6: *Given an undirected graph $G = (V, E)$, a set of random variables $W = (W_v)_{v\in V}$ indexed by V form a Markov random field with respect to G if they satisfy the following equivalent Markov properties:*

- *Pairwise Markov property: any two non adjacent variables are conditionally independent given all the other variable, or $W_r \perp\!\!\!\perp W_s | W_{V\setminus\{r,s\}}$;*
- *Local Markov property: a variable is conditionally independent of all the other variables given its neighbors, or $W_r \perp\!\!\!\perp W_{cl(r)} | W_{S(r)}$, where $S(r)$ is the set of neighbors of r and $cl(r) = \{r\} \cup S(r)$ is the closed neighborhood of r;*
- *Global Markov property: any two subsets of variables are conditionally independent given a separating subset, or $W_A \perp\!\!\!\perp W_B | W_P$, where P contains all the paths from a point in A to a point in B.*

Given the above definition of Markov random field, it is straightforward to make the following remark.

Remark 3.7: The generalized urn lattice model defines a Markov random field, that is to say a spatial model in which a set of random variables have a Markov property described by an undirected graph. This is not true if $S(v) = V$.

Markov random fields are very popular in spatial statistics, image analysis and statistical mechanics, where they are also known as particle models (Kindermann and Snell [38]).

Markov random fields show a lot of interesting properties and this justify their broad use in applications. One of the most useful one is related to the functional form of the joint probability distribution on the whole lattice, which we can indicate with P_V. In fact, it can be shown that P_V shows a factorized form

$$P_V(v) \propto \prod_{c \in C} f_c(v_c), \tag{12}$$

where C represents the set of cliques[6] c, the factor f_c depends only on the variable subset $v_c = \{v_i, i \in c\}$ and $\prod_c f_c$ is summable over its configuration set (see [52] for more details).

In some particular cases, the factorization property can be very useful to compute the joint distribution over the lattice. Especially when it is easy to compute the probability distributions over the different cliques.

Remark 3.8: In the simple urn lattice case of the previous subsection, at every time t, $\mathcal{F}_{t-1}(v)$ and hence $\mathcal{F}_{t-1}$ are known. Moreover, for the Markov properties, conditionally on their neighborhoods, every two urns on the lattice are independent. It follows that, at every time step, the joint distribution over the entire grid can be expressed as the product of conditionally independent Pólya-like urns. If we also assume that $F_t(v) = F$ is degenerate at $\alpha > 0$, both for white and black balls, the joint distribution can be well-approximated by a product-Dirichlet process (as shown in Cirillo [14]). The product-Dirichlet process arises from the interaction of conditionally independent Pólya urns. It has been introduced in Cifarelli and Regazzini [12] and recently re-discussed in Giudici et al. [33].

Anyway, in general, it may be not possible to derive the joint distribution in a closed form. In such a case, it is obviously necessary to get it through simulations, approximating it with techniques such as kriging and markov chain Monte Carlo.

Another interesting characteristic of the generalized urn lattice construction is partial exchangeability, which is in common with reinforced urn processes, as introduced by Muliere et al. [48]. In other words, it means that, conditionally on Z_∞, the sequence of colors generated by urn $U(v)$ is asymptotically i.i.d. according to a Bernoulli distribution G with parameter $Z_\infty(v)$. Moreover, the sequences of colors generated by distinct urns are asymptotically independent. In this framework G is called de Finetti measure, as a consequence of the de Finetti representation theorem (de Finetti [21]).

For readers' convenience, we here recall two basic definitions related to partial exchangeability, as available in Aldous [2].

[6] In an undirected graph $G = (V, E)$ a clique is a subset c of the vertex set V, such that for every two vertices in c, there exists an edge connecting the two.

DEFINITION 3.9: *A process* $\{R_n\}_{n\geq 0}$ *on* $\{0,1\}^V$ *is partially exchangeable with de Finetti measure* $\mu(\Xi)$*, if there exist a random element* $\Xi \in \{0,1\}^V$ *such that, for all* $l \geq 1$ *and* $A_1, \ldots, A_l$ *belonging to the product sigma field of* $\{0,1\}^V$*, we have*

$$P[R_1 \in A_1, \ldots, R_l \in A_l] = E\left[\prod_{j=1}^{l} \mu(\Xi)(A_j)\right]. \tag{13}$$

DEFINITION 3.10: *A process* $\{R_n\}_{n\geq 0}$ *with state space* $\{0,1\}^X$ *is said asymptotically partially exchangeable if,* $\forall l \geq 1$ *and for* $n \to \infty$*, the following convergence in distribution holds:* $\mathcal{L}(R_{n+1}, \ldots, R_{n+l}) \to^d \mathcal{L}(U_1, \ldots, U_l)$*, where* $\{U_k\}_{k=1}^l$ *is a partially exchangeable process on* $\{0,1\}^X$*.*

PROPOSITION 3.11: *The process generated by the urn lattice model is asymptotically partially exchangeable with de Finetti measure* $\mu(Z_\infty)$*.*

To prove this proposition we need a lemma given in Aldous [2] that we state here without proof.

LEMMA 3.12 (Extensions of Exchangeability, Aldous [2]): *Let* $\{V_n\}$ *be an infinite exchangeable sequence directed by G. Let* $\{Y_n\}$ *be an infinite sequence and let* G_n *be a regular conditional distribution for* $Y_{n+1}|Y_1, \ldots, Y_n$*. Suppose that* $G_n \to_{a.s.} G$*.*

Then

$$(Y_{n+1}, Y_{n+2}, \ldots) \to^d (V_1, V_2, \ldots) \quad \text{as } n \to \infty.$$

PROOF OF PROPOSITION 3.11: We have two possibilities.

