American University of Beirut

Asymptotic behaviour of probabilistic cellular automata

Project supported by the University Research Board
(2021–22 academic year)

PI: Siamak Taati

Abstract

Probabilistic cellular automata (PCA) are mathematical models of systems involving a large number of simple components that interact locally and synchronously via probabilistic rules. They are studied in connection with fundamental problems in non-equilibrium statistical physics, and as models of computation in presence of noise. The aim of this project is to tackle a few related open questions that are meant to clarify the concepts of ergodicity and approach to equilibrium in PCA. Tools from probability theory, ergodic theory, symbolic dynamics and analysis will be employed.

1 Introduction

1.1 Motivation and background

One of the challenges in modern mathematical sciences is to extend the high success of mathematics in describing physical reality to complex systems, that is, systems that involve a large number of interacting components. Starting from the middle of the last century, cellular automata and related models emerged simultaneously in several branches of mathematics, physics and computer science, as paradigmatic models of complex systems [16, 29, 27, 17, 8, 4, 30, 15, 14, 1, 20, 19].

A *cellular automaton* (CA) consists of identical components arranged on the sites of an infinite lattice such as $\mathbb{Z}^2$. Each component has a finite number of possible states. The component (often called *cells*) change states synchronously in discrete time steps according to a *local rule* that takes into account the current state of the component and those in its close vicinity. The overall evolution of the system can be viewed as a continuous dynamical system on a compact, metrizable, totally-disconnected topological space [15] or as an abstract model of massively parallel computation [8].

Probabilistic cellular automata (PCA) are variants of (deterministic) CA in which the local rule is allowed to be probabilistic. The evolution of a PCA can be understood as a Markov process with uncountable state space. PCA and their continuous-time counter-parts (known as *interacting particle systems*; IPS) are intensively studied in probability theory to investigate fundamental problems of non-equilibrium statistical mechanics [16, 27, 19] and the problem of reliable computation in the presence of noise [6, 7, 21, 26].

This project aims to answer a few related open questions regarding the basic concept of ergodicity in PCA. A PCA is said to be *ergodic* if its distribution converges to a unique equilibrium distribution irrespective of the initial condition. From the point of view of statistical physics, ergodicity corresponds to lack of phase transitions. From the point of view of computer science, ergodicity means that the system forgets all the information about its initial condition. The questions to be considered in the project are meant to clarify the concept of ergodicity: whether the convergence is always uniform, what are the possible rates of convergence, and when does the uniqueness of equilibrium distribution imply ergodicity. These are far from the most important questions in the field, but are instead non-trivial questions that need to be settled to complete the picture. They appear to have a reasonable level of difficulty for talented graduate/undergraduate students interested in starting research in this field.

1.2 Specific aims of the project

In this project, we plan to clarify the notion of ergodicity in PCA by addressing a few related open questions. These questions can informally be described as follows (see the next section for more details):

- Is the convergence in an ergodic PCA always uniform in the initial condition? Note that in the case of an ergodic finite-state Markov chain, the convergence is always uniform and exponentially fast, but this is not true in the case of countable-state Markov chains, even for those with positive transition probabilities [25, 5, 13].
- Under what conditions can the rate of convergence in an ergodic PCA be sub-exponential? Are all ergodic one-dimensional PCA exponentially ergodic? Slow convergence in an ergodic PCA can be associated to a form of *metastability* [23, 2].
- For one-dimensional PCA with positive transition probabilities, does the uniqueness of the equilibrium measures imply ergodicity (i.e., convergence towards equilibrium)? Without the positivity assumption, this is known to be false [3].

We will be to consider the first and the third questions in connection with analogous questions concerning purely deterministic CA. These analogous questions have been answered in the context of topological and symbolic dynamics and ergodic theory [9, 24, 28].

