Electronic Journal of Probability

Electron. J. Probab. **28** (2023), article no. 25, 1–38.
ISSN: 1083-6489 https://doi.org/10.1214/23-EJP917

On the Besicovitch-stability of noisy random tilings*

Gayral Léo† Sablik Mathieu‡

Abstract

In this paper, we introduce a framework for studying a subshift of finite type (SFT) with noise, allowing some amount of forbidden patterns to appear. Using the Besicovitch distance, which permits a global comparison of configurations, we then study the closeness of measures on noisy configurations to the non-noisy case as the amount of noise goes to 0. Our first main result is the full classification of the (in)stability in the one-dimensional case. Our second main result is a stability property under Bernoulli noise for higher-dimensional periodic SFTs, which we finally extend to an aperiodic example through a variant of the Robinson tiling.

Keywords: subshift of finite type; stability; Besicovitch distance; percolation; Robinson tiling.
MSC2020 subject classifications: Primary 37B51;37A50, Secondary 60K35;82B43.
Submitted to EJP on September 10, 2021, final version accepted on February 7, 2023.
Supersedes arXiv:2104.09885v2.
Supersedes HAL:hal-03203745.

1 Introduction

A subshift of finite type (usually called a *SFT*) is the set of tilings induced by a finite tileset with local rules. More formally, given a finite set of forbidden patterns on a finite alphabet $\mathcal{A}$, the corresponding SFT is the set of $\mathcal{A}$-colourings of $\mathbb{Z}^d$ where no forbidden pattern appears. In the last decades, there has been numerous studies on how local constraints affect the global structure of tilings in non-trivial ways. The most important property studied is aperiodicity: there exists some local rules which impose that any tiling (or configuration) of the SFT is not periodic [6, 22, 16]. Such SFTs are said to be *aperiodic*.

Aperiodic SFTs received a strong wave of interest when non-periodic crystals (now called *quasicrystals*) were discovered by the chemists Dan Shechtman *et al.* [24]. The

*This work was supported by the ANR "Difference" project (ANR-20-CE40-0002) and the CIMI LabEx "Computability of asymptotic properties of dynamical systems" project (ANR-11-LABX-0040). The second author wishes to thank Siamak Taati for many fruitful references on the stability of tilings.

†University Toulouse III – Paul Sabatier, France.
E-mail: leo.gayral@math.cnrs.fr lgayral.pages.math.cnrs.fr
‡University Toulouse III – Paul Sabatier, France.
E-mail: msablik@math.univ-toulouse.fr math.univ-toulouse.fr/~msablik

connection with tilings was indeed quickly made, with tiles representing atom clusters, and forbidden patterns modelling constraints on the way these atoms can fit together, *e.g.* finite range energetic interactions between them [18]. In computer science, tilings have been used as static geometrical models of computation, ever since Berger proved the undecidability of the so-called *domino problem* by implementing Turing machines in aperiodic tilings [6]. More recently, simulations of Turing machines have been implemented in several different ways in order to construct *complex* tilings [14, 3, 9]. Yet, in each construction, the aperiodic structure of the tiling is the key element to embed computations.

A natural question is whether aperiodic structures in SFTs survive in the presence of some amount of noise, considering how real-life quasicrystals can have some defects. In other words, we want to know if a configuration with few mistakes is structurally close to a generic configuration of the SFT up to a small amount of changes (ideally proportional to the number of mistakes). A first step in this direction is the construction of a three-dimensional model with infinite-range interactions which, at low positive temperature, enforces the Thue-Morse sequence along one direction [26]. Empirical evidence, obtained through Monte Carlo simulations, also suggests that even in two dimensions, using Ammann's aperiodic tileset (famous for using only 16 tiles), valid tilings remain stable at sufficiently low positive temperatures [1]. In a more formal way, there is a four-dimensional model with finite-range interactions which, at low but positive temperatures, admits Gibbs measures that are perturbations of Ammann's aperiodic tiling along two directions [25].