From one side, we can show that the product probability $\mu(Z_t) \times F$ represents the conditional probability distribution of (X_t, Y_{t+1}) given $\mathcal{F}_t$. Then, since $\{Z_t\}$ is a martingale we have that, with probability one, $\mu(Z_t) \times F$ converges to $\mu(Z_\infty) \times F$ on a subset of the probability space of the process. Thanks to Lemma 3.12 the proof is then completed.

Alternatively, we may use the results of Berti et al. [7] about conditionally identically distributed random variables. In this case the proof of the proposition is immediate, view that $\{Z_t\}$ is a martingale. ■

The fact that $U(v)$ generates partially exchangeable random variables has important implications from a Bayesian point of view (e.g. Ghosh and Ramamoorthi [34]). In particular it guarantees the possibility of developing Bayesian inference, even if this goes beyond the scope of the present paper.

For a complete introduction on exchangeability, its generalizations and its applications we refer to Aldous [2]. In Liggett et al. [39] it is possible to see some implications and uses of exchangeability in statistical mechanics.

4. SIMULATION ANALYSIS

In order to give a practical idea of the theoretical urn lattice model we have studied, we now present a simulation experiment of a simple version of the model. In particular, we analyse the behavior of the case in which neighborhoods are modeled as Pólya hyperurns (Sections 3.1 and 3.2), also assuming very simple forms for the distributions of random reinforcements. Such a modeling could represent a power grid or a system of interacting failing systems.

It goes without saying that it is not possible to simulate the generalized model of Section 3.3 without specifying its components, such as for example the neighborhood sets, the distributions of random reinforcements, etc.

For the experiment we use a square lattice of 60×60 urns characterized by a first order neighborhood system. In order to avoid barrier problems on the edges of the square, and to guarantee that all the nodes have degree 4, we imagine that the graph is folded to form a discrete torus.

As far as the evolution of the urns is concerned, we assume the following simplified rule for every $v \in V$:

$$b_t(v) = b_{t-1}(v) + X_{t-1}(v)Y_t^b(v)(1 - D_{t-1}(v)), \tag{14}$$

$$w_t(v) = w_{t-1}(v) + (1 - X_{t-1})Y_t^w(v)(1 - D_{t-1}(v)). \tag{15}$$

In other words, $Y_t^b(v)$ represents the strength of reinforcement for black balls, and hence its impact on the probability of defaulting, while $Y_t^w(v)$ represents recovery. In our simulations we will analyze both a deterministic case in which $Y_t^b(v) = \alpha$ and $Y_t^w(v) = \beta$, with $\alpha, \beta \in \mathcal{R}^+$, and a stochastic case, in which $Y_t^b(v)$ and $Y_t^w(v)$ are uniformly distributed random variables with support in $\mathcal{R}^+$. The choice of naive uniformly distributed random variables is due to the fact that, in this way, it is simpler for us to compare the results with those of the deterministic case.

For the deterministic case, a simple but realistic assumption is that $\alpha > \beta$. In modeling a disease, this means that its virulence is definitely stronger than the recovery rate. In the case of firms, on the contrary, a similar configuration of parameters can be summarized as "it is simpler to go bankrupt than to restructure". Consider that, for $\alpha = \beta$, every urn is a balanced Pólya urn.

In the stochastic case, as long as $E[Y_t^b(v)] \geq E[Y_t^w(v)]$, the same results of the deterministic scenario still hold. In general, the main difference is that the average time for the grid to collapse slightly decreases. If, on the contrary, $E[Y_t^b(v)] < E[Y_t^w(v)]$, the entire grid very rarely collapses, because the recovery process is too strong and it is able to cope very well with negative shocks. Naturally, the initial composition of the different urns have a strong influence on the behavior of the whole system. The higher the average number of black balls in the urns, the easier the grid can implode, notwithstanding the reinforcement rule.

Another important aspect of our simulations is that we assume that once an urn has collapsed it is no more reinforced, but it is considered dead. Nevertheless, it continues to influence its neighbors through Eqs. (14) and (15), even after default.

Table 1. Model Parameters

Reinf.	V	$w_0(v)$	$b_0(v)$	α	β	ξ	γ	limit sim. time
Det.1	$[1,60]^2$	95	5	4	1	0.9	500	10,000
Det.2	$[1,60]^2$	95	5	9	1	0.9	500	10,000
Stoc.	$[1,60]^2$	95	5	*Unif*$(2,6)$	*Unif*$(0.5,1.5)$	0.9	500	10,000

As far as the initialization of the process is concerned, Table 1 contains the values of the parameters for two deterministic (Det.1 and Det.2) and a stochastic (Stoc.) case. In particular, for the deterministic ones, we have used $\alpha = 4, 9$ and $\beta = 1$[7],while in the stochastic one the reinforcement is represented by two uniform distributions, *Unif*$(2,6)$ and *Unif*$(0.5,1.5)$, so that $Y_t^b(v) = 4$ and $Y_t^w(v) = 1$. This allows us to better compare, at this first stage, the impact of randomness on the system.

To simplify the analysis of the simulation results, we only show the situation in which all urns are initialized with the same amount of five black and 95 white balls,[8] but most of the results, for example, the distribution of grid failure times and clusters' size distribution, seem to be quite robust even with more initial heterogeneity, as long as the number of white balls is initially considerably higher than the number of black balls. This condition is necessary to avoid that the grid collapses too quickly, given a too high proportion of black balls.

The total number of urns on the lattice is 3,600 and we assume that the entire grid collapses when $\gamma = 500$ urns show a proportion of black balls greater than or equal to $\xi = 0.9$ (notice that the initial proportion of black balls in every urn is 0.05). It is evident that, cœteris paribus, the higher is γ, the longer is the average time to grid default.

The maximum number of iterations is set to be 10,000, but every simulation, on a total of 10,000 replications, has always stopped before this limit time, because of the collapse of the grid.

