Phase transitions of Cellular Automata

Franco Bagnoli and Raúl Rechtman

Abstract We explore some aspects of phase transitions in cellular automata. We start recalling the standard formulation of the Monte Carlo approach for a discrete system. We then formulate the cellular automaton problem using simple models, and illustrate different types of possible phase transitions: density phase transitions of first and second order, damage spreading, dilution of deterministic rules, asynchronism-induced transitions, synchronization phenomena, chaotic phase transitions and the influence of the topology.

1 Introduction: Monte Carlo simulations

The main results of statistical mechanics is that of expressing the probability distribution of a statistical ensemble in terms of its constraints. Just to be concrete, let us consider a discrete system that can be described by N Boolean variables $x_i \in \{0,1\}$, $i = 1,\ldots,N$ located in sites connected by a graph defined by an adjacency matrix $a_{ij} = 1$ if i is connected to j and zero otherwise. A configuration of the system is expressed as $x = (x_1, x_2, \ldots, x_N)$.

Let us denote by $E(x)$ is energy of such a configuration, and with $P(x)$ the probability of observing it.

The simplest way of deriving the equilibrium probability distribution is that following the principle of maximum entropy. One has to maximize the entropy

$$S = -\sum_x P(x)\log(P(x))$$

Franco Bagnoli
Department of Physics and Astronomy and CSDC, University of Florence, via G. Sansone 1 50019 Florence Italy; also INFN, sez. Firenze e-mail: franco.bagnoli@unifi.it

Raúl Rechtman
Instituto de Energías Renovables, Universidad Nacional Autónoma de México, Apdo. Postal 34, 62580 Temixco Mor., Mexico e-mail: rrs@ier.unam.mx

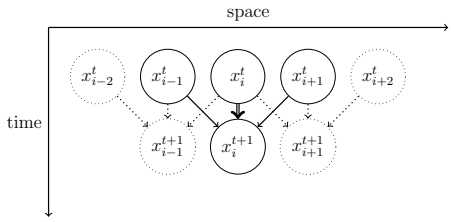


Fig. 1 Monte Carlo neighbourhood (2 cells plus the same-cell link).

with the given constraints. In the so-called "canonical ensemble" (a system in contact with a heat bath that keeps the temperature constant), the average energy

$$U = \sum_x E(x)P(x)$$

is kept constant by the heat bath. It is straightforward to derive the probability distribution $P(x)$,

$$P(x) = \frac{1}{Z}\exp(-\beta E(x))$$

where β corresponds to inverse temperature and Z the normalization constant (partition function). In principle this solves the problem of computing the average value $\langle A\rangle$ of an observable $A(x)$,

$$\langle A\rangle = \sum_x A(x)P(x).$$

The problem is that in general the number of configurations is huge, and therefore a brute-force evaluation of this sum is not feasible.

The Monte-Carlo technique allows one to compute the time-average $\overline{A}$ of the observable over a fictitious trajectory $x(t)$

$$\overline{A} = \frac{1}{T}\sum_{t=0}^{T} A(x(t)).$$

The trajectory is obtained by defining the conditional probability $M(x|y)$ of getting $x \equiv x(t+1)$ given $y \equiv x(t)$ (i.e., defining a Markov chain), as a function of the difference in energy of x and y. Clearly, speaking of numeric simulations, we refer to finite chains.

The main requirement is that the Markov chain has to be ergodic, i.e., each state can be reached from any other state in a finite number of steps (not being trivially periodic), and that the iteration of the procedure leads to a unique probability distribution.

Considering now the stochastic sampling, this property assures that a long enough trajectory visits all states a number of times proportional to the asymptotic distribution.

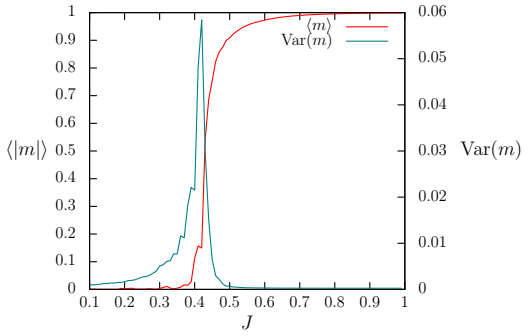


Fig. 2 Phase transition for the Ising model in 2D with nearest neighbour interactions. Average magnetization $|\langle m \rangle|$ and variance as a function of the rescaled coupling J for $H = 0$. Size 40×40, $T = 4000$, transient $4 \cdot 10^4$.

We shall now consider the temporal evolution of the probability distribution, denoted as $P(x,t)$. We have

$$P(x,t+1) = \sum_y M(x|y)P(y,t),$$

or, in vectorial terms

$$P(t+1) = MP(t).$$

$M(x|y)$ is in general decomposed into a series of N stochastic "local" moves that occur with probability $\tau(x_i|Y_i)$, which is the probability of getting x_i given its neighbourhood $Y_i = \{y_j : a_{ij} = 1\}$. One can visualize the Monte Carlo procedure as the evolution of a time-space graph, in which the connections are such that $(i,t+1)$ is connected to (j,t) if $a_{ij} = 1$, and in any case a cell at time $t+1$ is connected to the cell in the same location at time t (see Fig. 1). Each Monte Carlo time step is decomposed in N microscopic steps, in which just one random site is updated, and the others are copied ($x_i(t+1) = x_i(t)$).

The actual trajectory is computed by drawing, for each site i and time t, uniformly distributed random numbers $r_i(t)$ and computing Boolean quantities like $[r_i(t) < \tau(x_i|Y_i)]$, where $[\cdot] = 1$ if $\cdot$ is true and zero otherwise. If one thinks of extracting all the random numbers before the simulation, the trajectories are deterministic over the random field $r_i(t)$.

1.1 An example: the Ising model

Let us illustrate these concepts with the Ising model. Given the coupling J and the magnetic field H, the energy $E(s)$ of a spin configuration $s = (s_1, s_2, \ldots, s_N)$ ($s_i = 2x_i - 1 = \pm 1$) is given by

$$E(x) = -J\sum_{i,j} a_{ij}s_i s_j - H\sum_i s_i. \tag{1}$$

The Ising probability distribution is

$$P(s) = \frac{1}{Z}\exp\left(-\beta\left(J\sum_{i,j} a_{ij}s_i s_j + H\sum_i s_i\right)\right), \tag{2}$$

and we can absorb the inverse temperature β in the parameters J and H (control parameters).

The magnetization m is defined as

$$m = m(J,H) = \sum_s \left(P(s)\frac{1}{N}\sum_i s_i\right).$$

It constitutes a suitable observable for this problem, as also its variance. From Onsager solution in 2D and zero magnetic field [21], we should observe a phase transition at $J_c \simeq 0.44$, with a transition from $m = 0$ to $m \neq 0$ and the divergence of its variance, see Fig. 2 for a numerical simulation.

