Ergodicity of cellular automata subject to noise

Siamak Taati

Bernoulli Institute, University of Groningen

UNAM, Mexico City — May 2019

Computing in presence of noise

Challenge in building computers

When implementing computation using physical components,

- Transient errors due to thermal noise are inevitable.
- The smaller the scale, the more important the effect of noise.
- In a lengthy computation, the errors may propagate.

Computing in presence of noise

Challenge in building computers

When implementing computation using physical components,

- Transient errors due to thermal noise are inevitable.
- The smaller the scale, the more important the effect of noise.
- In a lengthy computation, the errors may propagate.

Computing in presence of noise

Challenge in building computers

When implementing computation using physical components,

- Transient errors due to thermal noise are inevitable.
- The smaller the scale, the more important the effect of noise.
- In a lengthy computation, the errors may propagate.

Computing in presence of noise

Challenge in building computers

When implementing computation using physical components,

- Transient errors due to thermal noise are inevitable.
- The smaller the scale, the more important the effect of noise.
- In a lengthy computation, the errors may propagate.

Computing in presence of noise

Shannon (1948):

Can we do reliable communication through a noisy channel?

Von Neumann (1952):

Can we do reliable computation using noisy components?

Computing in presence of noise

Shannon (1948):

Can we do reliable communication through a noisy channel?

Solution: Yes, if we use constant redundancy! [Shannon (1948)]

Von Neumann (1952):

Can we do reliable computation using noisy components?

Computing in presence of noise

Shannon (1948):

Can we do reliable communication through a noisy channel?

Solution: Yes, if we use constant redundancy! [Shannon (1948)]

Von Neumann (1952):

Can we do reliable computation using noisy components?

Solution:

Computing in presence of noise

Shannon (1948):

Can we do reliable communication through a noisy channel?

Solution: Yes, if we use constant redundancy! [Shannon (1948)]

Von Neumann (1952):

Can we do reliable computation using noisy components?

Solution:

- For logic circuits: Yes, if we use logarithmic redundancy!
  [Von Neumann (1956), Dobrushin and Ortyukov (1977), Pippenger (1985)]

Computing in presence of noise

Shannon (1948):

Can we do reliable communication through a noisy channel?

Solution: Yes, if we use constant redundancy! [Shannon (1948)]

Von Neumann (1952):

Can we do reliable computation using noisy components?

Solution:

- For logic circuits: Yes, if we use logarithmic redundancy!
  [Von Neumann (1956), Dobrushin and Ortyukov (1977), Pippenger (1985)]
- . . .

Computing in presence of noise

Shannon (1948):

Can we do reliable communication through a noisy channel?

Solution: Yes, if we use constant redundancy! [Shannon (1948)]

Von Neumann (1952):

Can we do reliable computation using noisy components?

Solution:

- For logic circuits: Yes, if we use logarithmic redundancy!
  [Von Neumann (1956), Dobrushin and Ortyukov (1977), Pippenger (1985)]
- . . .
- For cellular automata: Yes, but the known solution is very very sophisticated!
  [Toom (1974, 1980), Gács and Reif (1988), Gács (1986, 2001)]

Computing in presence of noise

Shannon (1948):

Can we do reliable communication through a noisy channel?

Solution: Yes, if we use constant redundancy! [Shannon (1948)]

Von Neumann (1952):

Can we do reliable computation using noisy components?

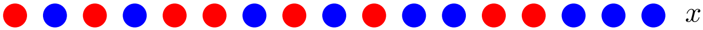
Solution:

- For logic circuits: Yes, if we use logarithmic redundancy!
  [Von Neumann (1956), Dobrushin and Ortyukov (1977), Pippenger (1985)]
- . . .
- For cellular automata: Yes, but the known solution is very very sophisticated!
  [Toom (1974, 1980), Gács and Reif (1988), Gács (1986, 2001)]

Cellular automata (CA)

A simple model of (parallel) computation [cf. Turing machine]

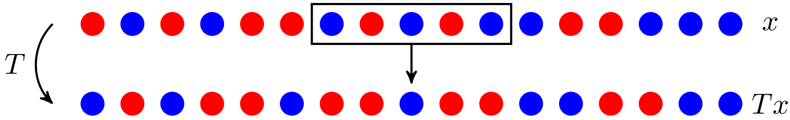
Cellular automata (CA)



A simple model of (parallel) computation [cf. Turing machine]

- Data is represented as an array of symbols (or colors) arranged on a (d-dimensional) lattice. [Finite or infinite]

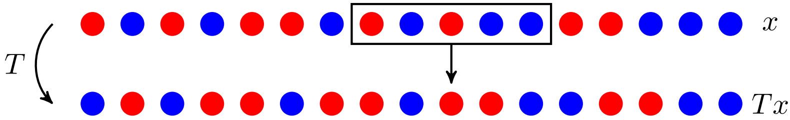
Cellular automata (CA)



A simple model of (parallel) computation [cf. Turing machine]

- Data is represented as an array of symbols (or colors) arranged on a (d-dimensional) lattice. [Finite or infinite]
- Each step of computation consists in updating the symbols simultaneously using a local rule. [Same local rule at every site!]

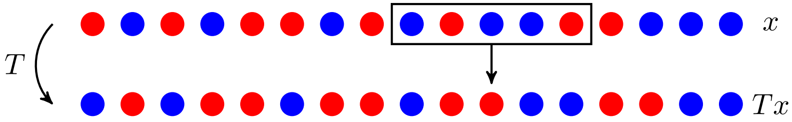
Cellular automata (CA)



A simple model of (parallel) computation [cf. Turing machine]

- Data is represented as an array of symbols (or colors) arranged on a (d-dimensional) lattice. [Finite or infinite]
- Each step of computation consists in updating the symbols simultaneously using a local rule. [Same local rule at every site!]

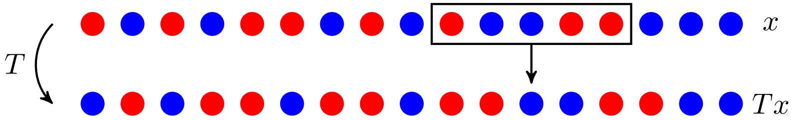
Cellular automata (CA)



A simple model of (parallel) computation [cf. Turing machine]

- Data is represented as an array of symbols (or colors) arranged on a (d-dimensional) lattice. [Finite or infinite]
- Each step of computation consists in updating the symbols simultaneously using a local rule. [Same local rule at every site!]

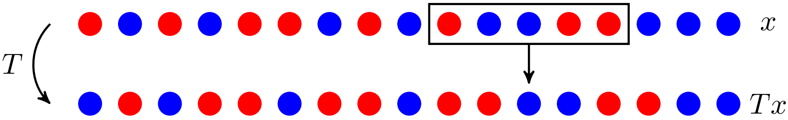
Cellular automata (CA)



A simple model of (parallel) computation [cf. Turing machine]

- Data is represented as an array of symbols (or colors) arranged on a (d-dimensional) lattice. [Finite or infinite]
- Each step of computation consists in updating the symbols simultaneously using a local rule. [Same local rule at every site!]

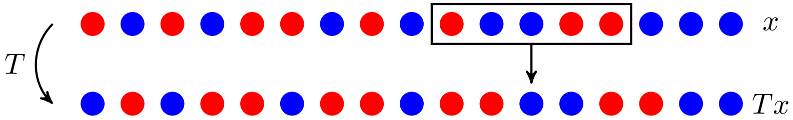
Cellular automata (CA)



A simple model of (parallel) computation [cf. Turing machine]

- Data is represented as an array of symbols (or colors) arranged on a (d-dimensional) lattice. [Finite or infinite]
- Each step of computation consists in updating the symbols simultaneously using a local rule. [Same local rule at every site!]
- Iterate! [Same local rule at every time step!]