Figure 1 shows an example of simulation for the urn model for cascading failures. In particular, it represents the state of the lattice when the entire grid collapses. Every little square embodies an urn. The gray scale helps in reading the graph: the darker the color, the higher $Z_t(v)$. Figure 2 displays, for the same simulation, the defaulted urns, that is, those urns for which $Z_t(v) \geq \xi$.

It is worth nothing that the emergence of clear clusters of defaults. This behavior is essentially due to the spatial contagion, for which the probability of defaulting of every urn is also a function of the compositions of its neighbors.

Figure 3 contains a realization of the joint bivariate distribution of the proportions of black balls over the lattice. Its highly irregular shape gives an idea of the complexity of finding a closed form for it, even in the simple case of constant reinforcements α

[7] More in general we have performed the analysis with $\alpha, \beta \in [0, 20]$. Here we only show some of the results. Figures and tables are available upon request.

[8] Anyway it is interesting to note how heterogeneity raises thanks to the model. This could be interesting from the point of view of Delli Gatti et al. [22].

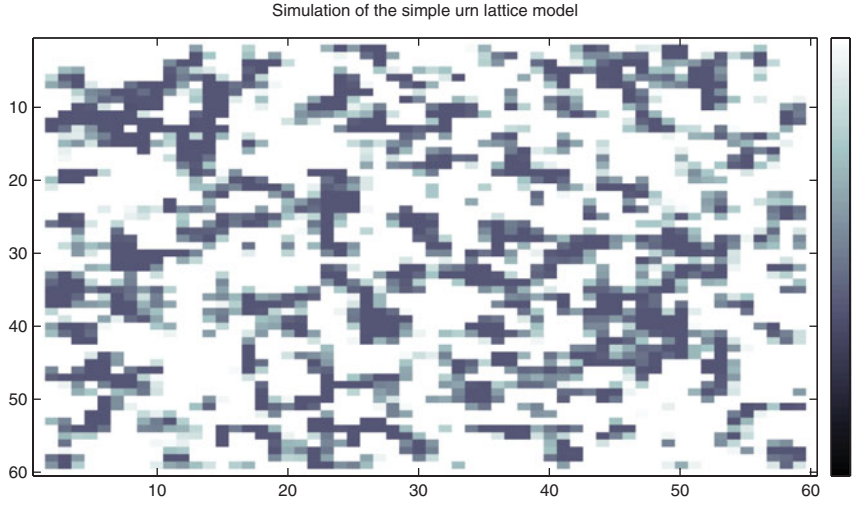

FIGURE 1. (Color online) Example of lattice configuration at the end of a simulation: $n = 60$, $\alpha = 4$, $\beta = 1$ and $\gamma = 500$.

and β. We suspect that, at least in this particular simple case, the joint distribution may be well approximated by a generalized version of the product-Dirichlet process of Cifarelli and Regazzini [12], but further studies are needed.

If we analyze clusters' size, we find an interesting result about its distribution. Figure 4 shows an evident discrete power law behavior. This holds true for all the initial settings of parameters and it is consistent with the statistical mechanic literature (e.g Aoki [4], Murri and Pinto [49] and Costantini et al. [19]).

As known, a random variable q is said to follow a power law if its probability distribution $f(q)$ is such that

$$f(q) \propto Aq^{-\delta}, \tag{16}$$

where A is some constant.

To estimate the parameter δ we use MLE. In particular, if the data follow a discrete power law distribution for $q \geq q_0$, the maximum likelihood exponent is the solution of the transcendental equation:

$$\frac{\zeta'(\hat{\delta}, q_0)}{\zeta(\hat{\delta}, q_0)} = -\frac{1}{n}\sum_{i=1}^{n} \log \frac{q_i}{q_0}, \tag{17}$$

where $\zeta(\cdot, \cdot)$ is the incomplete zeta function and q_i's are the ordered values. For further details on the estimation of the power law index with integer-valued data points, we refer to Clauset et al. [18].

Obviously, in order to have a good estimate of δ, we need to correctly identify the lower bound q_0 that represents the value above which the power law behavior

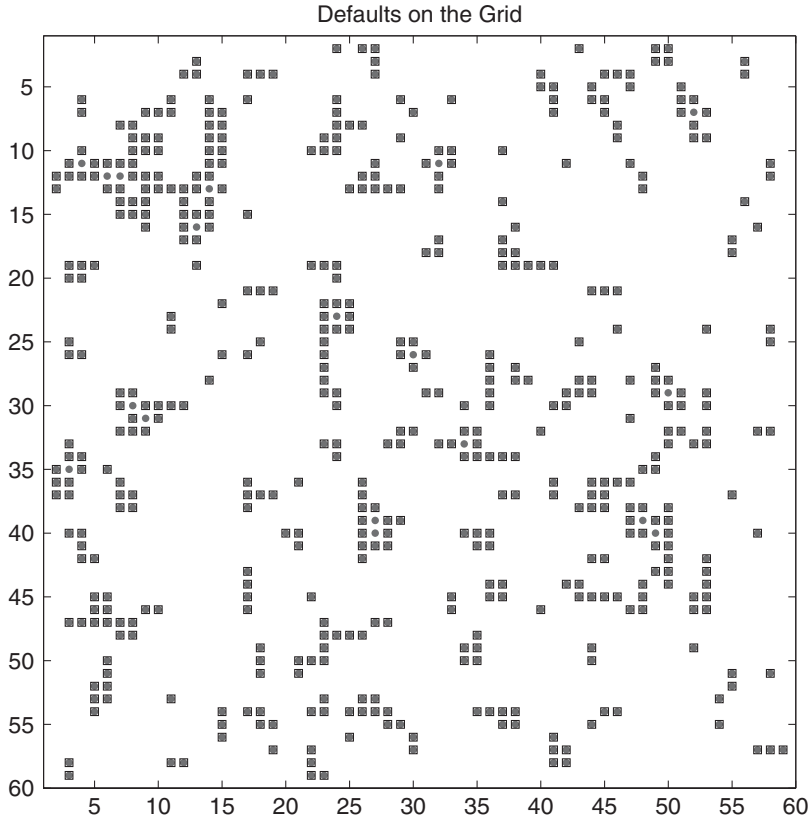

FIGURE 2. Example of the total number of defaults at grid's collapse: $n = 60$, $\alpha = 4$, $\beta = 1$ and $\gamma = 500$. Squares indicate urns that are on the borders of a cluster.

