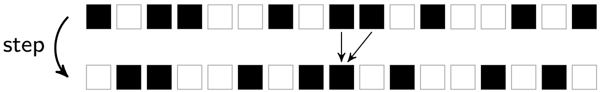
Probabilistic cellular automata

ergodicity and phase transitions

Siamak Taati
Department of Mathematics, American University of Beirut

Baby Steps Beyond the Horizon
Będlewo, September 2026

Example 1: Stavskaya model



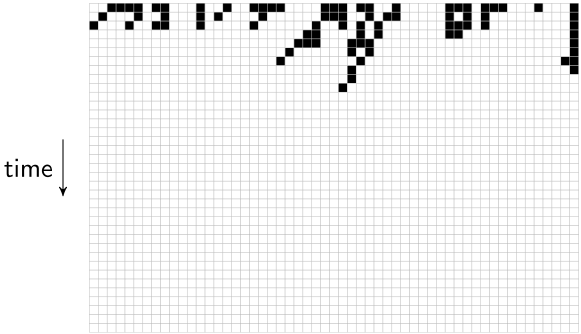
$$\square\square \longmapsto \square$$

$$\square\blacksquare \longmapsto \begin{cases} \blacksquare & \text{with prob. } p, \\ \square & \text{with prob. } 1-p. \end{cases}$$

$$\blacksquare\square \longmapsto \begin{cases} \blacksquare & \text{with prob. } p, \\ \square & \text{with prob. } 1-p. \end{cases}$$

$$\blacksquare\blacksquare \longmapsto \begin{cases} \blacksquare & \text{with prob. } p, \\ \square & \text{with prob. } 1-p. \end{cases}$$

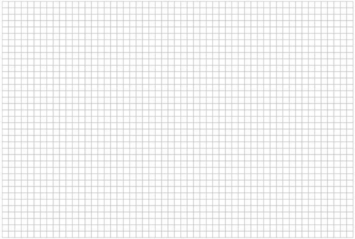
Example 1: Stavskaya model



$p = 0.4$

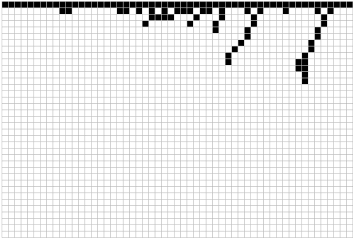
Example 1: Stavskaya model

Starting from all-0



the all-0 configuration is stationary

Starting from all-1



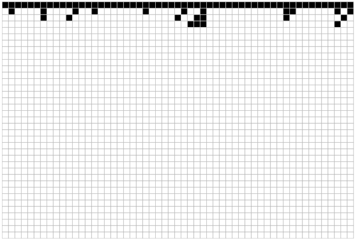
for $p = 0.4$, it still dies out

The monotonicity of this rule implies that

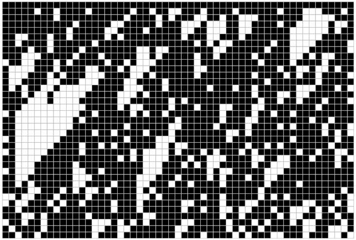
If the process starting from all-1 dies out, then every initial configuration dies.

... so we need to focus on the all-1 initial condition.

Example 1: Phase transition in the Stavskaya model



$p = 0.2$



$p = 0.8$

Theorem

There exists a critical value $p_* \in (0, 1)$ such that

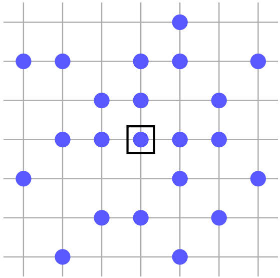
(i) if $p < p_*$, then almost surely everything dies out;

(ii) if $p > p_*$, then there is an equilibrium state other than the point mass at all-0 [equilibrium state = stationary probability measure]