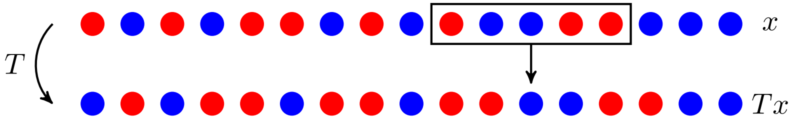
Cellular automata (CA)



CA are discrete-time dynamical systems

- The set of all configurations x is a compact metric space!
- The global transformation $x \mapsto Tx$ is continuous!

Cellular automata (CA)

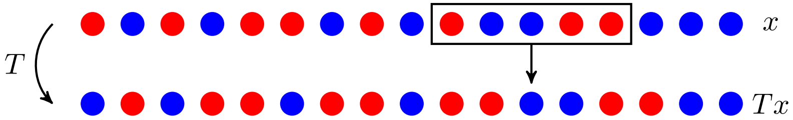


CA are discrete-time dynamical systems

- The set of all configurations x is a compact metric space!
- The global transformation $x \mapsto Tx$ is continuous!

$\Longrightarrow$ We can exploit the machinery of dynamical systems and ergodic theory!

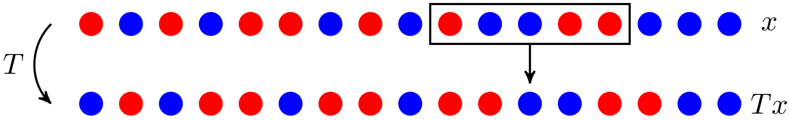
Cellular automata (CA)



CA have "physics-like" features

- Finite number of possible states at each site
- Local interactions [No action at a distance!]
- Reversibility and conservation laws can be easily implemented.

Cellular automata (CA)

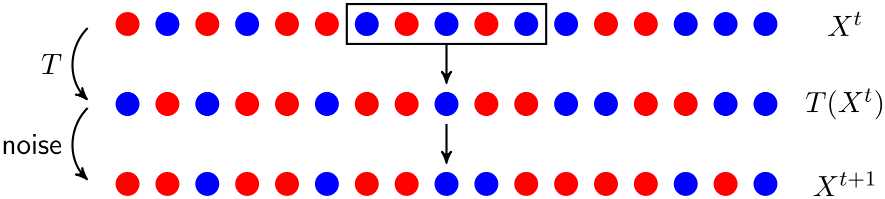


CA have "physics-like" features

- Finite number of possible states at each site
- Local interactions [No action at a distance!]
- Reversibility and conservation laws can be easily implemented.

$\Longrightarrow$ Convenient for mathematical reasoning about physical implementations of computation.

Cellular automata (CA)



Cellular automata subject to noise

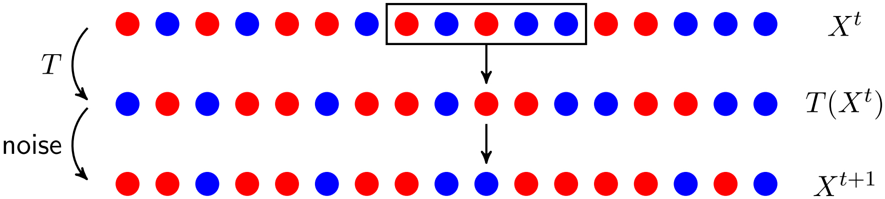
At each step,

a) first, apply the deterministic CA,

b) then, add noise independently at each site.

[Various models of noise possible!]

Cellular automata (CA)



Cellular automata subject to noise

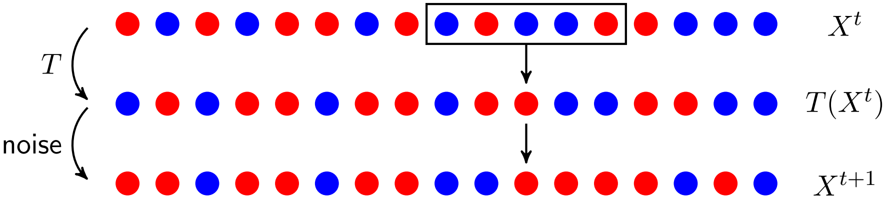
At each step,

a) first, apply the deterministic CA,

b) then, add noise independently at each site.

[Various models of noise possible!]

Cellular automata (CA)



Cellular automata subject to noise

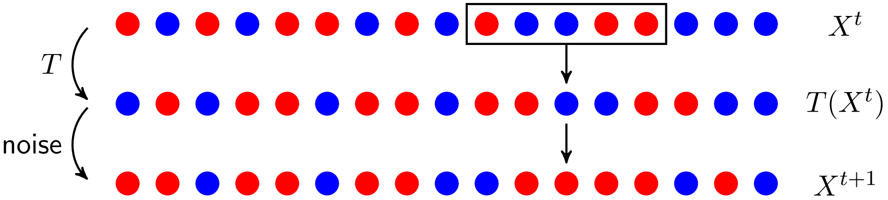
At each step,

a) first, apply the deterministic CA,

b) then, add noise independently at each site.

[Various models of noise possible!]

Cellular automata (CA)



Cellular automata subject to noise

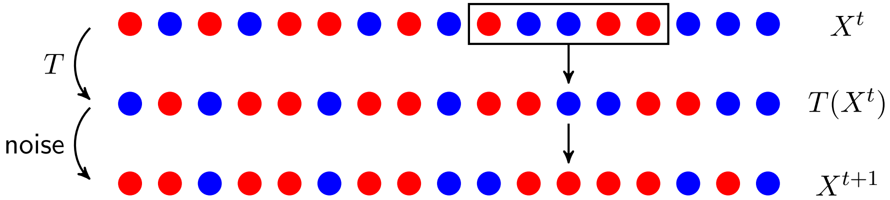
At each step,

a) first, apply the deterministic CA,

b) then, add noise independently at each site.

[Various models of noise possible!]

Cellular automata (CA)



Cellular automata subject to noise

At each step,

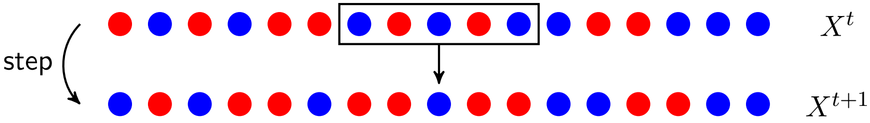
a) first, apply the deterministic CA,

b) then, add noise independently at each site.

[Various models of noise possible!]

$\rightsquigarrow$ A special type of probabilistic cellular automaton (PCA).

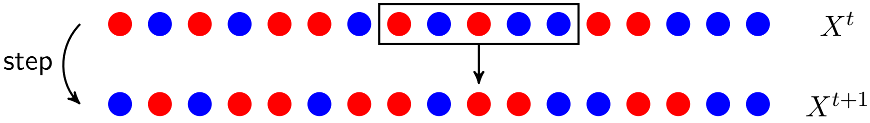
Probabilistic cellular automata (PCA)



PCA are similar to CA, except that

- The local rule is probabilistic! [Described by a stochastic matrix]
- Symbols at different sites are updated independently.

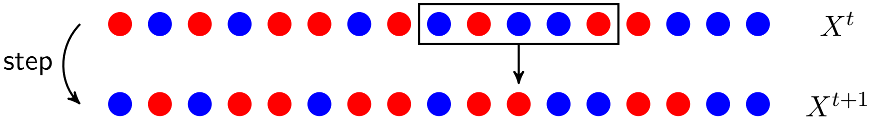
Probabilistic cellular automata (PCA)



PCA are similar to CA, except that

- The local rule is probabilistic! [Described by a stochastic matrix]
- Symbols at different sites are updated independently.