There are many possible recipes for the Monte Carlo implementation, the one that we examine is the *heat bath* dynamics, for which the probability that spin i takes value s_i' is

$$\begin{aligned}\tau(s_i'|S_i) &= \frac{\exp(s_i(H + J\sum_j a_{ij}s_j)}{\exp(\beta s_i(H + J\sum_j a_{ij}s_j)) + \exp(-s_i(H + J\sum_j a_{ij}s_j))} \\ &= \frac{1}{1+\exp(-2s_i(H + J\sum_j a_{ij}s_j))}.\end{aligned} \tag{3}$$

In practice, each element of the Markov matrix is $M(s'|s) = \tau(s_i'|S_i)$, all other spins remaining the same.

1.2 Equilibrium phase transitions

There is a vast literature about phase transition in equilibrium statistical physics. We want here just recall some properties that can be useful for extending the concept to arbitrary systems, not necessarily in equilibrium, and therefore we only refer to Monte Carlo investigations.

Phase transitions are characterized by a change of the value of some observable, say the magnetization $m(J,H)$, in correspondence of a precise value of a control parameter. In practice we can say that the dynamics of the system changes its structure in correspondence of a phase transition, for instance the phase space may effectively break in two zones that do not communicate at all. This is equivalent to say that the system is no more ergodic, and we speak of *ergodicity breaking*.

If we consider the point of view of deterministic trajectories over a random field, the phase transition can be seen as a bifurcation from a single to multiple attractors.

However, we have a kind of contradiction here: we chose the Monte Carlo dynamics to be ergodic, so how can ergodicity breaking occur? Actually, this breaking only manifests itself in a limit procedure: for a finite system (finite N), and long enough time, all the phase space is visited (it is finite), and therefore the average of observables takes a unique value. However, near the phase transition, the observables (say, the magnetization in the Ising model) maintain the same value for very long periods, with occasional switches from one extreme to another. So, while its average value has a certain value (say, zero), one never observes such value! The time that the system spends on one phase become longer as we approach the critical value of the control parameter and (exponentially) as we increase the system size.

If we take first the limit of infinite system size and then that of infinite time, we observe the ergodicity breaking. In practice, it is sufficient to use a large enough system. In the language of stochastic trajectories, there are two low-energy valley separated by a high (energy) and/or large (entropy) barrier. in order to connect the two valleys, a path should climb the separating saddle, and the associated probability becomes smaller and smaller with the system size, in the vicinity of the phase transition and above.

In the language of Markov processes, we always have an irreducible transition matrix (since the dynamics is ergodic), but in the previous limit the time-product of matrices (denoted as M) effectively breaks in two (or more) sub-matrices, that do not communicate

$$M = \begin{pmatrix} M_1 & \varepsilon \\ \varepsilon & M_2 \end{pmatrix} \xrightarrow{N\to\infty} \begin{pmatrix} M_1 & 0 \\ 0 & M_2 \end{pmatrix},$$

where the ε denote the paths that connects the two valleys. The asymptotic distribution $P^{\mathrm{eq}}(x)$ is proportional to the eigenvector of M with eigenvalue 1. At phase transition this eigenvalue becomes degenerate and we have two or more asymptotic distributions, with different "basins".

We can introduce the correlation function

$$C(\rho,\tau) = \left(\sum_{i=1}^{N}\sum_{t=1}^{T} x_i(t)s_{i+\rho}(t+\tau)\right) - \left(\sum_{i=1}^{N}\sum_{t=1}^{T} s_i(t)\right)\left(\sum_{i=1}^{N}\sum_{t=1}^{T} s_{i+\rho}(t+\tau)\right)$$

The observables can be defined in terms of the correlation function.

The correlation function is expected to decrease exponentially

$$C(\rho,\tau) \sim \exp\left(-\frac{\rho}{\xi_\perp}\right)\exp\left(-\frac{\tau}{\xi_\parallel}\right).$$

defining the correlation lengths $\xi_\perp$ (with respect to space) and $\xi_\parallel$ (with respect to time).

At a phase transition (non-analytical behaviour of some observables like discontinuities, divergence or angular points) the correlation lengths can stay finite (first-order phase transitions) or diverge (second-order phase transitions). In the latter

Table 1 Transition probabilities of the Domany-Kinzel model.

$S = s_{-1} + s_{+1}$	$X = x_{-1} + x_{+1}$	$\tau(1\|S)$	$\tau(0\|S) = 1 - \tau(1\|S)$	bond percolation	site percolation
-1	0	w	$1-w$	0	0
0	1	p	$1-p$	p_b	p_s
1	2	q	$1-q$	$p_b(2-p_b)$	p_s

case,

$$\xi(J,H;N) \sim N^{\alpha}\tilde{\xi}\left(\frac{J}{N^{\gamma}},\frac{H}{N^{\delta}}\right),$$

where α,γ,δ are critical exponents. Also observables like the magnetization exhibit similar scaling behaviour. This phenomenology extends to systems defined directly by stochastic transition probabilities.

2 Probabilistic Cellular Automata

In many cases we are looking for the asymptotic properties of a system that is just defined in terms of the local transition probabilities, of which Probabilistic Cellular Automata (PCA) are prototypical examples.

Cellular automata are defined in a way similar to the previous Monte Carlo time-space evolution, allowing for generic transition probabilities and parallel evolution of all cells at the same time. PCA are therefore Markov chains for which the matrix elements are given by the product of the local transition probabilities (generally uniform),

$$M(x|y) = \prod_i \tau(x_i|Y_i).$$

Again, we can define stochastic trajectories (or deterministic trajectories over a stochastic field)

$$x_i(t+1) = [r_i(t) < \tau(x_i|Y_i)].$$

Deterministic Cellular Automaton (DCA) can be considered as limit cases of PCA, where the transition probabilities τ are either zero or one.

2.1 Parallel Ising model

For instance, we can define a parallel version of the Ising model, for which

$$M(s'|s) = \prod_i \tau(s'_i|S_i),$$

with τ given by Eq. (3).

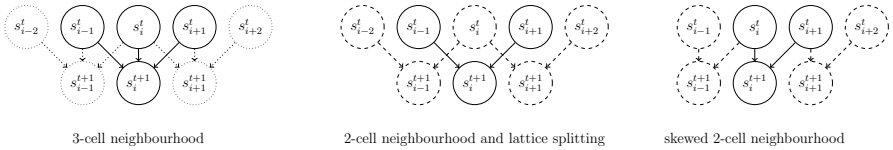


Fig. 3 2-cell neighbourhood.

In this case we can still have an asymptotic probability distribution if the interactions are symmetric (here they are for definition), but the asymptotic distribution is now [11]

$$P^{\mathrm{eq}}(s) = \frac{1}{Z}\prod_i e^{\beta H s_i}\cosh\left(\sum_j \beta(H + J\sum_j a_{ij}s_j)\right),$$

where Z is again the normalization constant.