approximately holds. For this purpose, we use a methodology similar to the one proposed in Cirillo and Hüsler [16], searching for the value $\widehat{q}_0$ which minimizes the Anderson–Darling statistic

$$AD = -N - \sum_{i=1}^{N} \frac{2i-1}{N} \left[\ln F(q_i) + \ln(1 - F(q_{N+1-i})) \right], \tag{18}$$

where $q_1, \ldots, q_N$ are once again ordered data.

In our simulation experiments, cluster size goes from a minimum of 1 to a maximum of 63 neighboring urns. Interestingly, there seems to be a slightly positive relationship between the value of α and the average size of the clusters. This can be explained by the fact that the impact of failed neighbors on the evolution of every urn increases with α (and decreases with β). Randomness, at least in our formulation, does not introduce relevant modifications.

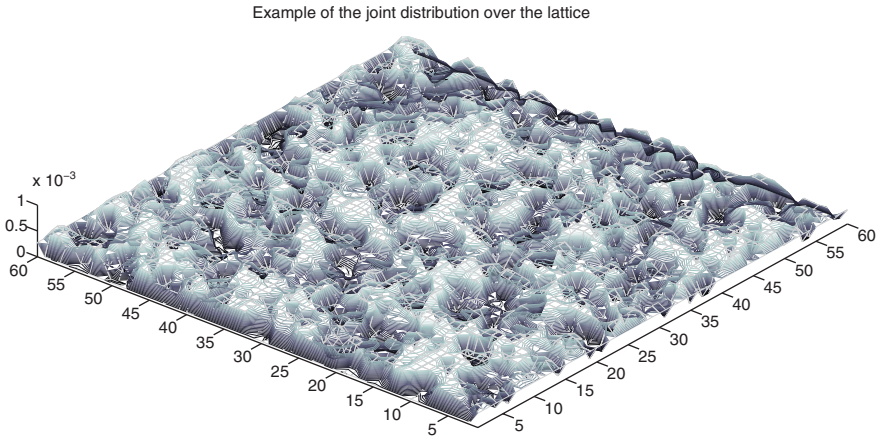

FIGURE 3. (Color online) Realization of the joint distribution of the proportions of black balls over the lattice: $n = 60$, $\alpha = 4$, $\beta = 1$ and $\gamma = 500$.

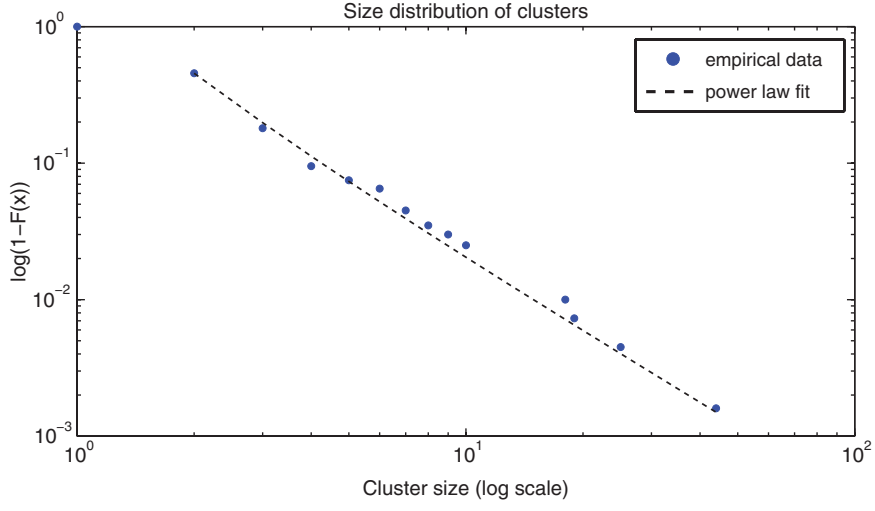

FIGURE 4. (Color online) Distribution of defaults' cluster size for $\alpha = Unif\,(2, 6)$ and $\beta = Unif\,(0.5, 1.5)$.

Regarding the estimates, we find that, in all the simulations, $\hat{\delta} \in (3, 3.5]$, indicating a particularly fat tail. For the simulation represented in Figure 4, $\hat{\delta} = 3.17$ with a standard error equal to 0.7742, while the threshold value is $\hat{q}_0 = 2$.

Analyzing the number of clusters, we have noted that, apart from a few exceptions (<0.2%), their total number is always in the range [37,49], for all the 10,000 replications we have performed for the different parameters' configurations of Table 1.

