

Exponential Decay of Correlations for Strongly Coupled Toom Probabilistic Cellular Automata

Augustin de Maere · Lise Ponselet

Abstract We investigate the low-noise regime of a large class of probabilistic cellular automata, including the North-East-Center model of Toom. They are defined as stochastic perturbations of cellular automata belonging to the category of monotonic binary tessellations and possessing a property of erosion. We prove, for a set of initial conditions, exponential convergence of the induced processes toward an extremal invariant measure with a highly predominant spin value. We also show that this invariant measure presents exponential decay of correlations in space and in time and is therefore strongly mixing.

Keywords Probabilistic Cellular Automata · Phase Transition · Toom Models · Monotonic Binary Tessellations

1 Introduction

Probabilistic cellular automata (PCA) are a type of models of multicomponent systems whose time evolution is driven by local interactions blurred with some noise. More precisely, they are discrete-time stochastic processes with the Markov property, made of lattices of components whose individual states take values in a finite set and are simultaneously updated at every time step. The transition rules involve interactions between neighbouring components.

The long-time limit of these processes has been the subject of many numerical and theoretical results in the last fifty years and a lot of questions remain open - see for instance the surveys [14, 15]. Despite the apparent simplicity of their discrete configuration space and merely local interactions, PCA exhibit a variety of macroscopic

A. de Maere · L. Ponselet
Université catholique de Louvain, IRMP, Chemin du cyclotron 2, bte L7.01.03, B-1348 Louvain-la-Neuve, Belgium.
Tel.: +32-10-473279
Fax: +32-10-472414
E-mail: lise.ponselet@uclouvain.be

tendencies. Moreover, due to their basic definition, these models provide a favorable field of study in order to give firm foundations to the growing knowledge about non-equilibrium phenomena. In particular, interest concentrates on the existence of phase transitions in this context of non-equilibrium statistical physics. For a non-negligible subset of the parameter set of certain PCA models, the limiting states of the processes can be non-unique and depend strongly on the initial conditions, thus providing examples of systems which keep remembering part of the data from their remote past [1].

Here we consider the PCA resulting from small random perturbations of the deterministic cellular automata in the class of *monotonic binary tessellations* (MBT) examined in [12, 13]. In an MBT, the *spin* variables disposed on the sites of a $\mathbb{Z}^d$ lattice can take only two different values and are all simultaneously updated according to the same monotonic function of their neighbours' states. In [12, 13], Toom gave a criterion for an MBT to be an *eroder*, i.e. to erase in a finite time any finite island of impurities in a predominantly homogeneous configuration of one of the two types of spins. The completely homogeneous space-time configuration with this spin value everywhere is then said to be an *attractive* trajectory of that MBT. He also proved in [13] that this erosion condition implies the *stability* of the homogeneous trajectory under the introduction of a small error rate: the corresponding PCA admits a stationary measure for which the low probability of finding an impurity is preserved. Another proof of this stability theorem in a more general context was later given in [3], using renormalization group methods.

For example, the one-dimensional percolation PCA of Stavskaya and Toom's two-dimensional North-East-Center voting PCA fall within this class of perturbations of MBT which are eroders. They both present a phase transition as non-uniqueness of their invariant probability measures in the low-noise regime follows from this stability theorem, while, when the noise is high enough, the processes converge toward a unique invariant measure, regardless of the initial state [5].

Beside this stability result, very little is known about the long-time asymptotics of this class of PCA in the nearly deterministic regime. It has already been explored by means of simulations [1, 4, 9, 10, 16]. Here we address the problem from a theoretical point of view. By virtue of the ergodic theorem, the asymptotic behavior of such a PCA is actually given by its ergodic invariant probability measures. We focus on the statistical properties of the invariant measure with a predominant spin value which is obtained as a perturbation of the stable homogeneous trajectory. Reference [2] established exponential convergence to this stationary measure, and therefore exponential decay of its correlations, for the Stavskaya PCA and a class of multidimensional generalizations of it.

In this paper, we use a different method allowing us to extend this result to the North-East-Center model and all the other MBT satisfying Toom's erosion condition, with a restriction to MBT without memory, where the state of the system at any time t depends only on its previous state at time $t - 1$ and not directly on earlier states. We prove rigorously that for some well-chosen initial conditions, the process converges exponentially fast toward the invariant measure under consideration. We also show that this measure has exponential decay of correlations in space and in time. In

particular, it is ergodic and therefore extremal. The proof is based on a perturbative expansion, with paths and graphs, which combines a technique of decoupling in the pure phases previously introduced and developed for coupled map lattices [6,8] with graphs constructed for a Peierls argument in the proof of stability in [13].

2 Formalism and Results

We consider the cellular automata in the class of monotonic binary tessellations described in Section IV of [13] except that we restrict ourselves to MBT without memory. Some of our notations are similar to the notations of [13] in order to facilitate the use of the graphs constructed there but they are not identical because we don't need as general a framework here.

Let $S = \{-1, 1\}$ be the binary state space for a component of the system, which we will call a *spin*, and let $X = S^{\mathbb{Z}^d}$ be the configuration space for the whole system. For any spin configuration $\underline{\omega} \in X$, ω_x will denote the value of $\underline{\omega}$ at site x and $\underline{\omega}_\Lambda$ will denote the vector $\{\omega_x\}_{x \in \Lambda}$ for any $\Lambda \subseteq \mathbb{Z}^d$. We will also use the notation $(\underline{\omega}_{\neq x}, a)$ for the configuration obtained from $\underline{\omega}$ by replacing the spin ω_x at site x with the value $a \in S$.

In a tessellation, the evolution law for the state of the spin at a site x of the $\mathbb{Z}^d$ lattice involves its neighbours, which are defined as the elements of $U(x) = x + U$ for a fixed finite set $U = \{u_1, \dots, u_R\} \subset \mathbb{Z}^d$. Starting from a configuration $\underline{\omega} \in X$, the simultaneous updating of every spin at every time step consists in transforming the spin ω_x at site x into the value

$$\phi_x(\underline{\omega}) = \phi(\omega_{x+u_1}, \dots, \omega_{x+u_R})$$

for a given function $\phi : S^R \rightarrow S$. From now on, as we consider monotonic binary tessellations, ϕ will be monotonic in the sense that if $\omega_u \leq \omega'_u$ for all $u \in U$, then $\phi(\underline{\omega}_U) \leq \phi(\underline{\omega}'_U)$. We also reject the trivial case of a constant function ϕ . Note that these two assumptions imply that $\phi(\pm 1, \dots, \pm 1) = \pm 1$ and then that the configurations $\{+1\}_{x \in \mathbb{Z}^d}$ and $\{-1\}_{x \in \mathbb{Z}^d}$ are left invariant by the deterministic time evolution. They generate completely homogeneous trajectories.

In addition, we suppose that ϕ verifies the Toom erosion criterion. It is expressed in terms of the *plus sets*: the subsets $Z \subseteq U$ such that if $\omega_u = +1$ for all $u \in Z$, then $\phi(\underline{\omega}_U) = +1$. When applied to the particular case of an MBT without memory, the criterion given in Proposition 1 of [12] and in Theorem 6 of [13] states that the intersection of all convex hulls of plus sets in $\mathbb{R}^d$ must be empty.

We now describe a way to introduce some noise in such an eroding MBT to generate a PCA. The system follows the same updating rule as in the deterministic case but, at each site of the lattice and at each time step, an error can occur with a probability of order ε , for some ε in $[0, 1]$, the spin then taking the opposite value. Occurrence of an error at a site is independent from occurrence of errors at other sites. The process resulting from the sequence of simultaneous updatings of all spins becomes a stochastic process. The local transition probabilities $p(\xi_x | \underline{\omega})$ of that process only depend on

$\underline{\omega}_{U(x)}$. We suppose that they satisfy the following assumptions, for some given values of the parameters $\varepsilon \in [0, 1]$ and $\alpha \geq 0$:

- (A1) if $\xi_x \neq \phi_x(\underline{\omega})$, then $p(\xi_x|\underline{\omega}) \leq \varepsilon$;
- (A2) if $a = \phi_x(\underline{\omega})$, then $|p(\xi_x|\underline{\omega}) - p(\xi_x|\underline{\omega}_{\neq y}, a)| \leq \alpha p(\xi_x|\underline{\omega})$ for all $y \in \mathbb{Z}^d$.

Assumption (A1) is the low-noise condition: it ensures that, at each site of the lattice, the deterministic rule is followed with a probability of at least $1 - \varepsilon$. Note that we make no restriction about the possible bias in favor of errors of a specific spin sign. Assumption (A2), used in conjunction with the monotonicity of ϕ , ensures a form of decoupling in the pure phases: if a spin was flipped without changing the deterministic prescription given by the value of ϕ_x , it would be of little consequence to the involved transition probabilities. The local transition probabilities define a transition probability kernel on X , which can be formally written as

$$\mathbb{P}[\underline{\xi}|\underline{\omega}] = \prod_{x \in \mathbb{Z}^d} p(\xi_x|\underline{\omega}). \quad (1)$$

Assume that S is endowed with the discrete topology and X with the product topology. Let $\mathcal{C}(X)$ be the set of continuous functions from X into $\mathbb{R}$ with the norm: $|\varphi|_\infty = \sup_{\underline{\omega} \in X} |\varphi(\underline{\omega})|$. Since X is compact, $\mathcal{C}(X)$ with the norm $|\cdot|_\infty$ is a Banach space, and its dual is $\mathcal{M}(X)$, the space of signed Borel measures on X with the norm:

$$|\mu| = \sup\{\mu(\varphi) \mid \varphi \in \mathcal{C}(X), |\varphi|_\infty \leq 1\}.$$

For PCA on infinite spaces, it is often useful to consider the following semi-norm on $\mathcal{C}(X)$: if $\delta_x \varphi(\underline{\omega}) = \varphi(\underline{\omega}) - \varphi(\underline{\omega}_{\neq x}, -\omega_x)$, we define:

$$\|\varphi\| = \sum_{x \in \mathbb{Z}^d} |\delta_x \varphi|_\infty. \quad (2)$$

Let $T : \mathcal{M}(X) \rightarrow \mathcal{M}(X)$ be the transfer operator associated to the transition probability kernel of (1), explicitly defined as:

$$T\mu(\varphi) = \int d\underline{\omega} \int d\underline{\xi} \varphi(\underline{\xi}) \mathbb{P}[\underline{\xi}|\underline{\omega}] \mu(\underline{\omega}), \quad (3)$$

for any φ in $\mathcal{C}(X)$.