Notice that the transition probabilities of Eq. (3) do not depend on the previous value of the site s_i. If we apply them in parallel to all sites, at least in one dimension and with nearest-neighbour interactions, the lattice decouples in two noninteracting sublattices (for even N, see Fig. 3), so that $s'_i = f(s^t_{i-1} + s_{i+1}(t), r_i(t))$ It is an example of a totalistic PCA, that has been studied by Kinzel [25] and shows no phase transition.

2.2 Domany-Kinzel model. Absorbing states.

We can extend the parallel Ising example to a general case, on the same two-neighbours network (Fig. 3), defining three independent totalistic transition probabilities, as shown in Table 1. This model has been studied by Domany and Kinzel [12, 25], and can be considered the simplest model showing a phase transition.

For generic values of w ($\tau(1|0)$), p ($\tau(1|1)$) and q ($\tau(1|2)$), this model can be mapped onto a parallel Ising model with a plaquette term [25] (we need another control parameter in addition to H and J since here we have three free probabilities),

$$E(S) = -\sum_i s_i\left(H + J(s_{i-1} + s_{i+1}) + K s_{i-1}s_{i+1}\right).$$

Denoting $h = \exp(-2H)$, $j = \exp(-4J)$, $k = \exp(-2K)$, we have $w = 1/(1 + hk/j)$, $p = 1/(1 + h/k)$, $q = 1/(1 + hjk)$ and therefore

$$H = \frac{1}{6}\log\frac{wpq}{(1-w)(1-p)(1-q)},\ J = \frac{1}{8}\log\frac{(1-w)q}{w(1-q)},\ K = \frac{1}{6}\log\frac{w(1-p)q}{(1-w)p(1-q)}.$$

However, this model does not show any phase transition.

If we set $w = 0$ (by letting the coupling take infinite values with suitable limits), we leave the equilibrium condition. In this limit the configuration $s = -1$ becomes an absorbing state. We can also switch to the Boolean representation by setting $x_i = (s_i + 1)/2$. In this representation the absorbing state is the configuration $x = 0$. It is called absorbing since it cannot be left by the dynamics once entered. The order parameter is here the "density" of ones

$$c = \frac{1}{N}\sum_i x_i.$$

We can reformulate the phase transition in this new language: for finite N there is always a probability $M(0|y)$ that brings any configuration to the absorbing state in one step. In the limit $N \to \infty$ and for a suitable value of the parameters p and q this probability goes to zero and the Markov matrix becomes reducible. It is composed by a submatrix M_1 that maps states "near" to 0 into 0 in a few time steps, and a set of states with a non-vanishing density c

Again, one can speak of deterministic trajectories one that the stochastic field has been laid out. The evolution equation of the system is

$$x_i' = [r_i^{(1)}(t) < p](x_{i-1}(t) \oplus x_{i+1}(t)) \oplus [r_i^{(2)}(t) < q]x_{i-1}(t)x_{i+1}(t)$$

where $\oplus$ is the XOR operation (sum modulus two). Notice that the two random numbers $r_i^{(1)}(t)$ and $r_i^{(2)}(t)$ may be the same or not, since the two conditions $(x_{i-1}(t) \oplus x_{i+1}(t)$ and $x_{i-1}(t)x_{i+1}(t)$ are never true at the same time (but this makes a difference for damage spreading, Section 2.5).

In the language of trajectories, one can say that there are two attractors, the fixed point 0 and a "chaotic" attractor with $d > 0$, each one with its own basin. More on absorbing phase transition can be found in Ref. [19].

For $w = q = 0$ and $p = 1$ we have the deterministic rule 90 in Wolfram's notation [30], so the line $q = w = 0$ corresponds to the dilution of rule 90.

2.3 Mean-field approximation

In order not to use a heavy notation, let us apply this approximation using the DK model, assuming that a site i at time $t+1$ is connected to sites i and $i+1$ at time t (i.e., using the skewed lattice of Fig. 3).

The evolution equation for the probability distribution is

$$P(x_1, x_2, \ldots, x_N; t+1) = \sum_{y_1, y_2, \ldots, y_N} \left(\prod_i \tau(x_i|y_i, y_{i+1} \right) P(y1, y_2, \ldots, y_N; t), \quad (4)$$

considering appropriate boundary conditions (e.g., periodic). We can obtain the reduced probabilities $\pi_\ell(x_i, \ldots, x_\ell; t)$ by summing $P(x_1, x_2, \ldots, x_N; t+1)$ over all $i > \ell$.

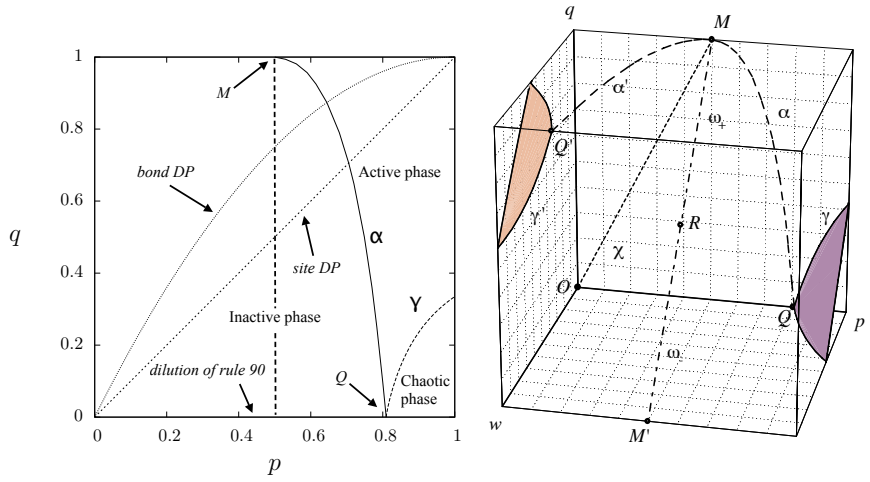


Fig. 4 The phase diagram of the Domany-Kinkel model, α marks the density transition and γ the damage transition. Left: the phase diagram for $w = 0$. The dashed line marks the transition line for the simplest mean-field approximation. Right: the complete phase diagram. The curves labelled α and α' belong to planes $w = 0$ and $w = 1$ resp., and correspond to the density phase transitions. The solid curves correspond to the intersection of the damage critical surface (shaded) γ and γ' with the boundaries of the cube. The dotted-dashed lines labelled ω_+ and ω_- correspond to the existence line for the parallel Ising model for positive and negative temperatures, resp. The points labelled M and M' to the critical points of the parallel Ising model at zero temperature (compact DP), and the point labelled R to infinite temperature. The dotted line labelled χ corresponds to the damage in the parallel Ising model.

