Probabilistic cellular automata with Andrei Toom

Peter Gács

June 27, 2023

Abstract

Andrei Toom, who died in September 2022, contributed some of the most fundamental results on probabilistic cellular automata. We want to acquaint the reader with these and will also try to give the reader a look at the environment in which they were born. Toom was an original and strong personality, and other aspects of his life (education, literature) will also deserve mention.

Andrei Toom, a key developer of the theory of probabilistic cellular automata, died in September 2022. In this space I will describe his most important results, adding some evaluation of their significance. In a closing section I will also refer to Andrei's life and the environment in which he acted.

1 Probabilistic cellular automata

Cellular automata are an attractive mathematical structure: simple to define, they give rise to highly complex behavior. They can model—in a qualitative way—a number of phenomena in physics, biology, society. And they offer a number of natural, and at the same time very challenging, mathematical problems. One way to think of them is as of a discrete generalization of partial differential equations.

A cellular automaton has a finite or countable number of units, its cells, typically arranged on a finite-dimensional lattice, its set of *sites* Λ; they interact with their neighbors in a way that is uniform in space and time. Here Λ will always be either the set $\mathbb{Z}^d$ of points in the d-dimensional space with integer coordinates, or a finite version of it, $\mathbb{Z}_m^d$, where $\mathbb{Z}_m$ is the set of remainders modulo m. Each cell has some *state*, belonging to some finite set $\mathbb{S}$. A *configuration* is a function $\xi : \Lambda \to \mathbb{S}$, that is $\xi(x)$ is the state of the cell sitting at site x. We will also use the following notation: if $L = (x_1, \dots, x_k)$ is a list of sites then

$$\xi(L) = (\xi(x_1), \dots, \xi(x_k)). \tag{1}$$

The system develops in discrete time: our *space-time* is given by $\Lambda \times \mathbb{Z}_+$. A *history* is a function $\eta : \Lambda \times \mathbb{Z}_+ \to \mathbb{S}$, so $\eta(x, t)$ is the state of cell x at time t.

The interaction between cells is local: the state of a cell at site x at time $t + 1$ is only influenced by the state of its *neighbors* $x + u_1, \dots, x + u_n$ at time t. The list $u_1, \dots, u_n$ is the same for all x. In the case of finite space $\Lambda = \mathbb{Z}_m^d$ the addition is taken modulo m (this is called *periodic boundary conditions*). Example for $d = 2$: the list

$$(0, 0), (-1, 0), (1, 0), (0, -1), (0, 1) \tag{2}$$

is called the *von Neumann neighborhood*. The most important part of the definition of cellular automata is the kind of constraints made on the history $\eta(x, t)$. We obtain a *deterministic* cellular automaton by specifying a *transition function* $g : \mathbb{S}^n \to \mathbb{S}$ and requiring that for each time t, each cell get its state at time $t + 1$ from the state of its neighbors at time t as follows:

$$\eta(x, t + 1) = g(\eta(x + u_1, t), \dots, \eta(x + u_n, t)). \tag{3}$$

For example, for $d = 1$ and the neighborhood $\{-1, 0, 1\}$ the requirement would be

$$\eta(x, t + 1) = g(\eta(x - 1, t), \eta(x, t), \eta(x + 1, t)).$$

Another way to express this is to say that there is a function, or "operator", $D : \mathbb{S}^\Lambda \to \mathbb{S}^\Lambda$ taking configurations to configurations such that

$$(D\xi)(x) = g(\xi(x + u_1), \dots, \xi(x + u_n)).$$

A history satisfying the requirement (3) will be called a *trajectory* of the cellular automaton having the transition function g. The initial configuration $\eta(\cdot, 0)$ and the transition D completely determines the trajectory: $\eta(\cdot, t) = D^t \eta(\cdot, 0)$.

Apparently the first deterministic cellular automata were introduced by von Neumann and Ulam in the 1940's with the intention of modelling biological systems. Von Neumann's explorations were cut short by his death, though their record is available in [60]. As for the mathematical theory, soon after their introduction it has been understood that even 1-dimensional cellular automata can simulate Turing machines (just choose the appropriate transition function), therefore it is hard to find interesting mathematical questions about them that are not undecidable. (We will see that Andrei Toom succeeded in this.)

It is natural to generalize cellular automata by allowing the cells to make their transition in a stochastic way. Then $\eta(x,t)$ becomes a random process. The role of the transition function g is taken over by a set of *transition probabilities* $\theta : \mathbb{S}^{n+1} \to [0,1]$. The value $\theta(s \mid r_1, \dots, r_n)$ shows the probability of transitioning into state s at time $t+1$ provided the neighbors at time t have states $r_1, \dots, r_n$ at time t. Of course

$$\sum_{s \in \mathbb{S}} \theta(s \mid r_1, \dots, r_n) = 1.$$

Example 1.1 For $d = 1$, $\mathbb{S} = \{0,1\}$, the neighborhood $\{-1,0,1\}$ and some $\varepsilon \in [0,1]$ let the transition be the following: first take the majority of the states of the three neighbors, then with probability ε change it to the opposite value. Thus

$$\theta(\mathrm{maj}(a,b,c) \mid a,b,c) = 1 - \varepsilon.$$

⌟

Call a *cylinder set* in the set of histories any set of the form

$$\{\eta : \eta(x_1,t_1) = s_1, \dots, \eta(x_k,t_k) = s_k\}.$$

The set of histories will be equipped with the discrete topology: the cylinder sets form a basis of its open sets. The same works for the set of configurations. A *random history* is defined by a probability measure on the Borel sets of the histories. If f is a measurable function over the probability space and μ is a proability measure then we will denote by μf the expected value (integral) of f by μ.

It is assumed that the cells make their choices independently, so the probability distribution $\mathrm{Prob}\{\cdot\}$ satisfies for any t, $x_1, \dots, x_k$:

$$\begin{aligned} &\mathrm{Prob}\{\eta(x_1,t+1) = s_1, \dots, \eta(x_k,t+1) = s_k \mid \eta(\cdot,t'), t' \le t\} \\ &\quad = \theta(s_1 \mid \eta(U(x_1),t)) \cdot \dots \cdot \theta(s_k \mid \eta(U(x_k),t)). \end{aligned} \tag{4}$$

Just like given a transition function of a deterministic cellular automaton and an initial configuration $\eta(\cdot,0)$ the trajectory is completely defined, given the transition probabilities and a probability distribution over the initial configurations $\eta(\cdot,0)$, the random process η is completely defined. (The initial probability distribution can, in particular, be the special one, δ_ξ, concentrated on a single configuration ξ.) In fact, the transition probabilities $\theta(\cdot \mid \cdot)$ define a linear operator P on the set of measurable functions over $\mathbb{S}^\Lambda$ as follows:

$$(Pf)(\eta(\cdot,t)) = \mathsf{E}\{f(\eta(\cdot,t+1)) \mid \eta(\cdot,t)\}$$

where $\mathsf{E}\{\cdot \mid \cdot\}$ is the conditional expected value. Functions f of particular interest are

$$e_{x,a}(\xi) = \delta_{\xi(x),a} = \begin{cases} 1 & \text{if } \xi(x) = a \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

This also defines a linear operator $\mu \mapsto P\mu$ on the space of probability measures over $\mathbb{S}^\Lambda$ via $(P\mu)f = \mu(Pf)$. If we denote the probability distribution of $\eta(\cdot,t)$ by μ_t, then $\mu_{t+1} = P\mu_t$. Clearly, the values $P\delta_\xi$ over all configurations ξ define P completely: without confusion we can write $P\xi = P\delta_\xi$. The transition operator D of a deterministic cellular automaton can be seen as a special case via the following extension of its domain: $D\delta_\xi = \delta_{D\xi}$. The distribution μ will be called *stationary* (or also, *invariant*) if $P\mu = \mu$. It is well-known and not difficult to prove that there is always at least one stationary distribution.