The question is also natural in the context of tilings as static models of computation. Indeed, for a given model of computation, we want to know whether the model is robust to errors, as in the case of cellular automata [12] or Turing machines [2]. We consider here ε-Bernoulli noises, such that all the cells may violate local rules with probability ε, independently from each other. In the case of general tilings, fixed-point methods allow us to construct an aperiodic SFT that is robust in the presence of Bernoulli noises [9]. In that case, robustness means that for any proportion $\alpha > 0$, for small-enough values of $\varepsilon > 0$, any typical noisy configuration is close to some valid configuration of the SFT, up to a subset of cells of density at most α in $\mathbb{Z}^2$. The key of the proof is to obtain a tileset where we can repair *islands* of errors, delimited areas containing forbidden patterns with a large neighbourhood without any other violation of local rules. Such matters will be discussed again in Section 6.

In this paper, we provide a quantitative formalism of error robustness for a given choice of local rules. In other words, we allow a small proportion of cells of $\mathbb{Z}^d$ to violate the local rules, and we want to quantify how close a generic noisy configuration is to a generic non-noisy configuration. To do so, it is easier to use a distance on the associated measure spaces. Formally, we will say that an SFT is f-stable for a given distance on the probability measure space if any shift-invariant measure with a proportion (at most) ε of errors is at distance at most $f(\varepsilon)$ of the set of non-noisy invariant measures on the SFT. This formalism will be precisely introduced in Section 2, up to the notion of stability in the beginning of Section 3.

Closeness in the weak-* topology is not enough, because it only characterises a high probability of agreement on a large *finite* box around the origin, but does not say anything about the symbols that are arbitrarily far from the origin. Thus, when the proportion of errors goes to zero, the shift-invariant noisy measures must necessarily converge to non-noisy measures of the SFT, which we demonstrate in Subsection 2.2. Yet, we can exhibit generic noisy configurations, with an arbitrarily small proportion of errors ε, which we cannot superpose with any configuration of the SFT on a high-density subset of $\mathbb{Z}^d$.

In order to compare configurations of $\mathcal{A}^{\mathbb{Z}^d}$ in a global way, the Hamming-Besicovitch pseudo-distance [7] – the density of the set of differences between two configurations in $\mathbb{Z}^d$ – is a natural approach, as it gives the same importance to all cells. It is possible to transpose this pseudo-distance into a genuine distance on the set of measures, in a similar fashion to the Kantorovich metric [27]. This process is detailed in Subsection 3.1.

In this paper, after introducing the general framework of stability, we prove in Subsection 3.2 that this notion is an invariant of conjugacy. Thus the stability of a tiling does not depend on the local rules of the tileset used to define it. In Section 4 we characterise which one-dimensional SFTs are stable (Theorem 4.8). It is well known that a one-dimensional SFT is represented as the set of bi-infinite paths in a labelled transition graph [21], also called an *automaton*. If this automaton has an aperiodic structure, then it is possible to locally correct the mistakes to obtain a linear $O(\varepsilon)$-stability. On the contrary, if it has a periodic structure, it is impossible to repair a mistake at the interface between two phases misaligned within the automaton. These one-dimensional (un)stable examples can then be extended to (un)stable SFTs in any dimension with Corollary 4.13.

Unlike the one-dimensional case, we prove in Section 5 that bi-dimensional strongly periodic SFTs are linearly stable (Theorem 5.7). The main idea here is that, if local rules are respected on some region of $\mathbb{Z}^2$, then this region is the restriction of a periodic configuration, except maybe on the boundary of the region. A percolation argument then allows us to prove the uniqueness of such an *infinite* region, what's more with a linear control on its density. This linear $O(\varepsilon)$-stability result is to put into perspective with the $O\left(1/\sqrt{\ln(1/\varepsilon)}\right)$-stability obtained in Section 6, using the strategy for tilesets with robust combinatorial properties described by Durand *et al.* [9], which holds in particular for periodic tilings [5].

In Section 7 we consider the famous Robinson aperiodic tiling [22] and, up to some modifications, we show that it is $O\left(\sqrt[3]{\varepsilon}\right)$-stable (Theorem 7.10). The key idea here is that Robinson configurations are almost periodic, up to a low-density grid of cells, so that we may adapt the periodic percolation argument from Section 5. This result is interesting for two reasons. First, the Robinson tileset is not combinatorially robust in the sense of Durand, Romashchenko and Shen [9], so it provides a new, perhaps simpler example of stable aperiodic tiling (recall that stability is a metric property). Second, the speed obtained here, though not linear, is still polynomial, thus much faster than the one one from Section 6. The question of whether we can achieve linear stability for some aperiodic SFT remains open.

