Positive-rate PCA with Bernoulli invariant measures

Siamak Taati
Bernoulli Institute, University of Groningen

Joint work with Irène Marcovici

Créteil — April 2019

Main point of this talk

Theorem (discrete time)

Every positive-rate PCA with a Bernoulli invariant measure is ergodic! [. . . with exponentially fast convergence!]

Theorem (continuous time)

Every positive-rate IPS with a Bernoulli invariant measure is ergodic! [. . . with exponentially fast convergence!]

Motivation

- Statistical mechanics
  - Q What about Gibbs/Markov invariant measures?
- Computer science
  - Q Can we do reversible computing with noisy components?

Main point of this talk

Theorem (discrete time)

Every positive-rate PCA with a Bernoulli invariant measure is ergodic! [. . . with exponentially fast convergence!]

Theorem (continuous time)

Every positive-rate IPS with a Bernoulli invariant measure is ergodic! [. . . with exponentially fast convergence!]

Motivation

- Statistical mechanics
  - Ⓠ What about Gibbs/Markov invariant measures?
- Computer science
  - Ⓠ Can we do reversible computing with noisy components?

Main point of this talk

Theorem (discrete time)

Every positive-rate PCA with a Bernoulli invariant measure is ergodic! [... with exponentially fast convergence!]

Theorem (continuous time)

Every positive-rate IPS with a Bernoulli invariant measure is ergodic! [... with exponentially fast convergence!]

Motivation

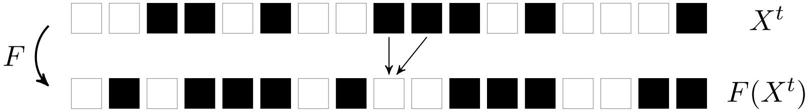
- Statistical mechanics
  - Q What about Gibbs/Markov invariant measures?
- Computer science
  - Q Can we do reversible computing with noisy components?

An example (XOR CA + noise)

X^t

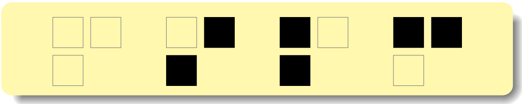
A discrete-time Markov process

An example (XOR CA + noise)



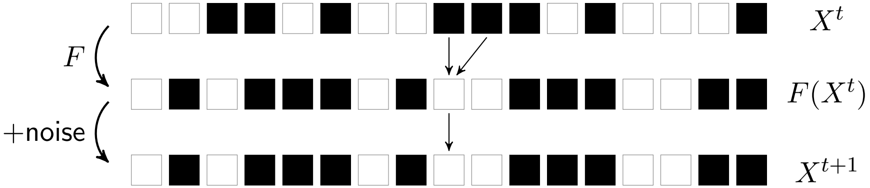
A discrete-time Markov process

I. Apply XOR transformation $x \mapsto F(x)$



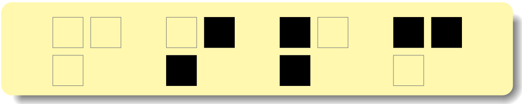
$$F(x)_i = x_i + x_{i+1} \pmod 2$$

An example (XOR CA + noise)



A discrete-time Markov process

I. Apply XOR transformation $x \mapsto F(x)$

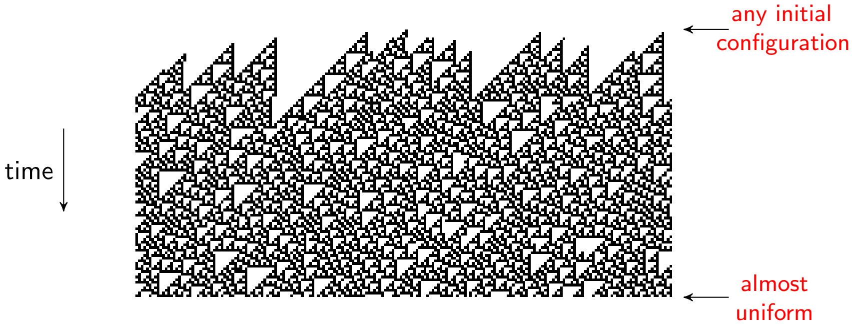


$$F(x)_i = x_i + x_{i+1} \pmod 2$$

II. Add an independent Bernoulli(ε) noise to each site

An example (XOR CA + noise)

Ergodicity (Vaserstein, 1969)



I. The uniform Bernoulli measure is invariant.

II. The Markov process is ergodic:

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

Probabilistic cellular automata (PCA)

X^t

A discrete-time Markov process

- A state of the process is a configuration $x : \mathbb{Z}^d \to \Sigma$.

Probabilistic cellular automata (PCA)

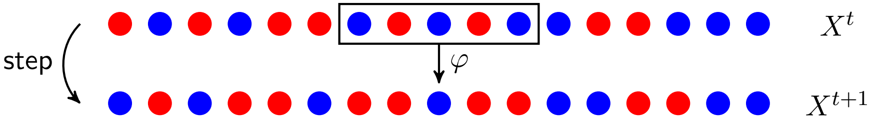
X^t

A discrete-time Markov process

a finite alphabet

- A state of the process is a configuration $x : \mathbb{Z}^d \to \Sigma$.