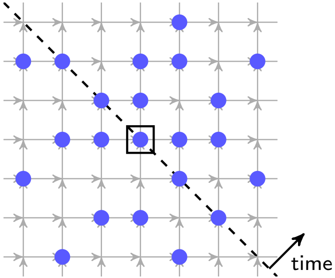
Numerically, $p_* \approx 0.7055$ [see Mendonça (2011)]

Stavskaya vs. oriented percolation

Site percolation



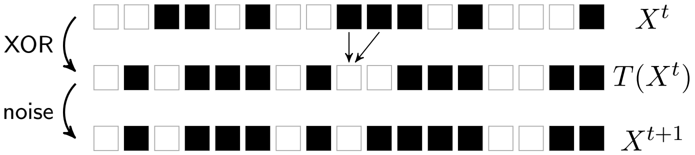
Oriented site percolation



Question

Is there an infinite open path starting from the origin?

Example 2: XOR CA + noise

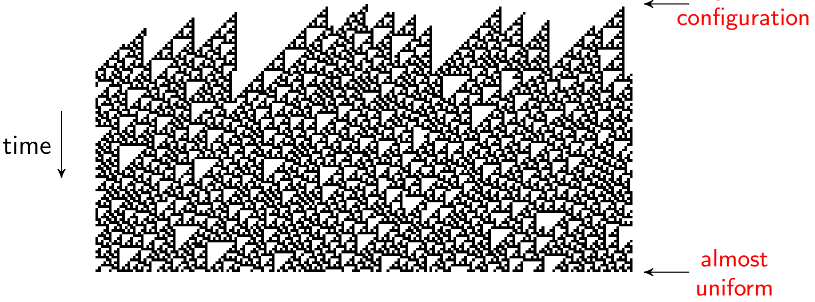


At each step,

I. Apply the XOR rule at every site:

II. Flip each value with probability ε, independently of the others.

Example 2: XOR CA + noise



Theorem (Vaserstein, 1969)

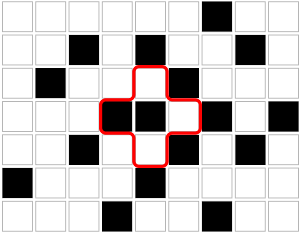
(i) The uniform Bernoulli measure is stationary.

(ii) For every $\varepsilon \in (0,1)$, the process is exponentially ergodic:

$$X^t \xrightarrow[t\to\infty]{} \text{uniform Bernoulli.}$$

The system very rapidly forgets its initial condition!

Example 3: Majority rule + noise



At every step,

I. Majority rule: Apply majority at each position.

$$\begin{array}{ccc} & \boxed{u} & \\ \boxed{l} & \boxed{c} & \boxed{r} \\ & \boxed{d} & \end{array} \longmapsto \mathrm{majority}(c, r, u, l, d)$$

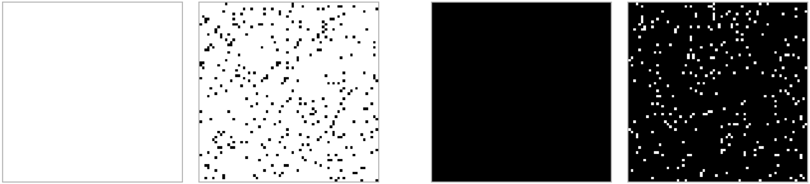
II. Noise: Flip each value independently with probability ε.

Question

Will the noise make the system forget its initial configuration?

Example 3: Majority rule + noise

Starting from all-0 | Starting from all-1



time 0 | time 1000 | time 0 | time 1000

This has not been proven!

Conjecture

For sufficiently small $\varepsilon > 0$, the system has two distinct stationary measures, characterized by their density of black sites.

A fundamental question

Question

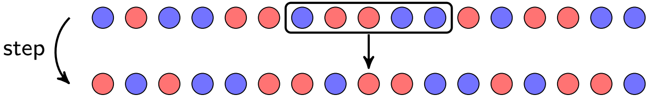
Can a fully probabilistic PCA remember some information about its initial condition indefinitely?