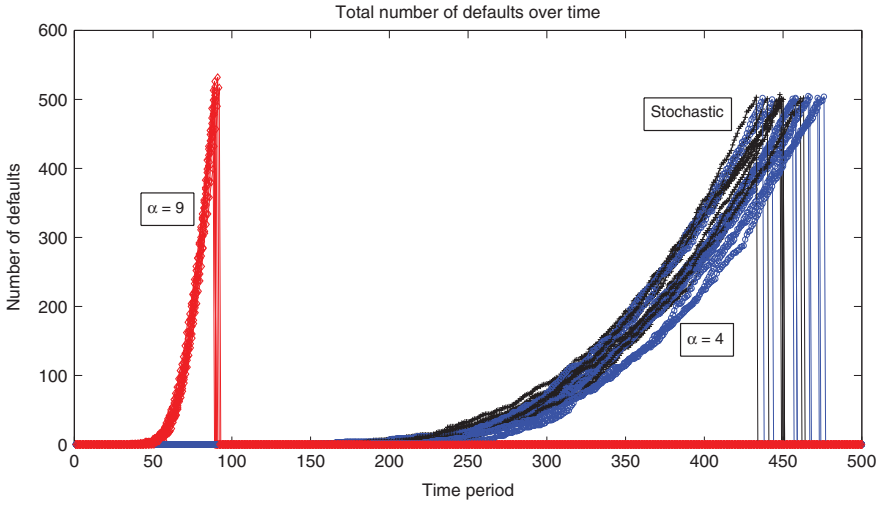

FIGURE 5. (Color online) The number of defaults on the grid over time in 10 different simulations (randomly extracted from a total of 10,000 replications) for the deterministic case with $\alpha = 4$ and 9, and for the stochastic one. The number of defaults grows until the entire grid collapses ($n.defaults \geq 500$).

This is interesting since, in multivariate analysis, there is a rule of thumb according to which the optimal number of clusters (n_{clust}) for cluster analysis can be determined with the simple formula $n_{clust} \simeq (\frac{n.obs}{2})^{-1/2}$, where $n.obs$ is the number of observations (see, for e.g., Johnson and Wichern [37]). In our case, $n.obs = 3{,}600$, that is to say the number of urns considered, so $n_{clust} \simeq 42.42$. It goes without saying that this approximation seems to be particularly good for us[9].

Let us now consider defaults. The number of defaults on the grid is clearly dependent on several quantities. For example, the higher the threshold value ξ, the more time is needed for every urn to collapse. Moreover, the larger is α, the quicker the urn defaults. The opposite is true if we consider β. Figure 5 displays the number of defaults on the grid over time in 10 different simulations (randomly selected from all the replications) for the deterministic and the stochastic cases. As expected, the higher α the faster the number of defaults increases. Anyway, it is particularly interesting to notice how the deterministic (for $\alpha = 4$) and the stochastic cases behave similarly: in general the stochastic reinforcement only slightly anticipates grid's default. This is probably due to the small variability of the stochastic reinforcement rules we have chosen. In fact, if we increase the range of the support of the two distributions of $Y_t^b(v)$ and $Y_t^w(v)$, always assuming $E[Y_t^b(v)] \geq E[Y_t^w(v)]$, the difference between the stochastic and the determinist models becomes more visible. In particular, if we increase the variability of $Y_t^b(v)$, not modifying the support of $Y_t^w(v)$, the curve of the

[9] Among the exceptions the maximum number of clusters is 73.

number of defaults over the grid gets similar to the deterministic case with $\alpha = 9$, moving to the left (the number of defaults grows in a faster way). If we augment the variability of $Y_t^w(v)$ only, on the contrary, the number of defaults increases more slowly and the curve shifts to the right. Further, if we modify both supports, there is no unique behavior for the system and the impact of randomness is more evident.

Finally, if we analyze the failure time of the entire grid, we find that its distribution is well approximated by a generalized extreme value (GEV) distribution. This evidence is persistent under the different initial parameters' configurations.

A random variable q is GEV distributed if, for all $1 + \gamma(q - \mu)/\sigma > 0$, its cumulative distribution function is

$$F(q) = \exp\left(-\left(1 - \lambda\frac{(q-\mu)}{\sigma}\right)^{-1/\lambda}\right), \tag{19}$$

where $\mu \in \mathbb{R}$, $\lambda \in \mathbb{R}$ and $\sigma > 0$.

For the estimation of the parameters of a GEV distribution using maximum likelihood, probability weighted moments and other methods), we refer to any text in extreme value analysis, for example Falk et al. [29] or Embrechts et al. [28]. However, note that the failure time is a discrete r.v., whereas the GEV are absolutely continuous distributions. Hence the estimation may incur some problems that, nevertheless, can be successfully treated. We refer to Embrechts [28] for further details.

Table 2 contains our estimates for the two deterministic cases with $\alpha = 4$ and 9, and for the stochastic one with uniformly distributed reinforcement. As expected, for $\alpha = 9$ the distribution of grid failure times is more concentrated on smaller values, while for $\alpha = 4$ and with random reinforcement the average default time for the grid increases. Once again, it is worth to underline how the stochastic reinforcement rules we have chosen do not substantially modify the results with $\alpha = 4$, indicating that this type of randomness does not add much disturbances in the simulations. In the future we surely aim to investigate other stochastic reinforcement rules to study their impact on the evolution of the system.

Figure 6 contains the fitting given by the GEV distribution, which is also supported by nonparametric goodness-of-fit tests such as the Anderson–Darling. As said, this test is based on the statistic introduced in Eq. (18), and it makes use of the fact that, if the tested data come from the hypothesized distribution, they can be transformed to uniformly distributed random variables, and then tested for uniformity with a distance test. For further details see Anderson and Darling [3]. In our case, the goodness-of-fit

TABLE 2. GEV Distribution Estimates for the Time of Grid Default with Standard Errors.

sim	μ	σ	λ	sample mean	sample var
$\alpha = 4$	454.08 ± 0.23	10.63 ± 0.16	-0.21 ± 0.01	458.37	123.91
$\alpha = 9$	89.40 ± 0.03	1.27 ± 0.19	-0.23 ± 0.01	89.89	1.71
random	449.42 ± 1.26	10.50 ± 0.83	-0.20 ± 0.06	453.74	123.14