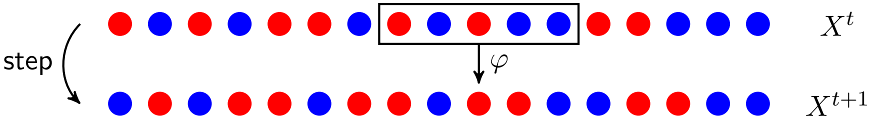
Probabilistic cellular automata (PCA)



A discrete-time Markov process

- A state of the process is a configuration $x : \mathbb{Z}^d \to \Sigma$. (a finite alphabet)
- At each step
  - $\longrightarrow$ Each symbols is updated according to a local rule φ.
  - $\longrightarrow$ The updates at different sites are independent.

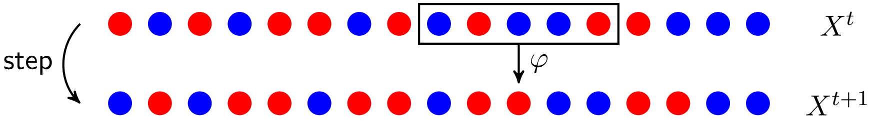
Probabilistic cellular automata (PCA)



A discrete-time Markov process

- A state of the process is a configuration $x : \mathbb{Z}^d \to \Sigma$. (a finite alphabet)
- At each step
  - $\longrightarrow$ Each symbols is updated according to a local rule φ.
  - $\longrightarrow$ The updates at different sites are independent.

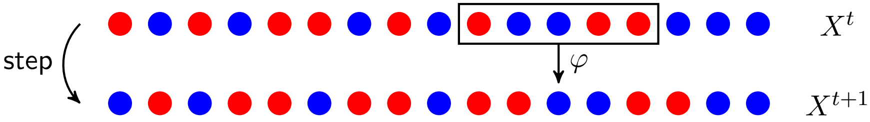
Probabilistic cellular automata (PCA)



A discrete-time Markov process

- A state of the process is a configuration $x : \mathbb{Z}^d \to \Sigma$. (a finite alphabet)
- At each step
  - $\longrightarrow$ Each symbols is updated according to a local rule φ.
  - $\longrightarrow$ The updates at different sites are independent.

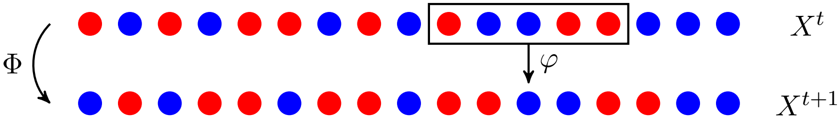
Probabilistic cellular automata (PCA)



A discrete-time Markov process

- A state of the process is a configuration $x : \mathbb{Z}^d \to \Sigma$. (a finite alphabet)
- At each step
  - ⟶ Each symbols is updated according to a local rule φ.
  - ⟶ The updates at different sites are independent.

Probabilistic cellular automata (PCA)



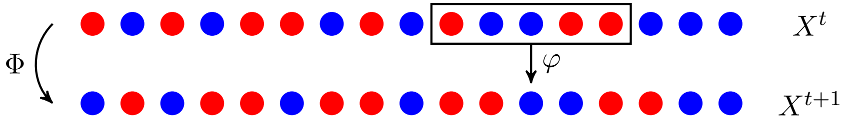
A discrete-time Markov process

- A state of the process is a configuration $x : \mathbb{Z}^d \to \Sigma$. (a finite alphabet)
- At each step
  - $\longrightarrow$ Each symbols is updated according to a local rule φ.
  - $\longrightarrow$ The updates at different sites are independent.

$$X^0, X^1, X^2, \dots$$

$$\mathbb{P}\big(X^{t+1} \in \cdot \,|\, X^0, X^1, \dots, X^t\big) = \Phi(X^t, \cdot) \qquad \text{a.s.}$$

Probabilistic cellular automata (PCA)



Positive-rate PCA

When φ is strictly positive, we say that Φ has positive rates.

Ergodicity

A PCA Φ is ergodic if

I. Φ has a unique invariant measure λ;

II. For every measure μ, we have

$$\mu\Phi^t \to \lambda \qquad \text{as } t \to \infty.$$

(i.e., the distribution of X^t converges to λ irrespective of the choice of X^0.)

Back to the results

Theorem (discrete time)

Every positive-rate PCA with a Bernoulli invariant measure is ergodic! [. . . with exponentially fast convergence!]

Theorem (continuous time)

Every positive-rate IPS with a Bernoulli invariant measure is ergodic! [. . . with exponentially fast convergence!]

Question

1. Which PCA have Bernoulli invariant measures?
2. Which IPS have Bernoulli invariant measures?

Back to the results

Theorem (discrete time)

Every positive-rate PCA with a Bernoulli invariant measure is ergodic! [. . . with exponentially fast convergence!]

Theorem (continuous time)

Every positive-rate IPS with a Bernoulli invariant measure is ergodic! [. . . with exponentially fast convergence!]

Question

1. Which PCA have Bernoulli invariant measures?
2. Which IPS have Bernoulli invariant measures?
   $\longrightarrow$ Ask Jean-François Marckert!

Back to the results

Theorem (discrete time)

Every positive-rate PCA with a Bernoulli invariant measure is ergodic! [. . . with exponentially fast convergence!]

Theorem (continuous time)

Every positive-rate IPS with a Bernoulli invariant measure is ergodic! [. . . with exponentially fast convergence!]

Question

1. Which PCA have Bernoulli invariant measures?
2. Which IPS have Bernoulli invariant measures?
   $\longrightarrow$ Ask Jean-François Marckert!