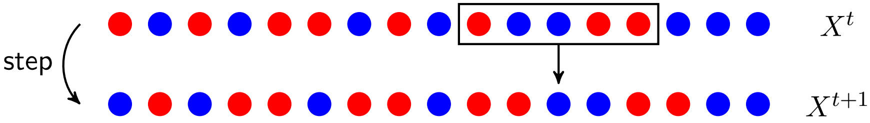
Probabilistic cellular automata (PCA)



PCA are similar to CA, except that

- The local rule is probabilistic! [Described by a stochastic matrix]
- Symbols at different sites are updated independently.

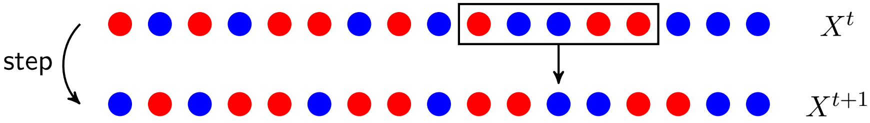
Probabilistic cellular automata (PCA)



PCA are similar to CA, except that

- The local rule is probabilistic! [Described by a stochastic matrix]
- Symbols at different sites are updated independently.

Probabilistic cellular automata (PCA)



PCA are similar to CA, except that

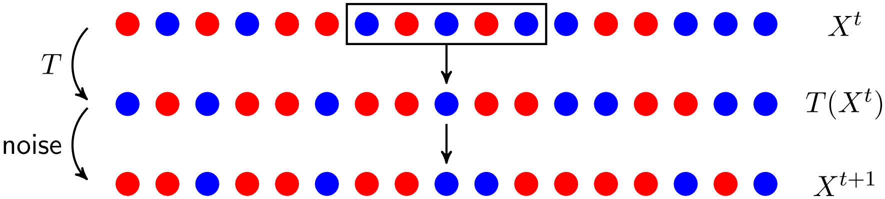
- The local rule is probabilistic! [Described by a stochastic matrix]
- Symbols at different sites are updated independently.

PCA are discrete-time Markov processes

- The state at time t is a random configuration X^t.
- The transition kernel has the Feller property.

[Discrete-time variants of interacting particle systems]

Computing with noisy CA



Problem (Reliable simulation)

Can we "simulate" a CA T with another CA S that is "reliable against sufficiently weak noise"?

$$\begin{tikzcd} x \arrow[r] \arrow[d, "\text{encode}"'] & Tx \\ y \arrow[r] & Sy \arrow[u, "\text{decode}"'] \end{tikzcd}$$

Computing with noisy CA

Problem (Reliable simulation)

Can we “simulate” a CA T with another CA S that is “reliable against sufficiently weak noise”?

Computing with noisy CA

Problem (Reliable simulation)

Can we "simulate" a CA T with another CA S that is "reliable against sufficiently weak noise"?

A simpler prerequisite:

Problem (Remembering a bit)

Find a CA that, in presence of sufficiently weak noise is cable of "remembering" at least 1 bit of information indefinitely!

Computing with noisy CA

Problem (Reliable simulation)

Can we "simulate" a CA T with another CA S that is "reliable against sufficiently weak noise"?

A simpler prerequisite:

Problem (Remembering a bit)

Find a CA that, in presence of sufficiently weak noise is cable of "remembering" at least 1 bit of information indefinitely!

Precise formulation in the language of Markov processes:

Problem (Ergodicity of noisy CA)

Find a CA that, in presence of sufficiently weak noise remains non-ergodic!

[Ergodicity: having a unique stationary measure that attracts every trajectory]

Ergodicity of noisy CA

Problem (Ergodicity of noisy CA)

Find a CA that, in presence of sufficiently weak noise remains non-ergodic!

Important earlier works

Ergodicity of noisy CA

Problem (Ergodicity of noisy CA)

Find a CA that, in presence of sufficiently weak noise remains non-ergodic!

Important earlier works

- Toom (1974, 1980): a broad family of examples of CA in two and higher dimensions that remain non-ergodic in presence of noise.

Ergodicity of noisy CA

Problem (Ergodicity of noisy CA)

Find a CA that, in presence of sufficiently weak noise remains non-ergodic!

Important earlier works

- Toom (1974, 1980): a broad family of examples of CA in two and higher dimensions that remain non-ergodic in presence of noise.
- Gács and Reif (1988): every d-dimensional CA can be reliably simulated by a $(d+2)$-dimensional CA.

Ergodicity of noisy CA

Problem (Ergodicity of noisy CA)

Find a CA that, in presence of sufficiently weak noise remains non-ergodic!

Important earlier works

- Toom (1974, 1980): a broad family of examples of CA in two and higher dimensions that remain non-ergodic in presence of noise.
- Gács and Reif (1988): every d-dimensional CA can be reliably simulated by a $(d+2)$-dimensional CA.

[3d reliable computer not practical!]

Ergodicity of noisy CA

Problem (Ergodicity of noisy CA)

Find a CA that, in presence of sufficiently weak noise remains non-ergodic!

Important earlier works

- Toom (1974, 1980): a broad family of examples of CA in two and higher dimensions that remain non-ergodic in presence of noise.
- Gács and Reif (1988): every d-dimensional CA can be reliably simulated by a $(d+2)$-dimensional CA.

  [3d reliable computer not practical!]
- Gács (1986, 2001): a one-dimensional CA that is non-ergodic in presence of sufficiently weak noise.

Ergodicity of noisy CA

Problem (Ergodicity of noisy CA)

Find a CA that, in presence of sufficiently weak noise remains non-ergodic!

Important earlier works

- Toom (1974, 1980): a broad family of examples of CA in two and higher dimensions that remain non-ergodic in presence of noise.
- Gács and Reif (1988): every d-dimensional CA can be reliably simulated by a $(d+2)$-dimensional CA.

  [3d reliable computer not practical!]
- Gács (1986, 2001): a one-dimensional CA that is non-ergodic in presence of sufficiently weak noise.
  In fact, Gács's example is universal: it reliably simulates any 1d CA!

Ergodicity of noisy CA

Problem (Ergodicity of noisy CA)

Find a CA that, in presence of sufficiently weak noise remains non-ergodic!

Important earlier works

- Toom (1974, 1980): a broad family of examples of CA in two and higher dimensions that remain non-ergodic in presence of noise.
- Gács and Reif (1988): every d-dimensional CA can be reliably simulated by a $(d+2)$-dimensional CA.

  [3d reliable computer not practical!]
- Gács (1986, 2001): a one-dimensional CA that is non-ergodic in presence of sufficiently weak noise.
  In fact, Gács's example is universal: it reliably simulates any 1d CA!

  [Very sophisticated construction with astronomical number of symbols!]

Ergodicity of noisy CA

Idea: Approach the problem from the other side in order to narrow down the search.

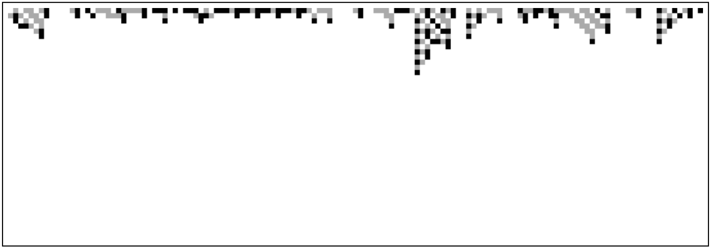
Problem (Sufficient conditions for ergodicity)

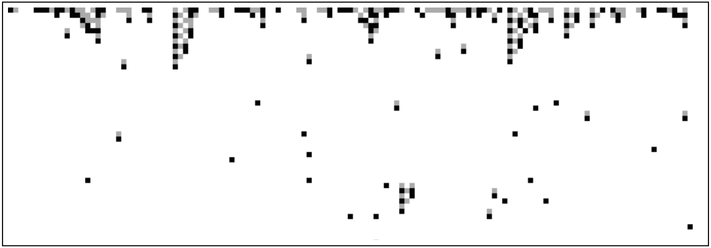
Identify dynamical/combinatorial properties for the CA that ensure the ergodicity of the noisy version.