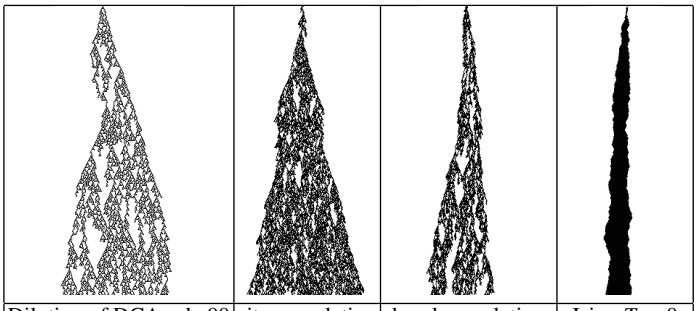
If the system is translation-invariant, one obtains the same result summing the elements of any set of consecutive variables. Since $\sum_{x_i} \tau(x_i|y_i,y_{i+1} = 1$ for all x_i, we can then sum over $y_{i+2},\ldots,y_N$, obtaining

$$\begin{aligned}
\pi_1(x_1,t+1) &= \tau(x_1|y_1,y_2)\pi_2(y_1,y_2;t),\\
\pi_2(x_1,x_2,t+1) &= \tau(x_1|y_1,y_2)\tau(x_2|y_2,y_3)\pi_3(y_1,y_2y_2;t),\\
&\ldots
\end{aligned}$$

i.e., a hierarchy of equations that are equivalent to Eq. 4.

If the correlation length ξ is less than N, two cell separated by a distance greater that ξ are practically independent. The system acts like a collection of subsystems each of length ξ (this is why ergodicy and selfaveraging holds far from the transition). Since ξ is not known a priori, one assumes a certain correlation length ℓ and computes the quantity of interest. By comparing the values of these quantities with increasing ℓ generally a clear scaling law appears, allowing to extrapolate the results to the case $\ell \to \infty$.

The very first step is to assume $\ell = 1$. In this case we can simply factorize $\pi_2(x_1,x_2) = \pi_1(x_1)\pi_1(x_2)$. By calling $c = \pi_1(1;t)$ $(1-c = \pi_1(0;t))$, $c' = \pi_1(1;t+1)$



Dilution of DCA rule 90	site percolation	bond percolation	Ising $T=0$
$p=0.81, q=0$	$p=q=0.71$	$p=0.64, q=0.87$	$p=0.5, q=1$

Fig. 5 Typical patterns of the DK model. Space runs horizontally and time vertically, from top to bottom.

and using the transition probabilities of Table 1 with $w=0$, one gets

$$c' = 2pc(1-c) + qc^2.$$

Notice that in the mean-field approximation, the evolution of the system is given by a deterministic equation for the average value of observables. In this approximation, a phase transition corresponds to a bifurcation (change of stability of the attractors) of the map.

The fixed points ($c'=c$) are $c=0$ and $c=2p/(2p-q)$. There is a change of stability from $c=0$ (the absorbing state) to $c>0$ for $p_c=1/2$. As shown in Fig. 4-left, this approximation is quite rough.

The DK model includes the Directed Percolation (DP) one [24], which can be formulated thinking to an infection process: an individual i at time t can get infected by its infected neighbours at the previous time step, with a probability that depends on the number of infected neighbours (bond percolation) or not (site percolation), see Table 1. In the mean-field approximation, we have for the bond percolation the line $q=p(2-p)$, and for the site percolation the line $q=p$, as shown in Fig. 4-left.

There are two ways of extending the above approximation. The first one is still to factorize the cluster probabilities at single site level but to consider more time steps, for instance obtaining $\pi_1(t+2)$ in terms of $\pi_3(t)$ and then factorizing π_3 in terms of π_2. The map is still expressed as a polynomial of the density c. The advantage of this method is that we still work with a scalar (the density), but in the vicinity of a phase transition the convergence towards the thermodynamic limit is very slow.

The second approach, sometimes called *local structure approximation* [18], is a bit more complex. Let us start from the generic ℓ cluster probabilities π_ℓ. We generate the $\ell-1$ cluster probabilities $\pi_{\ell-1}$ from π_ℓ by summing over one variable,

$$\pi_{\ell-1}(x_1,\ldots,x_{\ell-1}) = \sum_{x_\ell} \pi_\ell(x_1,\ldots,x_{\ell-1},x_\ell).$$

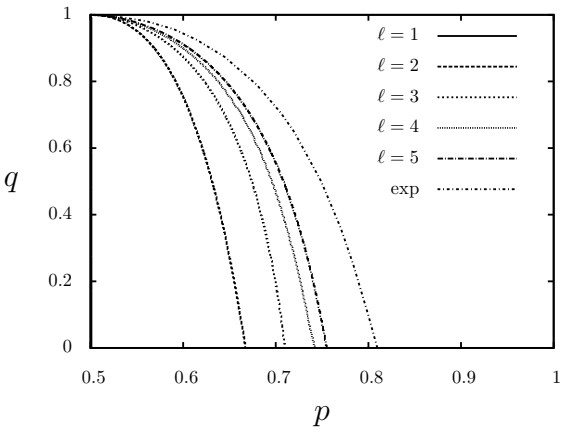


Fig. 6 Local structure approximation for the DK model, with several values of length ℓ. The case $\ell = 1$ is the simplest mean-field approximation and corresponds to the line $p = 1/2$. The line marked 'exp' corresponds to numerical simulations as in Fig. 4.

The $\ell+1$ cluster probabilities are generated by using a Bayesian estimation

$$\pi_{\ell+1}(x_1, x_2, \ldots, x_\ell, x_{\ell+1}) = \frac{\pi_\ell(x_1, \ldots, x_\ell)\pi_\ell(x_2, \ldots, x_{\ell+1})}{\pi_{\ell-1}(x_2, \ldots, x_\ell)}.$$

Finally, one is back to the ℓ cluster probabilities by applying the transition probabilities

$$\pi'(x_1, \ldots, x_\ell) = \sum_{y_1, \ldots, y_{\ell+1}} \prod_{i=1}^{l} \tau(x_i | y_i, y_{i+1}).$$

This last approach has the disadvantage that the map lives in a high-dimensional (2^ℓ) space, but the results converges much better in the whole phase diagram.

This mean-field technique can be considered an application of the transfer matrix concept to the calculation of the the eigenvector (asymptotic probability distribution) corresponding to the maximum eigenvalue (fundamental or ground state), by means of the iteration of the matrix.

2.4 Asynchronism of DCA

An unexpected phase transition occurs with an increasing level of asynchronism of some DCA rule [13, 14]. Let us denote by $f(x_{i-1}, x_i, x_{i+1})$ the deterministic rule. The evolution equation of its dilution is

$$x'_i = x_i \oplus [r_i(t) < (1-p)]\big(x_i \oplus f(x_{i-1}, x_i, x_{i+1})\big).$$

With probability $1-p$ the site follows the rule f, and with probability p it keeps its old value.