Recall

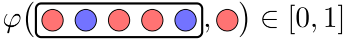
Every fully probabilistic finite-state Markov chain is ergodic:

- it has a unique stationary distribution;
- it converges to it irrespective of its initial state.

Probabilistic cellular automata (PCA)



The local rule can be represented by a stochastic matrix:

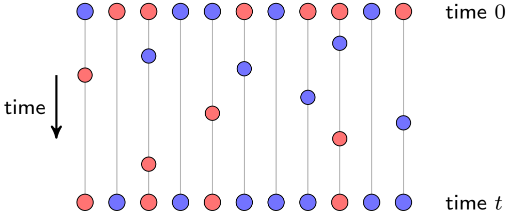


Points of view

1. Markov process on a compact metric space.
2. Computer science: computation in the presence of noise.
3. Statistical physics: non-equilibrium aspects of lattice models.

PCA vs. IPS

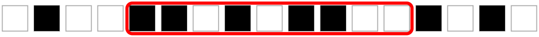
The continuous-time variants of PCA are called interacting particle systems (IPS).



Each site has an independent Poisson process that triggers transitions according to the local rule.

Ergodicity vs. non-ergodicity

An equilibrium state is represented by a probability measure on all configurations that is stationary.



Proposition

Every PCA has at least one stationary measure.

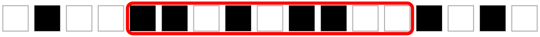
Three possibilities

(A) Ergodic: A unique stationary measure that attracts every initial configuration. [The system forgets its initial condition.]

(C) Multiple equilibrium states: More than one stationary measure.

Ergodicity vs. non-ergodicity

An equilibrium state is represented by a probability measure on all configurations that is stationary.



Proposition

Every PCA has at least one stationary measure.

Three possibilities

(A) Ergodic: A unique stationary measure that attracts every initial configuration. [The system forgets its initial condition.]

(B) In between: A unique stationary measure, but it does not attract every configuration.

(C) Multiple equilibrium states: More than one stationary measure.

Ergodicity vs. non-ergodicity

Three possibilities

(A) Ergodic: A unique stationary measure that attracts every initial configuration. [The system forgets its initial condition.]

(B) In between: A unique stationary measure, but it does not attract every configuration. [An example by Chassaing and Mairesse (2010)]

(C) Multiple equilibrium states: More than one stationary measure.

Open question

Can a fully probabilistic PCA have a unique stationary measure but be non-ergodic?

[See Mountford (1995) for one-dimensional IPS.]
[See Jahnel–Külske (2015) for a related construction.]

Fully probabilistic means the local rule is strictly positive.

Back to the fundamental question

- Example 1 (for $p < p_*$) and Example 2 are ergodic.
- Example 1 (for $p > p_*$) and Example 3 (without proof) are non-ergodic.

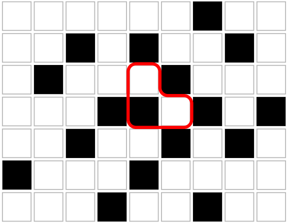
Question

Can a fully probabilistic PCA be non-ergodic?

Theorem (high noise regime)

Any perturbation of a CA with a sufficiently strong noise is ergodic.

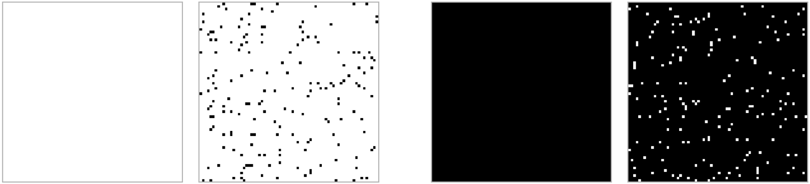
Example 4: Toom's NEC rule



$$\begin{array}{ll} \boxed{n} & \\ \boxed{c} & \boxed{e} \end{array} \longmapsto \mathrm{majority}(n, e, c)$$