Let $\mathbb{1}_{-\Lambda}$ denote the indicator function of the subset $\{\underline{\omega} \in X \mid \omega_x = -1 \ \forall x \in \Lambda\}$ and, for any measurable set $Y \subseteq X$, let $\mathbb{1}_Y : \mathcal{M}(X) \rightarrow \mathcal{M}(X)$ be the operator defined as:

$$\mathbb{1}_Y \mu(\varphi) = \mu(\varphi \mathbb{1}_Y) \quad \forall \varphi \in \mathcal{C}(X).$$

As the MBT defined above satisfies the Toom erosion condition, the results of [13] apply to it. Theorems 5 and 6 of [13] imply the stability of the homogeneous trajectory with the spin value $+1$ everywhere. More precisely, let $\delta^{(+)}$ be the Dirac measure concentrated on the configuration $\{+1\}_{x \in \mathbb{Z}^d}$. Then, using these theorems and Definition 1 of a stable trajectory in [13], we have

$$\lim_{\varepsilon \rightarrow 0} \sup_{\substack{n \in \mathbb{N} \\ x \in \mathbb{Z}^d}} T^n \delta^{(+)} (\mathbb{1}_{-\{x\}}) = 0. \quad (4)$$

Now consider the following sequence of measures:

$$\left(\frac{1}{n} \sum_{k=0}^{n-1} T^k \delta^{(+)} \right)_{n \in \mathbb{N}_0}.$$

We can always extract from it a weakly-* convergent subsequence. Indeed, the Banach-Alaoglu theorem states that the unit ball of $\mathcal{M}(X)$ is compact in the weak-* topology. The associated limit $\mu_{\text{inv}}^{(+)}$ is explicitly given by

$$\mu_{\text{inv}}^{(+)}(\varphi) = \lim_{j \rightarrow \infty} \left(\frac{1}{n_j} \sum_{k=0}^{n_j-1} T^k \delta^{(+)}(\varphi) \right) \quad \forall \varphi \in \mathcal{C}(X), \quad (5)$$

for a certain subsequence $(n_j)_{j \in \mathbb{N}}$ of increasing positive integers. The measure $\mu_{\text{inv}}^{(+)}$ is obviously an invariant probability measure, in the sense that $T\mu_{\text{inv}}^{(+)} = \mu_{\text{inv}}^{(+)}$, and equation (4) implies that

$$\lim_{\varepsilon \rightarrow 0} \sup_{x \in \mathbb{Z}^d} \mu_{\text{inv}}^{(+)}(\mathbb{1}_{-, \{x\}}) = 0. \quad (6)$$

Example 1 (The Stavskaya model) The percolation PCA of Stavskaya [11] is one of the first PCA for which the existence of a phase transition has been rigorously proven - see an outline of the proof using the method of contours in Chapter 1 of [15]. It has $d = 1$, $U = \{0, 1\} \subset \mathbb{Z}$ and $\phi(\omega_0, \omega_1) = -1$ if and only if $\omega_0 = \omega_1 = -1$. Errors are strongly biased in the sense that they can only transform a spin $+1$ into a -1 but not a -1 into a $+1$. The Dirac measure $\delta^{(-)}$ is then always a stationary state. We will call it $\mu_{\text{inv}}^{(-)}$. Besides, the Toom criterion is satisfied by these U and ϕ since the intersection of the plus sets $\{0\}$ and $\{1\}$ is empty, therefore equation (6) carries the existence of an $\varepsilon_c > 0$ such that for all $\varepsilon \leq \varepsilon_c$, the process admits a second different invariant measure $\mu_{\text{inv}}^{(+)}$.

Example 2 (The Toom model) The North-East-Center voting model has $d = 2$ and $U = \{(0,0); (1,0); (0,1)\} \subset \mathbb{Z}^2$ and ϕ returns the majority spin among these three neighbours' states. This model verifies the Toom erosion condition too, because $\{(0,0); (1,0)\}$, $\{(1,0); (0,1)\}$ and $\{(0,1); (0,0)\}$ are plus sets whose convex hulls have an empty intersection. Moreover, the model is symmetric under the interchange of positive and negative spin signs. Combining this symmetry with definition (5) gives two invariant measures, $\mu_{\text{inv}}^{(+)}$ which satisfies equation (6) and $\mu_{\text{inv}}^{(-)}$ which satisfies the symmetric counterpart of equation (6), thus revealing that they differ as long as ε is small enough.

Then, in such models, if ε is below a critical threshold ε_c , there exists an infinite number of invariant probability measures, since any convex combination of $\mu_{\text{inv}}^{(+)}$ and $\mu_{\text{inv}}^{(-)}$ is also an invariant probability measure. But, among these invariant measures, some have interesting properties such as ergodicity and exponential decay of correlations. In [2], the low-noise regime of a class of PCA is examined, including the Stavskaya model and a multidimensional generalization of it, but not the Toom model.

Exponential convergence toward equilibrium of the processes with the initial condition $\delta^{(+)}$ is proven for these PCA by constructing a cluster expansion. Exponential decay of correlations in space and in time follows for the non-trivial natural invariant measure $\mu_{\text{inv}}^{(+)}$ that appears at the phase transition of the Stavskaya model.

Our argument extends the results of [2] to the whole class of PCA defined above. It relies on an expansion which isolates the influence of each space-time site on each other site in its future, along several paths of influence. These paths pass through the dominant phase with positive spins almost everywhere and Assumption (A2) of decoupling in the pure phases can be used in order to bound the influence of spins one on another. The paths seldom encounter a negative spin. Whenever they do, in view to evaluate how improbable that negative spin is, we will attach to it a graph borrowed from a Peierls argument in the proof of stability in [13] and bring Assumption (A1) into play. The fact that paths select one site at a time will allow us to make use of the one-dimensional Toom graphs by associating them to single sites separately. As for the cluster expansion in [2], it requires contours which can enclose blocks of several spins, as the Stavskaya contours do, but not the most general Toom graphs, as far as we can see.

In Sections 3 to 6, we will show that any initial probability measure in a suitable basin of attraction of $\mathcal{M}(X)$ converges exponentially fast toward $\mu_{\text{inv}}^{(+)}$. We will consider the sets

$$\mathcal{B}^{(+)}(K, \varepsilon') = \left\{ \mu \in \mathcal{M}(X) \mid |\mathbb{1}_{-\Lambda} \mu| \leq K \varepsilon'^{|\Lambda|} \quad \forall \Lambda \subseteq \mathbb{Z}^d \right\},$$

with $K \geq 0$ and $\varepsilon' \in [0, 1]$ and prove the following result:

Theorem 1 *For any MBT without memory characterized by a non-constant monotonic function ϕ verifying the Toom erosion condition, there exist $\alpha_* > 0$ and a positive and strictly decreasing function ε_* defined on $[0, \alpha_*[$ such that, for all $\alpha \in [0, \alpha_*[$ and $\varepsilon \in [0, \varepsilon_*(\alpha)]$, the following assertion is true for any PCA satisfying the corresponding Assumptions (A1) and (A2). For any probability measure μ in $\mathcal{B}^{(+)}(K, \varepsilon')$ with $K \geq 0$ and $\varepsilon' < \varepsilon_*(\alpha)$, there exist some constants $C < \infty$ and $\sigma < 1$ such that, for all $\varphi \in \mathcal{C}(X)$ and all $n \in \mathbb{N}$,*

$$\left| T^n \mu(\varphi) - \mu_{\text{inv}}^{(+)}(\varphi) \right| \leq C \|\varphi\| \sigma^n.$$

Note that for PCA which are symmetric under the interchange of the two types of spins, such as the North-East-Center model, the symmetric result for μ in the class $\mathcal{B}^{(-)}(K, \varepsilon')$ and for $\mu_{\text{inv}}^{(-)}$ is also valid.

Eventually, in Section 7, we will prove that $\mu_{\text{inv}}^{(+)}$ has exponential decay of correlations in space and in time and is, consequently, strongly mixing.

Corollary *Assume that $\alpha < \alpha_*$ and that $\varepsilon < \varepsilon_*(\alpha)$, with α_* and ε_* as given by Theorem 1. Then there exist some constants $C' < \infty$ and $\eta < 1$ such that for any φ, ψ in $\mathcal{C}(X)$ with $\|\varphi\| < \infty$ and $\|\psi\| < \infty$, and with a positive Manhattan distance $d(\varphi, \psi)$ between their supports, we have*

$$\left| \mu_{\text{inv}}^{(+)}(\varphi\psi) - \mu_{\text{inv}}^{(+)}(\varphi)\mu_{\text{inv}}^{(+)}(\psi) \right| \leq C' (\|\varphi\| \|\psi\|_\infty + \|\varphi\|_\infty \|\psi\|) \eta^{d(\varphi, \psi)}.$$

For the exponential decay of correlations in time, we define the operator $T : \mathcal{C}(X) \rightarrow \mathcal{C}(X)$:

$$T\varphi(\underline{\omega}) = \int d\underline{\xi} \varphi(\underline{\xi}) \mathbb{P}[\underline{\xi} | \underline{\omega}],$$

which is simply the dual of the transfer operator acting on signed measures.

Theorem 2 *Assume that $\alpha < \alpha_*$ and that $\varepsilon < \varepsilon_*(\alpha)$ with α_* and ε_* as given by Theorem 1. Then, for any continuous functions φ, ψ with finite supports, there exist some constants $C_{\varphi, \psi} < \infty$ and $\sigma < 1$ such that, for all $n \in \mathbb{N}$,*

$$\left| \mu_{\text{inv}}^{(+)}(\varphi T^n \psi) - \mu_{\text{inv}}^{(+)}(\varphi) \mu_{\text{inv}}^{(+)}(\psi) \right| \leq C_{\varphi, \psi} \sigma^n.$$

3 Path Expansion

In this section, we will introduce a path expansion which is essentially equivalent to the Dobrushin criterion [5], using here a formalism which was originally introduced by Keller and Liverani for Coupled Map Lattices [6].