Which PCA have Bernoulli invariant measures?

Dimension one

- A necessary and sufficient condition
  for PCA with binary alphabet and neighbourhood of size 2
  [Mairesse and Marcovici, 2014]

  $\longrightarrow$ in terms of a system of linear equations

- A sufficient condition for the general case
  [Mairesse and Marcovici, 2014]

Higher dimensions

- Vasilyev's sufficient condition [Vasilyev, 1978]
  [Mityushin and Piatetski-Shapiro]
- Surjective CA $+$ additive noise [Marcovici, Sablik, T., 2017]
- The general identification problem is undecidable!
  [...so no hope of a constructive characterization]

Which PCA have Bernoulli invariant measures?

Dimension one

- A necessary and sufficient condition
  for PCA with binary alphabet and neighbourhood of size 2
  [Mairesse and Marcovici, 2014]

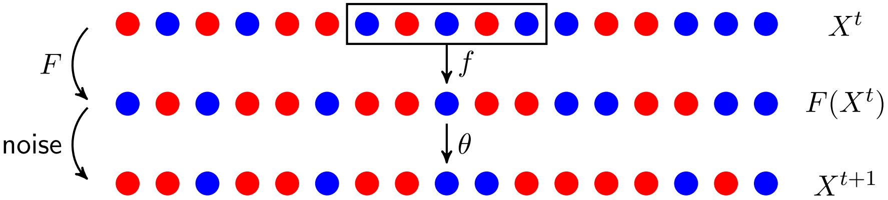
  $\longrightarrow$ in terms of a system of linear equations

- A sufficient condition for the general case
  [Mairesse and Marcovici, 2014]

Higher dimensions

- Vasilyev's sufficient condition [Vasilyev, 1978]
  [Mityushin and Piatetski-Shapiro]
- Surjective CA $+$ additive noise [Marcovici, Sablik, T., 2017]
- The general identification problem is undecidable!
  [...so no hope of a constructive characterization]

Cellular automata (CA) subject to noise



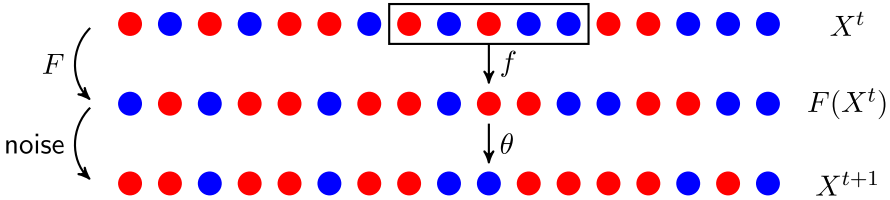
A class of PCA

At each step,

a) first, apply the deterministic CA,

b) then, add zero-range noise independently at each site.

Cellular automata (CA) subject to noise



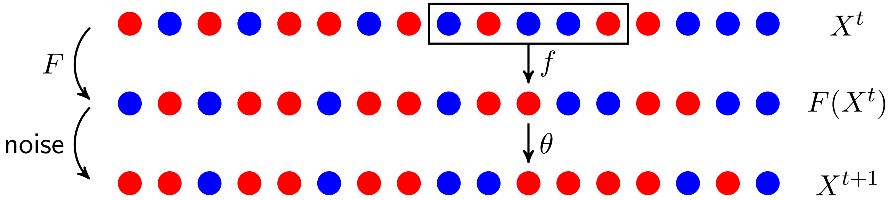
A class of PCA

At each step,

a) first, apply the deterministic CA,

b) then, add zero-range noise independently at each site.

Cellular automata (CA) subject to noise



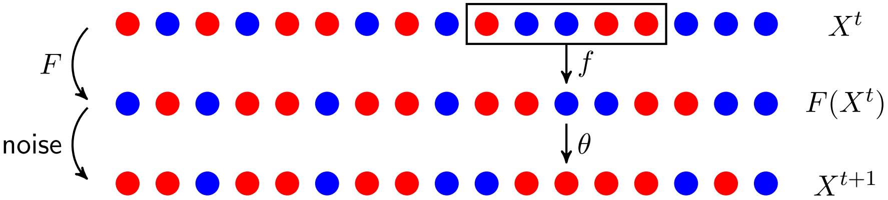
A class of PCA

At each step,

a) first, apply the deterministic CA,

b) then, add zero-range noise independently at each site.

Cellular automata (CA) subject to noise



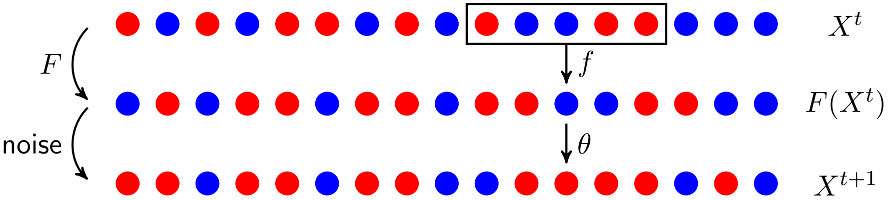
A class of PCA

At each step,

a) first, apply the deterministic CA,

b) then, add zero-range noise independently at each site.