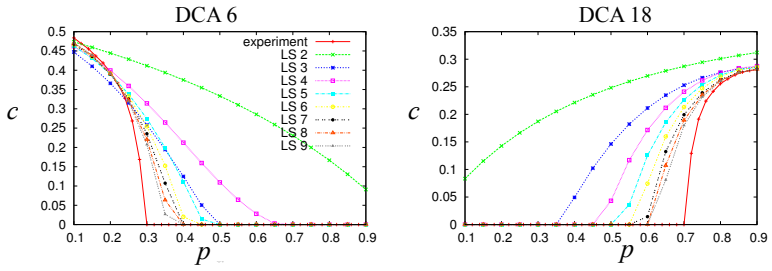


Fig. 7 DCA dilution phase transition for two Elementary Cellular Automaton rule 6 and 18 in Wolfram notation [30], comparisons between numerical simulations and the local structure approximation (from Ref. [15]). In the y axis the asynchronism parameter p.

Examples of phase transitions are shown in Fig 7.

An unexpected fact is that the simplest mean-field approximation completely fails for this problem. Indeed, we have

$$c' = pc + (1-p) \sum_{a,b,c=0}^{1} f(a,b,c) c^{a+b+c} (1-c)^{3-a-b-c}$$

and for the stationary state $c' = c$ one gets

$$c = \sum_{a,b,c=0}^{1} f(a,b,c) c^{a+b+c} (1-c)^{3-a-b-c}$$

i.e., the mean-field approximation of the deterministic rule, without any dependence on p. Increasing the order of the mean-field approximation (local structure approximation), one can approximate the actual phase transition behaviour [15], as shown in Fig. 7.

2.5 Damage spreading

We have said that the large-time distribution $x(T)$ depends in general on the random field and the initial conditions $x(0)$, although, for large N, the observables like the density does not depend on them due to ergodicity and self-averaging. Actually, we can check the dependence on the initial conditions by considering the evolution of an initial difference between two replicas, evolving on the same random field, and looking at the difference (or damage) $z_i = x_i \oplus y_i$,

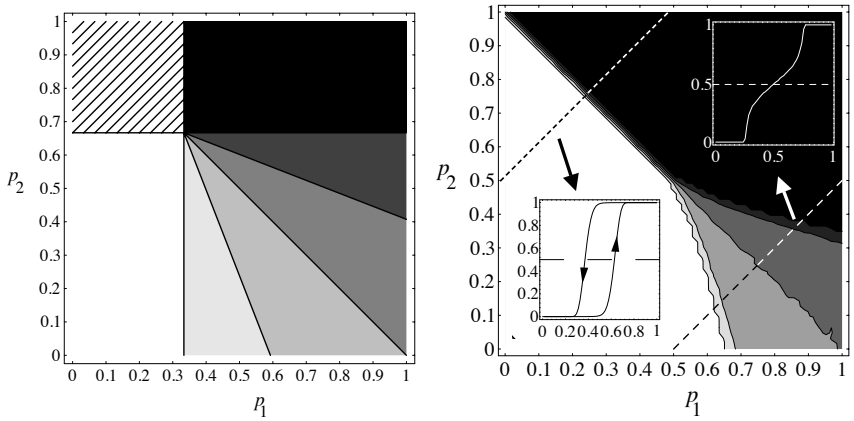


Fig. 8 Phase transition diagrams of the BBR model (colour code: white=0, black=1). Left: mean-field phase diagram for the density. Right: numerical phase diagram of the density, in the inset the variation of the density when cutting the phase diagram; the hysteresis inset at bottom right is obtained by setting $w = 10^{-4}$, $T = 500$. Numerical simulations with $N = T = 10^4$.

$$
\begin{aligned}
x_i' &= [r_i^{(1)}(t) < p]\big(x_{i-1}(t) \oplus x_{i+1}(t)\big) \oplus [r_i^{(2)}(t) < q] x_{i-1}(t) x_{i+1}(t), \\
y_i' &= [r_i^{(1)}(t) < p]\big(y_{i-1}(t) \oplus y_{i+1}(t)\big) \oplus [r_i^{(2)}(t) < q] y_{i-1}(t) y_{i+1}(t), \\
z_i' &= x_i' \oplus y_i' = [r_i^{(1)}(t) < p]\big(z_{i-1}(t) \oplus z_{i+1}(t)\big) \oplus [r_i^{(2)}(t) < q] \cdot \\
&\quad \big((z_{i-1}(t) z_{i+1}(t) \oplus z_{i-1}(t) x_{i+1}(t) \oplus x_{i-1}(t) z_{i+1}(t) \oplus x_{i-1}(t) x_{i+1}(t)\big).
\end{aligned}
$$

Since now the two conditions can occur at the same time, there is a difference in the evolution if one uses one or two random numbers per site (or if they are otherwise correlated). Looking only at the evolution of the difference z, the evolution of the x replica (which is not affected by z) is just another field (although it is not fully random). The quantity z shows another phase transition (Fig. 4) that characterizes the dependence on the initial condition: in one phase the difference goes to zero, meaning that all initial conditions will follow after a transient time the same trajectory, only depending on the stochastic field. In the other phase, the system maintains forever some memory of the initial condition.

This phase transition also belongs to the directed percolation universality class. It is possible to approximately map the density phase transition onto the damage one [2].

2.6 A richer phase diagram: the BBR model

The DK model is quite useful for studying nonequilibrium phase transitions due to its simplicity. In order to explore other types of transitions beyond DP, let us

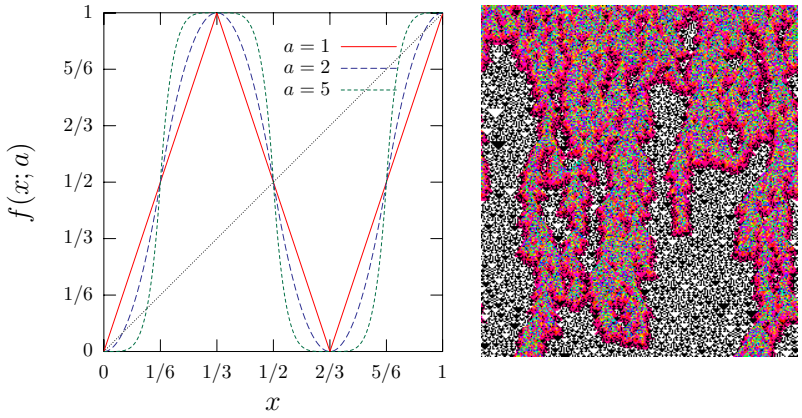


Fig. 9 (left) The graph of $f(x;a)$ for three values of a. (right) Space time pattern of the CML of Eq. (5) with $a = 1.9$ and $N = 256$ drawn horizontally for a total time of $T = 300$ time steps drawn vertically from top to bottom. The initial configuration $x(0)$ is chosen randomly. The color code assigns white (black) whenever $x_i(t) = 0(1)$ and a rainbow color scale for other values of $x_i(t)$ starting with red for values near zero. Patches of CA behaviour (rule 150) appear after a short transient and will eventually fill the whole pattern.

introduce the BBR model [10], that is a 3-input cellular automata with two absorbing states. It is a totalistic automaton, meaning the the transition probability depends on the sum S of the states in the neighbourhood, with $0 \le S \le 3$. The BBR transition probabilities $\tau(x'|S)$ are $\tau(1|0) = w$, $\tau(1|1) = p_1$, $\tau(1|2) = p_2$, $\tau(1|3) = 1-w$ By setting $w = 0$, the states 0 and 1 are absorbing, and on the line $p_1 = 1 - p_2$ the system is symmetric for the inversion $1 \leftrightarrow 0$.