Let $\prec$ be any well-ordering of $\mathbb{Z}^d$. The operator $\Pi_x : \mathcal{C}(X) \rightarrow \mathcal{C}(X)$ is defined as:

$$\Pi_x \varphi(\underline{\omega}) = \varphi(\underline{\omega}_{\prec x, \underline{a}_{\prec x}}) - \varphi(\underline{\omega}_{\succ x, \underline{a}_{\prec x}}),$$

where, from now on, $\underline{a}$ will denote the configuration for which $a_x = a = +1$ for all $x \in \mathbb{Z}^d$. $(\underline{\omega}_{\prec x, \underline{a}_{\prec x}})$ is the configuration obtained from $\underline{\omega}$ by replacing the spins at all sites $y \prec x$ with the value a . Using telescopic sums, we can check that, for any continuous function φ ,

$$\varphi(\underline{\omega}) = \varphi(\underline{a}) + \sum_{x \in \mathbb{Z}^d} \Pi_x \varphi(\underline{\omega}). \quad (7)$$

With a slight abuse of notation, let us denote by $\Pi_x : \mathcal{M}(X) \rightarrow \mathcal{M}(X)$ the dual of the operator $\Pi_x : \mathcal{C}(X) \rightarrow \mathcal{C}(X)$. The image of $\mathcal{M}(X)$ under this operator is actually included in the set

$$\mathcal{M}_x = \{\mu \in \mathcal{M}(X) \mid \mu(\varphi) = 0 \text{ if } \varphi(\underline{\omega}) \text{ is independent of } \omega_x\} \quad (8)$$

whose elements verify the following property:

$$\mu \in \mathcal{M}_x \quad \Rightarrow \quad |\mu(\varphi)| \leq |\mu| |\delta_x \varphi|_{\infty}. \quad (9)$$

With equation (7), we can see that any signed measure of zero mass $\mu \in \mathcal{M}(X)$ with $\mu(1) = 0$ admits the following decomposition:

$$\mu = \sum_{x \in \mathbb{Z}^d} \Pi_x \mu. \quad (10)$$

While $\Pi_x \mu$ belongs to $\mathcal{M}_x$, it is no longer the case for $T \Pi_x \mu$. Nevertheless, since the interactions are local, we will see that $T \Pi_x \mu$ can be expressed as the sum of R signed measures: a first one in $\mathcal{M}_{x-u_1}$, a second one in $\mathcal{M}_{x-u_2}$, ... , and a last one in $\mathcal{M}_{x-u_R}$.

For this, consider an arbitrary measure $\mu_x \in \mathcal{M}_x$. Using definitions (3) and (8) of T and $\mathcal{M}_x$, together with our hypothesis that $p(\xi_y|\underline{\omega})$ only depends on $\underline{\omega}_{y+U}$, it is easy to check that

$$T\mu_x = \sum_{i=1}^R T^{(x-u_i, x)} \mu_x, \quad (11)$$

where R new operators have been defined:

$$T^{(x-u_i, x)} \mu(\varphi) = \int d\underline{\omega} \int d\underline{\xi} \varphi(\underline{\xi}) \tau^{(x-u_i, x)}(\underline{\xi}, \underline{\omega}) \mu(\underline{\omega}) \quad \text{for } i = 1, \dots, R \quad (12)$$

with the kernels

$$\begin{aligned} \tau^{(x-u_i, x)}(\underline{\xi}, \underline{\omega}) &= \left(\prod_{y \notin x-U} p(\xi_y|\underline{\omega}) \right) \left(\prod_{j=1}^{i-1} p(\xi_{x-u_j}|\underline{\omega}_{\neq x}, a) \right) \cdot \\ &\quad \cdot \left(p(\xi_{x-u_i}|\underline{\omega}) - p(\xi_{x-u_i}|\underline{\omega}_{\neq x}, a) \right) \left(\prod_{j=i+1}^R p(\xi_{x-u_j}|\underline{\omega}) \right). \end{aligned} \quad (13)$$

We notice that, for all i , the image of $\mathcal{M}(X)$ under the operator $T^{(x-u_i, x)}$ is included in $\mathcal{M}_{x-u_i}$.

Let now μ be a signed measure of zero mass and consider $T^n \mu$. Using the decomposition (10) and applying (11) iteratively, we find

$$T^n \mu = \sum_{x_0 \in \mathbb{Z}^d} \sum_{x_0-x_1 \in U} \dots \sum_{x_{n-1}-x_n \in U} T^{(x_n, x_{n-1})} \dots T^{(x_1, x_0)} \Pi_{x_0} \mu.$$

This sum can be rewritten as a sum over paths. Indeed, if we introduce

$$P_x = \left\{ \gamma: \{0, \dots, n\} \rightarrow \mathbb{Z}^d \mid \gamma_n = x \text{ and } \gamma_{t-1} - \gamma_t \in U \ \forall t \right\},$$

it is equivalent to the following compact expression:

$$T^n \mu = \sum_{x \in \mathbb{Z}^d} \sum_{\gamma \in P_x} \prod_{t=1}^n T^{(\gamma_t, \gamma_{t-1})} \Pi_{\gamma_0} \mu, \quad (14)$$

where the operators $T^{(\gamma_t, \gamma_{t-1})}$ have to be applied in chronological order.

In the case of a weakly interacting system, this sum over paths can be used to prove the existence of a unique invariant probability measure, under the assumptions of the Dobrushin criterion [5]. The system we are considering here is certainly not weakly interacting. However, in order to prove Theorem 1, the idea will be to take advantage of the decoupling in the pure phases conveyed in Assumption (A2). This assumption will provide upper bounds of order α on the couplings $T^{(\gamma_t, \gamma_{t-1})}$ in the positive phase. Indeed, as $a = +1$ in (13), those bounds will be obtained for the instants t such that the considered space-time configuration presents at time $t-1$ the value $+1$ for $\phi(\underline{\omega}_{\gamma_t+U})$. As for the negative phase, it will be shown to be infrequent enough, for a suitable choice of initial condition.

4 Pure Phase Expansion

For fixed $x \in \mathbb{Z}^d$ and $\gamma \in P_x$, if γ denotes both the function and its trajectory $\{(t, \gamma) \mid t = 0, \dots, n\}$, let us partition the trajectory into $\gamma^{(+)} \subseteq \gamma \setminus \{\gamma_0\}$ and $\gamma^{(-)} = (\gamma \setminus \{\gamma_0\}) \setminus \gamma^{(+)}$. We define, for $t = 0, \dots, n$, the sets $F(\gamma^{(+)}, t) \subseteq X$ with the following indicator functions:

$$\mathbb{1}_{F(\gamma^{(+)}, t)}(\underline{\omega}) = \prod_{x: (t+1, x) \in \gamma^{(+)}} \mathbb{1}_+(\phi_x(\underline{\omega})) \prod_{x: (t+1, x) \in \gamma^{(-)}} \mathbb{1}_-(\phi_x(\underline{\omega})).$$

These subsets lead to a partition of the configuration space $X^{\{0, \dots, n\}}$. Inserting this partition in (14),

$$T^n \mu = \sum_{x \in \mathbb{Z}^d} \sum_{\gamma \in P_x} \sum_{\gamma^{(+)} \subseteq \gamma \setminus \{\gamma_0\}} \prod_{t=1}^n \left(\mathbb{1}_{F(\gamma^{(+)}, t)} T^{(\gamma, \gamma_{t-1})} \right) \mathbb{1}_{F(\gamma^{(+)}, 0)} \Pi_{\gamma_0} \mu,$$

where $\mathbb{1}_{F(\gamma^{(+)}, n)}$ is nothing but the identity operator. Using property (9) together with the fact that the image of $\mathcal{M}(X)$ under $T^{(x, \gamma_{n-1})}$ is included in $\mathcal{M}_x$, we have, for all $\varphi \in \mathcal{C}(X)$,

$$|T^n \mu(\varphi)| \leq \sum_{x \in \mathbb{Z}^d} \sum_{\gamma \in P_x} \sum_{\gamma^{(+)} \subseteq \gamma \setminus \{\gamma_0\}} \left| \prod_{t=1}^n \left(\mathbb{1}_{F(\gamma^{(+)}, t)} T^{(\gamma, \gamma_{t-1})} \right) \mathbb{1}_{F(\gamma^{(+)}, 0)} \Pi_{\gamma_0} \mu \right| |\delta_x \varphi|_\infty. \quad (15)$$

Next, for given $x \in \mathbb{Z}^d$, $\gamma \in P_x$ and $\gamma^{(+)} \subseteq \gamma \setminus \{\gamma_0\}$, we define the set

$$\mathcal{E}(\gamma^{(+)}) = \left\{ (\underline{\omega}_t)_{t=0}^n \in X^{\{0, \dots, n\}} \mid \underline{\omega}_t \in F(\gamma^{(+)}, t) \quad \forall t \in \{0, \dots, n\} \right\}.$$

We will now build a graph using the construction in [13] in order to control the extent of the negative phase. We will then rewrite the expansion (15) as a cluster expansion in terms of a combination of paths and graphs.

5 Graphs

As we included the Toom erosion condition in the hypotheses of Theorem 1, we know that the intersection of all convex hulls of plus sets of the considered MBT is empty. Using Lemma 12 of [13], it implies the existence of an integer $q \in \mathbb{N}_0$, of q linear functions $L_1, \dots, L_q$ from $\mathbb{R}^{d+1}$ to $\mathbb{R}$ and of a positive r such that the hypotheses of Theorem 1 of [13] are verified by the MBT and by the homogeneous space-time configuration with positive spins everywhere. In particular, we can use the structures which are defined in the Peierls argument proving this stability theorem in Section II of [13]. We will not rewrite the whole construction here but only extract the objects of use for our purpose and their properties. We will also adapt the general formalism to the more restricted class of models under consideration here. We refer the reader to [13] for all details and, for an introduction to that very general proof, to the review given in

Appendix A of [7] in the particular case of the North-East-Center model. A good exercise could also be to consider the case of the Stavskaya model. One then obtains the contours described in Chapter 1 of [15].