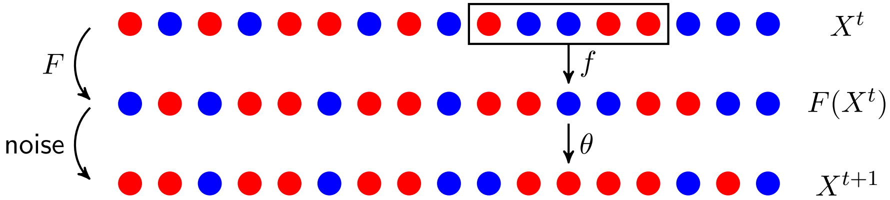
Cellular automata (CA) subject to noise



Motivation (statistical mechanics)

- These are PCA that are close to being deterministic!
- low noise $\longleftrightarrow$ low temperature

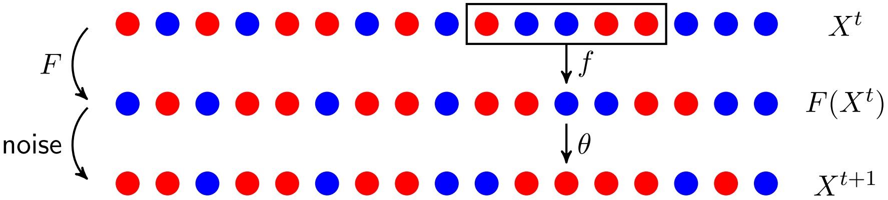
Cellular automata (CA) subject to noise



Motivation (computer science)

- CA is a model of massively parallel computation
- CA + noise models computation with noisy components

Cellular automata (CA) subject to noise

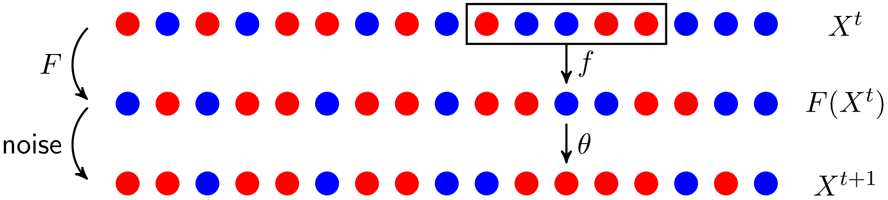


Motivation (computer science)

- CA is a model of massively parallel computation
- CA + noise models computation with noisy components
  - Ⓠ How to do reliable computation with noisy components?

  [Von Neumann (1956), Dobrushin and Ortyukov (1977), ...]

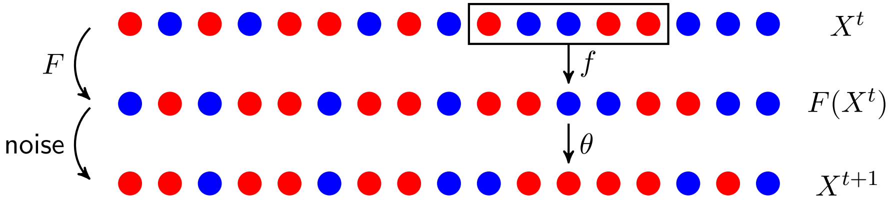
Cellular automata (CA) subject to noise



Motivation (computer science)

- CA is a model of massively parallel computation
- CA + noise models computation with noisy components
  - Ⓠ How to do reliable computation with noisy components?
    [Von Neumann (1956), Dobrushin and Ortyukov (1977), ...]
  - Ⓠ Which CA remain non-ergodic in presence of noise?
    [Toom (1974, 1980), Gács and Reif (1988), Gács (1986, 2001)]

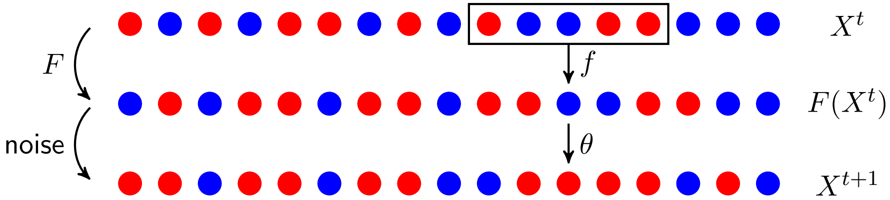
Cellular automata (CA) subject to noise



Observation

The resulting PCA has a Bernoulli invariant measure if both the CA and the noise preserve the same Bernoulli measure.

Cellular automata (CA) subject to noise



Observation

The resulting PCA has a Bernoulli invariant measure if both the CA and the noise preserve the same Bernoulli measure.

Corollary

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

Cellular automata (CA) subject to noise

Corollary

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

Cellular automata (CA) subject to noise

Corollary

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

- Invariance under noise
  $\longrightarrow$ Bernoulli with marginal q is preserved iff $q\theta = q$.

Cellular automata (CA) subject to noise

Corollary

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

- Invariance under noise
  $\longrightarrow$ Bernoulli with marginal q is preserved iff $q\theta = q$.
- Invariance under a CA
  $\longrightarrow$ The CA has to be surjective!

Cellular automata (CA) subject to noise

Corollary

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

- Invariance under noise
  $\longrightarrow$ Bernoulli with marginal q is preserved iff $q\theta = q$.
- Invariance under a CA
  $\longrightarrow$ The CA has to be surjective!
  $\longrightarrow$ Every surjective CA preserves the uniform Bernoulli measure.

Cellular automata (CA) subject to noise

Corollary

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

- Invariance under noise
  $\longrightarrow$ Bernoulli with marginal q is preserved iff $q\theta = q$.
- Invariance under a CA
  $\longrightarrow$ The CA has to be surjective!
  $\longrightarrow$ Every surjective CA preserves the uniform Bernoulli measure.