[A reliable CA should avoid those!]

Ergodicity of noisy CA

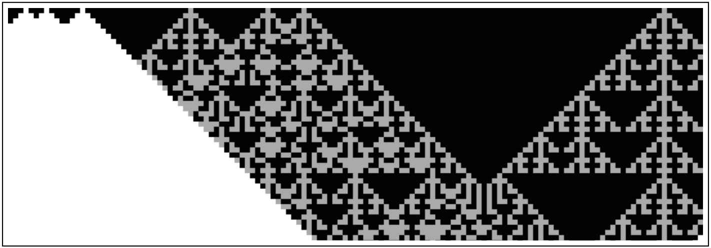
Example 1 (A nilpotent CA)

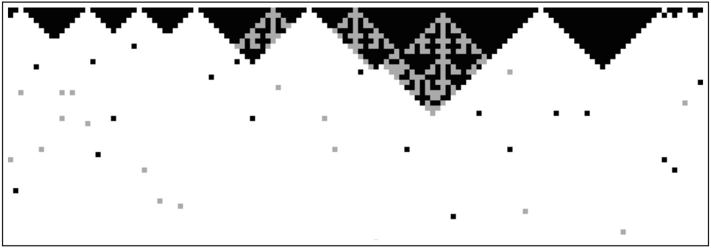




Ergodicity of noisy CA

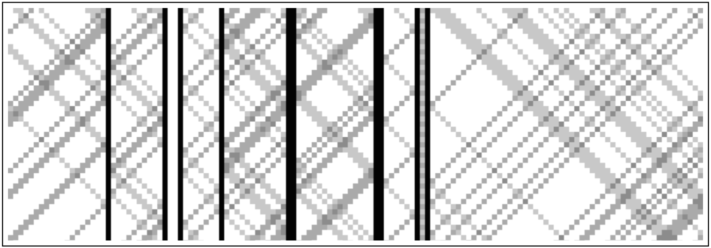
Example 2 (A CA with spreading symbol)

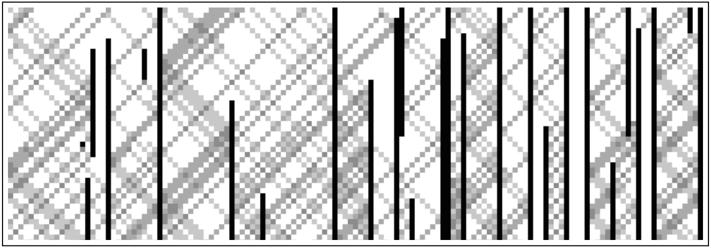




Ergodicity of noisy CA

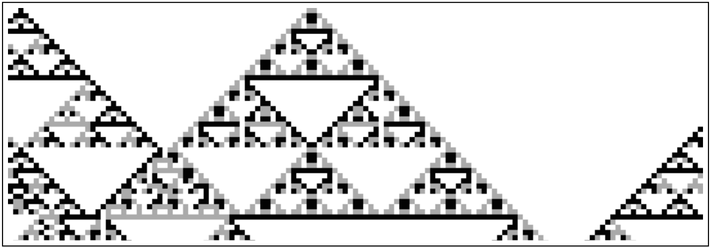
Example 3 (An almost equicontinuous CA)

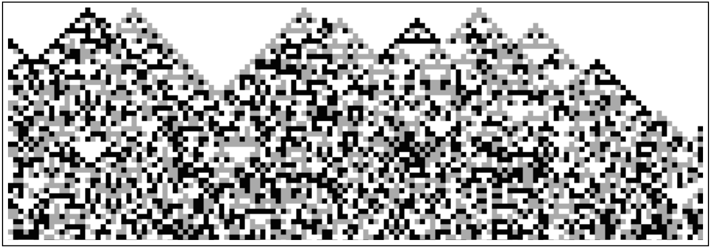




Ergodicity of noisy CA

Example 4 (A surjective CA)





Ergodicity of noisy CA

Idea: Approach the problem from the other side in order to narrow down the search.

Problem (Sufficient conditions for ergodicity)

Identify dynamical/combinatorial properties for the CA that ensure the ergodicity of the noisy version.

[A reliable CA should avoid those!]

Remarks

Ergodicity of noisy CA

Idea: Approach the problem from the other side in order to narrow down the search.

Problem (Sufficient conditions for ergodicity)

Identify dynamical/combinatorial properties for the CA that ensure the ergodicity of the noisy version.

[A reliable CA should avoid those!]

Remarks

- Most CA expected to be ergodic in presence of positive noise.

Ergodicity of noisy CA

Idea: Approach the problem from the other side in order to narrow down the search.

Problem (Sufficient conditions for ergodicity)

Identify dynamical/combinatorial properties for the CA that ensure the ergodicity of the noisy version.

[A reliable CA should avoid those!]

Remarks

- Most CA expected to be ergodic in presence of positive noise.
- Nevertheless, proving/disproving ergodicity often quite difficult. [$\leftrightsquigarrow$ equilibrium statistical mechanics at low temperature]

[algorithmically undecidable?]

Ergodicity of noisy CA

Idea: Approach the problem from the other side in order to narrow down the search.

Problem (Sufficient conditions for ergodicity)

Identify dynamical/combinatorial properties for the CA that ensure the ergodicity of the noisy version.

[A reliable CA should avoid those!]

Remarks

- Most CA expected to be ergodic in presence of positive noise.
- Nevertheless, proving/disproving ergodicity often quite difficult. [$\leftrightsquigarrow$ equilibrium statistical mechanics at low temperature]

  [algorithmically undecidable?]
- Different mechanisms for ergodicity.

Ergodicity of noisy CA

Summary of results [Marcovici, Sablik, T. (2017)]

	Type of CA	**Type of noise**
I	Any CA	High noise
II	Nilpotent	Small perturbation
III	CA with spreading symbol	Memoryless noise
IV	———— " ———— (1d with $\mathcal{N} = \{0, 1\}$)	Small positive perturbation
V	Gliders with annihilation	Birth-death noise
VI	Simple gliders with reflecting walls	——— " ———
VII	Permutive	Permutation noise
VIII	Surjective	Additive noise
IX	XOR	Zero-range
X	Binary CA with spreading symbol	Zero-range (75% of parameter range)

Ergodicity of noisy CA

Summary of results [Marcovici, Sablik, T. (2017)]

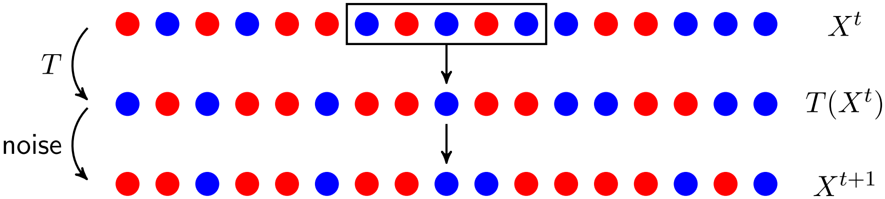
		Type of CA	Type of noise
coupling	I	Any CA	High noise
	II	Nilpotent	Small perturbation
	III	CA with spreading symbol	Memoryless noise
	IV	——— " ——— (1d with $\mathcal{N} = \{0,1\}$)	Small positive perturbation
	V	Gliders with annihilation	Birth-death noise
	VI	Simple gliders with reflecting walls	——— " ———
	VII	Permutive	Permutation noise
entropy	VIII	Surjective	Additive noise
Fourier analysis	IX	XOR	Zero-range
	X	Binary CA with spreading symbol	Zero-range (75% of parameter range)

Ergodicity of noisy CA

Summary of results [Marcovici, Sablik, T. (2017)]