In order to bound the probability of finding a negative spin at a given time t_- and at a given site x_- of the lattice, when the initial measure is $\delta^{(+)}$, [13] associates a structure called *truss* to each space-time configuration presenting this negative spin. Its iterative construction is given in Section II, Part 3 of [13] where it is denoted by Π . This truss consists of an unordered sequence of *edgers*, which are maps from $\{1, \dots, q\}$ to $\{0, \dots, t_-\} \times \mathbb{Z}^d$ whose image contains exactly two points. Here we use, much like in the review of [7], an equivalent reformulation. We view an edger as an unoriented edge that links two points of $\{0, \dots, t_-\} \times \mathbb{Z}^d$ and bears an extra attribute, namely a partition of the set $\{1, \dots, q\}$ of *poles* between its two vertices. This natural bijection between edgers and edges extends to trusses which we treat as graphs on $\{0, \dots, t_-\} \times \mathbb{Z}^d$ composed of these edges corresponding to their edgers. We name them *Toom graphs*.

The Toom graph G thus associated to the space-time configuration identifies part of the errors which lead to the negative spin at (t_-, x_-) : its set of vertices V_G admits only points where the given configuration presents negative spins and it contains (t_-, x_-) . It traces the source of this negative spin, via their propagation according to the deterministic rule conveyed by the function ϕ , back to some previous errors (when $\phi_x(\omega_{t-1}) = +1$ but $\omega_{t,x} = -1$). Such errors, classified in a distinguished subset $\hat{V}_G$ of V_G , all carry a probability smaller than ε , because of Assumption (A1). $\hat{V}_G$ is noted S in [13] and constructed iteratively parallel to the truss. The large number of different graphs associated to all possible space-time configurations with a negative spin at (t_-, x_-) is then an ‘entropic’ factor which has to be counterbalanced by an ‘energetic’ factor, i.e. the probability of the graphs, which is bounded by the probability of making all the identified errors, $\varepsilon^{|\hat{V}_G|}$.

Now, using our reformulation of trusses in terms of Toom graphs, the construction algorithm of the trusses in Section II, Part 3 of [13] guarantees the following properties:

- (P1) The graphs are connected and contain (t_-, x_-) , which is their only vertex at time coordinate t_- . We will call it their *origin*. Moreover, their edges are either *arrows*, with a displacement vector between their vertices of the form $\pm(-1, u)$ for a u in U , or *forks*, with a displacement vector $(0, u^{(1)} - u^{(2)})$ where $u^{(1)}$ and $u^{(2)}$ are both elements of U . Therefore, taking into account the 2^q possible allocations of the q poles to the two vertices of an edge, for any given point of $\{0, \dots, t_-\} \times \mathbb{Z}^d$ at most $2^q(R^2 + 2R)$ different types of edges can have that point as vertex. For any such connected graph, there always exists a walk which starts from (t_-, x_-) , passes along every edge exactly twice and then comes back to its departure point. We choose such a walk and consider the sequence that records, at each of its steps, the displacement vector and the pole distribution of the travelled edge, and call it the *Eulerian walk* associated to the graph. This correspondence is injective. Consequently, the number of different graphs grows only exponentially with their number of

edges, $|E_G|$, as the number of Eulerian walks of length $2|E_G|$ is less than $[2^q(R^2 + 2R)]^{2|E_G|}$. This is Lemma 3 of [13].

- (P2) The number $|E_G|$ of edges is in turn related to the number $|\hat{V}_G|$ of identified errors:

$$\frac{1}{1 + 2qr^{-1}} |E_G| + 1 \leq |\hat{V}_G|. \quad (16)$$

This is due to a principle which presides over the graph construction in [13] so as to take advantage of the erosion property: the truss is *overall even* and the number of forks equals the number of identified errors $|\hat{V}_G|$ minus one. Using Lemma 2 of [13], one deduces the upper bound (16).

Reference [13] uses these properties to prove that the ‘energetic’ factor wins as long as ε is small enough. The probability of finding a negative spin at a given time and a given site can be made as small as desired by taking ε appropriately small.

Here in our treatment of the space-time configurations that present, for all t such that $(t, \gamma) \in \gamma^{(-)}$, a negative value of the deterministic prescription given by ϕ evaluated at time $t - 1$ in $\gamma + U$, we will hardly need to modify Toom’s construction, but we will associate to each space-time configuration a collection of Toom graphs instead of only one, as we are interested in the source of a collection of negative spins, corresponding to the different elements of $\gamma^{(-)}$.

We define a map

$$\begin{aligned} g : \mathcal{E}(\gamma^{(+)}) &\rightarrow \mathcal{G}(\gamma^{(+)}) = g(\mathcal{E}(\gamma^{(+)})) \\ (\underline{\omega}_t)_{t=0}^n &\mapsto G = g((\underline{\omega}_t)_{t=0}^n) \end{aligned} \quad (17)$$

where G is a graph made of a disconnected collection of Toom graphs whose construction, for a given $(\underline{\omega}_t)_{t=0}^n \in \mathcal{E}(\gamma^{(+)})$, consists in the following steps.

1. We pick t_1 , the largest t such that $(t_1, \gamma_1) \in \gamma^{(-)}$. We know that $\phi_{\gamma_1}(\underline{\omega}_{t_1-1}) = -1$. We consider $(\tilde{\omega}_t)_{t=0}^n$ which is obtained from $(\underline{\omega}_t)_{t=0}^n$ by replacing ω_{t_1, γ_1} with the value -1 . We construct the Toom graph G_1 having the negative spin at (t_1, γ_1) as origin and which is associated to $(\tilde{\omega}_t)_{t=0}^n$ according to the algorithm in Section II, Part 3 of [13]. The whole algorithm remains valid here, the only difference being that no assumption about the initial condition forbids the presence of negative spins at $t = 0$ in $\tilde{\omega}_0$ while, in [13], the initial condition is $\delta^{(+)}$ and therefore the graph is always contained in $\{1, \dots, t_-\} \times \mathbb{Z}^d$ since the points where the configuration has positive spins can’t be vertices. A rule is then added to the algorithm: when the graph under construction reaches negative spins at $t = 0$, they are considered equivalent to errors and classed as elements of $\hat{V}_{G_1}$. In other words, this amounts only to a one-unit shift of the initial condition $\delta^{(+)}$ along the time axis. So we end up with a connected graph G_1 . Its set of vertices V_{G_1} contains (t_1, γ_1) and at all points $(t, x) \in V_{G_1} \setminus \{(t_1, \gamma_1)\}$, the space-time configuration $(\underline{\omega}_t)_{t=0}^n$ takes the value $\omega_{t,x} = -1$. It has a distinguished subset of vertices $\hat{V}_{G_1} \subseteq V_{G_1}$ where the space-time configuration presents errors:

$$\forall (t, x) \in \hat{V}_{G_1} \text{ such that } t \neq 0, \phi_x(\underline{\omega}_{t-1}) = +1.$$

It has a set E_{G_1} of edges connecting its vertices and satisfying (16).

G_1 is the first of the Toom graphs whose union will form the graph G .

2. We pick the next maximum $t_2 < t_1$ such that $(t_2, \gamma_{t_2}) \in \gamma^{(-)}$. We perform exactly the same construction as before and associate to (t_2, γ_{t_2}) the graph G_2 with properties analogous to those of G_1 . If $V_{G_1} \cap V_{G_2} \neq \emptyset$ then we discard G_2 , as we can't count any error twice. Otherwise the union of G_1 and G_2 will be part of G .
3. We repeat the same process up to the lowest time t such that $(t, \gamma_t) \in \gamma^{(-)}$, discarding any Toom graph G_i which intersects one of the previous retained ones.
4. G is defined as the union of all the retained G_i , $i \in \{1, \dots, |\gamma^{(-)}|\}$. V_G is the union of the retained V_{G_i} , $\hat{V}_G$ is the union of the retained $\hat{V}_{G_i}$ and E_G is the union of the retained E_{G_i} .

The map g of (17) is then completely defined. It is usually not injective, as there are many configurations corresponding to one graph. But $\mathcal{E}(\gamma^{(+)})$ can always be written as:

$$\mathcal{E}(\gamma^{(+)}) = \bigcup_{G \in \mathcal{G}(\gamma^{(+)})} g^{-1}G. \quad (18)$$

Defining, for all $G \in \mathcal{G}(\gamma^{(+)})$ and for $t = 0, \dots, n$, the subsets $E(G, t) \subseteq X$:

$$\mathbb{1}_{E(G, t)}(\underline{\omega}) = \prod_{x: (t, x) \in V_G \setminus \gamma^{(-)}} \mathbb{1}_-(\omega_x) \prod_{x: (t+1, x) \in \hat{V}_G} \mathbb{1}_+(\phi_x(\underline{\omega})), \quad (19)$$

we have, by construction of the map g ,

$$\mathbb{1}_{g^{-1}G}((\underline{\omega}_t)_{t=0}^n) \leq \prod_{t=0}^n \mathbb{1}_{E(G, t)}(\underline{\omega}_t). \quad (20)$$

Let us consider any $G \in \mathcal{G}(\gamma^{(+)})$. G is the union of c individually connected but pairwise disconnected Toom graphs $(G_{i_1}, \dots, G_{i_c})$. From this point on, they will be noted $(G_1, \dots, G_i, \dots, G_c)$. They all possess Properties (P1) and (P2). We show that G itself inherits similar properties.

First we find an upper bound on the number of different graphs with given numbers of edges and of connected parts. We define a map w on $\mathcal{G}(\gamma^{(+)})$. For $G \in \mathcal{G}(\gamma^{(+)})$, $w(G)$ consists of two elements: first, the list of the origins of $G_1, \dots, G_c$; second, the sequence given by the concatenation of the Eulerian walks associated to $G_1, \dots, G_c$ in Property (P1). This map $w : \mathcal{G}(\gamma^{(+)}) \rightarrow w(\mathcal{G}(\gamma^{(+)}))$ is injective.