The study of probabilistic cellular automata was started in the early 1960's by a group of mathematicians around Ilya Piatetsky-Shapiro. Toom recalls the circumstances vividly in [53]. The atmosphere of political thaw and wave of optimism, lasting just a few years, spurned new scientific initiatives, partly by opening areas that were off-limits before, being considered bourgeois science (like theoretical biology and "cybernetics"). In the department headed by I. I. Piatetski-Shapiro in I. M. Gelfand's laboratory of applied mathematics at the Moscow State University, a diverse and dynamic group of young scientists explored a variety of potential applications. For the few mathematicians among then (a minority), it took a while to focus on the kind of nontrivial questions where an answer with mathematical rigor could be expected. This is how the model of probabilistic cellular automata emerged; soon it became clear that even the simplest examples and the simplest natural questions about them posed worthy challenges.

In Moscow, two other groups lead by prominent mathematicians worked on related problems of theoretical statistical mechanics: those of Roland L. Dobrushin and Yakov G. Sinai. There was a fertile interaction between these three groups. A closely related model of cellular automata, where time is continuous, has been mostly developed in the United States under the name of *interacting particle systems*, and has the good fortune of a monograph [21] by Thomas Liggett to refer to. The development of the discrete-time and continuous-time models was largely independent of each other, probably somewhat due to the limited possibilities of contact between their researchers. Andrei Toom participated in the writing of at least two useful introductions: [55] and [43].

One of those simplest natural questions is whether the operator P defining some probabilistic cellular automaton has more than one stationary distribution. The transition (4) defines a Markov process. If the space Λ is finite then this is a finite-state Markov chain; the theory of these is well-developed. The operator P can then be represented by a matrix. The chain has a single stationary distribution if and only if every state is reachable from every other state by a sequence of positive-probability transitions. A related question is whether for every initial distribution μ the sequence $P^t\mu$ converges to this stationary distribution. The answer is yes if and only if an additional condition is satisfied: that there is a t such that every element of the matrix P^t is positive (this excludes the possibility of "cycling"). Such a Markov chain is called *ergodic*. An informal way to express the meaning of ergodicity is to say that an ergodic process eventually *forgets everything* about its initial state.

The question is therefore new only for infinite cellular automata—here the state space is uncountable. Ergodicity is defined for this Markov process in the same way: requiring that for every initial distribution μ the sequence $P^t\mu$ converge to one and the same (stationary) distribution. It is natural to choose the sense of convergence here to be that of *weak* convergence, which in this case says that $(P^t\mu)(C)$ converges for every cylinder set C. The problem of giving sufficient criteria of ergodicity has been given attention at the very beginning, also because it is related to the question of phase transition in the models of equilibrium statistical mechanics developed by Dobrushin, Lanford and Ruelle. One of these sufficient conditions can be found in [56].

The first interesting question posed to the group was to find an example of a non-trivial *non-ergodic* probabilistic cellular automaton. Cellular automata are universal computing devices, therefore it is hard to ask non-trivial questions about them—like this one—that are decidable. Kurdyumov showed in [19] (strengthened in [48]) that the ergodicity question for probabilistic cellular automata is undecidable, even if only 0, 1, 1/2 are allowed as local transition probabilities. Toom in [49] proved a similar result for continuous-time systems.

When speaking about non-ergodicity we generally expect that there are at least two stationary distributions. (I am not aware of any non-trivial example of an infinite probabilistic cellular automaton P with only one stationary distribution but where $P^t\mu$ does not always converge to it. "Non-trivial" is, of course, important here, as we could just combine infinitely many identical copies of a finite example.)

2 The Stavskaya model

A candidate example [28] emerged from computer experiments carried out by O. N. Stavskaya. It is one-dimensional, with state space $\mathbb{S} = \{0, 1\}$. I will describe it by switching the states from 0, 1 to 1, 0, to make it analogous to the interacting particle system called the *contact process*. There is a parameter $\varepsilon \in [0, 1]$. The neighborhood is $\{0, 1\}$, and transition probabilities are

$$\theta(1 \mid a, b) = \begin{cases} 0 & \text{if } a + b = 0, \\ 1 - \varepsilon & \text{otherwise.} \end{cases} \tag{6}$$

In words: call a site *healthy* if its state is 0, and *sick* otherwise. A sick site will be healed "spontaneously" with probability ε. A healthy site can only become infected by a sick right neighbor; this will happen with probability $1 - \varepsilon$: one can say that infection happens with certainty but then spontaneous healing is applied to the site immediately. Let $\mathbf{0}$ be the configuration consisting of all 0's and $\mathbf{1}$ the one consisting of all 1's. The distribution $\delta_{\mathbf{0}}$ is trivially stationary. It should be clear that when ε is close to 1 then $P^t\mu$ converges to $\delta_{\mathbf{0}}$ for every μ. Simulations suggested the conjecture that for small ε there are also other stationary distributions because they found that if $\eta(\cdot, 0) = \delta_{\mathbf{1}}$ then the values $\mathrm{Prob}\{\eta(x, t) = 0\}$ appeared to be bounded by a constant $c < 1$. In this case one could say that the system exhibits a *phase transition*. This conjecture was proved in 1968, by M. A. Shnirman in [27] and by Toom in [31].

Toom's proof is simpler and more useful in the long run, as it exploits an important observation: that the random space-time history of this model can be viewed as a percolation. Indeed, consider a graph on the points of the space-time history $\mathbb{Z} \times \mathbb{Z}_+$, where edges are from each point (x, t) to $(x, t + 1)$ and $(x - 1, t + 1)$. A space-time point is *closed* if spontaneous healing happens there: this occurs with probability ε independently for each (x, t). Otherwise it is *open*. Suppose that we start from the initial configuration $\eta(\cdot, 0)$ in which every site is sick. Then $\eta(x, t) = 1$ if and only if there is an open path from time 0 to (x, t).

The proof that for small ε there is an upper bound to $\mathrm{Prob}\{\eta(x, t) = 0\}$ independent of x, t uses an argument that became called the "Peierls argument", or the "contour argument", and it goes by the following steps.

1. Assuming an unfavorable event at space-time point (x, t) find a finite set B of space-time points in the past responsible for it.

2. Find some combinatorial structure $\Gamma(\eta, x, t)$ (typically a "contour") defining a nonempty subset $B'(\Gamma) \subseteq B$.
3. Bound the number of structures Γ with $|B'(\Gamma)| = k$ by c^k for some constant c.

This proves $\text{Prob}\{\eta(x,t) = 0\} \leq \sum_k (\varepsilon c)^k$, which is small for small ε. In a later paper [32], Toom proved that an invariant measure different from $\delta_{\mathbf{1}}$ cannot be very simple (like a Markov chain).