		Type of CA	**Type of noise**
coupling	I	Any CA	High noise
	II	Nilpotent	Small perturbation
	III IV	CA with spreading symbol ———— ” ———— (1d with $\mathcal{N} = \{0, 1\}$)	Memoryless noise Small positive perturbation
	V VI	Gliders with annihilation Simple gliders with reflecting walls	Birth-death noise ———— ” ————
	VII	Permutive	Permutation noise
entropy	VIII	Surjective	Additive noise
Fourier analysis	IX	XOR	Zero-range
	X	Binary CA with spreading symbol	Zero-range (75% of parameter range)

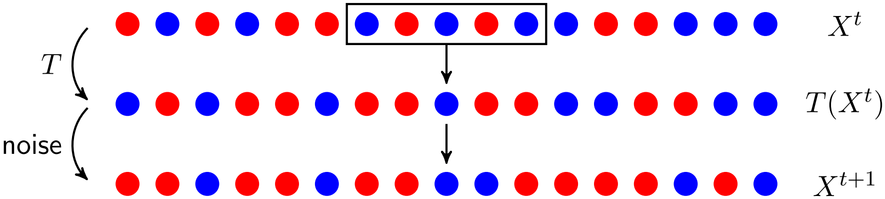
Surjective CA + additive noise



Terminology

- *Surjective CA:* The global map T is onto.
- *Additive noise:* Noise adds a random value to current value, independently at each site. [modulo $|\Sigma|$]

Surjective CA + additive noise



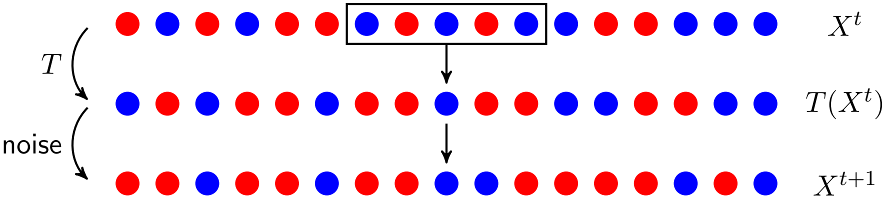
Terminology

- *Surjective CA:* The global map T is onto.
- *Additive noise:* Noise adds a random value to current value, independently at each site. [modulo $|\Sigma|$]

Remark

- Both a surjective CA and an additive noise preserve the uniform Bernoulli measure.

Surjective CA + additive noise



Theorem [Marcovici, Sablik, T. (2017) and Markovici, T. (2018)]

Every perturbation of a surjective CA with a positive additive noise is ergodic with the uniform Bernoulli measure as its invariant measure. [Convergence is exponentially fast!]

Surjective CA $+$ additive noise

Theorem [Marcovici, Sablik, T. (2017) and Markovici, T. (2018)]

Every perturbation of a surjective CA with a positive additive noise is ergodic with the uniform Bernoulli measure as its invariant measure. [Convergence is exponentially fast!]

Remarks

Surjective CA + additive noise

Theorem [Marcovici, Sablik, T. (2017) and Markovici, T. (2018)]

Every perturbation of a surjective CA with a positive additive noise is ergodic with the uniform Bernoulli measure as its invariant measure. [Convergence is exponentially fast!]

Remarks

- Surjective CA include all reversible CA.

Surjective CA + additive noise

Theorem [Marcovici, Sablik, T. (2017) and Markovici, T. (2018)]

Every perturbation of a surjective CA with a positive additive noise is ergodic with the uniform Bernoulli measure as its invariant measure. [Convergence is exponentially fast!]

Remarks

- Surjective CA include all reversible CA.
- Computing with reversible components has been suggested as a way to control heat production during the computation. [Landauer (1961), Bennett (1973), Fredkin and Toffoli (1982)]

Surjective CA + additive noise

Theorem [Marcovici, Sablik, T. (2017) and Markovici, T. (2018)]

Every perturbation of a surjective CA with a positive additive noise is ergodic with the uniform Bernoulli measure as its invariant measure. [Convergence is exponentially fast!]

Remarks

- Surjective CA include all reversible CA.
- Computing with reversible components has been suggested as a way to control heat production during the computation. [Landauer (1961), Bennett (1973), Fredkin and Toffoli (1982)]

Interpretation of the corollary

A reversible CA-like computer subject to noise forgets all the information in its input/software exponentially fast!

Reversible computing with noisy components

Interpretation of the corollary

A reversible CA-like computer subject to noise forgets all the information in its input/software exponentially fast!

In fact:

Reversible computing with noisy components

Interpretation of the corollary

A reversible CA-like computer subject to noise forgets all the information in its input/software exponentially fast!

In fact:

- The state of any region of size n mixes in $O(\log n)$ steps.

Reversible computing with noisy components

Interpretation of the corollary

A reversible CA-like computer subject to noise forgets all the information in its input/software exponentially fast!

In fact:

- The state of any region of size n mixes in $O(\log n)$ steps.
- A finite parallel reversible computer with n noisy components mixes in $O(\log n)$ steps. [Very limited computational power!]

[see Aharonov, Ben-Or, Impagliazzo, Nisan (1996)]

Reversible computing with noisy components

Interpretation of the corollary

A reversible CA-like computer subject to noise forgets all the information in its input/software exponentially fast!

In fact:

- The state of any region of size n mixes in $O(\log n)$ steps.
- A finite parallel reversible computer with n noisy components mixes in $O(\log n)$ steps. [Very limited computational power!]

[see Aharonov, Ben-Or, Impagliazzo, Nisan (1996)]

Practical implication

In order to implement noise-resilient (CA-like) computers, some degree of irreversibility is necessary.

[see Bennett (1982) and Bennett and Grinstein (1985)]

Surjective CA + additive noise

Theorem [Marcovici, Sablik, T. (2017) and Markovici, T. (2018)]

Every perturbation of a surjective CA with a positive additive noise is ergodic with the uniform Bernoulli measure as its invariant measure. [Convergence is exponentially fast!]

Proof idea.

Ergodicity is due to the accumulation of information.
Use entropy to measure the amount of information. □

Surjective CA + additive noise

Theorem [Marcovici, Sablik, T. (2017) and Markovici, T. (2018)]

Every perturbation of a surjective CA with a positive additive noise is ergodic with the uniform Bernoulli measure as its invariant measure. [Convergence is exponentially fast!]

Proof idea.

Ergodicity is due to the accumulation of information.
Use entropy to measure the amount of information. □

The entropy of a discrete random variable A is

$$H(A) \coloneqq -\sum_{a} \mathbb{P}(A = a) \log \mathbb{P}(A = a) \ .$$

It measures the average information content of A.

Surjective CA + additive noise

Theorem [Marcovici, Sablik, T. (2017) and Markovici, T. (2018)]

Every perturbation of a surjective CA with a positive additive noise is ergodic with the uniform Bernoulli measure as its invariant measure. [Convergence is exponentially fast!]

Proof ingredients.

a) A surjective CA does not "erase" entropy, only "diffuses" it.

b) Additive noise increases entropy. [Sharp estimate needed!]

For each finite set of sites J and each time step $t \geq 0$, we find

$$H(X_J^t) \geq \left[1 - (1-\kappa)^t\right] |J| \, \hbar - O(|\partial J|)$$

where $\hbar := \log |\Sigma|$ is the maximum capacity of a single site.

c) A bootstrap lemma

□

Surjective CA + zero-range noise

Theorem [Marcovici, T. (2018)]

A perturbation of a surjective CA with a positive zero-range noise is ergodic provided that both the CA and the noise preserve the same Bernoulli measure.

Proof idea.

Use pressure instead of entropy.
Use a characterization of when a surjective CA preserves a Bernoulli measure [Kari, T. (2015)]. □

Surjective CA + zero-range noise

Theorem [Marcovici, T. (2018)]

A perturbation of a surjective CA with a positive zero-range noise is ergodic provided that both the CA and the noise preserve the same Bernoulli measure.

Proof idea.