Corollary

[Marcovici, Sablik, T., 2017; Marcovici and 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!]

Cellular automata (CA) subject to noise

Corollary

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

- Invariance under noise
  $\longrightarrow$ Bernoulli with marginal q is preserved iff $q\theta = q$.
- Invariance under a CA
  $\longrightarrow$ The CA has to be surjective!
  $\longrightarrow$ Every surjective CA preserves the uniform Bernoulli measure.
  $\longrightarrow$ A necessary and sufficient condition in the general case
  [Kari and T., 2015]

Corollary [Marcovici, Sablik, T., 2017; Marcovici and 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

Corollary

[Marcovici, Sablik, T., 2017; Marcovici and 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

Corollary [Marcovici, Sablik, T., 2017; Marcovici and 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

Corollary [Marcovici, Sablik, T., 2017; Marcovici and 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

Corollary [Marcovici, Sablik, T., 2017; Marcovici and 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!]

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!]

Practical implication

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

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

Proof ideas

Corollary [Marcovici, Sablik, T., 2017; Marcovici and 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. □

Proof ideas

Corollary [Marcovici, Sablik, T., 2017; Marcovici and 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 □

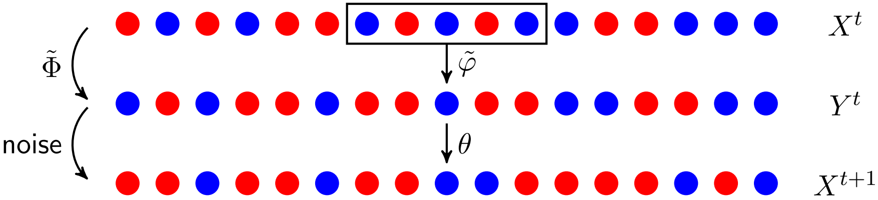
Proof ideas

Theorem (discrete time) [Marcovici and T., 2018]

Every positive-rate PCA with a Bernoulli invariant measure is ergodic! [. . . with exponentially fast convergence!]

Proof idea.

Write the positive-rate PCA Φ as a composition of another PCA $\tilde{\Phi}$ and a zero-range noise both preserving the same Bernoulli measure.



Follow the pattern of the previous proof. $\square$

Entropy method for Markov processes

Some earlier works

- 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

  Some other works:

  $\longrightarrow$ Dawson (1974), Higuchi and Shiga (1974), Sullivan (1976), Moulin Ollagnier and Pinchon (1977), Georgii (1979), Vanheuverzwijn (1981), Künsch (1984), Yaguchi (1990, 1998), Handa (1996), Sakagawa (1999), Dai Pra, Louis and Rœlly (2002), Jahnel and Külske (2015, 2018), . . .
- 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.]

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) $H(A, B) = H(A) + H(B \,|\, A)$.

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) $H(A,B) = H(A) + H(B \mid A)$.

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

[If not, use pressure instead of entropy!]

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.

[If not, use pressure instead of entropy!]

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.

[If not, use pressure instead of entropy!]

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.

[If not, use pressure instead of entropy!]

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.

[If not, use pressure instead of entropy!]

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.

[If not, use pressure instead of entropy!]

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.

[If not, use pressure instead of entropy!]

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.

[If not, use pressure instead of entropy!]

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.

[If not, use pressure instead of entropy!]

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

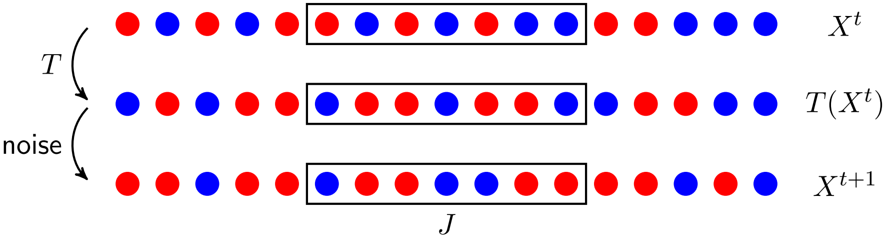
$$H(B) \geq \kappa \log|\Sigma| + (1-\kappa) H(A) \ .$$

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



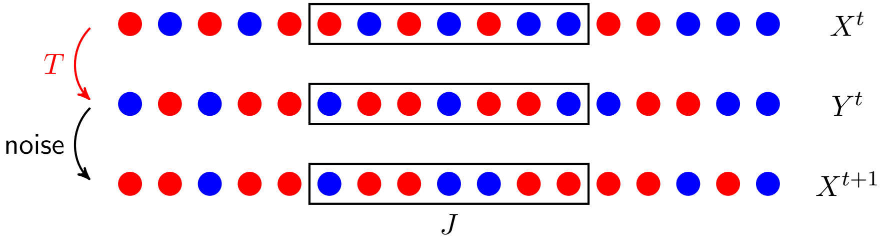
Note

- 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

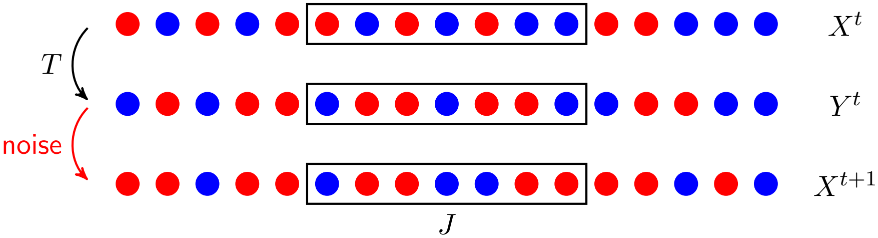


Effect of a surjective CA

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:

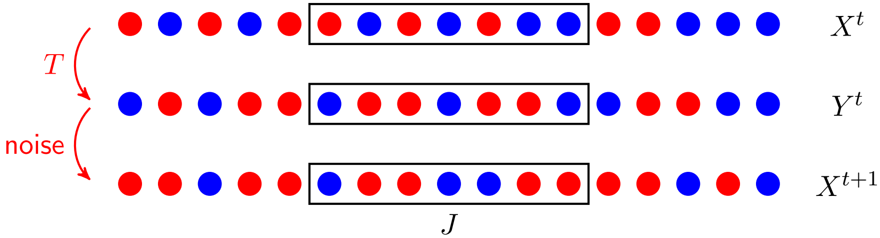
$$H(Y_J^t) \geq H(X_J^t) - O(|\partial J|)$$

Effect of additive noise

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

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

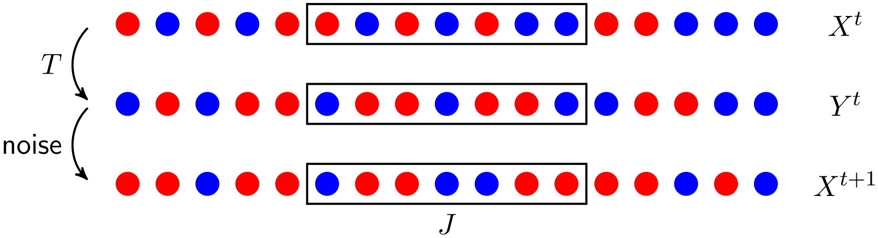
Entropy method for surjective CA + additive noise



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

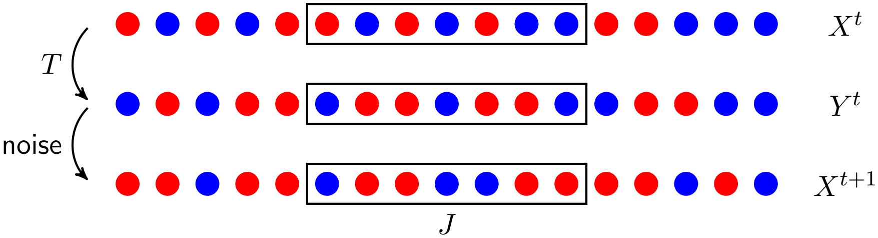
$$H(X_J^{t+1}) \geq \kappa\,|J|\,\hbar + (1-\kappa)H(X_J^t) - O(|\partial J|)\ .$$

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

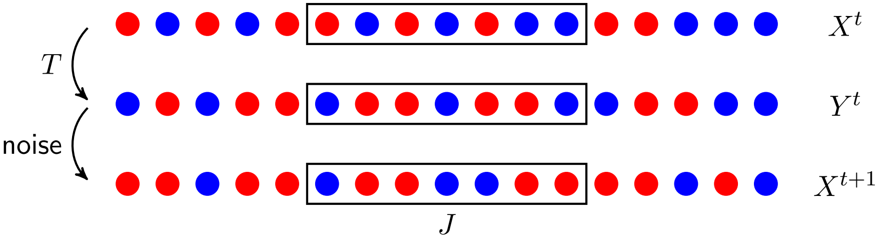
$$H(X_J^{t+1}) \geq \kappa\,|J|\,\hbar + (1-\kappa)H(X_J^t) - O(|\partial J|)\ .$$

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

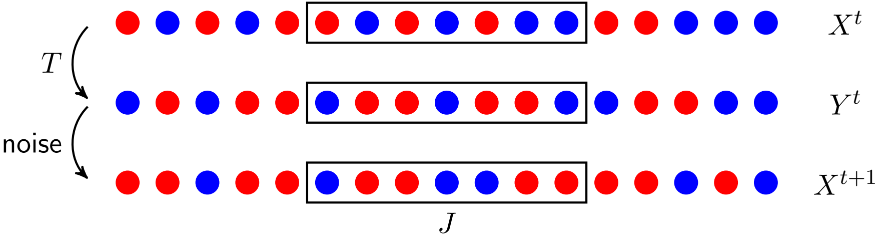
$$H(X_J^{t+1}) \geq \kappa\,|J|\,\hbar + (1-\kappa)H(X_J^t) - O(|\partial J|)\ .$$

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 $t \geq 0$.