The monotonicity properties based on the partial order of measures introduced by Mityushin in [23] and the above result imply a "phase transition": there is a critical value $\varepsilon^* \in (0,1)$ such that for $\varepsilon < \varepsilon^*$ the system is non-ergodic and for $\varepsilon > \varepsilon^*$ it is ergodic. The result raises another natural question: how many stationary measures are there in the non-ergodic case? Of course, the convex combination of any two stationary measures is also stationary, so what we are really asking about is whether there are only two *extremal* stationary measures. Three papers have proved this, under some conditions. Their methods are different, and each is instructive in its own right. (One of the conditions is the *translation-invariance* of the stationary measures. Formally, for a configuration ξ let us define the configuration $T_v\xi$ translated by the vector v as follows: $(T_v\xi)(x) = \xi(x - v)$. The translation of a function $f : \mathbb{S}^\Lambda \to \mathbb{R}$ is defined then by $(T_v f)(\xi) = f(T_v\xi)$, and a translation of a measure μ by $(T_v\mu)f = \mu(T_v f)$. The measure μ is translation-invariant if $T_v\mu = \mu$ for all v.)

In 1970, Vasiliev in [58] uses the technique of correlation equations developed in statistical physics. It requires ε to be small. In the same year, Vasershtein and Leontovich in [57] prove the result for all $\varepsilon < \varepsilon^*$, using an elegant algebraic representation, for translation-invariant measures. Finally Toom's proof in 1998 in [46], requiring both small ε and translation-invariance, combines three-way coupling with a contour argument in percolation. Each of the three papers generalizes the model in a different way.

3 Positive rates

In the Stavskaya model the invariance of the measure μ_0 is guaranteed by the requirement that in (6) a healthy cell with healthy neighbors never becomes sick: such a transition is *prohibited*. For a number of years, the conjecture was considered that a probabilistic cellular automaton with no prohibited transitions—that is when all local transition probabilities $\theta(\cdot \mid \cdot)$ are positive—is always ergodic. In the corresponding models in continuous time we would talk of positive *rates* in place of local transition probabilities. Toom's best-known result is a refutation of this conjecture: he provided a whole family of non-ergodic probabilistic cellular automata with no prohibited transitions.

Let us call a probabilistic cellular automaton with all-positive local transition probabilities *noisy*. It is not unjustifed to view the goal of defining a non-ergodic cellular automaton that is also noisy as a goal of "error-correction". An ergodic automaton would erase eventually all information about the initial configuration, while a noisy non-ergodic one would preserve some of it, despite the presence of noise. In fact all members of the family of examples defined by Toom can be viewed as follows: a deterministic cellular automaton is given performing some local error-correcting action, but then its action is "perturbed" by changing each transition with some (small but positive) probability to something else. A formal way of looking at this is the following. Recall the definition (5). For some value $\varepsilon \in [0, 1]$, we say that transition operator N is an ε-bounded *noise operator* if for all $x \in \Lambda$,

$$(N\delta_{\xi})e_{x,\xi(x)} > 1 - \varepsilon. \tag{7}$$

In words, it changes the value $\xi(x)$ only with probability $< \varepsilon$. If D is our deterministic error-correcting operation then the full transition would be ND: the action of D is "perturbed" by the noise N. What would be some candidate actions D for local error-correction? Let our automaton have just two local states, $\mathbb{S} = \{0, 1\}$. We would like to see a noisy automaton with the property that, for example, for both $j \in \{0, 1\}$, if $\eta(x, 0) = j$ for all x then $\mathrm{Prob}\{\eta(x, t) \neq j\} < 1/3$ for all x, t. In one dimension let the neighbors be $-1, 0, 1$, and in two dimensions take the neighborhood (2). A good candidate error-correcting action seems to be taking the majority of all neighbor states.

In one dimension, this is doomed to failure. A rigorous (complex) proof of ergodicity is given in 1987 by Gray in [14] (for all monotonic two-state nearest-neighbor rules in one dimension). For continuous time, Gray gave in 1982 a simpler but still non-trivial proof in [13]. Here is an informal argument. In a cellular automaton with $0 \in \mathbb{S}$, in some configuration ξ let us call an *island* a finite set S such that $\xi(x) = 0$ if and only if $x \notin S$. Let $j = 0$, so we are starting from a "see of zeros". The noise will, occasionally, create a large (say, size 10) island. The local majority vote rule will not be too helpful in eliminating the island. Indeed, it cannot do much about the ends. Noise will make them fluctuate, essentially perform a random walk. The size of the island, the difference of these random walks, is also a random walk. It will eventually return to 0, but only in infinite expected time; in the meantime many other islands arise.

In two dimensions the situation is better, but not ideal. In the absence of noise the majority rule would not shrink an island, say a square of size n, at all. Simulations (and informal arguments) show that in "unbiased" noise it will be eliminated in about $O(n^2)$ steps. Even this is not proved, but here is a non-ergodic noisy version, see for example [17], and the more elaborate [20]. Let $\mathbb{S} = \{-1, 1\}$. With $U = \{(0,1), (0,-1), (1,0), (-1,0)\}$, let us define the process P by the following relation: If s is the sum of $\eta(x,t)$ over the four neighbors $\{x + v : v \in U\}$ of x then

$$\text{Prob}\{\eta(x, t+1) = -1\} = \varepsilon^s \text{Prob}\{\eta(x, t+1) = 1\}.$$

So for example if there are 3 neighbors with $+1$ and 1 neighbor with -1 then the next state is ε^{-2} times more likely to become $+1$ than -1. For $j = 0, 1$ let for $x = (x_1, x_2)$:

$$\Lambda_j = \{(x,t) \in \Lambda \times \mathbb{Z}_+ : x_1 + x_2 + t \equiv j \pmod 2\}.$$

We turn Λ_j into a graph, connecting (x,t) and $(y, t+1)$ by an edge whenever x and y differ by 1 in just one coordinate. Then the process $\{\eta(x,t) : (x,t) \in \Lambda_0\}$ is a Gibbs state of the Ising model of equilibrium statistical mechanics (of course the same holds for Λ_1). It is known that for small ε (corresponding to "low temperature") the Ising model has more than one Gibbs state, which makes for several invariant distributions for the process P. This is a very delicate process, though. If the transition probabilities are changed ever so slightly to prefer the 1's over the -1's, the process becomes ergodic. In the continuous-time models called interacting particle systems, the corresponding model is called the *stochastic Ising model*, see [21].

Simulation of the perturbed majority rule shows a similar behavior. If the noise prefers the 1's over the 0's, then a large island will *grow*! The $1/n$ speed of shrinking provided (on average) by the majority rule is overpowered by the constant speed of growth. "Unbiased" is understood here with respect to the states 0 and 1. More precisely, let us define the flipping operation F for a function f over $\{0,1\}^\Lambda$ as follows:

$$(Ff)(\xi) = f(1 - \xi). \tag{8}$$

We say that the noise operator N is *unbiased* with respect to flips if $NF = FN$.