Use pressure instead of entropy.
Use a characterization of when a surjective CA preserves a Bernoulli measure [Kari, T. (2015)]. □

The pressure of a discrete random variable A w.r.t. an energy functional f is

$$\Psi_f(A) := H(A) - \mathbb{E}[f(A)] \, .$$

It can be thought of as a contorted version of entropy.

PCA with Bernoulli invariant measure

Theorem [Marcovici, T. (2018)]

Every positive-rate PCA that has a Bernoulli invariant measure is ergodic.

PCA with Bernoulli invariant measure

Theorem [Marcovici, T. (2018)]

Every positive-rate PCA that has a Bernoulli invariant measure is ergodic. [Same true for positive-rate IPS!]

PCA with Bernoulli invariant measure

Theorem [Marcovici, T. (2018)]

Every positive-rate PCA that has a Bernoulli invariant measure is ergodic. [Same true for positive-rate IPS!]

Remarks on related results

PCA with Bernoulli invariant measure

Theorem [Marcovici, T. (2018)]

Every positive-rate PCA that has a Bernoulli invariant measure is ergodic. [Same true for positive-rate IPS!]

Remarks on related results

- This simultaneously extends:
  - i) The above result on the ergodicity of surjective CA + noise
  - ii) An earlier partial result by Vasilyev (1978)

PCA with Bernoulli invariant measure

Theorem [Marcovici, T. (2018)]

Every positive-rate PCA that has a Bernoulli invariant measure is ergodic. [Same true for positive-rate IPS!]

Remarks on related results

- This simultaneously extends:
  - i) The above result on the ergodicity of surjective CA + noise
  - ii) An earlier partial result by Vasilyev (1978)
- The entropy method goes back to Boltzmann.
  Its applications for lattice systems were pioneered by:
  - $\longrightarrow$ Holley (1971), Holley and Stroock (1976) for IPS
  - $\longrightarrow$ Kozlov and Vasilyev (1980) for PCA

PCA with Bernoulli invariant measure

Theorem [Marcovici, T. (2018)]

Every positive-rate PCA that has a Bernoulli invariant measure is ergodic. [Same true for positive-rate IPS!]

Remarks on related results

- This simultaneously extends:
  i) The above result on the ergodicity of surjective CA + noise
  ii) An earlier partial result by Vasilyev (1978)
- The entropy method goes back to Boltzmann.
  Its applications for lattice systems were pioneered by:
  ⟶ Holley (1971), Holley and Stroock (1976) for IPS
  ⟶ Kozlov and Vasilyev (1980) for PCA
- With the exception of Holley and Stroock (1976), the entropy method has been limited to shift-invariant starting measures.

[Our result doesn't have this limitation.]

PCA with Bernoulli invariant measure

Theorem [Marcovici, T. (2018)]

Every positive-rate PCA that has a Bernoulli invariant measure is ergodic. [Same true for positive-rate IPS!]

Conjecture 1

Every positive-rate PCA that has a Gibbs invariant measure converges to the set of Gibbs measures with the same specification.

PCA with Bernoulli invariant measure

Theorem [Marcovici, T. (2018)]

Every positive-rate PCA that has a Bernoulli invariant measure is ergodic. [Same true for positive-rate IPS!]

Conjecture 1

Every positive-rate PCA that has a Gibbs invariant measure converges to the set of Gibbs measures with the same specification.

Conjecture 2

Every (local) positive-rate IPS that has a Gibbs invariant measure converges to the set of Gibbs measures with the same specification.

Entropy method for Markov processes

As a warm-up, consider the . . .

Convergence theorem of Markov chains

A finite-state Markov chain is ergodic
provided that it is irreducible and aperiodic.

[Convergence is exponentially fast!]

Different proofs

- Using Perron–Frobenius theory
- Using a coupling argument
- . . .
- Entropy method [Goes back to Boltzmann!]

Entropy method for Markov processes

As a warm-up, consider the . . .

Convergence theorem of Markov chains

A finite-state Markov chain is ergodic
provided that it is irreducible and aperiodic.

[Convergence is exponentially fast!]

Different proofs

- Using Perron–Frobenius theory
- Using a coupling argument
- . . .
- Entropy method [Goes back to Boltzmann!]

Entropy (review)

The entropy of a discrete random variable A taking values in a finite set Σ is

$$H(A) := -\sum_{a\in\Sigma} \mathbb{P}(A=a) \log \mathbb{P}(A=a) \ .$$

It measures the average information content of A.

Entropy (review)

The entropy of a discrete random variable A taking values in a finite set Σ is

$$H(A) := -\sum_{a \in \Sigma} \mathbb{P}(A = a) \log \mathbb{P}(A = a) \ .$$

It measures the average information content of A.

Important properties of entropy

Entropy (review)

The entropy of a discrete random variable A taking values in a finite set Σ is

$$H(A) := -\sum_{a\in\Sigma} \mathbb{P}(A = a) \log \mathbb{P}(A = a) \ .$$

It measures the average information content of A.

Important properties of entropy

- (positivity) $H(A) \geq 0$.

Entropy (review)

The entropy of a discrete random variable A taking values in a finite set Σ is

$$H(A) := -\sum_{a\in\Sigma} \mathbb{P}(A=a)\log \mathbb{P}(A=a)\ .$$

It measures the average information content of A.

Important properties of entropy

- (positivity) $H(A) \geq 0$.
- (capacity) $H(A)$ is maximized precisely when $A \sim \mathrm{unif}(\Sigma)$.

Entropy (review)

The entropy of a discrete random variable A taking values in a finite set Σ is

$$H(A) := -\sum_{a\in\Sigma} \mathbb{P}(A=a)\log \mathbb{P}(A=a)\ .$$

It measures the average information content of A.

Important properties of entropy

- (positivity) $H(A) \geq 0$.
- (capacity) $H(A)$ is maximized precisely when $A \sim \mathrm{unif}(\Sigma)$.
- (chain rule) $H(A,B) = H(A) + H(B \mid A)$.

Entropy (review)

The entropy of a discrete random variable A taking values in a finite set Σ is

$$H(A) := -\sum_{a\in\Sigma} \mathbb{P}(A=a)\log\mathbb{P}(A=a)\ .$$

It measures the average information content of A.

Important properties of entropy

- (positivity) $H(A) \geq 0$.
- (capacity) $H(A)$ is maximized precisely when $A \sim \mathrm{unif}(\Sigma)$.
- (chain rule) $H(A,B) = H(A) + H(B \mid A)$.

[. . . for a suitable definition of conditional entropy $H(B \mid A)$]

Entropy (review)

The entropy of a discrete random variable A taking values in a finite set Σ is

$$H(A) := -\sum_{a \in \Sigma} \mathbb{P}(A = a) \log \mathbb{P}(A = a) \ .$$

It measures the average information content of A.

Important properties of entropy

- (positivity) $H(A) \geq 0$.
- (capacity) $H(A)$ is maximized precisely when $A \sim \mathrm{unif}(\Sigma)$.
- (chain rule) $H(A, B) = H(A) + H(B \,|\, A)$.

  [. . . for a suitable definition of conditional entropy $H(B \,|\, A)$]
- (continuity) $H(A)$ is continuous.

  [. . . as a function of the distribution of A]

Entropy method for finite-state Markov chains

Let $X^0, X^1, \ldots$ be a Markov chain with finite state space Σ and transition matrix $\theta : \Sigma \times \Sigma \to [0, 1]$.

Entropy method for finite-state Markov chains

Let $X^0, X^1, \ldots$ be a Markov chain with finite state space Σ and transition matrix $\theta : \Sigma \times \Sigma \to [0, 1]$.
For simplicity, assume $\mathrm{unif}(\Sigma)$ is stationary.

Entropy method for finite-state Markov chains

Let $X^0, X^1, \ldots$ be a Markov chain with finite state space Σ and transition matrix $\theta : \Sigma \times \Sigma \to [0, 1]$.
For simplicity, assume $\mathrm{unif}(\Sigma)$ is stationary.