Example 4: Toom's NEC rule + noise

Starting from all-0 | Starting from all-1



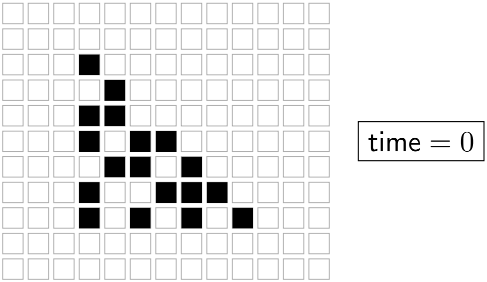
time 0 | time 1000 | time 0 | time 1000

Theorem (Toom, 1980)

For sufficiently small $\varepsilon > 0$, Toom's NEC rule with noise has two distinct stationary measures, characterized by their density of black sites.

In fact, this remains true even when the noise is biased.

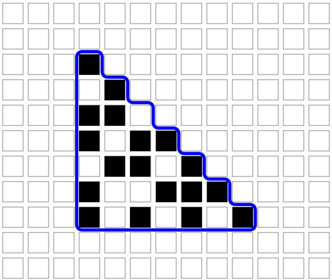
Cellular automata + independent noise



Cleaning finite islands

A finite island of black in a sea of white is quickly cleaned and vice versa! [A triangle of faults shrink.]

Cellular automata + independent noise

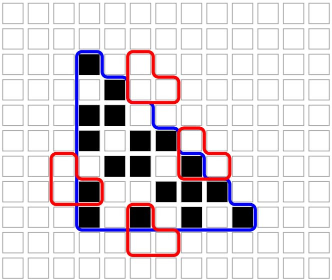


time $= 0$

Cleaning finite islands

A finite island of black in a sea of white is quickly cleaned and vice versa! [A triangle of faults shrink.]

Cellular automata + independent noise



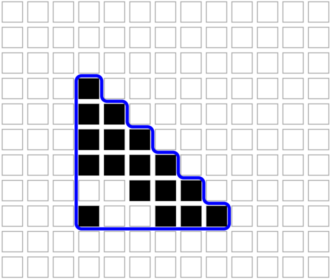
time $= 0$

Cleaning finite islands

A finite island of black in a sea of white is quickly cleaned and vice versa!

[A triangle of faults shrink.]

Cellular automata + independent noise

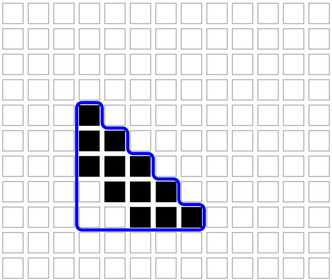


time $= 1$

Cleaning finite islands

A finite island of black in a sea of white is quickly cleaned and vice versa! [A triangle of faults shrink.]

Cellular automata + independent noise

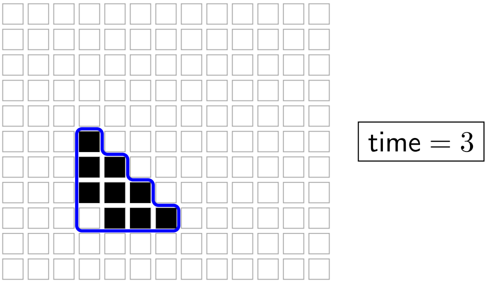


time $= 2$

Cleaning finite islands

A finite island of black in a sea of white is quickly cleaned
and vice versa! [A triangle of faults shrink.]

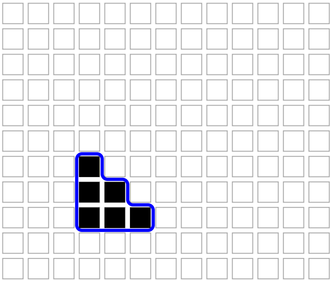
Cellular automata + independent noise



Cleaning finite islands

A finite island of black in a sea of white is quickly cleaned and vice versa! [A triangle of faults shrink.]