Let us now turn to Toom's best-known example, in two dimensions, called the *Toom rule*:

$$(D_{\text{Toom}}\xi)(0,0) = \text{maj}(\xi(0,0), \xi(0,1), \xi(1,0)).$$

The transition rule is a majority vote over three neighbors: north, east, self—which is then perturbed with some small probability ε. The novelty is just

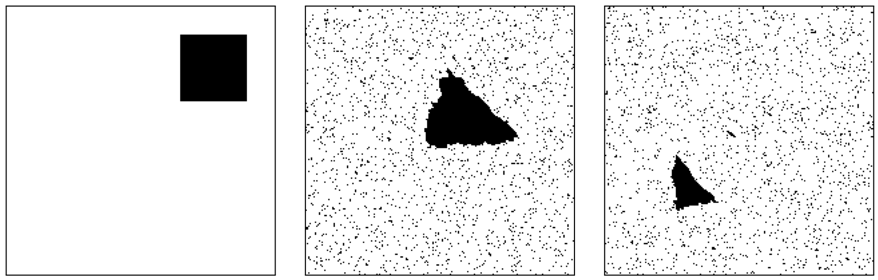

Figure 1: An island shrinking under Toom's rule (simulation). The faults have probability 0.04 and are completely biased towards the 1's.

that the neighborhood over which the majority is taken is *not (central-) symmetric*—but this makes a big difference. Imagine an island. Enclose it into a triangle with vertices $(a, b), (a + c, b), (a, b + c)$. If $c \geq 1$ then without noise, one application of the Toom rule will squeeze the island into the triangle $(a, b), (a + c - 1, b), (a, b + c - 1)$. Thus, the island will shrink with *constant speed*. Small noise *(even if it is biased!)* can slow down this shrinking, but still leaves its speed constant (see Figure 1). This last argument is not a proof; we will say more about the (not easy) proof in the next section.

How small must be the probability ε for non-ergodicity? The existing proofs don't give an explicit bound, though for example (a rather bad) one can easily be computed from the version of the proof in [7]. Simulations suggest that the upper bound 0.06 is sufficient.

4 Reliable computation

We pointed above to a connection of non-ergodicity with error-correction. In its original form, non-ergodicity is only asking to safeguard at least one bit of information—a minimal form of information conservation in the presence of local faults. But the solutions found have much wider application. Von Neumann in [59] addressed the question of reliable computation with unreliable components. His computation model is now called a Boolean circuit (for example with logic gates AND and NOT), computing a Boolean function with, say, a one-bit output. For a sufficiently small ε, for any Boolean circuit C of size N he constructed another one, C', of size $O(N \log N)$ that computes the same output as C with probability $1 - O(\varepsilon)$ even though each gate of C' is allowed to malfunction with probability $< \varepsilon$ (independently from the others). The construction multiplies each wire and each gate some $O(\log N)$ times: in the absence of faults, each wire in a bundle carries the same bit. The key addition is to insert into each such wire bundle a little circuit called the *restoring organ*. The role of this organ is that if, say, the fraction of its input wires carrying faulty information is $< 5\varepsilon$ then after its application it will be reduced to $< 2\varepsilon$: so it is about protecting one bit of information! Von Neumann used random permutations for the restoring organ; for constructive (but still not local) solutions, see the survey [25].

As a model of computations, Turing machines are in several ways better than Boolean circuits. They were introduced in the 1930's to formalize the theory of computability, and it has long been accepted that every function (with, say, strings in some alphabet as input and output) computable in an intuitive sense can also be computed on an appropriate Turing machine. There is a *universal* Turing machine, one that can simulate every other Turing machine.

Every Turing machine (in particular also the universal ones) can be simulated by an appropriate one-dimensional cellular automaton, so cellular automata can also serve well as a model of arbitrary computation. When asking for a reliable computer, it makes sense therefore to ask for a (noisy) probabilistic cellular automaton capable of performing arbitrary computations (defined by its input). Our 1988 construction with John Reif in [10], built on Toom's rule, does this. It takes an arbitrary (deterministic) one-dimensional cellular automaton A (say, a universal one), say with neighborhood $\{-1, 0, 1\}$, and builds a *three-dimensional* noisy cellular automaton A' simulating A. Let $\zeta(x, t)$ be a trajectory of A. The intended history $\eta(x, y, z, t)$ of A' would be $\eta(x, y, z, t) = \zeta(x, t)$: so ideally in A' each cell of the whole plane $\{x\} \times \mathbb{Z}^2$ has the same state as the symbol at position x of A. If D_A is the transition rule of A then the rule of A' says: in order to obtain your state at time $t+1$, first apply the Toom rule in each plane defined by fixing the first coordinate—call this $D_{\text{Toom},2,3}$. Then, apply rule of A on each line obtained by fixing the second and third coordinates—call this $D_{A,1}$. Finally, apply an ε-bounded noise operator N, so the complete transition of A' is $N D_{A,1} D_{\text{Toom},2,3}$.

The reliability of the automaton A' is proved in essentially the same way as the non-ergodicity of the (perturbed) Toom rule; see below for remarks on the proof.

It seems unrealistic for the automaton A' to store each symbol of information $\zeta(x, t)$ with *infinite redundancy*, as the array of all values $\eta(x, y, t)$. But if it is known that the computation of automaton A uses only S cells and runs for only T steps then it can be run on the space $\Lambda = \mathbb{Z}_S$, and can be simulated reliably (failing only with probability $O(\varepsilon)$) on the space $\Lambda' = \mathbb{Z}_N \times \mathbb{Z}_R^2$ where $R = O(\log(ST))$: so repeating each symbol only $O(\log^2(ST))$ times. See the paper [1] by Berman and Simon, which improves on [10].

5 Eroders

Toom called a deterministic cellular automaton an *eroder* if it eliminates any island. From now on we will tacitly require $\mathbb{S} = \{0, 1\}$ and that the rule the rule be *monotonic*. The Toom rule D_{Toom} leads to non-ergodicity because both it and its *dual* are eroders, where the dual $FD_{\text{Toom}}F$ is defined using (8) (it happens that D_{Toom} is self-dual).

What rules are eroders? Mityushin and Toom in [29] proved that it is undecidable about an arbitrary one-dimensional monotonic binary cellular automaton whether it will erase the island $\ldots 001100 \ldots$. Mityushin and Toom also define there a one-dimensional monotonic binary cellular automaton A for which given an island it is undecidable whether A will erase it.

These statements make us appreciate Toom's elegant formula of [34] given below—in any dimension d—that decides about an arbitrary monotonic binary cellular automaton A whether it is an eroder. Let $\Lambda_A = \mathbb{Z}^d$, with neighborhood $U_A \subset \Lambda$, and transition rule $g_A : \mathbb{S}^{U_A} \to \mathbb{S}$. Recall that $\mathbb{S} = \{0, 1\}$. We call a subset $S \subseteq U_A$ a *null set* of A if whenever in a configuration ξ we have $\xi(x) = 0$ for all $x \in S$, this implies $(g_A\xi)(0) = 0$. For example, in the Toom rule, the minimal null subsets are $\{(0, 0), (0, 1)\}$, $\{(0, 0), (1, 0)\}$, $\{(1, 0), (1, 1)\}$.

Theorem 1 *For a given cellular automaton A let*

$$\sigma(A) = \bigcap_{\textit{null sets } S} \operatorname{conv}(S)$$

where $\operatorname{conv}(S)$ *is the convex hull taken in the Euclidean space $\mathbb{R}^d$. Then A is an eroder if and only if $\sigma(A)$ is empty.*

It is important here that the the operation above happens in $\mathbb{R}^d$. Consider for example the automaton A in two dimensions with the transition

$$(D_A\xi)(0, 0) = (\xi(0, 0) \vee \xi(1, 1)) \wedge (\xi(0, 1) \vee \xi(1, 0)).$$

The minimal nullsets are $\{(0, 0), (1, 1)\}$ and $\{(0, 1), (1, 0)\}$, so $\sigma(A) = \{(1/2, 1/2)\}$, which is not in $\mathbb{Z}^2$. By Toom's theorem, this is not an eroder.