2 Noisy framework and weak-* stability

2.1 Noisy framework

Definition 2.1 (Configuration Space)**.** *Consider the lattice $\mathbb{Z}^d$ with positive dimension $d \in \mathbb{N}^* := \{1, 2, 3 \dots\}$, and a finite alphabet $\mathcal{A}$. The full-shift configuration space is $\Omega_{\mathcal{A}} = \mathcal{A}^{\mathbb{Z}^d}$.*

We endow this space with the discrete product topology. In this framework, the clopen cylinders $[w] = \{\omega \in \Omega_{\mathcal{A}}, \omega|_I = w\}$ form a countable base of the topology, where $w \in \mathcal{A}^I$ is a finite pattern *over a* window *$I \subsetneq \mathbb{Z}^d$, a* finite *set of cells. Consequently, we use the induced Borel algebra on this space.*

For any vector $k \in \mathbb{Z}^d$, let us define the shift $\sigma_k : \Omega_{\mathcal{A}} \to \Omega_{\mathcal{A}}$ such that $\sigma_k(\omega)_l = \omega_{k+l}$. Likewise, we can apply σ_k to any finite pattern or even a range of cells.

Let us denote $(e_i)_{1\leq i\leq d}$ the canonical basis of $\mathbb{Z}^d$. Then $(\Omega_{\mathcal{A}}, \sigma_{e_1}, \dots, \sigma_{e_d})$ forms a commutative dynamical system.

Definition 2.2 (Subshift of Finite Type)**.** *A* subshift *of $\Omega_{\mathcal{A}}$ is a σ-invariant subset,* i.e. *a* *subset stable under the action of any shift σ_k with $k \in \mathbb{Z}^d$.*

Let $\mathcal{F}$ be a finite set of forbidden *(finite) patterns $w \in \mathcal{A}^{I(w)}$. A SFT is the subshift $\Omega_{\mathcal{F}}$ induced by such a set $\mathcal{F}$ as follows:*

$$\Omega_{\mathcal{F}} := \left\{\omega \in \Omega_{\mathcal{A}}, \forall w \in \mathcal{F}, \forall k \in \mathbb{Z}^d, \sigma_k(\omega)|_{I(w)} \neq w\right\}.$$

In other words, the configurations of the SFT are the ones where no forbidden pattern occurs anywhere. Then $(\Omega_{\mathcal{F}}, \sigma_{e_1}, \ldots, \sigma_{e_d})$ is a commutative dynamical subsystem of the full-shift $\Omega_{\mathcal{A}}$.

To be exact, we should specify *both* the alphabet $\mathcal{A}$ and the set of forbidden patterns $\mathcal{F}$ when talking about an SFT, so that $\Omega_{\mathcal{F}} := \Omega_{\mathcal{A},\mathcal{F}}$ and in particular $\Omega_{\mathcal{A}} := \Omega_{\mathcal{A},\emptyset}$ for the full-shift. We chose purposefully the shortened notations for brevity's sake. The same remark holds for the measure spaces in Definition 2.5. In this article we will consistently use *tileset* to refer to local (combinatorial) behaviours and *tiling* to refer to global (dynamical) behaviours.

Definition 2.3 (Locally and Globally Admissible)**.** *A pattern or configuration on $\mathcal{A}$ will be called* locally admissible *whenever it contains no forbidden clear pattern from $\mathcal{F}$. It will be called* globally admissible *when it is the restriction of an actual configuration $\omega \in \Omega_{\mathcal{F}}$.*

In general, a *locally admissible pattern* is not necessarily a *globally admissible* one. The nuance between the two notions is highlighted in the following remark.