Facts

Entropy method for finite-state Markov chains

Let $X^0, X^1, \ldots$ be a Markov chain with finite state space Σ and transition matrix $\theta : \Sigma \times \Sigma \to [0,1]$.
For simplicity, assume $\mathrm{unif}(\Sigma)$ is stationary.

Facts

I) If $A \xrightarrow{\theta} B$, then $H(B) \geq H(A)$.

Entropy method for finite-state Markov chains

Let $X^0, X^1, \ldots$ be a Markov chain with finite state space Σ and transition matrix $\theta : \Sigma \times \Sigma \to [0, 1]$.
For simplicity, assume $\mathrm{unif}(\Sigma)$ is stationary.

Facts

I) If $A \xrightarrow{\theta} B$, then $H(B) \geq H(A)$.

II) Suppose $\theta > 0$.
If $A \xrightarrow{\theta} B$, then $H(B) \geq H(A)$ with equality iff $A \sim \mathrm{unif}(\Sigma)$.

Entropy method for finite-state Markov chains

Let $X^0, X^1, \ldots$ be a Markov chain with finite state space Σ and transition matrix $\theta : \Sigma \times \Sigma \to [0,1]$.
For simplicity, assume $\mathrm{unif}(\Sigma)$ is stationary.

Facts

I) If $A \xrightarrow{\theta} B$, then $H(B) \geq H(A)$.

II) Suppose $\theta > 0$.
If $A \xrightarrow{\theta} B$, then $H(B) \geq H(A)$ with equality iff $A \sim \mathrm{unif}(\Sigma)$.

Proof of the convergence theorem.

We can assume $\theta > 0$.

□

Entropy method for finite-state Markov chains

Let $X^0, X^1, \ldots$ be a Markov chain with finite state space Σ and transition matrix $\theta : \Sigma \times \Sigma \to [0,1]$.
For simplicity, assume $\mathrm{unif}(\Sigma)$ is stationary.

Facts

I) If $A \xrightarrow{\theta} B$, then $H(B) \geq H(A)$.

II) Suppose $\theta > 0$.
If $A \xrightarrow{\theta} B$, then $H(B) \geq H(A)$ with equality iff $A \sim \mathrm{unif}(\Sigma)$.

Proof of the convergence theorem.

We can assume $\theta > 0$.
Since $H(X^0), H(X^1), \ldots$ is increasing and bounded from above, it converges to a value $M \leq \log |\Sigma|$.

□

Entropy method for finite-state Markov chains

Let $X^0, X^1, \ldots$ be a Markov chain with finite state space Σ and transition matrix $\theta : \Sigma \times \Sigma \to [0,1]$.
For simplicity, assume $\mathrm{unif}(\Sigma)$ is stationary.

Facts

I) If $A \xrightarrow{\theta} B$, then $H(B) \geq H(A)$.

II) Suppose $\theta > 0$.
If $A \xrightarrow{\theta} B$, then $H(B) \geq H(A)$ with equality iff $A \sim \mathrm{unif}(\Sigma)$.

Proof of the convergence theorem.

We can assume $\theta > 0$.
Since $H(X^0), H(X^1), \ldots$ is increasing and bounded from above, it converges to a value $M \leq \log|\Sigma|$.
If $M < \log|\Sigma|$, then by compactness and continuity, we can find $A \xrightarrow{\theta} B$ with $H(A) = H(B) < \log|\Sigma|$, a contradiction. □

Entropy method for finite-state Markov chains

Let $X^0, X^1, \ldots$ be a Markov chain with finite state space Σ and transition matrix $\theta : \Sigma \times \Sigma \to [0,1]$.
For simplicity, assume $\mathrm{unif}(\Sigma)$ is stationary.

Facts

I) If $A \xrightarrow{\theta} B$, then $H(B) \geq H(A)$.

II′) Suppose $\theta > 0$. Then, $\exists$ constant $0 < \kappa \leq 1$ s.t.
If $A \xrightarrow{\theta} B$, then

$$H(B) \geq \kappa \log |\Sigma| + (1 - \kappa) H(A) \; .$$

Entropy method for finite-state Markov chains

Let $X^0, X^1, \ldots$ be a Markov chain with finite state space Σ and transition matrix $\theta : \Sigma \times \Sigma \to [0,1]$.
For simplicity, assume $\mathrm{unif}(\Sigma)$ is stationary.

Facts

I) If $A \xrightarrow{\theta} B$, then $H(B) \geq H(A)$.

II′) Suppose $\theta > 0$. Then, $\exists$ constant $0 < \kappa \leq 1$ s.t. If $A \xrightarrow{\theta} B$, then

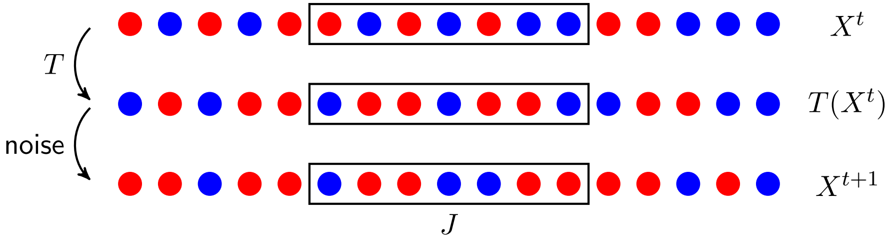
$$H(B) \geq \kappa \log|\Sigma| + (1-\kappa) H(A) \, .$$

Proof of exponential convergence.

It follows from Fact II′ that

$$H(X^t) \geq \log|\Sigma| - \underbrace{(1-\kappa)^t \big[\log|\Sigma| - H(X^0)\big]}_{\to 0} \, . \qquad \square$$

Entropy method for surjective CA + additive noise



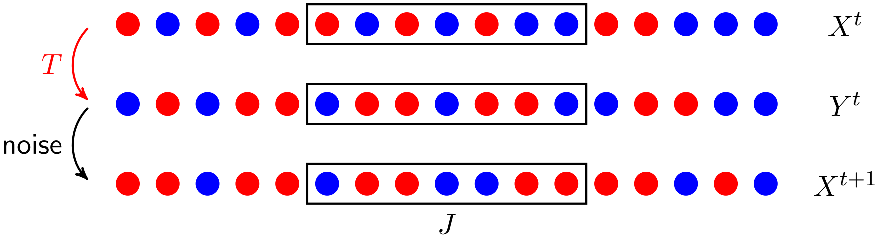
Note

- The uniform Bernoulli measure is stationary.
- In order to prove ergodicity, it is enough to show that for every finite set of sites J,

$$H(X_J^t) \to |J|\,\hbar \qquad \text{as } t \to \infty$$

where $\hbar := \log|\Sigma|$ is the maximum capacity of each site.