Cellular automata + independent noise

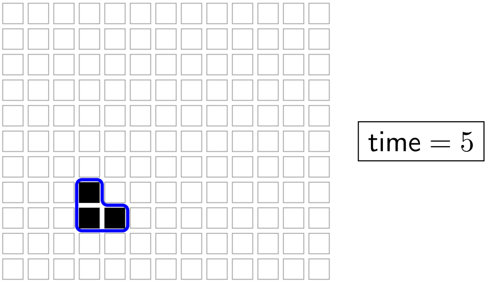


time $= 4$

Cleaning finite islands

A finite island of black in a sea of white is quickly cleaned and vice versa! [A triangle of faults shrink.]

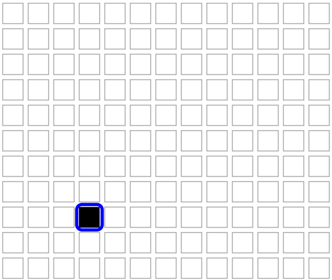
Cellular automata + independent noise



Cleaning finite islands

A finite island of black in a sea of white is quickly cleaned and vice versa! [A triangle of faults shrink.]

Cellular automata + independent noise

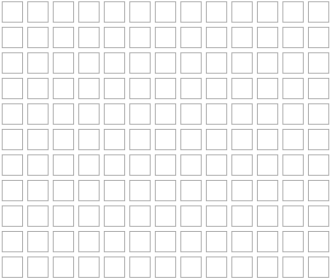


time $= 6$

Cleaning finite islands

A finite island of black in a sea of white is quickly cleaned and vice versa! [A triangle of faults shrink.]

Cellular automata + independent noise



time $= 7$

Cleaning finite islands

A finite island of black in a sea of white is quickly cleaned
and vice versa! [A triangle of faults shrink.]

Toom's stability theorem

Eroders

A deterministic CA is an eroder on a configuration x if it cleans every finite island of perturbation on x.

Monotonic CA

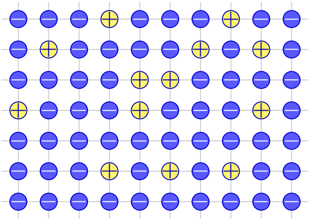
A two-state CA is monotonic if $T(x) \preceq T(y)$ whenever $x \preceq y$.

Theorem (Toom, 1980)

Let T be a monotonic CA, and suppose that T is an eroder on all-0. Then all-0 is stable against sufficiently small (possibly biased) noise.

In particular, if a monotonic CA is both a 0-eroder and a 1-eroder, then ($T +$ weak noise) has two distinct stationary measures.

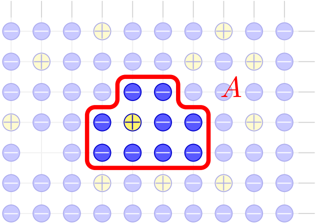
Example 5: The Ising model



States		Interaction energies
$\oplus$:	upward spin	$\Phi(\oplus\,\oplus) = \Phi(\ominus\,\ominus) = -1$
$\ominus$:	downward spin	$\Phi(\oplus\,\ominus) = 1$

The energy of a finite region A is the sum of the interaction energies involving the spins in A.

Example 5: Gibbs measures for the Ising model



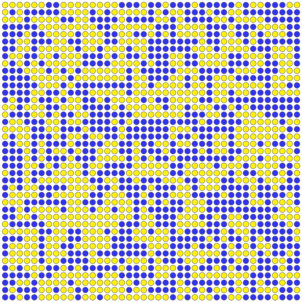
Conditioned on the configuration outside A, the pattern inside A has the Boltzmann distribution determined by the boundary.

$$\mathbb{P}(X_A = x_A \mid X_{A^c}) \sim \mathrm{e}^{-\beta E_{A|A^c}(x_A \vee X_{A^c})}$$

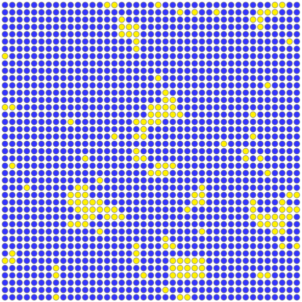
A probability measure with this property for every finite region A is called a Gibbs measure.