How about one dimension? There are monotonic binary eroders in one dimension as well; however, none whose dual would also be an eroder.

By the theorem of Toom given below, an eroder remains an eroder also in noise. It turns out that for this theorem we don't even need space-time uniformity or a strict Markov property of the noise, so we can relax the requirement (7) further, using a formulation borrowed from [38]. Let A be an arbitrary deterministic cellular automaton, and $\eta(x,t)$ a random history of A. Let the random *fault set* $\mathcal{E} \subseteq \Lambda \times \mathbb{Z}_+$ be the subset of the space-time in which η violates the transition rule g_A. We will say that the distribution of η is an ε*-perturbation* of A if for every finite subset S of space-time,

$$\mathrm{Prob}\{S \subseteq \mathcal{E}\} \le \varepsilon^{|S|}. \tag{9}$$

A trajectory $\zeta(x,t)$ of automaton A is called *stable* if there is a function $\delta(\varepsilon)$ with $\delta(\varepsilon) \to 0$ as $\varepsilon \to 0$ such that for all ε-perturbations η of A starting from the same initial configuration: $\eta(\cdot,0) = \zeta(\cdot,0)$, and for all x,t we have $\mathrm{Prob}\{\eta(x,t) \ne \zeta(x,t)\} < \delta(\varepsilon)$. Call an eroder *stable* if the history ζ with $\zeta(x,t) = 0$ for all x,t is a stable trajectory for it.

Theorem 2 *Every binary eroder is stable.*

As mentioned above, this theorem implies that the perturbed Toom rule is non-ergodic.

In continuous time, a simple characterization of eroders similar to Theorem 1 is not available. But Gray gave a sufficient condition, and proved in [15] that transition rates corresponding to the Toom rule in continuous time are non-ergodic. This result is also robust with respect to the kind of faults allowed: they can be biased.

6 On the proofs

Toom proved Theorem 2 using a kind of "contour argument" as outlined in Section 2. The paper [38] contains detailed proofs of both theorems 1 and 2, in a somewhat more general setting. The part of the contour argument defining an appropriate structure $\Gamma(\eta,x,t)$ is quite sophisticated. A somewhat simplified version of this proof of just the stability of Toom's rule, following [1], can be found in [7].

Toom's first examples of noisy non-ergodic cellular automata in [33] used some special two-dimensional eroders other than the north-east-center voting; their dual is also an eroder. The contour argument in the proof of their stability is simpler, the structure Γ in question is indeed only a kind of one-dimensional contour. Here is one of these eroders, call it A, with transition operator

$$(D_A\xi)(0,0) = (\xi(0,0) \vee \xi(1,0)) \wedge (\xi(1,0) \vee \xi(1,1)).$$

Given any island, this rule squeezes it in every step into a narrower and narrower horizontal stripe, eventually erasing it. Its dual will squeeze an island in every step into a narrower and narrower vertical stripe.

Our paper [10] contains a completely different kind of proof, based on a hierarchical structure discoverable in the set of independent faults. It is in some ways more intuitive as it is based directly on the picture of shrinking triangles (see Figure 1), but leads to somewhat worse estimates. The paper [2] by Bramson and Gray develops this method more systematically, introducing a hierarchy of random processes built up from each other by a "decoding" operation. Its method is also used in the proof of the continuous-time "Toom theorem" in [15].

7 One dimension

With Kurdyumov and Levin we defined a simple one-dimensional (not monotonic) two-state eroder (call it K) in [9] (it is sometimes referred to as the GKL rule). Its dual is also an eroder; more precisely, recall the flip operation F in (8), and define the *reflection* operation R defined as $(R\xi)(x) = \xi(-x)$ which commutes with flip. Then $KFR = RFK$. Alas, this rule (along with other simple ones similar to it) is not stable, at least not in strongly biased noise, as shown in Kihong Park's thesis [24].

As said above, there are no eroders in one dimension whose dual is also an eroder. More generally, as also said, monotonic two-state one-dimensional probabilistic cellular automata are ergodic: as proved by Lawrence Gray for continuous time in [13], and for discrete time in [14]. Based on this and also on the fact that in one-dimensional lattice equilibrium systems (like the Ising model) there is no phase transition, and that therefore the corresponding reversible Markov processes are ergodic, the probabilistic cellular automata community formulated a hypothesis called the "positive rate conjecture", saying that all noisy one-dimensional cellular automata are ergodic.

Georgii Kurdyumov outlined in [18] a one-dimensional cellular automaton that should be non-ergodic. Its cells were meant to implement a hierarchical structure dealing with larger and larger groups of faults. However, details of the proposal did not follow. I worked these out in [6], defining a cellular automaton that simulates a similar one; this gives rise to an infinite hierarchy of more and more reliable cellular automata all of which but the first one "live" in simulation. A much more structured—and longer—paper [8] extended the result also to continuous time. Its method uses the idea of Bramson and Gray in [2] of a hierarchy of random processes derived from each other by a "decoding" operation.

These systems refute the positive rate conjecture. As they are not monotonic and the corresponding Markov processes are not reversible either, they don't contradict the motivating examples of the conjecture.

The one-dimensional constructions and proofs are very complex; the cells have a very large number of states. (One can say that as in one dimension the geometry does not help, all error-correction must rely on "organization".) The challenge to find simpler examples is still standing.

8 Multi-level eroders

After the beautiful characterization of two-state monotonic eroders in Theorem 1, Toom asked the question whether a similar characterization exists also when the set of states is a finite ordered set, say $\{0, 1, \ldots, n\}$. The situation turns out to be more complicated. Galperin characterized in [12] the one-dimensional eroders in terms of the running speeds of the ends of islands of various levels. But Toom showed in [35] that some of these eroders, even with just three levels, are not stable, so the analogue of Theorem 2 does not hold. With Ilkka Törmä in [11] we characterized all *stable* one-dimensional multilevel eroders.

The question of characterizing multi-level eroders is still open in dimensions greater than one. In one dimension even an unstable eroder erases islands in linear time. In two dimensions, this is not always true. Toom's paper [3] with his student Lima gives an example of a two-dimensional three-level eroder that erodes some islands only in quadratic time. In one-sided noise, it becomes ergodic. This example can easily be modified to an eroder that erases some islands only in exponential time. At this point it is not known whether the question of which three-level monotonic two-dimensional cellular automata are eroders is decidable at all. On the other hand, our method used in [11] seems generalizable to several dimensions, using Toom's substantial sharpening in [39] of his main stability result. Therefore it is likely decidable about a monotonic multi-level cellular automaton whether it is a stable eroder, while it might remain undecidable whether it is just an eroder.

9 Other work on cellular automata

Toom was one of the leading figures in Russian research on biologically and physically inspired systems with local interaction throughout the 1970's and 80's. He was one of the organizers of a regular conference in the biological research center in Pushchino, and one of the editors, mostly along with R. L. Dobrushin and V. I. Kryukov, of the proceedings. Unfortunately, most of these are in Russian, though some resulted also in an English-language publication: see Selecta Mathematica Sovietica and [4]. The book [55] is an important and hard-to-access survey of research on various aspects of probabilistic cellular automata.