Proof The induction procedure described in Section II, Part 3 of [13] for the construction of trusses starts from the vertex which we called the origin of the graph. One can check by inspection of the induction steps that the first step creates exactly q arrows leaving from this origin and then no other edge with this vertex will be drawn during the following steps of the construction. Consequently, we know that the Eulerian walk associated to a Toom graph G_i will leave from and come back to the origin of G_i exactly q times. Given any element of the set $w(\mathcal{G}(\gamma^{(+)}))$, its unique inverse image can then be deduced from the list of origins and the sequence of steps of the concatenated Eulerian walks by the following method. Starting from the first origin recorded in the

list, we add the steps of the sequence and redraw the first connected part of the graph, until the origin has been reached q times. Then we jump to the next origin recorded in the list and read on the sequence of steps until we again come back q times to this second departure point, and so on, until we have read the whole sequence. While redrawing a graph, we give back each travelled edge its pole distribution, which is recorded in the sequence.

For all $c \in \mathbb{N}$ and all $m \in \mathbb{N}$, let $\mathcal{G}(\gamma^{(+)}, c, m)$ be the subset of $\mathcal{G}(\gamma^{(+)})$ consisting of the graphs with c connected parts and m edges exactly. Any element of $w(\mathcal{G}(\gamma^{(+)}, c, m))$ is made of a list with c origins chosen among the elements of $\gamma^{(-)}$ and of a sequence with exactly $2m$ steps. Therefore Property (P1) is transferred from Toom graphs to graphs $G \in \mathcal{G}(\gamma^{(+)}, c, m)$:

$$|\mathcal{G}(\gamma^{(+)}, c, m)| \leq \binom{|\gamma^{(-)}|}{c} \cdot [2^q(R^2 + 2R)]^{2m} \leq 2^{|\gamma^{(-)}|} \cdot [2^q(R^2 + 2R)]^{2m}. \quad (21)$$

On the other hand, every connected part verifies Property (P2), that is to say, satisfies the analog of (16). Summing this inequality over all parts gives a similar property for G :

$$\frac{1}{1 + 2qr^{-1}} |E_G| + c \leq |\hat{V}_G|. \quad (22)$$

We also have

$$|\gamma^{(-)}| \leq |E_G| + c. \quad (23)$$

Proof It is sufficient to construct an injective map $f: \gamma^{(-)} \rightarrow E_G \cup \{1, \dots, c\}$. Let us consider any $(t, \gamma_t) \in \gamma^{(-)}$. If (t, γ_t) belongs to a connected part G_i of G which is contained in $\{0, \dots, t\} \times \mathbb{Z}^d$, then we know that (t, γ_t) is the unique origin of G_i . We assign to (t, γ_t) the image $f((t, \gamma_t)) = i \in \{1, \dots, c\}$. Otherwise, we know that the Toom graph with origin (t, γ_t) has been discarded. Then that Toom graph intersects at least one of the retained Toom graphs with origin (s, γ_s) , $s > t$. Now the discarded graph is contained in $\{0, \dots, t\} \times \mathbb{Z}^d$ and the time component of edges of Toom graphs has maximum absolute value 1, as can be seen in the definition of arrows and forks in Property (P1). Therefore there exists an edge of the retained graph which arrives onto some point (t, x) , $x \in \mathbb{Z}^d$. Such an edge belongs to E_G and we take it to be the image of (t, γ_t) under f . The map f thus defined is easily seen to be injective.

6 Exponential Convergence to Equilibrium

We now combine the collections of Toom graphs introduced in Section 5 with the paths of influence of Section 3 and their pure phase partition of Section 4. Inserting partition (18) in expansion (15) yields a sum over graphs. It introduces intricate couplings between the configurations at different times, due to the complexity of the Toom graph construction algorithm in Section II, Part 3 of [13]. Since $\mathbb{1}_{g^{-1}G}$ does not have the property of factorization over time, we use the upper bound (20) in the equivalent

form $\mathbb{1}_{g^{-1}G}((\underline{\omega}_t)_{t=0}^n) = \mathbb{1}_{g^{-1}G}((\underline{\omega}_t)_{t=0}^n) \cdot \prod_{t=0}^n \mathbb{1}_{E(G,t)}(\underline{\omega}_t)$, as indicator functions can only take values 0 or 1. Now, for all $(\underline{\omega}_t)_{t=1}^n$, we have

$$\begin{aligned} & \left| \Pi_{\gamma_0} \mu \left[\tau^{(\gamma_1, \gamma_0)}(\underline{\omega}_1, \cdot) \mathbb{1}_{F(\gamma^{(+)}, 0) \cap E(G, 0)}(\cdot) \mathbb{1}_{g^{-1}G}(\cdot; (\underline{\omega}_t)_{t=1}^n) \right] \right| \\ & \leq |\Pi_{\gamma_0} \mu| \left[\left| \tau^{(\gamma_1, \gamma_0)}(\underline{\omega}_1, \cdot) \right| \mathbb{1}_{F(\gamma^{(+)}, 0) \cap E(G, 0)}(\cdot) \right], \end{aligned}$$

where $|\Pi_{\gamma_0} \mu|$ is not the norm, but the absolute value of the measure $\Pi_{\gamma_0} \mu$. Consequently, if we define the operator $\tilde{T}^{(y, x)}$ by replacing $\tau^{(y, x)}(\underline{\xi}, \underline{\omega})$ with its absolute value $|\tau^{(y, x)}(\underline{\xi}, \underline{\omega})|$ in (12), the graph expansion gives

$$\begin{aligned} & \left| \prod_{t=1}^n \left(\mathbb{1}_{F(\gamma^{(+)}, t)} T^{(\gamma_t, \gamma_{t-1})} \right) \mathbb{1}_{F(\gamma^{(+)}, 0)} \Pi_{\gamma_0} \mu \right| \\ & \leq \sum_{G \in \mathcal{G}(\gamma^{(+)})} \left| \prod_{t=1}^n \left(\mathbb{1}_{F(\gamma^{(+)}, t) \cap E(G, t)} \tilde{T}^{(\gamma_t, \gamma_{t-1})} \right) \mathbb{1}_{F(\gamma^{(+)}, 0) \cap E(G, 0)} |\Pi_{\gamma_0} \mu| \right|. \end{aligned} \quad (24)$$

This expansion will be the starting point of the proof of Theorem 1. Before, we need the following lemmas.

Lemma 1 *For all $x \in \mathbb{Z}^d$, $\gamma \in P_x$, $\gamma^{(+)} \subseteq \gamma \setminus \{\gamma_0\}$ and $G \in \mathcal{G}(\gamma^{(+)})$, for all $v \in \mathcal{M}(X)$ and for all $t \in \{1, \dots, n\}$,*

$$\begin{aligned} & \left| \mathbb{1}_{F(\gamma^{(+)}, t) \cap E(G, t)} \tilde{T}^{(\gamma_t, \gamma_{t-1})} \mathbb{1}_{F(\gamma^{(+)}, t-1) \cap E(G, t-1)} v \right| \\ & \leq \left| \mathbb{1}_{F(\gamma^{(+)}, t-1) \cap E(G, t-1)} v \right| \varepsilon^{|\hat{V}_{G,t}|} \alpha^{|\gamma_t^{(+)}|} 2^{|\gamma_t^{(-)}|}, \end{aligned}$$

where we introduced the notation $\gamma_t^{(+)} = \gamma^{(+)} \cap \{(t, x) | x \in \mathbb{Z}^d\}$ and the analogs $\gamma_t^{(-)}$ and $\hat{V}_{G,t}$.

Proof Using definitions (12) and (19) of $T^{(\gamma_t, \gamma_{t-1})}$ and $\mathbb{1}_{E(G, t)}$, we already have

$$\begin{aligned} & \left| \mathbb{1}_{F(\gamma^{(+)}, t) \cap E(G, t)} \tilde{T}^{(\gamma_t, \gamma_{t-1})} \mathbb{1}_{F(\gamma^{(+)}, t-1) \cap E(G, t-1)} v \right| \\ & \leq \left| \mathbb{1}_{F(\gamma^{(+)}, t-1) \cap E(G, t-1)} v \right| \cdot \\ & \quad \sup_{\underline{\omega} \in F(\gamma^{(+)}, t-1) \cap E(G, t-1)} \int d\underline{\xi} \prod_{x: (t, x) \in \hat{V}_{G,t} \setminus \gamma^{(-)}} \mathbb{1}_{-}(\xi_x) \left| \tau^{(\gamma_t, \gamma_{t-1})}(\underline{\xi}, \underline{\omega}) \right|. \end{aligned} \quad (25)$$

Now, keeping definition (13) of $\tau^{(\gamma_t, \gamma_{t-1})}$ in mind, the contributions of the spins ξ_x at different sites $x \in \mathbb{Z}^d$ to the integral in (25) are decoupled and factorize. We first consider the set $\{x \in \mathbb{Z}^d : (t, x) \in \hat{V}_{G,t} \setminus \gamma\}$. Because of Assumption (A1), its elements all contribute by a factor bounded by ε , since the supremum in (25) is taken over configurations $\underline{\omega}$ in $E(G, t-1)$. Everywhere else on $\mathbb{Z}^d \setminus \{\gamma_t\}$, the contribution is trivially bounded by 1. For $x = \gamma_t$, we need upper bounds on $|p(\xi_{\gamma_t} | \underline{\omega}) - p(\xi_{\gamma_t} | \underline{\omega}_{\neq \gamma_{t-1}}, a)|$:

- if $\phi_{\gamma_t}(\underline{\omega}) = +1$,
 - and if $\xi_{\gamma_t} = -1$, $|p(\xi_{\gamma_t}|\underline{\omega}) - p(\xi_{\gamma_t}|\underline{\omega}_{\neq \gamma_{t-1}}, a)| \leq \alpha p(\xi_{\gamma_t}|\underline{\omega}) \leq \alpha \varepsilon$;
 - and if $\xi_{\gamma_t} = +1$, $|p(\xi_{\gamma_t}|\underline{\omega}) - p(\xi_{\gamma_t}|\underline{\omega}_{\neq \gamma_{t-1}}, a)| \leq \alpha p(\xi_{\gamma_t}|\underline{\omega})$,
 where we used Assumptions (A1) and (A2);
- if $\phi_{\gamma_t}(\underline{\omega}) = -1$, we use the trivial bound $|p(\xi_{\gamma_t}|\underline{\omega}) - p(\xi_{\gamma_t}|\underline{\omega}_{\neq \gamma_{t-1}}, a)| \leq 1$.