Remark 2.4 (Reconstruction Function)**.** Consider $\varphi : \mathcal{P}_F\left(\mathbb{Z}^d\right) \to \mathbb{N}$ the reconstruction function defined on finite windows $I \subset \mathbb{Z}^d$ by:

$$\varphi(I) = \inf\left\{k \in \mathbb{N}, w \in \mathcal{A}^{I+B_k} \text{ is locally admissible } \Rightarrow w|_I \text{ is globally admissible}\right\},$$

with $B_k = [\![-k,k]\!]^d$ the ball of radius $k \in \mathbb{N}$ (for the $\|.\|_\infty$ norm), thus a $(2k+1)$-square.

A priori $\varphi(I)$ could be infinite. However, as $\mathcal{A}^I$ is finite, consider $L_I \subset \mathcal{A}^I$ the finite subset of locally admissible patterns that are *not* globally admissible. If $v \in L_I$ could be embedded into arbitrarily large admissible patterns, then by compactness it would be globally admissible. In other words, there exists a rank $k(v) \in \mathbb{N}$ after which there is no locally admissible $w \in \mathcal{A}^{I+B_k}$ such that $w|_I = v$. This holds for all the patterns in L_I, so that $\varphi(I) = \max_{v \in L_I} k(v) < \infty$.

As we can embed Turing machines into SFTs [22], the function $(I, \mathcal{F}) \mapsto \varphi_{\mathcal{F}}(I)$ is a sort of non-computable busy beaver. We will see later on specific choices of $\mathcal{F}$ for which φ is bounded. This function may seem anecdotal at first, but it will in fact appear in some way or another in most of the next sections, and be a fundamental tool in all of our main results.

For a measure μ on Ω and a measurable mapping $\theta : \Omega \to \Omega'$, we denote $\theta^*(\mu)$ the pushforward measure on Ω' such that $[\theta^*(\mu)](B) = \mu\left(\theta^{-1}(B)\right)$ for any measurable set B.

Definition 2.5 (Invariant Probability Measures)**.** *A measure μ is said to be σ-invariant if, for any $k \in \mathbb{Z}^d$, $\sigma_k^*(\mu) = \mu \circ \sigma_{-k}$ is equal to μ. For a given SFT induced by $\mathcal{F}$, we denote $\mathcal{M}_{\mathcal{F}}$ the set of σ-invariant probability measures on the space $\Omega_{\mathcal{F}}$, and $\mathcal{M}_{\mathcal{A}}$ for all the σ-invariant probability measures on the full-shift $\Omega_{\mathcal{A}}$.*

By compactness, $\mathcal{M}_{\mathcal{F}}$ is non-empty as long as $\Omega_{\mathcal{F}} \neq \emptyset$. From now on, We will always assume that $\Omega_{\mathcal{F}}$ is non-empty. Let us now introduce our noisy *clair-obscur* framework:

Definition 2.6 (Noisy SFT)**.** *Consider the alphabet $\widetilde{\mathcal{A}} = \mathcal{A} \times \{0,1\}$. For a given cell value $(a,b) \in \widetilde{\mathcal{A}}$, whenever $b = 0$ we say that the cell is* clear*, and when $b = 1$ we say that the cell is* obscured*. We may thus identify $\mathcal{A}$ to the clear subset $\mathcal{A} \times \{0\} \subsetneq \widetilde{\mathcal{A}}$. Formally, we will denote $\pi_1 : \widetilde{\mathcal{A}} \to \mathcal{A}$ and $\pi_2 : \widetilde{\mathcal{A}} \to \{0,1\}$ the canonical projections.*

By extension, we will call patterns $w \in \widetilde{\mathcal{A}}^I$ and configurations $\omega \in \Omega_{\widetilde{\mathcal{A}}}$ clear *when they are actually defined on the alphabet $\mathcal{A}$, as opposed to the* obscured *ones, that contain at least one* obscured *letter in $\widetilde{\mathcal{A}} \backslash \mathcal{A}$.*

Using the same identification, we can define the set of forbidden clear patterns as $\widetilde{\mathcal{F}} = \left\{ \left(w, 0^{I(w)}\right) \in \widetilde{\mathcal{A}}^{I(w)}, w \in \mathcal{F} \right\}$ and the corresponding SFT on the space $\Omega_{\widetilde{\mathcal{F}}} \subset \Omega_{\widetilde{\mathcal{A}}}$.