As can be seen in Fig. 8, we have here, for high-p_1 and low-p_2 value, two DP transitions reminiscent of the DK model. The two lines meet at about $p_1 = p_4 = 0.5$ ($p_1 = 1 - p_2 = 1/3$ in the mean-field approximation). In this point the universality class changes to that of parity conservation. In the low-p_1, high-p_2 part of the diagram, we have a first-order transition: the two absorbing states are stable (as predicted by the mean-field analysis) and we can investigate the nature of an hysteresis cycle. In order to do that, we have to remove the absorbing characteristic of the states 0 and 1. We do this by imposing that w ($\tau(1|0)$ is small but different from zero, so that in principle the system does not more show a true phase transition. Indeed, that states with high or low values of the density are now metastable, so we have to tune the simulation time with the value of w. This tuning is however not critical: for a large range of values of simulations times, we obtain an hysteresis diagram similar to that of Fig. 8.

2.7 Janssen-Grassberger's conjecture

The DP class is extremely robust with respect to the microscopic dynamic rules. The large variety and robustness of DP models led Janssen [23] and Grassberger [16] to the conjuncture that all systems with a single order parameter and a single absorbing state will belong to the universality class of the Directed Percolation (DP) model. More precisely, the requirements are

- The model displays a continuous phase transition from a fluctuating active phase into a dominant stable absorbing state.
- The transition is characterized by a positive one-component order parameter.
- The dynamic rules involve only short-range processes.
- The system has no unconventional attributes such as additional symmetries or quenched randomness.

We have already seen that the BBR model has two absorbing states. As far as their basins are different, the phase transition belongs to the DP universality class, on the symmetry line $p_1 = 1 - p_2$ it switches to the parity conservation class.

Another way of violating this condition is that of modifying the stability of the absorbing state. This can be easily realized in the synchronization scenario [17].

2.8 Synchronization

The idea of a replica synchronization is the following: take two replicas of a system, either driven by a deterministic or a stochastic dynamics (in the latter case, the random field is the same for the two systems). Let one system evolve by itself, and "push" the other towards the first. If the pushing is strong enough, the system will synchronize. A simple illustration is the following. Let's consider a continuous map $x' = f(x)$, and construct the synchronization mechanism

$$\begin{aligned} x' &= f(x), \\ y' &= (1-p)f(y) + pf(x), \end{aligned}$$

for $p = 0$ the two systems are completely disconnected, and if the map f is chaotic, they stay well separated. For $p = 1$ the two system are identically the same. There is a critical value p_c such that the distance $\delta = |x - y|$ goes to zero. For small distance, δ evolves as

$$\delta' = (1-p)|f(y) - f(x)| \simeq (1-p)\left|\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|\delta$$

and thus

$$\delta(t) = (1-p)^t \delta(0) \prod_{t'} \left| \frac{\mathrm{d}f(x(t'))}{\mathrm{d}x} \right| = (1-p)^t \delta(0) \exp\left(\sum_{t'} \log \left| \frac{\mathrm{d}f(x(t'))}{\mathrm{d}x} \right| \right)$$
$$= \delta_0 \exp((\log(1-p)+\lambda)t),$$

where λ is the Lyapunov exponent of the map. Thus, when $\delta(t) = \delta(0)$ (the synchronization threshold), $p_c = 1 - \exp(-\lambda)$, and this relates the synchronization threshold to the chaotic properties of the map.

This mechanism can be extended in several ways to extended systems (coupled map lattices and cellular automata). For reference, consider the following generic coupled system

$$x_i' = f(g(x_{i-1}, x_i, x_{i+1})),$$

where g defines the coupling. One can use a homogeneous "pushing", i.e., use the same p for all sites, or, at the other extreme, a all-or-none pushing, i.e., choose a fraction p of sites to be completely synchronized and leave the other unperturbed.

Using the first mechanism, one again relates the synchronization threshold to the Maximum Lyapunov Exponent (MLE) of the system. Chaotic systems are expected to amplify the distance between replicas. For a value of p slightly below the synchronization threshold, some patches may synchronize for some time, after which they will separate. This picture resembles that of a growing interface that may stay pinned to local traps. From field theory studies, such a behaviour is denoted multiplicative noise (MN) and is equivalent to the behaviour of the "bounded" Kardar-Parisi-Zhang equation, which describes the behaviour of a growing surface that tends to pin and is pushed from below [22, 26, 27]. On the other hand, stable systems have a negative MLE. So, replicas should naturally synchronize once their distance is (locally) below the threshold of validity of linear analysis. However, when the local difference is large, non-linear terms may maintain or amplify this distance. In this case synchronized patches may be destabilized only at the boundaries. Again, theoretical studies associate such a behaviour to that of directed percolation (DP) [24]

However, it is questionable if the MLE exponent really captures the chaotic properties of an extended system. For instance, let us take f chaotic and $g(a,b,c) = \varepsilon(a+c) + (1-\varepsilon)b$, i.e., a diffusive coupling. The Lyapunov exponent $\lambda(\varepsilon)$ in general decreases with ε, since the coupling acts like a constraint (a kind of surface tension). Thus λ takes its maximum values for $\varepsilon = 0$, but in this case the chaos does not spread on the lattice.

On the contrary, the all-or-none ("pinching") synchronization mechanism shows that the case in which synchronization is most difficult is for $\varepsilon \simeq 1/3$, which is what one intuitively expects. Moreover, we can apply this synchronization mechanism also to cellular automata, provided that the two replicas evolve using the same random field. It is possible to show that in this case one can develop a concept of Boolean derivative for such a discrete systems, and obtain an equivalent of the maximum Lyapunov exponent, which is related to the pinching synchronization threshold [1, 6].