The account in the present section of Toom's later work is incomplete, reflecting my own interests. The technique developed for proving the main stable eroder theorem appears in several later publications. The work [36], generalizes the cellular automata model in an unexpected direction: space-time is a multi-dimensional Euclidean space. However, the set of states is just $\{0, 1\}$, so a history is just a subset of space-time. Via the definition of a "monotonic evolution", the space-time set grows still in discrete steps—allowing a generalization of the original eroder question. Though the results are somewhat similar, there is a richer set of possibilities: some eroders will erode only in non-linear time. Systems that erode in linear time are characterized in a way similar to Theorem 1: via a set σ; in the present case of a space-time system, instead of being empty, σ has to consist of just the origin 0. Now there will be some eroders even with $\sigma \neq \{0\}$, just not working in linear time. Noise is introduced and it is shown again that a linear-time eroder is also stable, and if it is not eroding then it is not. But only Toom's much later paper [50] proves that the system is not a stable eroder when $\sigma \neq \{0\}$.

The papers [41, 42] strengthen the stable eroder results to cases where the set $\mathbb{S}$ of local states is not finite, but is the set of integers. Of course, new conditions are needed on the transition probabilities.

The paper [37] places some cellular automata results into the broader context of *tiling systems*, thus touching the area of symbolic dynamics. Generally tiling questions don't involve probability, but here a version of probabilistic perturbation with a condition similar to (9) is introduced, so that questions of stability can be examined, allowing an application of the stable eroder technique (among others). Similar conditions on random perturbations have much later been used by Durand, Romashchenko and Shen in [5], although for non-periodic tile sets, and a very different technique of "error-correction". The perturbation condition raises a new challenge still mostly unsolved: to find Gibbs distributions satisfying it. The paper gives a construction only in one dimension, but even this with a quite technical proof.

The problem of information conservation in one dimension continued to occupy Andrei Toom throughout his career; he was not satisfied with the overly complex constructions in [6, 8]. In two papers, he explored alternative models: in these, the cells have a continuous set of states (real numbers or real numbers modulo some $M > 0$). The local transition rule is a (linear) averaging operation over the neighborhood—with the intent of preserving a kind of continuity—plus some noise. (There is reference to a related idea of Hammersley's "harnesses" in [16].) The paper [45] gives a very detailed analysis of how the loss of continuity depends on the tail behavior of the noise variables. One idea for a finite system, also favored by physicists, is to store the information in some kind of non-local, topological characteristic, like rotation number. In [44], Toom explores this idea via a cellular automaton with a finite number L of cells, periodic boundary conditions, whose local state space is the set of real numbers modulo M. Given the finiteness of L, all information, including the rotation number, will eventually be lost; the main result of the paper lower-bounds the expected time for the loss of the rotation number by approximately $L^{-1}\exp(cM^3)$ (with a similar upper bound in some cases, without the L). So the information is preserved for a time growing "super-exponentially" as a function of the size M of the local state space. The dependence on the size L of the space is, alas, less nice. In the finite versions of Toom's two-dimensional non-ergodic models as well as of the complex one-dimensional models [6, 8], the lower bound on relaxation time *grows* exponentially with the system size L (number of cells) while here it decreases as L^{-1}.

The papers [41, 42] cited above can also be seen as surface growth models where despite the fact that noise drives the surface to grow exclusively in the upward direction, with probability 1 the surface height remains bounded. The physics literature calls this *pinning*. In [47], Toom returned to this problem in a one-dimensional model in which cells can have real number states. The local rule drives the state in each step towards the local minimum of the three neighbors, with a speed $1 - \alpha$, but again there is a small rate β of random growth added. The result is that this system has a property he calls *pseudo-pinning*: there is growth, but its velocity is bounded $C_1\alpha^{C_2/\beta}$, so it decreases exponentially in $1/\beta$. (There is a somewhat similar lower bound.) The proofs use multiple times a technique involving a majorizing operation $\prec$ in the following way: $AB \prec B'A'$ where operators A, B are replaced with operators A', B' having slightly different parameters.

10 Disappearing cells

It is a natural question to try to generalize the cellular automata model by allowing the birth and death of cells—in the sense that for example in one dimension when a cell dies its site disappears, and its nearest neighbors connect with each other. For a finite set of sites there are natural ways to do this in general form even for random evolutions, see Malyshev's [22]. However, it does not seem possible to talk about, for example, an infinite two-dimensional cellular automaton in which cells (sites) can be eliminated or added: we would have to jump to the case of general infinite graphs.

The one-dimensional case, the subject of Toom's pioneering [51], is an exception. The model is attractive as it promises a new kind of simple one-dimensional noisy "non-ergodicity". Recall that with a binary set of states $\mathbb{S} = \{0, 1\}$, we may want to remember, in low-level noise, whether we started from all 0's or all 1's. When in a sea of 0's a large island of 1's appears, we did not find any *simple* local rule capable of erasing it (in noise). But when cells can be removed there is such a rule: for example just remove every pair of cells containing 01! Repeated application of this rule would eat up the island—hopefully even in noise.

Right at the start, however, Toom is faced with a new kind of problem: how to define the probability space in question. If a cell is deleted then the position of cells over a whole half-line changes. The model he offers does not define a probability measure over histories, only considers measures over the space of configurations. Even here it is restricted to translation-invariant measures, so the history is just a sequence of such measures $\mu_0, \mu_1, \dots$. The main definition is that of a transition operator P with $\mu_{t+1} = P\mu_t$. It is more complex than in the ordinary case, and the operator obtained is *not linear*. The techniques relying on linearity are not available, so even a new proof of the existence of an invariant measure is needed. A generalization allowing arbitrary local (one-dimensional) substitutions appeared in the joint paper [26] with Toom's students Rocha and Simas.

In [51], Toom proves non-ergodicity of his variable-length medium only for one-sided noise. It is plausible that it would also hold for arbitrary small (independent) noise; Toom promised but did not live to deliver this result. The proof in the paper is a very detailed contour argument, in its general shape not unlike the one Toom introduced for the Stavskaya model in [31].

11 The life

(I posted an English translation of Toom's autobiographical notes in [52].) Toom's activities extended far beyond research in mathematics. In the latter he focused indeed mainly on probabilistic cellular automata, but to computer scientists his name is probably best known for the work he did as an undergraduate. Following the surprising discovery of Karatsuba that two n-digit numbers can be multiplied in $O(n^{\log_2 3})$ bit-operations (instead of the $O(n^2)$ steps of the thousand-years old school algorithm), he published in [30] an algorithm doing this in $O(n^{1+\varepsilon})$ steps. Stephen Cook found a similar algorithm independently at about the same time.

In the deep tradition of Russian mathematicians, among them many outstanding ones, he was very active in mathematical education—considering an honor to be able to contribute to the system he himself benefited from. We can get an impression of the extent of his activity from the corresponding parts of his homepage http://www.toomandre.com (which hopefully will remain alive for a while): /my-articles/ruseduc, /my-articles/engeduc. He wrote a number of articles in the magazine Kvant (addressed to interested and able high-school students), posing problems and giving little expositions. He was also a main organizer of and contributor to the so-called "School by correspondence". This institution addressed the need of motivated students who wanted to go beyond what their schools could offer but, living far from metropolitan areas (before the era of internet) had no access to other opportunities. They received challenging mathematical problems, had some time (say, a month) to work on them and send in the solutions. Professional mathematicians like Toom sent back the commented solutions, even giving the students a second chance to solve the problem correctly. In Moscow, Toom ran a popular computer club, which also turned out to be a greatly successful way of leading a number of youngsters towards professional mathematics and computer science.