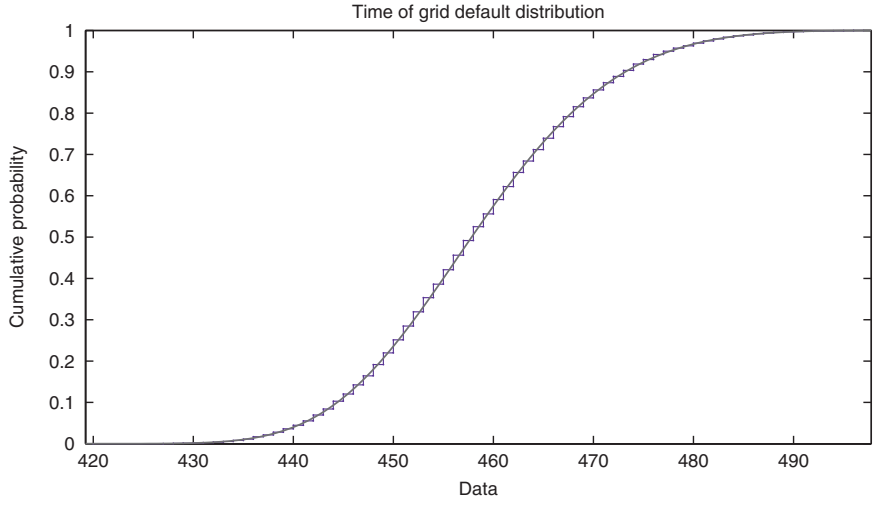

FIGURE 6. (Color online) Ecdf and GEV fit for the time of grid default with $\alpha = 4$, $\beta = 1$ and $\gamma = 500$.

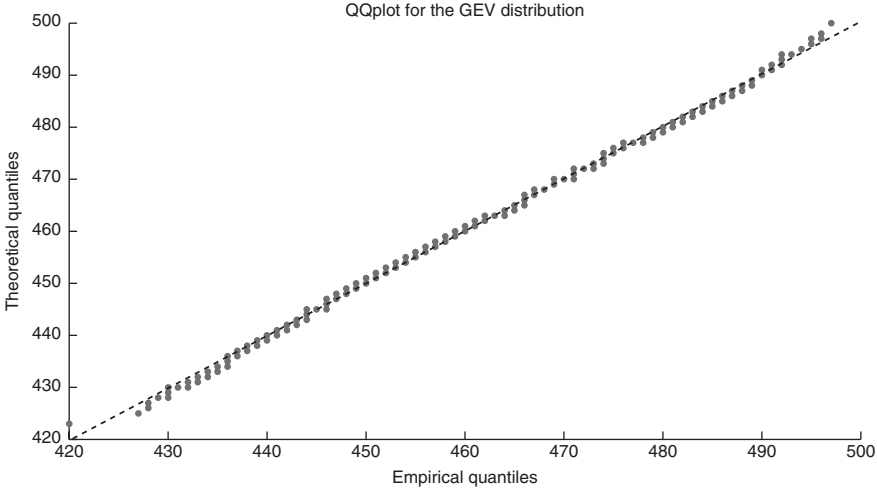

FIGURE 7. GEV QQplot for the distribution of the time of grid default with $\alpha = 4$, $\beta = 1$ and $\gamma = 500$.

is performed with estimated parameters, that is why we have to use the adjustments proposed in Ahmad [1] (but also in Drees et al. [25] and Choulakian and Stephens [11]).

Figures 7 and 8 display the quantile quantile and the probability plots, for $\alpha = 4$ and 9, of the distribution of the time of grid default, using the GEV as theoretical

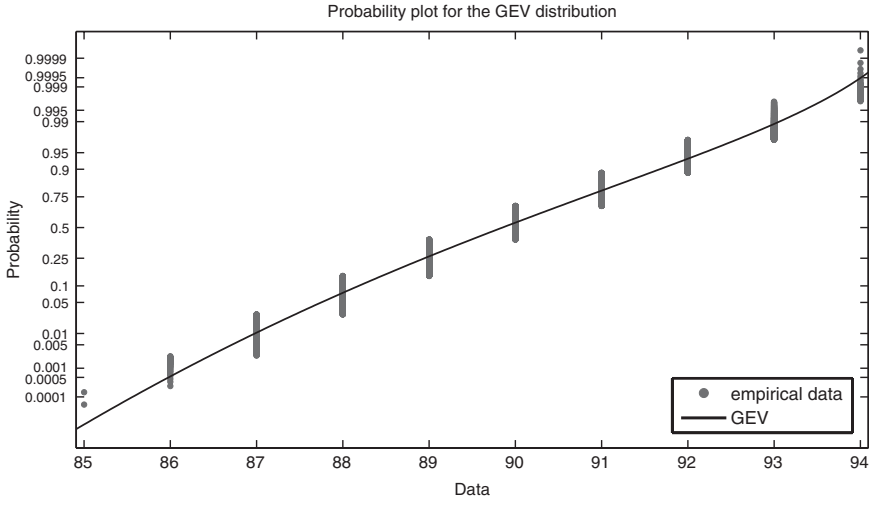

FIGURE 8. Probability plot of the GEV fit for the time of grid default with $\alpha = 9$, $\beta = 1$ and $\gamma = 500$.