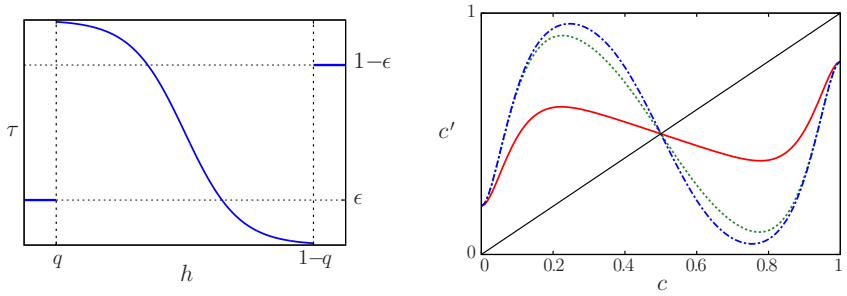


Fig. 10 (left) The transition probability $\tau(h)$ given by Eq. (6) with $J=-3$, $k=20$, $q=0.1$, and $\varepsilon=0.2$. (right) Graphs of the mean field map, Eq. (7) for different values of J and $k=20$. From bottom to top for $c<1/2$, $J=-0.5$ (red, lower line), $J=-3.0$ (green, middle line), and $J=-6.0$ (blue, upper line).

The synchronized state is an example of absorbing state, but clearly in real cases one rarely expect to find a complete synchronization: the evolution may be influenced by noise, or the two replicas can be slightly different.

We can test this hypothesis using the map

$$f(x;a)=\begin{cases}(6x)^a/2 & 0\le x<1/6,\\ 1-|6(1/3-x)|^a/2 & 1/3\le x<1/2,\\ |6(x-2/3)|^a/2 & 1/2\le x<5/6,\\ 1-(6(1-x))^a/2 & 5/6\le x<1,\end{cases} \tag{5}$$

where $1\le a<\infty$ (see Fig. 9-left), see Ref. [7]. This map that has the advantage of reducing to the DCA rule 150 for a large, and to a chaotic map from a small. For $a\gtrsim 1.81$ (stable chaos) one observes a transient chaos, with positive Lyapunov exponent, followed by a cellular automata pattern. One may wonder about the unpredictability of such map: in the chaotic phase an infinitesimal damage will amply, while in stable chaos phase infinitesimal damages are absorbed (and thus the word "stable") but finite ones spread (and thus the word "chaos"). The synchronization procedure applied to a lattice of such maps indeed shows that a certain effort is needed even in the "stable" phase to get the synchronization. In agreement with the Janssen-Grassberger conjecture, one finds the synchronization phase transition for $a<1$ do belongs to the MN universality class, while for $a\gtrsim 1.81$

Such a behaviour is not limited to systems that reduce to DCA, see Ref. [5] for an example.

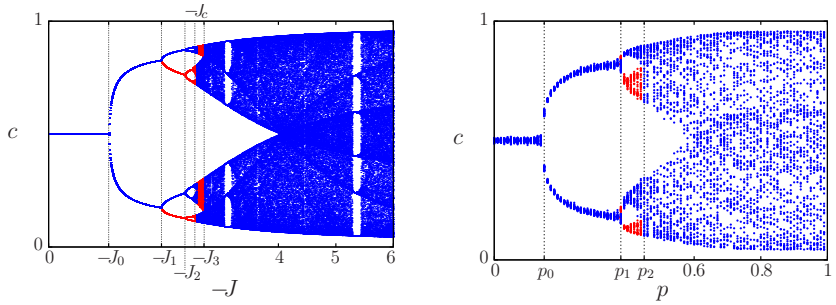


Fig. 11 (left) Bifurcation diagram of the mean-field map, Eq. (7), by varying J. (right) Small-world probabilistic bifurcation diagrams as functions of the long range probability p. The colors mark different attractors.

2.9 Topology and chaotic phase transitions

Up to now we have not investigated the influence of the topology, i.e., of the connections defined by the adjacency matrix a_{ij}. It is well known that if we replace a regular lattice with a random network of the same connectivity, the global behaviour becomes that of the mean-field, since in this way correlations are disrupted.

We can study the influence of the topology by adopting the Watts-Strogatz rewiring mechanism [29]: start with a regular lattice of connectivity k in 1D and, for each site, rewire at random a fraction p of incoming links.

In order to show the effects of the mechanism and also to present a new type of phase transition, let us consider a cellular automaton whose mean-field approximation is chaotic. This model has been developed originally as an opinion formation model [8].

The average local opinion or social pressure h_i, is defined by

$$h_i = \frac{\sum_j a_{ij} s_j}{k}.$$

The opinion of agent i changes in time according to the transition probability $\tau(s_i|h_i)$ that agent i will hold the opinion s_i at time $t+1$ given the local opinion h_i at time t. This transition probability, shown in Fig. 10-left, is given by

$$\tau(h) = \begin{cases} \varepsilon & \text{if } h < q, \\ \dfrac{1}{1+\exp(-2J(2h-1))} & \text{if } q \le h \le 1-q, \\ 1-\varepsilon & \text{if } h > 1-q, \end{cases} \tag{6}$$

with $\tau(h) = \tau(1|h)$.

The simplest mean-field description of the model is given by

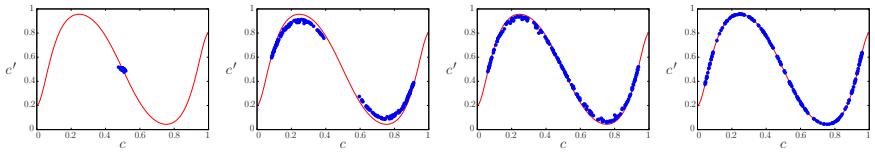


Fig. 12 (Color online) Return map of the average opinion c on small-world networks for several values of the long-range connection probability p with $J=-6$, $k=20$, $N=10^3$, and a transient of 10^3 time steps. The following 200 iterations are shown as (blue, darker) dots. The (red, lighter) continuous curve is Eq. (7). From left to right: $p=0.0$, $p=0.5$, $p=0.6$, and $p=1.0$.

$$c' = f(c) = \sum_{w=0}^{k} \binom{k}{w} c^w (1-c)^{k-w} \tau\left(\frac{w}{k}\right), \tag{7}$$

with $c' = c(t+1)$ and $c = c(t)$. The term in parenthesis on the *r.h.s* of this expression denotes the w-combinations from a set of k elements. In Fig. 10-right we show some graphs of f. The bifurcation digram of this map after varying J is shown in Fig. 11-left. The doubling bifurcation route to chaos ends at $J=J_c$. For $0>J\geq J_2$ and $J_3>J\geq 6$ there is only one attractor (blue, darker dots). For $J_2>J\geq J_c$ there are two, one corresponding to the lower branches that bifurcate up to J_c (red, lighter dots), and the other one to the upper branches (blue, darker dots). For $J_c>J\geq J_3$ there are two chaotic attractors, one corresponding to the lower branches (blue, darker dots), the other to the top branches (red, lighter dots). For every value of J, the dots are 64 iterates of the map after a transient of 10^3 time steps. For values of J with only one basin of attraction the orbits do not depend on the initial average opinion $c(t=0)$. For values of J that correspond to two attractors, one of them was found with $c(0)=0.1$, the other one with $c(0)=0.9$.