In 1989, Toom moved to the United States; among all the new social and existential challenges he had to face, his interest and activism in education never diminished. Ideally, given the depth of his experience and commitment, he should have thrived here, but over eight years of trying, he did not secure a tenured position in the United States. Paradoxically, his passionate educational interest worked against him, because of one handicap: a complete lack of the diplomatic gene. He was successful in the classroom, but his rather bluntly expressed opinions, starting with [40], made him unattractive to search committees. In Russia, he did criticize some buerocratic aspects of the educational system, but in America he mostly posed the Russian way of mathematics teaching as a superior example and attacked the core institutional principles, as expressed for example in some standards documents, of the American system (and the Brazilian one trying to follow it). One of his favorite topics was "word problems"; he argued their usefulness in great, inspiring and convincing detail (and also practiced in the classroom what he preached). Thanks to the respect he commanded in the strong probability-theory community of Brazil he ended up there (learning Portuguese at an advanced age). Settling at the University of Pernambuco in Recife he felt finally able to teach what he wanted and the way he wanted.

Andrei never considered himself merely a mathematician, and has always tried to apply his rigorous thinking to other areas. In the early 1970's he was drawn into the circle of the psychologist Vladimir Lefebvre (later a professor at UC Irvine), who introduced a formal, algebraic system of social interactions. Toom wrote some papers on the subject, for example applying those concepts to game theory. But more importantly, he took up the idea of "reflexion" as a useful tool of analyzing human behavior and interpreting literature. An interesting example is his study of the famous—and rather enigmatic—novel "A hero of our time" by the 19th century poet and writer Mikhail Lermontov. His last work, cut short by his death on his retirement to New York, was working, together with his wife Anna, on the rich legacy of his grandfather, a noted Russian poet Antokolsky; they published their new discoveries in [54].

Bibliography

[1] Piotr Berman and Janos Simon. Investigations of fault-tolerant networks of computers. In *Proc. of the 20-th Annual ACM Symp. on the Theory of Computing*, pages 66–77, 1988. 4, 6

[2] Maury Bramson and Lawrence F. Gray. A useful renormalization argument. In Richard Durrett and Harry Kesten, editors, *Random Walks, Brownian Motion and Interacting Particle Systems*, pages 113–152. Birkhäuser, Boston, 1991. 6, 7

[3] Moises Lima de Mendezes and Andrei L. Toom. A non-linear eroder in presence of one-sided noise. *Brazilian Journal of Probability and Statistics*, 20:1–12, 2006. 8

[4] R. L. Dobrushin, V. I. Kryukov, and A. L. Toom, editors. *Locally Interacting Systems and their Application in Biology.* Lecture Notes in Mathematics 653. Springer, 1976. Proceedings of a School-Seminar held in Puschchino,. 9

[5] Bruno Durand, Andrei E. Romashchenko, and Alexander Kh. Shen. Fixed-point tile sets and their applications. *Journal of Computer and System Sciences*, 78:731–764, 2012. 9

[6] Peter Gács. Reliable computation with cellular automata. *Journal of Computer System Science*, 32(1):15–78, February 1986. Conference version at STOC' 83. 7, 9

[7] Peter Gács. A new version of Toom's proof. Technical report, Department of Computer Science, Boston University, TR 95-009, Boston, MA 02215, 1995. arXiv:2105.05968. 3, 6

[8] Peter Gács. Reliable cellular automata with self-organization. *Journal of Statistical Physics*, 103(1/2):45–267, April 2001. See also arXiv:math/0003117 [math.PR] and the proceedings of STOC '97. 7, 9

[9] Peter Gács, Georgii L. Kurdyumov, and Leonid A. Levin. One-dimensional homogenuous media dissolving finite islands. *Problems of Inf. Transm.*, 14(3):223–226, 1978. Translation of the Russian version whose page numbers are 92-96. 7

[10] Peter Gács and John Reif. A simple three-dimensional real-time reliable cellular array. *Journal of Computer and System Sciences*, 36(2):125–147, April 1988. Short version in STOC '85. 4, 6

[11] Peter Gács and Ilkka Törmä. Stable multi-level monotonic eroders. *Theory of Computing Systems*, 66:322–353, 2022. 8

[12] Gregory A. Galperin. One-dimensional networks of automata with monotonic local interaction. *Problemi Peredachi Informatsii*, 12(4):74–87, 1976. Translation from Russian in Problems of Information Transmission. 8

[13] Lawrence F. Gray. The positive rates problem for attractive nearest neighbor spin systems on $\mathbb{Z}$. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 61:389–404, 1982. 3, 7

[14] Lawrence F. Gray. The behavior of processes with statistical mechanical properties. In Harry Kesten, editor, *Percolation Theory and Ergodic Theory of Infinite Particle Systems*, pages 131–167. Springer Verlag, 1987. 3, 7

[15] Lawrence F. Gray. Toom's stability theorem in continuous time. In Harry Kesten, Maury Bramson, and Rick Durrett, editors, *Perplexing Problems in Probability*, pages 331–353. Birkhäuser, Boston, 1999. 5, 6

[16] J. M. Hammersley. Harnesses. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume III, pages 89–117. University of California Press, Berkeley, 1967. 9

[17] O. Kozlov and N. Vasiliev. Reversible Markov chains with local interaction. In R. L. Dobrushin and Ya. G. Sinai, editors, *Multicomponent Random Systems*, volume 6 of *Advances in Probability and Related Topics*, pages 451–469. Marcel Dekker Inc., 1980. Contributing editor D. Griffeath, series editor P. Ney. Translated from Russian. 3

[18] G. L. Kurdyumov. An example of a nonergodic homogenous one-dimensional random medium with positive transition probabilities. *Soviet Mathematics Doklady*, 19(1):211–214, 1978. 7

[19] G. L. Kurdyumov. An algorithm-theoretic method for the study of uniform random networks. In R. L. Dobrushin and Ya. G. Sinai, editors, *Multicomponent Random Systems*, volume 6 of *Advances in Probability and Related Topics*, pages 471–503. Marcel Dekker Inc., 1980. Contributing editor D. Griffeath, series editor P. Ney. Translated from Russian. 1

[20] Joel L. Lebowitz, Christian Maes, and Eugene R. Speer. Statistical mechanics of stochastic cellular automata. *Journal of Statistical Physics*, 59:117–170, 1990. 3

[21] Thomas M. Liggett. *Interacting Particle Systems*, volume 276 of *Grundlehren der mathematischen Wissenschaften*. Springer Verlag, New York, 1985. Reprinted with new postface in 2005. 1, 3

[22] V. A. Malyshev. Random graphs and grammars on graphs. *Discrete Mathematics*, 10(2):30–44, 1998. In Russian. 10

[23] L. G. Mityushin. Nonergodicity of uniform threshold networks for small self-excitation. *Problemy Peredachi Informatsii*, 6:99–103, 1970. In Russian. 2

[24] Kihong Park. *Ergodicity and Mixing Rate of One-Dimensional Cellular Automata*. PhD thesis, Boston University, Boston, MA 02215, 1996. See https://cs-web.bu.edu/faculty/gacs/papers/1996-015-park-phdthesis.pdf. 7

[25] Nicholas Pippenger. Developments in "the synthesis of reliable organisms from unreliable components". In James Glimm, John Impagliazzo, and Isadore Singer, editors, *The Legacy of John von Neumann*, pages 311–324,

Hempstead, NY, 1988. American Mathematical Society. Proceedings of Symposia in Pure Mathematics, Vol 50. 4