2 Detailed description or the project

2.1 Terminology and notation

Configuration space. Let Σ be a finite alphabet with at least two elements, and let $d \geq 1$ be an integer. A *configuration* of the d-dimensional lattice $\mathbb{Z}^d$ is a map $x\colon \mathbb{Z}^d \to \Sigma$ that assigns a symbol x_i from Σ to each *site* i of $\mathbb{Z}^d$. The symbol x_i is referred to as the *state* of site i. The restriction of a configuration x to a set $A \subseteq \mathbb{Z}^d$ is denoted by x_A. A *pattern* is a partial configuration $p\colon D \to \Sigma$ where $D \subseteq \mathbb{Z}^d$ is finite. The set D is called the *shape* of the pattern p.

The set $\mathcal{X} \triangleq \Sigma^{\mathbb{Z}^d}$ of all configurations can be given the product topology. This is a compact, metrizable, totally-disconnected space. A *cylinder* set is a set of the form $[p] \triangleq \{x \in \mathcal{X} : x_D = p\}$ where p is a pattern with shape D. The cylinder sets form a basis for the product topology. Given $k \in \mathbb{Z}^d$, the *shift-by-k* operator is the map $\sigma^k\colon \mathcal{X} \to \mathcal{X}$ defined by $\sigma^k(x)_i \triangleq x_{k+i}$ for $i \in \mathbb{Z}^d$.

The space of Borel probability measures on $\mathcal{X}$ will be denoted by $\mathscr{P}(\mathcal{X})$. The Banach space of continuous real-valued functions on $\mathcal{X}$ with the uniform norm is denoted by $C(\mathcal{X})$.

Probabilistic cellular automata. A *probabilistic cellular automaton* (PCA) can be understood as a discrete-time, time- and space-homogeneous Markov process on the configuration space $\mathcal{X}$. At every time step, the state of every site is updated simultaneously and independently according to transition probabilities that take into account the state of the neighbouring sites.

More specifically, let $N \subseteq \mathbb{Z}^d$ be a finite set, indicating the relative position of the neighbouring sites. The *local* transition probabilities are prescribed by a stochastic matrix $\varphi \colon \Sigma^N \times \Sigma \to [0,1]$. This identifies a *global* transition kernel Φ on $\mathcal{X}$ defined by

$$\Phi\big(x, [y_A]\big) \triangleq \prod_{k \in A} \varphi(\sigma^k(x)_N, y_k)$$

for every configuration $x \in \mathcal{X}$ and every finite pattern $y_A \in \Sigma^A$. In words, for each configuration x, $\Phi(x, \cdot)$ is a Borel probability measure on $\mathcal{X}$, describing the distribution of the PCA after one step starting from x. The distribution of each site $k \in \mathbb{Z}^d$ is dictated by the pattern x_{k+N} via the local transition kernel φ. Different sites are updated independently of one another.

A *trajectory* of the PCA is a Markov process with transition kernel Φ, that is, is a sequence $X^0, X^1, X^2, \dots$ of random configurations satisfying

$$\mathbb{P}\left(X^{t+1} \in [y_A] \,\middle|\, X^0, X^1, \dots, X^t\right) = \Phi\big(X^t, [y_A]\big)$$

almost surely for each cylinder $[y_A]$ and every $t \geq 0$.

Note that in our setting, the PCA are assumed to have translation symmetry: $\Phi\big(\sigma^k(x), [p]\big) = \Phi\big(x, \sigma^{-k}([p])\big)$ for every configuration x and cylinder set $[p]$. In other words, a shift in the frame of reference does not change the law of the evolution of the PCA.

Interacting particle systems. The continuous-time variants of PCA are often called *interacting particle systems* (IPS). In an IPS, the updating of the states does not happen simultaneously for different site, but is rather driven by independent Poisson clocks. Namely, there is a Poisson process ξ_k with rate 1 associated to each site k, whose arrival times trigger the update of site k. The Poisson processes for different sites are independent.

While the theories of PCA and IPS were originally developed independently by two different communities [27, 16], there is much similarity between the two models. However, the distinction between synchronous and asynchronous updating leads to notable differences. In this project, the focus will be on the setting of PCA. PCA have the advantage of being less technical: no technical background needed regarding the construction of the process, and everything is discrete.