Therefore we obtain the following three types of upper bounds on the contribution of the spin at γ_t to the integral in (25): the contribution is bounded by $\alpha \varepsilon$ if $(t, \gamma_t) \in \gamma^{(+)} \cap \hat{V}_G$, by α if $(t, \gamma_t) \in \gamma^{(+)} \setminus \hat{V}_G$ or by 2 if $(t, \gamma_t) \in \gamma^{(-)}$. Inserting all these bounds in (25) proves Lemma 1.

Lemma 2 *For any MBT without memory characterized by a non-constant monotonic function ϕ verifying the Toom erosion condition, there exist $\alpha_* > 0$ and a positive and strictly decreasing function ε_* defined on $[0, \alpha_*[$ such that, for all $\alpha \in [0, \alpha_*[$ and $\varepsilon \in [0, \varepsilon_*(\alpha)[$, the following result is verified for any PCA satisfying the corresponding Assumptions (A1) and (A2). Let $K \geq 0$ and $\varepsilon' \in [0, \varepsilon_*(\alpha)[$. Let $\mu \in \mathcal{M}(X)$ be such that its absolute value $|\mu|$ is in the class $\mathcal{B}^{(+)}(K, \varepsilon')$. Then, there exist some constants $C < \infty$ and $\sigma < 1$ such that, for all $\varphi \in \mathcal{C}(X)$ and all $n \in \mathbb{N}$,*

$$\sum_{x \in \mathbb{Z}^d} \sum_{\gamma \in P_x} \sum_{\gamma^{(+)} \subseteq \gamma \setminus \{\gamma_0\}} \left| \prod_{t=1}^n \left(\mathbb{1}_{F(\gamma^{(+)}, t)} T^{(\gamma, \gamma_{t-1})} \right) \mathbb{1}_{F(\gamma^{(+)}, 0)} \Pi_{\gamma_0} \mu \right| |\delta_x \varphi|_{\infty} \leq C \|\varphi\| \sigma^n.$$

Proof Multiple uses of Lemma 1 on the RHS of (24) imply

$$\begin{aligned} & \left| \prod_{t=1}^n \left(\mathbb{1}_{F(\gamma^{(+)}, t) \cap E(G, t)} \tilde{T}^{(\gamma, \gamma_{t-1})} \right) \mathbb{1}_{F(\gamma^{(+)}, 0) \cap E(G, 0)} \Pi_{\gamma_0} \mu \right| \\ & \leq \varepsilon^{\sum_{t=1}^n |\hat{V}_{G, t}|} \alpha^{|\gamma^{(+)}|} 2^{|\gamma^{(-)}|} \left| \mathbb{1}_{F(\gamma^{(+)}, 0) \cap E(G, 0)} \Pi_{\gamma_0} \mu \right|. \end{aligned}$$

The important point is now that we assumed that $|\mu|$ belongs to $\mathcal{B}^{(+)}(K, \varepsilon')$. Indeed, since $\mathbb{1}_{E(G, 0)} \leq \mathbb{1}_{-, \hat{V}_{G, 0}}$ and $\mathbb{1}_{-, \hat{V}_{G, 0}}(\underline{\omega}_{\succeq \gamma_0}, \underline{a}_{\prec \gamma_0}) \leq \mathbb{1}_{-, \hat{V}_{G, 0}}(\underline{\omega})$, and using the definition of Π_{γ_0} , we obtain

$$\left| \mathbb{1}_{F(\gamma^{(+)}, 0) \cap E(G, 0)} \Pi_{\gamma_0} \mu \right| \leq 2K \varepsilon' |\hat{V}_{G, 0}|.$$

Consequently, with $\tilde{\varepsilon} = \max\{\varepsilon, \varepsilon'\}$,

$$\left| \prod_{t=1}^n \left(\mathbb{1}_{F(\gamma^{(+)}, t) \cap E(G, t)} \tilde{T}^{(\gamma, \gamma_{t-1})} \right) \mathbb{1}_{F(\gamma^{(+)}, 0) \cap E(G, 0)} \Pi_{\gamma_0} \mu \right| \leq 2K \tilde{\varepsilon}^{|\hat{V}_G|} \alpha^{|\gamma^{(+)}|} 2^{|\gamma^{(-)}|}. \quad (26)$$

Now for all $G \in \mathcal{G}(\gamma^{(+)})$ with c connected parts, (23) implies that the number of edges can be written as $|E_G| = |\gamma^{(-)}| - c + k$ with a certain $k \in \mathbb{N}$. Therefore $\mathcal{G}(\gamma^{(+)}) = \bigcup_{\substack{c \in \mathbb{N} \\ k \in \mathbb{N}}} \mathcal{G}(\gamma^{(+)}, c, |\gamma^{(-)}| - c + k)$ and, by virtue of the graph properties established above,

we have, using first equation (22) and then equation (21),

$$\begin{aligned}
\sum_{G \in \mathcal{G}(\gamma^{(+)})} \tilde{\varepsilon}^{|\hat{V}_G|} &\leq \sum_{c=0}^{\infty} \sum_{k=0}^{\infty} |\mathcal{G}(\gamma^{(+)}, c, |\gamma^{(-)}| - c + k)| \tilde{\varepsilon}^{\frac{1}{1+2qr^{-1}}(|\gamma^{(-)}| - c + k) + c} \\
&\leq \sum_{c=0}^{\infty} \sum_{k=0}^{\infty} 2^{|\gamma^{(-)}|} [2^q(R^2 + 2R)]^{2(|\gamma^{(-)}| - c + k)} \tilde{\varepsilon}^{\frac{1}{1+2qr^{-1}}(|\gamma^{(-)}| + 2qr^{-1}c + k)} \\
&\leq \frac{\left(2 \cdot [2^q(R^2 + 2R)]^2 \tilde{\varepsilon}^{\frac{1}{1+2qr^{-1}}}\right)^{|\gamma^{(-)}|}}{\left(1 - [2^q(R^2 + 2R)]^2 \tilde{\varepsilon}^{\frac{1}{1+2qr^{-1}}}\right) \left(1 - [2^q(R^2 + 2R)]^{-2} \tilde{\varepsilon}^{\frac{2qr^{-1}}{1+2qr^{-1}}}\right)},
\end{aligned} \tag{27}$$

provided that ε and ε' are such that $[2^q(R^2 + 2R)]^2 \tilde{\varepsilon}^{1/(1+2qr^{-1})} < 1$.

Combining (24), (26) and (27), we find

$$\begin{aligned}
&\sum_{x \in \mathbb{Z}^d} \sum_{\gamma \in P_x} \sum_{\gamma^{(+)} \subseteq \gamma \setminus \{\gamma_0\}} \left| \prod_{t=1}^n \left(\mathbb{1}_{F(\gamma^{(+)}, t)} T^{(\gamma, \gamma_{t-1})} \right) \mathbb{1}_{F(\gamma^{(+)}, 0)} \Pi_{\gamma_0} \mu \right| |\delta_x \varphi|_{\infty} \\
&\leq C \sum_{x \in \mathbb{Z}^d} \sum_{\gamma \in P_x} \sum_{\gamma^{(+)} \subseteq \gamma \setminus \{\gamma_0\}} \alpha^{|\gamma^{(+)}|} \left(4 \cdot [2^q(R^2 + 2R)]^2 \tilde{\varepsilon}^{\frac{1}{1+2qr^{-1}}} \right)^{|\gamma^{(-)}|} |\delta_x \varphi|_{\infty},
\end{aligned}$$

where

$$C = \frac{2K}{\left(1 - [2^q(R^2 + 2R)]^2 \tilde{\varepsilon}^{\frac{1}{1+2qr^{-1}}}\right) \left(1 - [2^q(R^2 + 2R)]^{-2} \tilde{\varepsilon}^{\frac{2qr^{-1}}{1+2qr^{-1}}}\right)}.$$

Here, we can use Newton's binomial formula: for any finite set A and any x and y in $\mathbb{R}$, $\sum_{B \subseteq A} x^{|B|} y^{|A \setminus B|} = (x + y)^{|A|}$. Finally, since $|P_x| = R^n$ and keeping definition (2) in mind, we get the desired upper bound if we take $\sigma = R(\alpha + 4 \cdot [2^q(R^2 + 2R)]^2 \tilde{\varepsilon}^{1/(1+2qr^{-1})})$. For σ to be lower than 1, the parameters must satisfy $\alpha < \alpha_* = 1/R$ and $\tilde{\varepsilon} = \max\{\varepsilon, \varepsilon'\} < \varepsilon_*(\alpha)$ where the function ε_* is strictly decreasing and positive on $[0, \alpha_*[$.

Using Lemmas 1 and 2 and the previous sections, we now prove Theorem 1.