[26] Andréa Rocha, Alexandre B. Simas, and Andrei L. Toom. Substitution operators. *Journal of Statistical Physics*, 143:585–618, 2011. 10

[27] M. G. Shnirman. To the question of the ergodicity of some Markov chain with an inifinite number of states. *Problemy Kibernetiki*, 20:115–122, 1968. In Russian. 2

[28] O. N. Stavskaya and Ilya Piatetski-Shapiro. Homogenous networks of spontaneously active elements. *Problemy Kibernetiki*, 20:91–106, 1968. In Russian. 2

[29] Adrei L. Toom and L. G. Mityushin. Two results resulting in undecidability for univariate cellular automata. *Problems of Information Transmission*, 12:69–75, 1974. In the translation journal, pages 135-141. See also http://www.toomandre.com/my-articles/engmat/MITUSH-E.PDF. 5

[30] Andrei L. Toom. The complexity of a scheme of functional elements realizing the multiplication of integers. *Soviet Mathematics - Doklady*, 3(1):714–716, 1963. Translation from Russian. 11

[31] Andrei L. Toom. On a family of uniform nets of formal neurons. *Soviet Mathematics - Doklady*, 9(6):1338–1341, 1968. 2, 10

[32] Andrei L. Toom. On invariant measures in nonergodic random media. *Moscow University Press*, 41:43–51, 1972. In Russian. 2

[33] Andrei L. Toom. Nonergodic multidimensional systems of automata. *Problems of Information Transmission*, 10(3):239–246, 1973. Translation from Russian. Pages 70-79 in the original. 6

[34] Andrei L. Toom. Monotone binary cellular automata. *Problems of Information Transmission*, 12(1):33–37, 1976. Translation from Russian. Pages 48-54 in the original. 5

[35] Andrei L. Toom. Unstable multicomponent systems. *Problems of Information Transmission*, 12(3):78–84, 1976. Pages 220-225 in the translated version. 8

[36] Andrei L. Toom. Monotonic evolutions in real spaces. In A. L. Toom R. L. Dobrushin, V. I. Kyukov, editor, *Localy Interacting Systems and their Applications in Biology*, Proceedings of a summer seminar held in Pushchino, 1976, pages 1–14. Springer, 1978. Translation from Russian. 9

[37] Andrei L. Toom. Stability in lattice systems with local prohibitions. In A. L. Toom R. L. Dobrushin, V. I. Kryukov, editor, *Interacting Markov Processes and their Applications in Biology*, Proceedings of a summer seminar held in Pushchino, Moscow Region, May 1978, pages 74–99. Biological research center, Pushchino, 1978. Translation from Russian in Sel. Math. Sov., 1985, v. 4, n.1, 73-92, see also in http://www.toomandre.com/my-articles/engmat/PROHIB-E.PDF. 9

[38] Andrei L. Toom. Stable and attractive trajectories in multicomponent systems. In R. L. Dobrushin, editor, *Multicomponent Systems*, volume 6 of *Advances in Probability*, pages 549–575. Dekker, New York, 1980. Translation from Russian. 5, 6

[39] Andrei L. Toom. Estimates for measures describing the behavior of stochastic systems with local interaction. In A. L. Toom R. L. Dobrushin, V. I. Kryukov, editor, *Interacting Markov processes and their applications to the mathematical modeling of biological systems*, Proceedings of a summer school held in Pushchino, 1981, pages 21–33. Biological research center, Pushchino, 1982. In Russian. 8

[40] Andrei L. Toom. A Russian teacher in America. *Journal of Mathematical Behavior*, 12(2):117–139, 1993. 11

[41] Andrei L. Toom. On critical phenomena in interacting growth systems. part i: General. *Journal of Statistical Physics*, 74(1/2):91–109, 1994. 9

[42] Andrei L. Toom. On critical phenomena in interacting growth systems. part ii: Bounded growth. *Journal of Statistical Physics*, 74(1/2):111–130, 1994. 9

[43] Andrei L. Toom. Cellular automata with errors: Problems for students of probability. In J. Laurie Snell, editor, *Topics in Contemporary Probability and its Applications*, Probability and Stochastic Series, pages 117–157. CRC Press, Boca Raton, FL, 1995. 1

[44] Andrei L. Toom. Simple 1-dimensional systems with superexponential relaxation time. *Journal of Statistical Physics*, 80(3/4):545–563, 1995. 9

[45] Andrei L. Toom. Tails in harnesses. *Journal of Statistical Physics*, 88(1/2):347–364, 1997. 9

[46] Andrei L. Toom. On periodic processes of percolation cellular automata. *Resenhas IME-USP*, 3(3):351–357, 1998. 2

[47] Andrei L. Toom. Pseudo-pinning in a growth model. *Markov Processes and Related Fields*, 4:101–106, 1998. 9

[48] Andrei L. Toom. Algorithmic unsolvability of the ergodicity problem for binary cellular automata. *Markov Processes and Related Fields*, 6(4):579–577, 2000. 1

[49] Andrei L. Toom. Algorithmic unsolvability of the ergodicity problem for locally interacting prcesses with continuous time. *Journal of Statistical Physics*, 98(1/2):495–501, 2000. 1

[50] Andrei L. Toom. On critical values for some random processes with local interaction in r2. In Vladas Sidoravicius, editor, *In and Out of Equilibrium: Probability with a Physics Flavor*, pages 309–319. Birkhäuser Boston, Boston, MA, 2002. 9

[51] Andrei L. Toom. Non-ergodicity in a 1-d particle process with variable length. *Journal of Statistical Physics*, 115:895–924, 2004. 10

[52] Andrei L. Toom. Autobiographical notes. `http://www.cs.bu.edu/fac/gacs/papers/Pa-longp-eng.pdf`, 2005. 11

[53] Andrei L. Toom. Ilya Piatetski-Shapiro in 1963-1973: Passion for Applied Mathematics, 2009. `http://www.toomandre.com/my-articles/engmat/ILYA-ENG.doc`. 1

[54] Andrei L. Toom and Anna Toom. *The Name of this Friendship: Poetic Brotherhood*. Samizdat, 2022. In Russian, `https://www.litres.ru/book/andrey-toom-32253688/imya-etoy-druzhby-poeticheskoe-bratstvo-68282777/`. 11

[55] Andrei L. Toom, N. B. Vasiliev, O. N. Stavskaya, L. G. Mityushin, G. L. Kurdyumov, 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*, pages 1–182. Manchester University Press, 1990. 1, 9

[56] L. N. Vasershtein. Markov processes over denumerable products of spaces. *Problems of Information Transmission*, 5(3):64–72, 1969. 1

[57] L. N. Vasershtein and A. M. Leontovich. On invariant measures of some Markov operators describing a homogenous random medium. *Problems of Information Transmission*, 6(1):71–80, 1970. In Russian. 2

[58] Nikolay B. Vasiliev. Correlation equations for the stationary measure of a Markov chain. *Theory of Probability and its Applications*, 15(3):521–525, 1970. Translated from Russian. 2

[59] John von Neumann. Probabilistic logics and the synthesis of reliable organisms from unreliable components. In C. Shannon and McCarthy, editors, *Automata Studies*. Princeton University Press, Princeton, NJ., 1956. 4

[60] John von Neumann. *The Theory of Self-Reproducing Automata*. University of Illinois Press, Urbana, IL, 1966. Edited by A. Burks. 1