Markov processes on compact metric spaces. A (probability) *transition kernel* Φ on a measurable space $(\mathcal{X}, \mathfrak{F})$ is a map $\Phi \colon \mathcal{X} \times \mathfrak{F} \to [0,1]$ such that (i) For each $x \in \mathcal{X}$, $\Phi(x, \cdot)$ is a probability measure on $\mathcal{X}$, (ii) For each $E \in \mathfrak{F}$, $\Phi(\cdot, E)$ is measurable. A Markov process with transition kernel Φ is a sequence $X^0, X^1, \dots$ of random points from $\mathcal{X}$ such that

$$\mathbb{P}\left(X^{t+1} \in \cdot \,\middle|\, X^0, X^1, \dots, X^t\right) = \Phi(X^t, \cdot)$$

almost surely for every $t \geq 0$. A *two-sided* Markov process $\dots, X^{-1}, X^0, X^1, \dots$ is defined analogously. A transition kernel Φ naturally defines two linear operators, one on the space $\mathscr{P}(\mathcal{X})$ of probability measures on $\mathcal{X}$ and the other on the space $BM(\mathcal{X})$ of bounded measurable real-valued functions on $\mathcal{X}$:

- A measure $\mu \in \mathscr{P}(\mathcal{X})$ is mapped to a measure $\mu\Phi$ defined with $(\mu\Phi)(E) \triangleq \int \Phi(x, E)\mu(\mathrm{d}x)$. In words, $\mu\Phi$ is the distribution of X^1 if $X^0 \sim \mu$.

- A function $f \in BM(\mathcal{X})$ is mapped to a function Φf defined by $(\Phi f)(x) \triangleq \int f(y)\Phi(x, \mathrm{d}y)$. In words, $(\Phi f)(x)$ is the expected value of $f(X^1)$ given $X^0 = x$.

A transition kernel Φ on a compact metric space $\mathcal{X}$ is said to have the *Feller property* if it satisfies either of the following equivalent conditions:

(a) The map $\mu \mapsto \mu\Phi$ is continuous (w.r.t. the weak topology on $\mathscr{P}(\mathcal{X})$).

(b) If $f \in C(\mathcal{X})$, then $\Phi f \in C(\mathcal{X})$.

For instance, if $F\colon \mathcal{X} \to \mathcal{X}$ is a continuous map, then the kernel Φ defined by $\Phi(x, E) \triangleq \mathbb{1}_E\big(F(x)\big)$ has the Feller property. The global transition kernel of every PCA has the Feller property.

We say that a Feller kernel Φ has *full support*, if for each $x \in \mathcal{X}$, the measure $\Phi(x, \cdot)$ has full support (i.e., assigns positive probabilities to all non-empty open sets). The global transition kernel of every PCA has full support if and only if its local transition probabilities (prescribed in the stochastic matrix φ) are all positive. In this case, the PCA is often said to have *positive rates*.

Invariant measures and ergodicity. A probability measure $\mu \in \mathscr{P}(\mathcal{X})$ is said to be *invariant* under a transition kernel Φ if $\mu\Phi = \mu$. In words, an invariant measure describes an equilibrium distribution of the process described by Φ. A process $X^0, X^1, \ldots$ with transition kernel Φ is *stationary* if the distribution of X^0 is invariant under Φ.

A standard argument shows that every Feller transition kernel on a compact metric space has at least one invariant measure. A Feller kernel Φ on a compact metric space $\mathcal{X}$ is said to be *ergodic* if it has a unique invariant measure π and furthermore, for every initial measure $\mu \in \mathcal{P}(\mathcal{X})$, we have $\mu\Phi^t \to \pi$ weakly as $t \to \infty$. If the convergence is uniform in μ, we say that Φ is *uniformly ergodic*. The following proposition provides some equivalent definitions of uniform ergodicity.