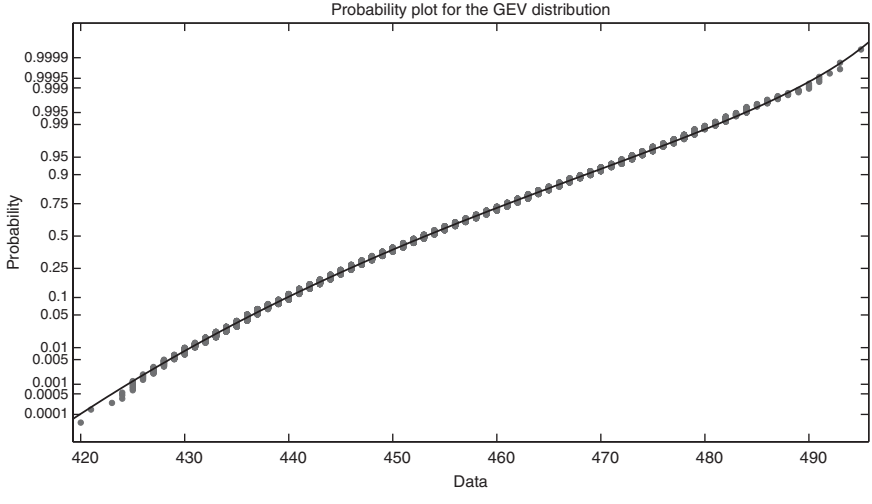

FIGURE 9. Probability plot of the GEV fit for the time of grid default with $\alpha = \mathit{Unif}(2, 6)$, $\beta = \mathit{Unif}(0.5, 1.5)$ and $\gamma = 500$.

distribution. Even in this case the goodness-of-fit is evident, although the problem of discreteness is quite manifest, especially with $\alpha = 9$, for which the data are more concentrated. Finally, Figure 9 shows a very good probability plot in the case of random reinforcement.

5. CONCLUSIONS

We have introduced a spatiotemporal urn model, in which several urns interact on a lattice. In particular, every Pólya-like urn shows a special reinforcement rule, which is a function of time (as in the standard Pólya), of the neighboring urns and their compositions (spatial dependence) and of a random variable, which can be equal for all the urns or not, representing exogenous systemic shocks and even fate.

The model shows interesting probabilistic properties, it is quite general and flexible, and it can be easily simulated. For these reasons, in the last part of the paper, we have shown a first simulation experiment, using both deterministic and stochastic reinforcement rules. In both cases we obtain interesting results concerning the number of defaults on the grid and the times of system's collapse.

We believe that our model can be used to analyze different situations in which spatiotemporal contagion and interacting defaults arise, causing cascading failures. A short list could be: power grids and power outages, firms' defaults in industrial districts, epidemic diseases, cancer spread through cells, etc.

The development and the study of new special cases and their applications will be investigated in the future. In particular, we are interested in finding models for which it is possible to derive results, for example the joint distribution over the lattice, in a closed form.

Similarly, it would be worth to compare in detail our construction with several existing models of cascading failures.

Acknowledgments

This work has been supported by the Swiss National Science Foundation. The authors are grateful to the Editor for his useful comments.

References

1. Ahmad, M.I. (1988). Applications of statistics in flood frequency analysis. PhD thesis, University of St. Andrews.
2. Aldous, D.J. (1985). Exchangeability and related topics. *Lecture Notes in Mathematics* vol. 117, New York: Springer.
3. Anderson, T.W. & Darling, D.A. (1952). Asymptotic theory of certain goodness-of-fit criteria based on stochastic processes. *Annals of Mathematical Statistics* 23: 193–212.
4. Aoki, M. (2000). Cluster size distributions of economic agents of many types in a market. *Journal of Mathematical Analysis and Applications* 249: 32–52.
5. Balakrishnan, N. (1997). *Advances in combinatorial methods and applications to probability and statistics*. Berlin: Birkhäuser.
6. Bernoulli, J. (1713). *Ars Conjectandi*. German version in Wahrscheinlichkeitsrechnung: Ars conjectandi. 1., 2., und 4. Teil (1998). Berlin: Harry Deutsch.
7. Berti, P., Pratelli, L. & Rigo, P. (2004). Limit theorems for a class of identically distributed random variables. *Annals of Probability* 32: 2029–2052.
8. Bhargava, S.C. & Mukherjee, A. (1994). Evolution and technological growth in a model based on stochastic cellular automata. In L., Leydesdorff, P., Van den Besselaar, (eds.), *Evolutionary economics and chaos theory: new directions in technology studies*. London: Pinter.