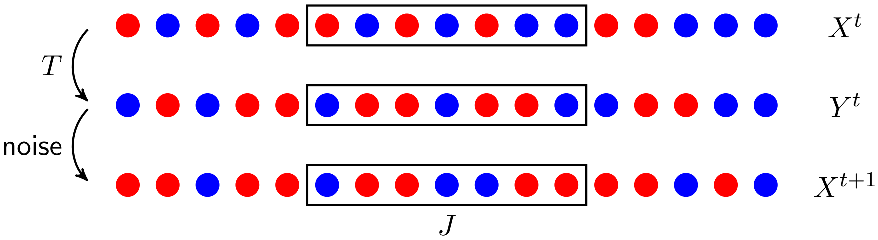
Entropy method for surjective CA + additive noise



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

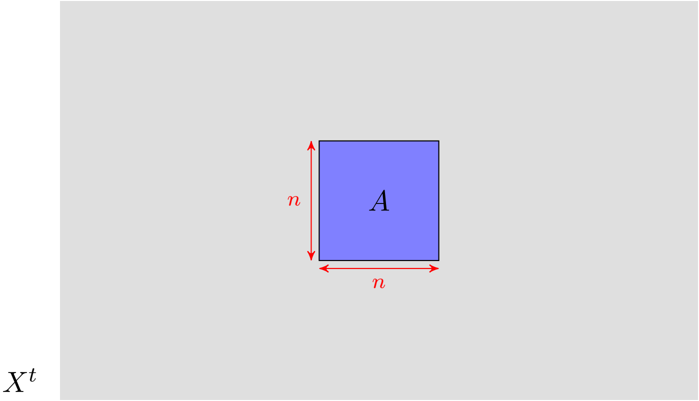
$$\boxed{H(X_J^t) \geq \left[1 - (1-\kappa)^t\right] |J|\, \hbar - O(|\partial J|)} .$$

In particular:

$$\underbrace{|J|\,\hbar - H(X_J^t)}_{\Xi(X_J^t)} \leq O(|\partial J|) \qquad \text{for all } t \geq a \log \frac{|J|}{O(|\partial J|)} + b$$

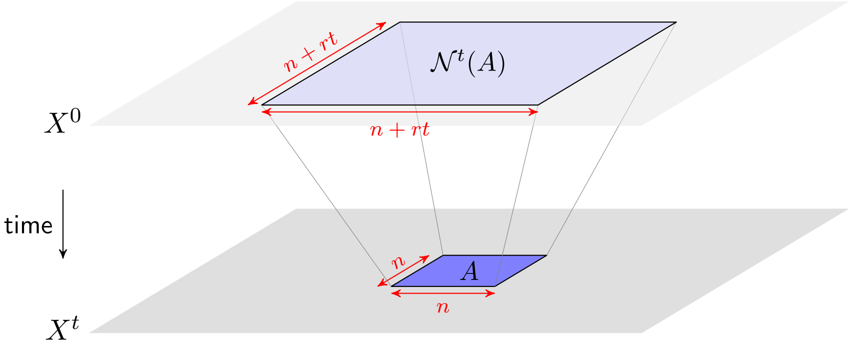
$\Xi(X_J^t)$ ← missing entropy

Bootstrapping



$$\Xi(X^t_A) \le O(n^{d-1}) \qquad \text{for all } t \ge O(\log n)$$

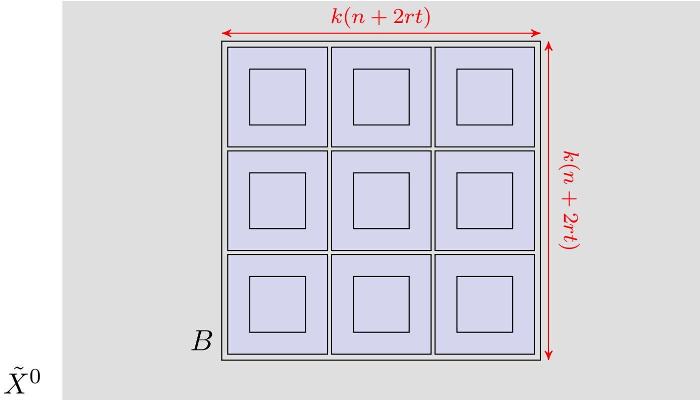
Bootstrapping



Note

The restriction of X^t to A depends only on
the restriction of X^0 to $\mathcal{N}^t(A)$,
where $\mathcal{N} = [-r, r]^d$ is the neighbourhood of the local rule.

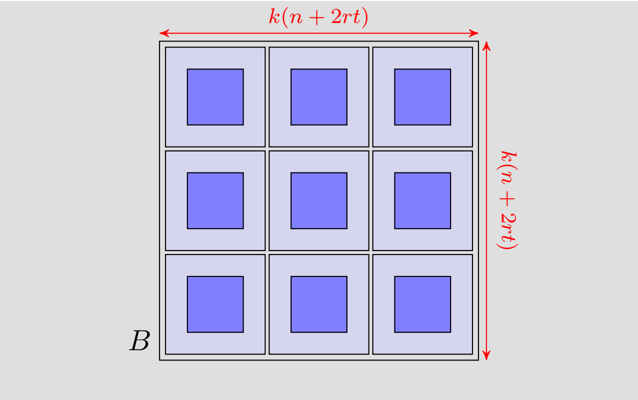
Bootstrapping



Choose $\tilde{X}^0$ such that

$\tilde{X}^0_B$ contains k^d independent copies of $X^0_{\mathcal{N}^t(A)}$.