Proposition 2.1 (Equivalent conditions for uniform ergodicity)**.** *Let Φ be a Feller transition kernel on a compact metric space $\mathcal{X}$. The following conditions are equivalent:*

(i) *Φ has a unique invariant measure π, and $\mu\Phi^t \to \pi$ uniformly for $\mu \in \mathscr{P}(\mathcal{X})$.*

(ii) *$\bigcap_{t=0}^{\infty} \mathscr{P}(\mathcal{X})\Phi^t$ is singleton.*

(iii) *The two-sided Markov process with transition kernel Φ is unique in distribution.*

(iv) *For each $f \in C(\mathcal{X})$, the sequence $\Phi^t f$ converges to a constant c_f.*

If the convergence in Condition (iv) above is exponentially fast, we say that Φ is *exponentially ergodic*.

Warning! Unfortunately, the terminology regarding ergodicity is not uniform in the literature. For instance, Liggett [16] uses the same terminology as ours (although he does not talk about uniform ergodicity), whereas Toom et al. [27] call uniformly ergodic kernels simply "ergodic". In addition, in some articles in the literature, "ergodicity" refers to the uniqueness of the invariant measure (no assumption on convergence). Even worse, the term ergodicity is used in ergodic theory in a still different sense.

2.2 Questions to be tackled

In an earlier paper together with Irène Marcovici and Mathieu Sablik [21], we raised the following question:

Question 1. *Is every ergodic PCA uniformly ergodic?*

This question was raised based on the observation that all the examples of ergodic PCA which we are aware of are in fact uniformly ergodic. Independently, Jeff Steif posed an abstract variant of the above question:

Question 1′ (Jeff Steif)**.** *Is every ergodic full-support Feller kernel on a compact metric space uniformly ergodic?*

The two questions are independent of each other, although clearly related. A positive answer to Question 1′ would settle Question 1 in the special case of PCA with positive transition probabilities.

In the same earlier paper [21], we also observed that in most examples of ergodic PCA, the convergence appears to be exponentially fast. We therefore asked:

Question 2. *Find examples of uniformly ergodic PCA which are not exponentially ergodic.*

A third related question was raised by Philippe Chassaing and Jean Mairesse [3]:

Question 3. *Is every positive-rate [one-dimensional] PCA with a unique invariant measure ergodic?*

Without the positivity assumption, the answer is negative: Chassaing and Mairesse provided an example of a PCA which has a unique invariant measure but also admits a periodic trajectory that does not converge to the unique invariant measure. Another example is provided by Ilkka Törmä [28].

3 References

[1] A. Adamatzky, editor. *Cellular Automata*. Encyclopedia of Complexity and Systems Science. Springer, 2009. `doi:10.1007/978-1-4939-8700-9`.

[2] A. Bovier and F. den Hollander. *Metastability: A Potential-Theoretic Approach*. Springer, 2015. `doi:10.1007/978-3-319-24777-9`.

[3] P. Chassaing and J. Mairesse. A non-ergodic probabilistic cellular automaton with a unique invariant measure. *Stochastic Processes and their Applications*, 121(11):2474–2487, 2011. `doi:10.1016/j.spa.2011.06.009`.

[4] B. Chopard and M. Droz. *Cellular Automata Modeling of Physical Systems*. Cambridge University Press, 1998. `doi:10.1017/CBO9780511549755`.

[5] T. Cox. Entrance laws for Markov chains. *The Annals of Probability*, 5(3):533–549, 1977. `doi:10.1214/aop/1176995759`.

[6] P. Gács. Reliable computation with cellular automata. *Journal of Computer and System Sciences*, 32(1):15–78, 1986. `doi:10.1016/0022-0000(86)90002-4`.

[7] P. Gács. Reliable cellular automata with self-organization. *Journal of Statistical Physics*, 103(1–2):45–267, 2001. `doi:10.1023/A:1004823720305`.

[8] M. Garzon. *Models of Massive Parallelism*. Springer, 1995. doi:10.1007/978-3-642-77905-3.

[9] P. Guillon and G. Richard. Nilpotency and limit sets of cellular automata. In *Proceedings of MFCS 2008*, volume 5162 of *LNCS*, pages 375–386. Springer, 2008. doi:10.1007/978-3-540-85238-4_30.

[10] R. Holley. Possible rates of convergence in finite range, attractive spin systems. In R. Durrett, editor, *Particle Systems, Random Media and Large Deviations*, volume 41 of *Contemporary Mathematics*, pages 215–234. American Mathematical Society, 1985. doi:10.1090/conm/041/814713.