9. Blume, L.E. & Durlauf, S.N. (2003). Equilibrium concepts for social interaction models. *International Game Theory Review* 5: 193–209.
10. Bottazzi, G. & Secchi, A. (2003). Why are distributions of firm growth rates tent-shaped? *Economics Letters* 80: 415–420.
11. Choulakian, V. & Stephens, M.A. (2001). Goodness-of-fit tests for the generalized Pareto distribution. *Technometrics* 43: 478–484.
12. Cifarelli, D.M. & Regazzini, E. (1978). Problemi statistici non parametrici in condizioni di scambiabilità parziale. Impiego di medie associative. In: Istituto di Matematica Finanziaria dell'Università di Torino, *Serie III*, No. 12. English translation available from http://www.unibocconi.it/wps/allegatiCTP/CR-Scamb-parz[1].20080528.135739.pdf
13. Cirillo, P. (2008). New urn approaches to shock and default models. PhD thesis, Milan: Bocconi University.
14. Cirillo, P. (2012). A simple model of spatially dependent urns. Working Paper, University of Bern.
15. Cirillo, P., & Hüsler, J. (2009). An urn-based approach to generalized extreme shock models. *Statistics and Probability Letters* 79: 969–976.
16. Cirillo, P., & Hüsler, J. (2009). On the upper tail of Italian firms size distribution. *Physica A Statistical Mechanics and Applications* 388: 1546–1554.
17. Cirillo, P., & Hüsler, J. (2011). Shock models for defaults: parametric and nonparametric approaches. In D.R., Hunter, D.S.P., Richards, J.L., Rosenberger (eds.) *Nonparametric statistics and mixture models*. Singapore: WSP.
18. Clauset, A., Shalizi, C.R. & Newman, M.E.J. (2009). Power-law distributions in empirical data. *SIAM Review* 51(4): 661–678.
19. Costantini, D., Donadio, S., Garibaldi, U. & Viarengo, P. (2005). Herding and clustering: Ewens vs. Simon-Yule models. *Physica A Statistical Mechanics and Applications* 355: 224–231.
20. Dai Pra, P, Runggaldier, W.J., Sartori, E. & Tolotti, M. (2009). Large portfolio losses: a dynamic contagion model. *Annals of Applied Probabability* 19: 347–394.
21. de Finetti, B. (1975). *Theory of probability*. New York: John Wiley and Sons.
22. Delli Gatti, D., Gallegati, M., Greenwald, B.C., Russo, A. & Stiglitz, J.E. (2009). Business fluctuations and bankruptcy avalanches in an evolving network economy. *Journal of Econimic Interaction and Coordination* 4: 195–212.
23. Dobson, I., Carreras, B.A., Lynch, V.E. & Newman, D.E. (2007). Complex systems analysis of series of blackouts: cascading failure, critical points, and self-organization. *Chaos* 17, 026103.
24. Dobson, I., Wierzbicki, K.R., Carreras, B.A., Lynch, V.E. & Newman, D. (2006). An estimator of propagation of cascading failures. *Thirty-ninth Hawaii International IEEE Conference on System Sciences Proceedings*.
25. Drees, H., de Haan, L. & Li, D. (2006). Approximations to the tail empirical distribution function with application to testing extreme value conditions. *Journal of Statistical Planning and Inference* 136: 3498–3538.
26. Durrett, R. (2010). Some features of the spread of epidemics and information on a random graph. *Proc. Natl. Acad. Sei.* 107: 4491–4498.
27. Eggenberger, F. & Pólya, G. (1923). Über die Statistik verketteter Vorgänge. *Zeitschrift für Angerwandte Mathematik and Mechanik* 1: 279–289.
28. Embrechts, P., Klüppelberg, C. & Mikosch, T. (1997). *Modelling extremal events for insurance and finance*. Berlin: Springer.
29. Falk, M., Hüsler, J. & Reiss, R-D. (2004). *Laws of small numbers: extremes and rare events*. Basel: Birkhäuser.
30. Flajolet, P., Dumas, P. & Puyhaubert, V. (2006). Some exactly solvable models of urn process theory. *Fourth Colloquium on Mathematics and Computer Science DMTCS* 59.
31. Frey, R. & Backhaus, J. (2008). Pricing and hedging of portfolio credit derivatives with interacting default intensities. *International Journal of Theoretical and Applied Finance* 11: 611–631.
32. Giesecke, K. & Weber, S. (2006). Credit contagion and aggregate losses. *Journal of Economic Dynamics and Control* 30: 741–767.

33. Giudici, P., Mezzetti, M. & Muliere, P. (2003). Mixtures of products of Dirichlet processes for variable selection in survival analysis. *Journal of Statistical Planning and Inference* 111: 101–115.
34. Ghosh, J.K. & Ramamoorthi, R.V. (2002). *Bayesian nonparametrics*. New York: Springer.
35. Hall, P. & Heyde, C.C. (1980). *Martingale limit theory and its applications*. San Diego: Academic Press.
36. Johnson, N.L. & Kotz, S. (1977). *Urn models and their applications*. New York: Wiley.
37. Johnson, R.A. & Wichern, D.A. (2007). *Applied multivariate analysis*. New Jersey: Prentice Hall.
38. Kindermann, R. & Snell, J.L. (1980). *Markov random fields and their applications*. San Francisco: AMS Press.
39. Liggett, T.M., Steif, J.E. & Tóth, B. (2007). Statistical mechanical systems on complete graphs, infinite exchangeability, finite extensions and a discrete finite moment problem. *Annals of Probability* 35: 867–914.
40. Liggett, T. (2009). *Stochastic interacting systems: Contact, voter and exclusion processes*. New York: Springer.
41. Lindley, V. & Singpurwalla, N.D. (2002). On exchangeable, causal and cascading failures. *Statistical Science* 17: 209–219.
42. Lorenz, J., Battiston, S. & Schweitzer, F. (2009). Systemic risk in an unifying framework for cascading processes on networks. *European Physical Journal B* 71: 441–460.
43. Mahmoud, H.M. (2009). *Polya Urn Models*. Boca Raton: CRP Press.
44. Marshall, A.W. & Olkin, I. (1993). Bivariate life distributions from Polya's urn model for contagion. *Journal of Applied Probability* 30: 497–508.
45. Marsili, M. & Valleriani, A. (1998). Self organization of interacting Polya urn. *European Physical Journal B* 3: 417–420.
46. May, C., Secchi, P. & Paganoni, A. (2002). On a two-color generalized Polya urn. *Metron* LXIII, 115–134.
47. Muliere, P., Paganoni, A. & Secchi, P. (2006). A randomly reinforced urn. *Journal of Statistical Planning and Inference* 136: 1853–1874.
48. Muliere, P., Secchi, P. & Walker, S.G. (2000). Urn schemes and reinforced random walks. *Stochastic Processes and their Applications* 88: 59–78.
49. Murri, N. & Pinto, N. (2002). Cluster size distribution in self-organized systems. *Physica B Condensed Matter* 321: 404–407.
50. Paganoni, A.M. & Secchi, P. (2004). Interacting reinforced-urn systems. *Advances in Applied Probability* 36: 791–804.
51. Pemantle, R. (2007). A survey of random processes with reinforcement. *Probability Surveys* 4: 1–79.
52. Pérez, P. (1998). Markov random fields and images. *CWI Quarterly* 11: 413–437.
53. Swift, A. (2008). Stochastic models of cascading failures. *Journal of Applied Probability* 45: 907–921.
54. Wang, J., Liu, Y. & Jiao, Y. (2009). A new cascading failure model with delay time in congested complex networks. *Journal of System Science and System Engineering* 18: 369–381.
55. Young, A.P. (1998). *Spin glasses and random fields*. Singapore: WSP.