Bootstrapping



$\tilde{X}^t$

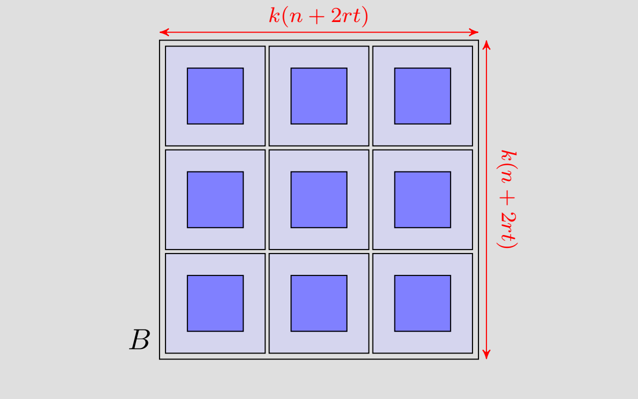
Choose $\tilde{X}^0$ such that

$$\tilde{X}^0_B \text{ contains } k^d \text{ independent copies of } X^0_{\mathcal{N}^t(A)}.$$

Then,

$$\tilde{X}^t \text{ will contain } k^d \text{ independent copies of } X^t_A \text{ inside } B.$$

Bootstrapping

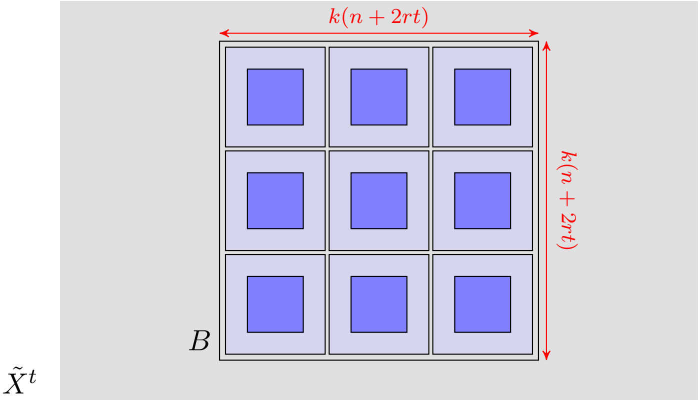


$\tilde{X}^t$

It follows that

$$k^d \, \Xi(X_A^t) \leq \Xi(\tilde{X}_B^t)$$

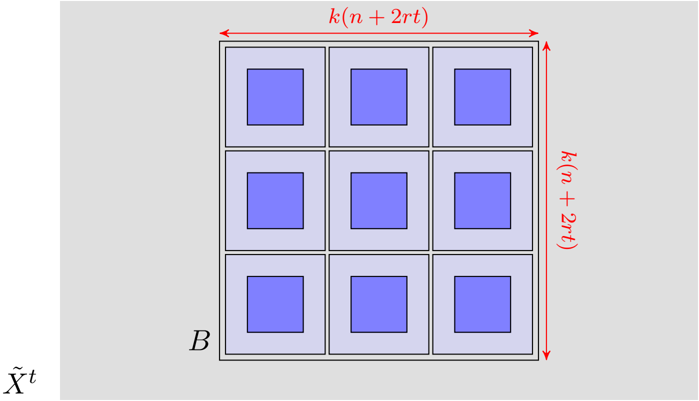
Bootstrapping



It follows that, if $t \geq O\Big(\log[k(n+2rt)]\Big)$,

$$k^d\,\Xi(X_A^t) \leq \Xi(\tilde{X}_B^t) \leq O\Big([k(n+2rt)]^{d-1}\Big)$$

Bootstrapping

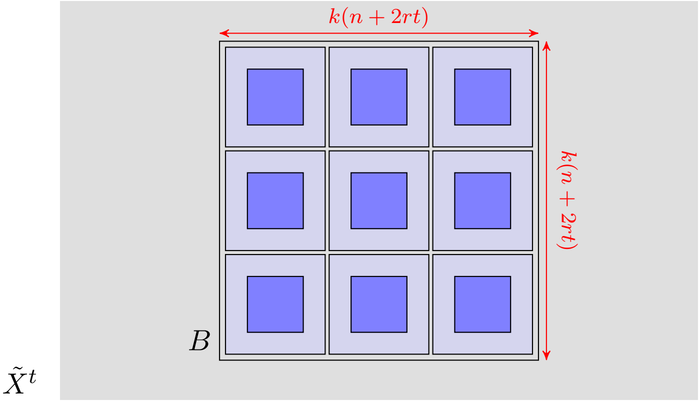


It follows that, if $t \geq O\big(\log[k(n+2rt)]\big)$,

$$k^d\,\Xi(X_A^t) \leq \Xi(\tilde{X}_B^t) \leq O\big([k(n+2rt)]^{d-1}\big)$$

Now, given $t \geq 0$, choose $k := \mathrm{e}^{ct}$ for $c > 0$ small.

Bootstrapping



Conclusion

For every $t \geq 0$ large enough,

$$\Xi(X^t_A) \leq \underbrace{O\left((n + 2rt)^{d-1} e^{-ct}\right)}_{\to 0}$$

□

Gibbs/Markov invariant measures

Conjecture 1 (discrete time)

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

Conjecture 2 (continuous time)

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

Gibbs/Markov invariant measures

Conjecture 1 (discrete time)

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

Conjecture 2 (continuous time)

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

What is known?

- Reversible dynamics
  $\longrightarrow$ Convergence starting from shift-invariant measures
  [Holley, 1971; Kozlov and Vasilyev, 1980]
  $\longrightarrow$ In 1d and 2d, all stationary measures are Gibbs!
  [Holley and Stroock, 1997]
- General dynamics
  $\longrightarrow$ All shift-invariant stationary measures are Gibbs!
  [Künsch, 1984; Dai Pra, Louis and Rœlly, 2002]

Open questions

Conjecture 1 (discrete time)

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

Conjecture 2 (continuous time)

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

Question 1

Can we relax the positive-rate condition?

Question 2

How much irreversibility is necessary for reliable computation in presence of noise?

Thank you for your attention!