Gibbs measures always exist, but they need not be unique.

Example 5: Phase transition in the Ising model



High temperature

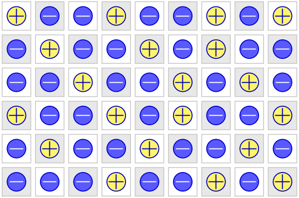


Low temperature

Phase transition

- At high temperature, there is a unique Gibbs measure.
- At sufficiently low temperature, there are multiple Gibbs measures.

Example 5: Gibbs sampler for the Ising model



At each step,

I. Resample the even sites independently according to their Gibbs conditional distribution.

II. Resample the odd sites independently according to their Gibbs conditional distribution.

Key fact

Every Ising Gibbs measure is stationary under each half-step, and hence under the full checkerboard Gibbs sampler.

At every positive temperature the updates are fully probabilistic, but at low temperature the sampler is non-ergodic.

Non-ergodicity in one dimension

Question

Can a one-dimensional fully probabilistic PCA be non-ergodic?

Evidence against

- There is no one-dimensional monotonic CA with the erosion property on both all-0 and all-1.
- Gibbs measures in one dimension are unique.

Positive-rate conjecture

Every one-dimensional fully probabilistic PCA/IPS is ergodic.

Non-ergodicity in one dimension

Question

Can a one-dimensional fully probabilistic PCA be non-ergodic?

Evidence against

- There is no one-dimensional monotonic CA with the erosion property on both all-0 and all-1.
- Gibbs measures in one dimension are unique.

Positive-rate conjecture

Every one-dimensional fully probabilistic PCA/IPS is ergodic.

False!

Non-ergodicity in one dimension

Theorem (Gács, 1986, 2001)

There exists a one-dimensional CA that remains non-ergodic in the presence of sufficiently weak noise, even if the noise is biased.

In fact

- Gács's CA is capable of performing any computation reliably, as long as the noise is sufficiently weak.
- The construction is quite flexible and can implement other things. [e.g., it also works in continuous time]

Downside

- The construction is very complex, with an astronomical number of symbols per site.
- For it to work, the noise has to be astronomically small.

Gács's construction: rough idea

Theorem (Gács, 1986, 2001)

There exists a one-dimensional CA that remains non-ergodic in the presence of sufficiently weak noise, even if the noise is biased.

The construction relies on the idea of self-reference in computation theory.

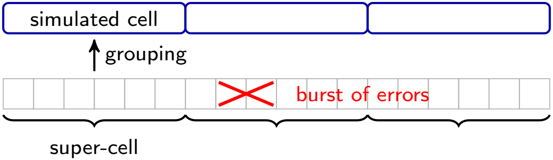
Starting ideas

- It is easy to hard-wire a simple error-correction mechanism in a CA that takes care of sufficiently sparse isolated errors.
- There are one-dimensional CA that are intrinsically universal: by grouping cells into super-cells, they can simulate any other one-dimensional CA.
- Construct a CA that corrects sparse isolated errors and simulates itself.

Gács's construction: rough idea

Theorem (Gács, 1986, 2001)

There exists a one-dimensional CA that remains non-ergodic in the presence of sufficiently weak noise, even if the noise is biased.



The self-simulation should allow sufficiently sparse larger bursts of error to be handled at higher scales of simulation.

Or would it?

Larger bursts of error can also disrupt the organization of cells into super-cells.

Non-ergodicity in one dimension

Theorem (Gács, 1986, 2001)

There exists a one-dimensional CA that remains non-ergodic in the presence of sufficiently weak noise, even if the noise is biased.

Open question

Is there a simple example of a non-ergodic fully probabilistic PCA in one dimension?