Proof (Theorem 1) We take the constant α_* and the function ε_* obtained in Lemma 2. Let us consider any PCA satisfying Assumptions (A1) and (A2) with parameters $\alpha < \alpha_*$ and $\varepsilon < \varepsilon_*(\alpha)$. We write $T^n \mu(\varphi) - \mu_{\text{inv}}^{(+)}(\varphi) = T^n \bar{\mu}(\varphi)$ where $\bar{\mu} = \mu - \mu_{\text{inv}}^{(+)} \in \mathcal{M}(X)$ is a signed measure of zero mass. Therefore we can apply the above path expansion and pure phase expansion to $T^n \bar{\mu}$ which then satisfies (15). Using definition (5), $\bar{\mu}$ can be rewritten as $\bar{\mu} = \lim_{j \rightarrow \infty}^* \frac{1}{n_j} \sum_{k=0}^{n_j-1} (\mu - T^k \delta^{(+)})$, so that $|T^n \bar{\mu}(\varphi)|$ is

bounded above by

$$\liminf_{j \rightarrow \infty} \frac{1}{n_j} \sum_{k=0}^{n_j-1} \sum_{x \in \mathbb{Z}^d} \sum_{\gamma \in P_x} \sum_{\gamma^{(+)} \subseteq \gamma \setminus \{\gamma_0\}} \left[\prod_{t=1}^n \left(\mathbb{1}_{F(\gamma^{(+)}, t)} T^{(\gamma, \gamma_{t-1})} \right) \mathbb{1}_{F(\gamma^{(+)}, 0)} \Pi_{\gamma_0} \mu \right] \\ + \left[\prod_{t=1}^n \left(\mathbb{1}_{F(\gamma^{(+)}, t)} T^{(\gamma, \gamma_{t-1})} \right) \mathbb{1}_{F(\gamma^{(+)}, 0)} \Pi_{\gamma_0} T^k \delta^{(+)} \right] |\delta_x \varphi|_{\infty}. \quad (28)$$

As the probability measure μ belongs to $\mathcal{B}^{(+)}(K, \varepsilon')$ with $\varepsilon' < \varepsilon_*(\alpha)$, Lemma 2 applies to the first part of (28):

$$\sum_{x \in \mathbb{Z}^d} \sum_{\gamma \in P_x} \sum_{\gamma^{(+)} \subseteq \gamma \setminus \{\gamma_0\}} \left| \prod_{t=1}^n \left(\mathbb{1}_{F(\gamma^{(+)}, t)} T^{(\gamma, \gamma_{t-1})} \right) \mathbb{1}_{F(\gamma^{(+)}, 0)} \Pi_{\gamma_0} \mu \right| |\delta_x \varphi|_{\infty} \leq C \|\varphi\| \sigma^n, \quad (29)$$

where $C < \infty$ and $\sigma < 1$ are given in Lemma 2.

As for the second part of (28), it can be bounded thanks to a slight modification of the same arguments. We don't know whether all considered models have the property that $T^k \delta^{(+)}$ belongs to some $\mathcal{B}^{(+)}(K, \varepsilon')$. Therefore, we will now extend the Toom graphs up to time $-k$ instead of 0. For fixed $k \in \mathbb{N}$, $x \in \mathbb{Z}^d$, $\gamma \in P_x$ and $\gamma^{(+)} \subseteq \gamma \setminus \{\gamma_0\}$, we define

$$\mathcal{E}_k(\gamma^{(+)}) = \left\{ (\underline{\omega})_{t=-k}^n \in X^{\{-k, \dots, n\}} \mid \underline{\omega}_t \in F(\gamma^{(+)}, t) \quad \forall t \in \{0, \dots, n\} \right\}.$$

A map g_k is defined, similarly to the map g above:

$$g_k : \mathcal{E}_k(\gamma^{(+)}) \rightarrow \mathcal{G}_k(\gamma^{(+)}) = g_k(\mathcal{E}_k(\gamma^{(+)}) \\ (\underline{\omega})_{t=-k}^n \mapsto G = g_k((\underline{\omega})_{t=-k}^n).$$

The only change in the graph construction algorithm takes place whenever a Toom graph G_i under construction reaches a negative spin at time $t = 0$. Instead of classing it into the set of identified errors $\hat{V}_{G_i}$, we carry on the construction of this branch of G_i , as for positive times, until we meet either an error or a negative spin at time $-k$, which is now considered the initial time, and class it in $\hat{V}_{G_i}$. Again, this merely amounts to a translation of the initial condition $\delta^{(+)}$ by k units along the time axis. At the end, we obtain a disconnected union of Toom graphs on $\{-k, \dots, n\} \times \mathbb{Z}^d$, with origins $(t, \gamma_t) \in \gamma^{(-)}$ and with the properties described above. In particular, all the previous results about $\mathcal{G}(\gamma^{(+)})$ still hold for $\mathcal{G}_k(\gamma^{(+)})$ and for the associated subsets $\mathcal{G}_k(\gamma^{(+)}, c, m)$ of graphs with c connected parts and m edges.

Extending definition (19) of $E(G, t)$ to graphs $G \in \mathcal{G}_k(\gamma^{(+)})$ and to negative times, the analogs for g_k of (18) and (20) lead to the following graph expansion for the second

part of (28):

$$\begin{aligned}
& \left| \prod_{t=1}^n \left(\mathbb{1}_{F(\gamma^{(+)}, t)} T^{(\gamma, \gamma_{-1})} \right) \mathbb{1}_{F(\gamma^{(+)}, 0)} \Pi_{\gamma_0} T^k \delta^{(+)} \right| \\
& \leq \sum_{G \in \mathcal{G}_k(\gamma^{(+)})} \left| \prod_{t=1}^n \left(\mathbb{1}_{F(\gamma^{(+)}, t) \cap E(G, t)} \tilde{T}^{(\gamma, \gamma_{-1})} \right) \right. \\
& \quad \left. \circ \mathbb{1}_{F(\gamma^{(+)}, 0) \cap E(G, 0)} \Theta_{\gamma_0} \prod_{t=-k+1}^{-1} \left(\mathbb{1}_{E(G, t)} T \right) \mathbb{1}_{E(G, -k)} \delta^{(+)} \right|,
\end{aligned} \tag{30}$$

where we defined the operator Θ_x :

$$\Theta_x \mu(\varphi) = \int d\omega \int d\xi \varphi(\xi) \left| (\Pi_x \mathbb{P}[\cdot | \omega]) \left(\frac{\xi}{\omega} \right) \right| \mu(\omega),$$

with Π_x acting on the measure $\mathbb{P}[\cdot | \omega]$.

Lemma 1 as well extends immediately to graphs $G \in \mathcal{G}_k(\gamma^{(+)})$. Applying it n times to the RHS of (30), we find that it is bounded above by

$$\sum_{G \in \mathcal{G}_k(\gamma^{(+)})} \varepsilon^{\sum_{t=1}^n |\hat{V}_{G,t}|} \alpha^{|\gamma^{(+)}|} 2^{|\gamma^{(-)}|} \left| \mathbb{1}_{F(\gamma^{(+)}, 0) \cap E(G, 0)} \Theta_{\gamma_0} \prod_{t=-k+1}^{-1} \left(\mathbb{1}_{E(G, t)} T \right) \mathbb{1}_{E(G, -k)} \delta^{(+)} \right|.$$

Then we can use again Assumption (A1) k times. Indeed, we know from definition (19) that $\mathbb{1}_{E(G, t)} \leq \mathbb{1}_{-, \hat{V}_{G,t}}$ and from Assumption (A1) that $\left| \mathbb{1}_{-, \hat{V}_{G,t}} T \mathbb{1}_{E(G, t-1)} \mathbf{v} \right| \leq \varepsilon^{|\hat{V}_{G,t}|} \left| \mathbb{1}_{E(G, t-1)} \mathbf{v} \right|$ for any probability measure $\mathbf{v}$. The operator Θ_{γ_0} can be handled in the same way as T , taking its definition into account, together with the fact that $\mathbb{1}_{-, \hat{V}_{G,0}}(\omega_{\succeq x}, a_{\prec x}) \leq \mathbb{1}_{-, \hat{V}_{G,0}}(\omega)$. It will simply introduce an extra factor of 2 due to the operator Π_{γ_0} . Lastly, by definition of $\delta^{(+)}$, $\delta^{(+)}(\mathbb{1}_{E(G, -k)}) = 0$ unless $\hat{V}_{G, -k}$ is empty, in which case $\delta^{(+)}(\mathbb{1}_{E(G, -k)}) = 1$. Consequently, the RHS of (30) is lower than

$$2 \sum_{G \in \mathcal{G}_k(\gamma^{(+)})} \varepsilon^{|\hat{V}_G|} \alpha^{|\gamma^{(+)}|} 2^{|\gamma^{(-)}|}.$$

And so, performing the same calculations as for the proof of Lemma 2, we obtain, for the second part of (28),

$$\begin{aligned}
& \sum_{x \in \mathbb{Z}^d} \sum_{\gamma \in P_x} \sum_{\gamma^{(+)} \subseteq \gamma \setminus \{\gamma_0\}} \left| \prod_{t=1}^n \left(\mathbb{1}_{F(\gamma^{(+)}, t)} T^{(\gamma, \gamma_{-1})} \right) \mathbb{1}_{F(\gamma^{(+)}, 0)} \Pi_{\gamma_0} T^k \delta^{(+)} \right| |\delta_x \varphi|_{\infty} \\
& \leq C_{\text{inv}} \|\varphi\| \sigma^n,
\end{aligned} \tag{31}$$

where

$$C_{\text{inv}} = \frac{2}{\left(1 - [2^q(R^2 + 2R)]^2 \varepsilon^{\frac{1}{1+2qr-1}} \right) \left(1 - [2^q(R^2 + 2R)]^{-2} \varepsilon^{\frac{2qr-1}{1+2qr-1}} \right)}.$$

Inserting (29) and (31) in (28), we conclude:

$$\left| T^n \mu(\varphi) - \mu_{\text{inv}}^{(+)}(\varphi) \right| \leq C \|\varphi\| \sigma^n,$$

where we renamed $C + C_{\text{inv}}$ to C for simplicity.

7 Exponential Decay of Correlations

We will now prove a well-known consequence of Theorem 1: the invariant measure $\mu_{\text{inv}}^{(+)}$ presents exponential decay of correlations in space.