[11] R. A. Holley and D. W. Stroock. L_2 theory for the stochastic ising model. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 35:87–101, 1976. doi:10.1007/BF00533313.

[12] B. Jahnel and C. Külske. A class of non-ergodic probabilistic cellular automata with unique invariant measure and quasi-periodic orbit. *Stochastic Processes and their Applications*, 125(6):2427–2450, 2015. doi:10.1016/j.spa.2015.01.006.

[13] S. Kalikow. An entrance law which reaches equilibrium. *The Annals of Probability*, 5(3):467–469, 1977. doi:10.1214/aop/1176995807.

[14] J. Kari. Theory of cellular automata: A survey. *Theoretical Computer Science*, 334:3–33, 2005. doi:10.1016/j.tcs.2004.11.021.

[15] P. Kůrka. *Topological and Symbolic Dynamics*, volume 11 of *Cours Spécialisés*. Société Mathématique de France, 2003.

[16] T. M. Liggett. *Interacting Particle Systems*. Springer, 1985. doi:10.1007/978-1-4613-8542-4.

[17] D. Lind and B. Marcus. *An Introduction to Symbolic Dynamics and Coding*. Cambridge University Press, 1995. doi:10.1017/CBO9780511626302.

[18] P.-Y. Louis. Ergodicity of PCA: Equivalence between spatial and temporal mixing conditions. *Electronic Communications in Probability*, 9:119–131, 2004. doi:10.1214/ECP.v9-1116.

[19] P.-Y. Louis and F. R. Nardi, editors. *Probabilistic Cellular Automata: Theory, Applications and Future Perspectives*. Springer, 2018. doi:10.1007/978-3-319-65558-1.

[20] J. Mairesse and I. Marcovici. Around probabilistic cellular automata. *Theoretical Computer Science*, 559:42–72, 2014. doi:10.1016/j.tcs.2014.09.009.

[21] I. Marcovici, M. Sablik, and S. Taati. Ergodicity of some classes of cellular automata subject to noise. *Electronic Journal of Probability*, 24(41):44pp, 2019. doi:10.1214/19-EJP297.

[22] T. S. Mountford. A coupling of infinite particle systems. *Journal of Mathematics of Kyoto University*, 35(1):43–52, 1995. doi:10.1215/kjm/1250518839.

[23] E. Olivieri and M. E. Vares. *Large Deviations and Metastability*. Cambridge University Press, 2005. doi:10.1017/CBO9780511543272.

[24] V. Salo. On nilpotency and asymptotic nilpotency of cellular automata. *Electronic Proceedings in Theoretical Computer Science*, 90:86–96, 2012. doi:10.4204/EPTCS.90.7.

[25] F. Spitzer. Phase transition in one-dimensional nearest-neighbor systems. *Journal of Functional Analysis*, 20(3):240–255, 1975. doi:10.1016/0022-1236(75)90043-9.

[26] S. Taati. Reversible cellular automata in presence of noise rapidly forget everything. In *Proceedings of AUTOMATA 2021*, volume 90 of *OASIcs*, pages 3:1–3:15. Schloss Dagstuhl, 2021. doi:10.4230/OASIcs.AUTOMATA.2021.3.

[27] A. L. Toom, N. B. Vasilyev, O. N. Stavskaya, L. G. Mityushin, G. L. Kuryumov, and S. A. Pirogov. Discrete local Markov systems. In R. L. Dobrushin, V. I. Kryukov, and A. L. Toom, editors, *Stochastic cellular systems: ergodicity, memory, morphogenesis*. Manchester University Press, 1990.

[28] I. Törmä. A uniquely ergodic cellular automaton. *Journal of Computer and System Sciences*, 81(2):415–442, 2015. doi:10.1016/j.jcss.2014.10.001.

[29] S. Wolfram, editor. *Theory and applications of cellular automata*. World Scientific, 1986.

[30] S. Wolfram. *A New Kind of Science*. Wolfram Media, 2002.