Back to Stavskaya PCA: Proof sketch

Theorem

There exists a critical value $p_* \in (0, 1)$ such that

(i) if $p < p_*$, then almost surely everything dies out;

(ii) if $p > p_*$, then there is an equilibrium state other than the point mass at all-0 [equilibrium state = stationary probability measure]

Back to Stavskaya PCA: Proof sketch

Theorem

There exists a critical value $p_* \in (0,1)$ such that

(i) if $p < p_*$, then almost surely everything dies out;

(ii) if $p > p_*$, then there is an equilibrium state other than the point mass at all-0 [equilibrium state = stationary probability measure]

Thanks to monotonicity

It is enough to show that:

(i) Ergodicity: for p close to 0, the trajectory starting from all-1 almost surely asymptotically dies out.

(ii) Non-ergodicity: for p close to 1, the trajectory starting from a single 1 survives with positive probability. [oriented percolation]

Exercise

Why does monotonicity reduce the theorem to these two statements?

Stavskaya PCA: Ergodicity for small p

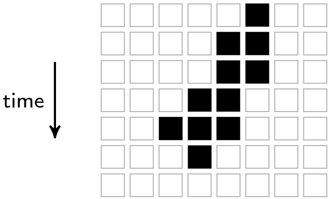
Claim

For p sufficiently close to 0, the trajectory starting from all-1 almost surely asymptotically dies out.

Stavskaya PCA: Ergodicity for small p

Claim

For p sufficiently close to 0, the trajectory starting from a single 1 almost surely asymptotically dies out.



The set of 1s at time t is dominated by a branching process with mean offspring $2p$.

Hence it almost surely dies out when $p < \frac{1}{2}$. $\square$

[equivalently: subcritical oriented percolation]

Stavskaya PCA: Ergodicity for small p

Claim

For p sufficiently close to 0, the trajectory starting from all-1 almost surely asymptotically dies out.

We use the idea of duality (a.k.a. coupling from the past).

It is enough to show that the probability that a given site is black goes to 0 as $t \to \infty$.

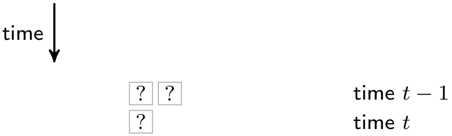
Stavskaya PCA: Ergodicity for small p

Claim

For p sufficiently close to 0, the trajectory starting from all-1 almost surely asymptotically dies out.

We use the idea of duality (a.k.a. coupling from the past).

It is enough to show that the probability that a given site is black goes to 0 as $t \to \infty$.



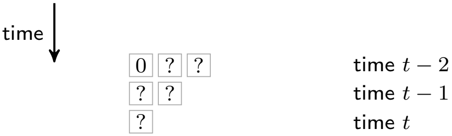
Stavskaya PCA: Ergodicity for small p

Claim

For p sufficiently close to 0, the trajectory starting from all-1 almost surely asymptotically dies out.

We use the idea of duality (a.k.a. coupling from the past).

It is enough to show that the probability that a given site is black goes to 0 as $t \to \infty$.



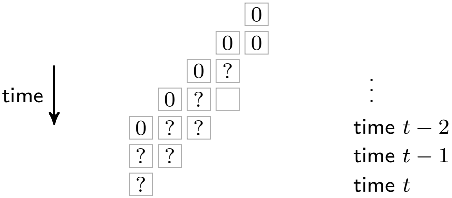
Stavskaya PCA: Ergodicity for small p

Claim

For p sufficiently close to 0, the trajectory starting from all-1 almost surely asymptotically dies out.

We use the idea of duality (a.k.a. coupling from the past).

It is enough to show that the probability that a given site is black goes to 0 as $t \to \infty$.



The backward dependency process has the same law as oriented percolation, and dies out when $p < \frac{1}{2}$. □

Stavskaya PCA: Non-ergodicity for large p

Before giving the proof ...