By varying the long-range probability p, we observe the transition towards the mean-field behaviour, as reported in Fig. 12. This induces a stochastic bifurcation diagram by varying p, Fig. 11-right that is quite similar to that obtained in the mean-field approximation by varying J, Fig. 11-left. For $p\lesssim p_0$ there are almost periodic orbits of period one and for $p_0\lesssim p\lesssim p_1$ of period two. For $p_1\lesssim p\lesssim p_2$ we find two attractors, one (in red, lighter) in the lower branches, the other one (in blue, darker) in the top ones.

Notice that up to now we have met phase transition that, in the mean-field description, implies the change of stability of fixed points, while here we observe a real bifurcation diagram with coexistence of basins, period-doubling and chaos.

3 Conclusions

The main aim of this presentation was that of discussing some characteristics of nonequilibrium phase transitions.

We have illustrated some aspects of phase transitions in probabilistic cellular automata, trying to show how such a problem arises in different contexts and some of the method used for its study.

The real-life problems are usually more complex than those faced here. However, a phase transition separates qualitatively different states, and since continuous phase transitions are related to the divergence of the correlation function, the general scenario is quite independent of the details of the model, so that the investigation of simplified models is justified.

The numerical simulations of phase transitions constitute also a challenge by itself: the need of approaching the limit of infinite space and time requires particular techniques and an efficient implementation.

This study can be complemented by an analysis based on the principles of the renormalization group (see for instance Ref. [28]), that in principle allows to group several models into a few universality classes.

The study of phase transitions in stochastic systems (and the related one of bifurcations in dynamical systems) can constitute a good training ground for both theoretical, computational and experimental students.

Acknowledgements This work was partially supported by EU projects 288021 (EINS – Network of Excellence in Internet Science) and project PAPIIT-DGAPA-UNAM IN109213.

References

1. Bagnoli, F.: Boolean derivatives and computation of cellular automata. Int. J. Mod. Phys. C **3**, 307–320 (1992) doi: 10.1142/S0129183192000257
2. Bagnoli, F., Cecconi, F.: Synchronization of non-chaotic dynamical systems. Phys. Lett. A **282** 9–17 (2001) doi: 10.1016/S0375-9601(01)00154-2
3. Bagnoli, F.: Rechtman, R.: Synchronization and maximum Lyapunov exponents of cellular automata. Phys. Rev. E **59** R107–R1310 (1999) doi: 10.1103/PhysRevE.59.R1307
4. Bagnoli, F., Rechtman, R.: Synchronization universality classes and stability of smooth coupled map lattices. Phys. Rev. E **73**, 026202 (2006) doi: 10.1103/PhysRevE.73.026202
5. Bagnoli, F., Rechtman, R.: Topological bifurcations in a model society of reasonable contrarians Phys. Rev. E **88** 062914 (2013) doi: 10.1103/PhysRevE.88.062914
6. Bagnoli, F., Baroni, L., Palmerini, P.: Synchronization and directed percolation in coupled map lattices. Phys. Rev. E. **59**, 409–416 (1999) doi: 10.1103/PhysRevE.59.409
7. Bagnoli, F., Boccara, N., Rechtman, R.: Nature of phase transitions in a probabilistic cellular automaton with two absorbing states. Phys. Rev. E **63**, 046116 (2001) doi: 10.1103/PhysRevE.63.046116
8. Derrida, B.: Dynamical Phase Transitions in Spin Models and Automata. In: Van Beijeren, H. (ed.) Fundamental Problems in Statistical Mechanics VII, pp. 273–309. Elsevier Science Publisher, Amsterdam (1990)
9. Domany, E., Kinzel, W.: Equivalence of cellular automata to Ising models and directed percolation. Phys. Rev. Lett. **53**, :311–314 (1984) doi:10.1103/PhysRevLett.53.311
10. Fatès, N.: Asynchronism Induces Second Order Phase Transitions in Elementary Cellular Automata. Journal of Cellular Automata **4**, (2009) 21–38
11. Fatès, N.: A Guided Tour of Asynchronous Cellular Automata, J. Cell. Autom. **9**, (2014) 387–416

12. Fukś, H. and Fatès, N.: Local structure approximation as a predictor of second order phase transitions in asynchronous cellular automata. Nat. Comp. **14**, 507-522 (2015) doi:10.1007/s11047-015-9521-6
13. Grassberger, P.: On phase transitions in Schlögl's second model. Z. Phys. B **47**, 365 (1982).
14. Grassberger, P.: Synchronization of coupled systems with spatiotemporal chaos. Phys. Rev. E **59**, R2520–R2524 (1999) doi: 10.1103/PhysRevE.59.R2520
15. Gutowitz, H. A., Victor, J. D., Knight, B. K.: Local Structure Theory For Cellular Automata Physica **28**D, 18–48 (1987)
16. Henkel, M., Hinrichsen, H., Lübeck, S.: Non-Equilibrium Phase Transitions Volume 1: Absorbing Phase Transitions, Springer Science, Dordrecht (2008)
17. Hinrichsen, H. Weitz, J. S. Domany, E.: An Algorithm-Independent Definition of Damage Spreading—Application to Directed Percolation. J. Stat. Phy. **88**, 617–636 (1997)
18. Huang, K.: Statistical Mechanics. Wiley and Sons, New York (1963)
19. Kardar, G. P. M. Zhang, Y.-C.: Dynamic scaling of growing interfaces. Phys. Rev. Lett. **56**, 889–892 (1986).
20. Janssen, H.K.: On the nonequilibrium phase transition in reaction-diffusion system with an absorbing stationary state. Z. Phys. B **42**, 151 (1981) .
21. Kinzel, W.: Directed Percolation. In: Adler, J., Zallen, R. Deutscher, G. (eds) Percolation structures and processes, Annals of the Israel Phys. Soc. **5**, 425 (1983)
22. Kinzel, E.: Phase Transition of Cellular Automata, Z. Phys. B **58**, 229–244 (1985) doi: 10.1007/BF01309255
23. Muñoz, M. A., Hwa, T.: On nonlinear diffusion with multiplicative noise. Europhys. Lett. **41**, 147–152 (1998).
24. Tu, Y. Grinstein, G., Muñoz, M. A.: Systems with multiplicative noise: critical behavior from KPZ equation and numerics. Phys. Rev. Lett. **78**, 274–277 (1997)
25. Tomé, T., de Oliveira, M. J.: Renormalization group of the Domany-Kinzel cellular automaton. Phys. Rev. E **55**, 4000–4004 (1997) doi: 10.1103/PhysRevE.55.4000
26. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. Nature **393**, 440-442 (1998) doi: 10.1038/30918
27. Wolfram, S.: Statistical Mechanics of Cellular Automata. Rev. Mod. Phys.**55**, 601–644 (1983) doi:10.1103/RevModPhys.55.601.