Entropy method for surjective CA + additive noise

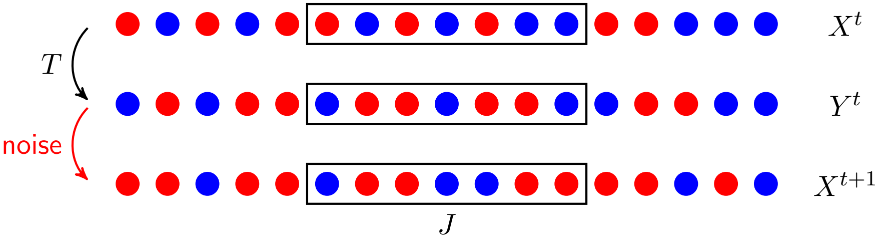


Effect of a surjective CA

A surjective CA does not "erase" entropy, only "diffuses" it:

$$H(Y^t_J) \geq H(X^t_J) - O(|\partial J|)$$

Entropy method for surjective CA + additive noise



Effect of a surjective CA

A surjective CA does not “erase” entropy, only “diffuses” it:

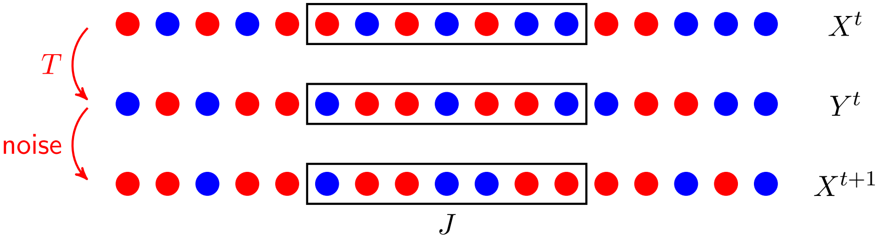
$$H(Y_J^t) \geq H(X_J^t) - O(|\partial J|)$$

Effect of additive noise

Additive noise increases entropy: $\exists$ constant $0 < \kappa \leq 1$ s.t.

$$H(X_J^{t+1}) \geq \kappa\,|J|\,\hbar + (1-\kappa)H(Y_J^t)$$

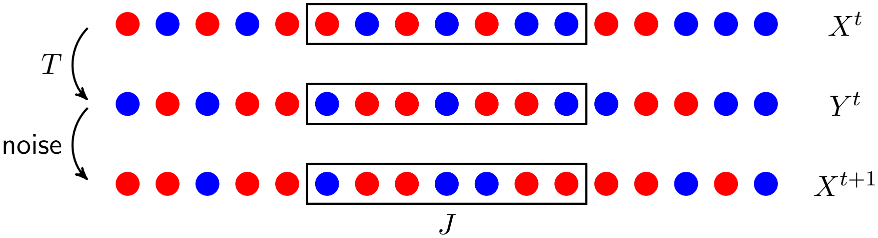
Entropy method for surjective CA + additive noise



Combined effect

$$H(X_J^{t+1}) \geq \kappa\,|J|\,\hbar + (1-\kappa)H(X_J^t) - O(|\partial J|)\ .$$

Entropy method for surjective CA + additive noise



Combined effect

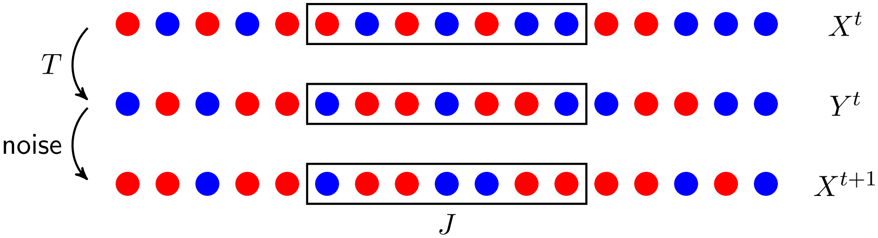
$$H(X_J^{t+1}) \geq \kappa\,|J|\,\hbar + (1-\kappa)H(X_J^t) - O(|\partial J|)\ .$$

which implies

$$H(X_J^t) \geq \left[1 - (1-\kappa)^t\right]|J|\,\hbar - O(|\partial J|)\ .$$

for each $t \geq 0$.

Entropy method for surjective CA + additive noise



Combined effect

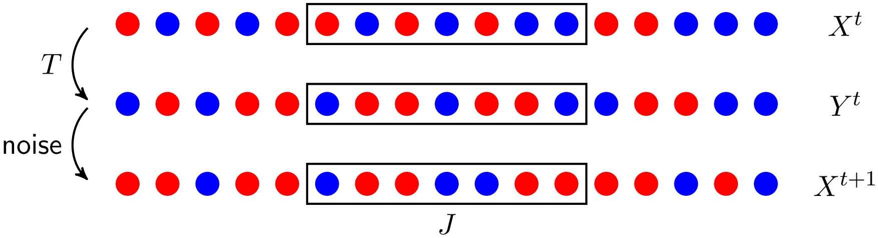
$$H(X_J^{t+1}) \geq \kappa\,|J|\,\hbar + (1-\kappa)H(X_J^t) - O(|\partial J|)\ .$$

which implies

$$H(X_J^t) \geq \underbrace{\left[1-(1-\kappa)^t\right]}_{\to 1} |J|\,\hbar - O(|\partial J|)\ .$$

for each $t \geq 0$.

Entropy method for surjective CA $+$ additive noise



Combined effect

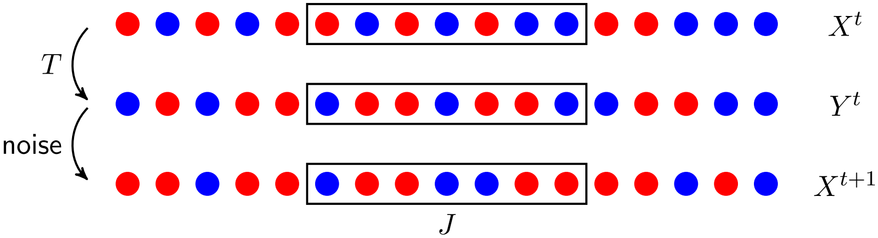
$$H(X_J^{t+1}) \geq \kappa\,|J|\,\hbar + (1-\kappa)H(X_J^t) - O(|\partial J|)\ .$$

which implies

$$H(X_J^t) \geq \underbrace{\left[1-(1-\kappa)^t\right]}_{\to 1} |J|\,\hbar - \overbrace{O(|\partial J|)}^{\text{relatively smaller}}\ .$$

for each $t \geq 0$.

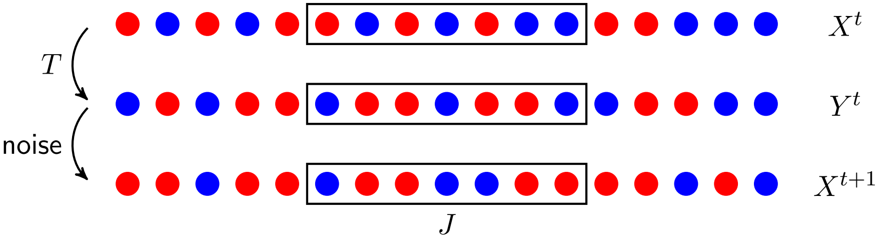
Entropy method for surjective CA + additive noise



Evolution of entropy

$$\boxed{H(X_J^t) \geq \left[1-(1-\kappa)^t\right] |J|\, \hbar - O(|\partial J|)}\,.$$

Entropy method for surjective CA + additive noise



Evolution of entropy

$$\boxed{H(X_J^t) \geq \left[1-(1-\kappa)^t\right] |J|\,\hbar - O(|\partial J|)}\ .$$

A bootstrap lemma

The above implies ergodicity!

Intuitively:
Addition of entropy is much faster than its diffusion.

$\Longrightarrow$ entropy accumulates!

Summary

Key points

- Can a CA perform reliable computation in presence of noise?
- Ergodicity $\equiv$ total forgetfulness
- “Reliable computation requires some degree of irreversibility!”
- Entropy method for Markov processes

Summary

Key points

- Can a CA perform reliable computation in presence of noise?
- Ergodicity $\equiv$ total forgetfulness
- “Reliable computation requires some degree of irreversibility!”
- Entropy method for Markov processes

A related project

- Noise in (combinatorial) tilings
- Stability of quasicrystals at positive temperature
- Using self-organization to generate patterns at the nano-scale

Summary

Key points

- Can a CA perform reliable computation in presence of noise?
- Ergodicity $\equiv$ total forgetfulness
- "Reliable computation requires some degree of irreversibility!"
- Entropy method for Markov processes

A related project

- Noise in (combinatorial) tilings
- Stability of quasicrystals at positive temperature
- Using self-organization to generate patterns at the nano-scale

Thank you for your attention!