Proof (Corollary) Since $\delta^{(+)}$ belongs to $\mathcal{B}^{(+)}(1, 0)$, Theorem 1 implies that for some $C < \infty$ and some $\sigma < 1$, we have, for all $k \in \mathbb{N}$,

$$\begin{cases} \left| T^k \delta^{(+)}(\varphi \psi) - \mu_{\text{inv}}^{(+)}(\varphi \psi) \right| \leq C \|\varphi \psi\| \sigma^k; \\ \left| T^k \delta^{(+)}(\varphi) - \mu_{\text{inv}}^{(+)}(\varphi) \right| \leq C \|\varphi\| \sigma^k; \\ \left| T^k \delta^{(+)}(\psi) - \mu_{\text{inv}}^{(+)}(\psi) \right| \leq C \|\psi\| \sigma^k. \end{cases}$$

But $\delta^{(+)}$ is a product measure and the interactions are local, so

$$T^k \delta^{(+)}(\varphi \psi) = T^k \delta^{(+)}(\varphi) T^k \delta^{(+)}(\psi)$$

as long as $2k \max_{u \in U} \|u\|_1 < d(\varphi, \psi)$, where $\|\cdot\|_1$ is the Manhattan norm. Since $d(\varphi, \psi) > 0$, we also have

$$\|\varphi \psi\| \leq \|\varphi\| \|\psi\|_\infty + |\varphi|_\infty \|\psi\| < \infty.$$

Hence, as long as $2k \max_{u \in U} \|u\|_1 < d(\varphi, \psi)$,

$$\left| \mu_{\text{inv}}^{(+)}(\varphi \psi) - \mu_{\text{inv}}^{(+)}(\varphi) \mu_{\text{inv}}^{(+)}(\psi) \right| \leq 2C (\|\varphi\| \|\psi\|_\infty + |\varphi|_\infty \|\psi\|) \sigma^k.$$

If we take $k = \lceil d(\varphi, \psi) / 2 \max_{u \in U} \|u\|_1 \rceil - 1$ and choose $C' = \frac{2C}{\sigma}$ and $\eta = \sigma^{\frac{1}{2 \max_{u \in U} \|u\|_1}}$, we obtain the desired inequality.

The invariant measure $\mu_{\text{inv}}^{(+)}$ also exhibits exponential decay of correlations in time.

Proof (Theorem 2) Theorem 1 applied to $\delta^{(+)} \in \mathcal{B}^{(+)}(1, 0)$ implies that, for all $k \in \mathbb{N}$,

$$\left| \mu_{\text{inv}}^{(+)}(\varphi T^n \psi) - T^k \delta^{(+)}(\varphi T^n \psi) \right| \leq C \|\varphi T^n \psi\| \sigma^k. \quad (32)$$

Now, remembering definition (2) of the semi-norm $\|\cdot\|$, we notice that

$$\|\varphi T^n \psi\| \leq 2 |\varphi|_\infty |\psi|_\infty |\text{supp}(\varphi T^n \psi)|.$$

But the conical structure of the space-time influence of spins implies

$$|\text{supp}(\varphi T^n \psi)| \leq |\text{supp}(\varphi)| + (\text{diam}(\text{supp}(\psi)) + 2n \max_{u \in U} \|u\|_1)^d,$$

so there exists a finite constant $C'_{\varphi, \psi}$ such that $\|\varphi T^n \psi\| \leq C'_{\varphi, \psi} n^d$ for all $n \in \mathbb{N}_0$.

Theorem 1 also yields

$$\left| \mu_{\text{inv}}^{(+)}(\varphi) \mu_{\text{inv}}^{(+)}(\psi) - \mu_{\text{inv}}^{(+)}(\varphi) T^{k+n} \delta^{(+)}(\psi) \right| \leq C |\varphi|_{\infty} \|\psi\| \sigma^{k+n}; \quad (33)$$

$$\left| \mu_{\text{inv}}^{(+)}(\varphi) T^{k+n} \delta^{(+)}(\psi) - T^k \delta^{(+)}(\varphi) T^{k+n} \delta^{(+)}(\psi) \right| \leq C |\psi|_{\infty} \|\varphi\| \sigma^k. \quad (34)$$

In order to find an upper bound for $|T^k \delta^{(+)}((\varphi - T^k \delta^{(+)}(\varphi)) T^n \psi)|$, we rewrite it as $|T^n \mu_{\varphi}^{(k)}(\psi)|$, where the signed measure $\mu_{\varphi}^{(k)} \in \mathcal{M}(X)$ is defined by

$$\mu_{\varphi}^{(k)}(\psi) = T^k \delta^{(+)}\left(\left(\varphi - T^k \delta^{(+)}(\varphi)\right) \psi\right) \quad \forall \psi \in \mathcal{C}(X).$$

$\mu_{\varphi}^{(k)}$ is a measure of zero mass so it satisfies (15):

$$\left| T^n \mu_{\varphi}^{(k)}(\psi) \right| \leq \sum_{x \in \mathbb{Z}^d} \sum_{\gamma \in P_x} \sum_{\gamma^{(+)} \subseteq \gamma \setminus \{\gamma_0\}} \left| \prod_{l=1}^n \left(\mathbb{1}_{F(\gamma^{(+)}, l)} T^{(\gamma, \gamma_{l-1})} \right) \mathbb{1}_{F(\gamma^{(+)}, 0)} \Pi_{\gamma_0} \mu_{\varphi}^{(k)} \right| |\delta_x \psi|_{\infty}.$$

The last expression is similar to the second term of (28) with $\mu_{\varphi}^{(k)}$ instead of $T^k \delta^{(+)}$ and we can then perform the same argument as in the proof of Theorem 1 in order to bound $|T^n \mu_{\varphi}^{(k)}(\psi)|$, extending again the collections of Toom graphs up to time $-k$.

The only change in these calculations is the presence of an extra factor $\varphi - T^k \delta^{(+)}(\varphi)$ whose supremum norm is bounded by $2|\varphi|_{\infty}$. Provided we keep track of this factor, the calculations which lead to (31) still hold here:

$$\left| T^k \delta^{(+)}\left(\left(\varphi - T^k \delta^{(+)}(\varphi)\right) T^n \psi\right) \right| \leq 2 C_{\text{inv}} |\varphi|_{\infty} \|\psi\| \sigma^n. \quad (35)$$

Combining (32), (33), (34) and (35) and taking the $k \rightarrow \infty$ limit ends the proof.

Theorem 2 implies that $\mu_{\text{inv}}^{(+)}$ is not only ergodic but also mixing, that is, for any continuous functions φ, ψ , we have

$$\lim_{n \rightarrow \infty} \left| \mu_{\text{inv}}^{(+)}(\varphi T^n \psi) - \mu_{\text{inv}}^{(+)}(\varphi) \mu_{\text{inv}}^{(+)}(\psi) \right| = 0.$$

Indeed, by definition of $\mathcal{C}(X)$, the set of continuous functions with finite support is dense in $\mathcal{C}(X)$.

Acknowledgements The authors would like to thank Jean Bricmont, Carlangelo Liverani and Christian Maes for helpful comments and discussions. They also thank the referees for their constructive remarks and suggestions.

Augustin de Maere was partially supported by the Belgian IAP (Interuniversity Attraction Pole) program P6/02.

Lise Ponselet was supported by a grant from the Belgian F.R.S.-FNRS (Fonds de la Recherche Scientifique) as 'Aspirant FNRS'.

This article is published in J. Stat. Phys. 147(3), 634-652 (2012). The final publication is available at www.springerlink.com. DOI 10.1007/s10955-012-0487-9

References

1. Bennett, C.H., Grinstein, G.: Role of irreversibility in stabilizing complex and nonergodic behavior in locally interacting discrete systems. *Phys. Rev. Let.* **55**(7), 657–660 (1985)
2. Berezner, S., Krutina, M., Malyshev, V.: Exponential convergence of Toom’s probabilistic cellular automata. *J. Stat. Phys.* **73**(5-6), 927–944 (1993). DOI 10.1007/BF01052816
3. Bramson, M., Gray, L.: A useful renormalization argument. In: R. Durrett, H. Kesten (eds.) *Random walks, Brownian motion, and interacting particle systems: a festschrift in honor of Frank Spitzer*, Progress in probability. Birkhäuser (1991)
4. Diakonova, M., MacKay, R.: Mathematical examples of space-time phases. *Int. J. Bifurcat. Chaos* **21**(8), 2297–2304 (2011)
5. Dobrushin, R.: Markov processes with a large number of locally interacting components: existence of a limit process and its ergodicity. *Probl. Inf. Transm.* **7**(2), 149–164 (1971)
6. Keller, G., Liverani, C.: Uniqueness of the SRB measure for piecewise expanding weakly coupled map lattices in any dimension. *Commun. Math. Phys.* **262**(1), 33–50 (2006)
7. Lebowitz, J., Maes, C., Speer, E.: Statistical mechanics of probabilistic cellular automata. *J. Stat. Phys.* **59**, 117–170 (1990)
8. de Maere, A.: Phase transition and correlation decay in coupled map lattices. *Commun. Math. Phys.* **297**(1), 229–264 (2010). DOI 10.1007/s00220-010-1041-8
9. Makowiec, D.: Toom probabilistic cellular automata stationary states via simulations. *Acta. Phys. Pol. B* **29**(6), 1599–1607 (1998)
10. Makowiec, D.: Stationary states of Toom cellular automata in simulations. *Phys. Rev. E* **60**(4), 3787–3796 (1999)
11. Stavskaya, O., Piatetski-Shapiro, I.I.: On homogeneous nets of spontaneously active elements. *Systems Theory Res.* **20**, 75–88 (1971)
12. Toom, A.: Monotonic binary cellular automata. *Probl. Inf. Transm.* **12**(1), 33–37 (1976)
13. Toom, A.: Stable and attractive trajectories in multicomponent systems. *Adv. Probab. Relat. Top.* **6**, 549–575 (1980)
14. Toom, A.: Cellular automata with errors: problems for students of probability. In: L. Snell (ed.) *Topics in contemporary probability and its applications*, Probability and stochastics series. CRC Press (1995)
15. Toom, A., Vasilyev, N.B., Stavskaya, O., Mityushin, L.G., Kurdyumov, G.L., Pirogov, S.A.: Discrete local Markov systems. In: R. Dobrushin, V. Kryukov, A. Toom (eds.) *Stochastic cellular systems: ergodicity, memory, morphogenesis*. Manchester University Press (1990)
16. Vasilyev, N.B., Petrovskaya, M.B., Piatetski-Shapiro, I.I.: Simulation of voting with random errors. *Automat. Remote Control* **30**(10), 1639–1642 (1969)