Stavskaya PCA as a noisy CA

$\square\square \longmapsto \square$

$\square\blacksquare \longmapsto \begin{cases} \blacksquare & \text{with prob. } p, \\ \square & \text{with prob. } 1-p. \end{cases}$

$\blacksquare\square \longmapsto \begin{cases} \blacksquare & \text{with prob. } p, \\ \square & \text{with prob. } 1-p. \end{cases}$

$\blacksquare\blacksquare \longmapsto \begin{cases} \blacksquare & \text{with prob. } p, \\ \square & \text{with prob. } 1-p. \end{cases}$

Stavskaya PCA can be viewed as a perturbation of a CA with one-sided noise.

At each step,

I. Apply the OR rule:

$\square\square \mapsto \square \qquad \square\blacksquare \mapsto \blacksquare \qquad \blacksquare\square \mapsto \blacksquare \qquad \blacksquare\blacksquare \mapsto \blacksquare$

II. Apply one-sided noise

$\blacksquare \xrightarrow{\varepsilon} \square \qquad \varepsilon = 1-p$

Stavskaya PCA as a noisy CA

Remark

The OR CA is a monotonic eroder on all-0, hence Toom's stability theorem applies.

There are two fundamentally different proofs of Toom's theorem:

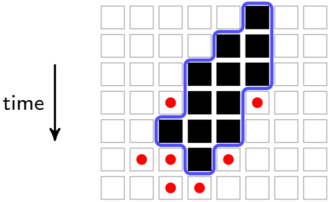
1) via a contour argument; [Toom, 1980]
2) using the sparseness of Bernoulli noise. [Gács–Reif, 1986]

We give a taste of the contour argument in this simple example.

Stavskaya PCA: Non-ergodicity for large p

Claim

For p sufficiently close to 1, the trajectory starting from a single 1 survives with positive probability.



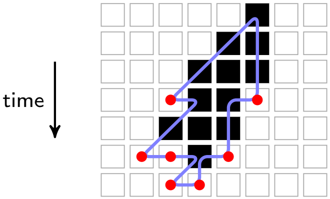
If the trajectory dies out, its 1s are encircled by a finite contour.

Each such contour requires a number of $1 \mapsto 0$ errors.

Stavskaya PCA: Non-ergodicity for large p

Claim

For p sufficiently close to 1, the trajectory starting from a single 1 survives with positive probability.



If the trajectory dies out, its 1s are encircled by a finite contour.

Each such contour requires a number of $1 \mapsto 0$ errors.

Stavskaya PCA: Non-ergodicity for large p

Claim

For p sufficiently close to 1, the trajectory starting from a single 1 survives with positive probability.

Let $\varepsilon = 1 - p$. If N_k denotes the number of possible contours requiring k errors, then

$$\mathbb{P}(\text{dying out}) \;\le\; \sum_{k=1}^{\infty} N_k \, \varepsilon^k.$$

The number N_k is bounded by an exponential function of k.

Therefore,

$$\mathbb{P}(\text{dying out}) < 1$$

for sufficiently small $\varepsilon > 0$.

Thus the process survives with positive probability, and the PCA is non-ergodic. □

Conclusions

Key points

- Ergodicity vs. non-ergodicity
  $\longrightarrow$ *Can a fully probabilistic PCA retain information indefinitely?*
  $\longrightarrow$ *Can a fully probabilistic PCA admit multiple equilibrium states?*
  $\longrightarrow$ *Can a noisy CA perform computation reliably?*
- Yes, even in one dimension!
- Many basic questions about PCA remain open.

Exercises

1) Prove the ergodicity of XOR CA + noise.
2) Prove ergodicity in the high-noise regime.

[Hint: coupling + Stavskaya/oriented percolation]

Some references

- R. L. Dobrushin, V. I. Kryukov, and A. L. Toom (eds.),
  Stochastic Cellular Systems: Ergodicity, Memory, Morphogenesis,
  Manchester University Press, 1990.

- J. Mairesse and I. Marcovici,
  "Around probabilistic cellular automata,"
  Theoretical Computer Science, 2014.

- P. Gács,
  "Probabilistic cellular automata with Andrei Toom,"
  Brazilian Journal of Probability and Statistics, 2024.

[historical narrative]

Thank you!

. . . and tune in for the next talk!