Note that, by filling the rest of the space with obscured cells, any locally admissible pattern for $\widetilde{\mathcal{F}}$ is globally admissible. Thus, when working on the noisy case on $\Omega_{\widetilde{\mathcal{F}}}$, We will purposefully set aside the term *globally admissible* for *clear patterns* that are actually globally admissible in $\Omega_{\mathcal{F}}$ instead.

Remark 2.7 (Noise vs. Impurities)**.** With the notion of noise defined above, comparing the *clear* configurations on $\mathcal{A}$ is a mere matter of projecting $\pi_1 : \Omega_{\widetilde{\mathcal{F}}} \to \Omega_{\mathcal{A}}$, which results in configurations that may have some amount of forbidden patterns.

Another way to define noise would be to add a blank symbol $\square \notin \mathcal{A}$ not already in the alphabet, without changing $\mathcal{F}$. The main difference in this case is that there is no natural way to project $\square$ into $\mathcal{A}$ so that we can compare *clear* configurations. The symbol $\square$ behaves less like a noise and more like an impurity in itself.

From the point of view of the entropy, this changes things up. Informally, when the binary noise is maximal, we can obtain the uniform measure on $\Omega_{\mathcal{A}}$, for which the entropy is maximal. In comparison, the only measure that maximizes the amount of impurities is the Dirac measure $\delta_{\square^{\mathbb{Z}^d}}$ which has a null entropy. Studying more precisely the behaviour of the entropy in either of these settings, as a function of the amount of noise, may yield interesting further results.

Just as before, one can consider noisy measures on $\widetilde{\mathcal{A}}$ with the space $\mathcal{M}_{\widetilde{\mathcal{F}}}$. However, by doing so, we have no control on the weight of obscured cells, which is why we introduce the following measure spaces.

Definition 2.8 (Noisy Probability Measures)**.** *Let $\varepsilon \in [0,1]$. A σ-invariant probability measure ν on $\{0,1\}^{\mathbb{Z}^d}$ ($\nu \in \mathcal{M}_{\{0,1\}}$) is called an ε-noise if the probability of a given cell being obscured is at most ε, that is $\nu([1]) \leq \varepsilon$.*

For a given class of noises $\mathcal{N} \subset \mathcal{M}_{\{0,1\}}$, we now define the measure space:

$$\widetilde{\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}}(\varepsilon) = \left\{ \lambda \in \mathcal{M}_{\widetilde{\mathcal{F}}}, \pi_2^*(\lambda) \in \mathcal{N} \text{ is an } \varepsilon\text{-noise} \right\}.$$

Likewise, we define the projection $\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}(\varepsilon) = \pi_1^* \left(\widetilde{M_{\mathcal{F}}^{\mathcal{N}}}(\varepsilon) \right)$, which consists of measures on $\Omega_{\mathcal{A}}$. If no class is written, it is implied that $\mathcal{N} = \mathcal{M}_{\{0,1\}}$, that we allow for any *noise.*

Definition 2.9 (Classes of Dependent Noises)**.** *We define $\mathcal{B} = \left\{ \mathcal{B}(\varepsilon)^{\otimes \mathbb{Z}^d}, \varepsilon \in [0,1] \right\}$ the class of independent Bernoulli noises, where each cell is obscured with probability ε independently of the other cells.*

More generally, we consider the class $\mathcal{D}_k$ of k-dependent noises, such that any two windows at distance at least k are independent. More formally, $\nu \in \mathcal{D}_k$ when, for any patterns $w \in \mathcal{A}^I$ and $w' \in \mathcal{A}^J$ such that $d_\infty(I,J) \geq k$, $\nu([w] \cap [w']) = \nu([w])\nu([w'])$. For any rank k, we naturally have $\mathcal{D}_k \subset \mathcal{D}_{k+1}$. In particular, $\mathcal{D}_1 = \mathcal{B}$.

A direct consequence of this definition is that, on any class $\mathcal{N}$, for $\varepsilon < \delta$, we have the increasing inclusion $\widetilde{\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}}(\varepsilon) \subset \widetilde{\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}}(\delta)$, which naturally still holds for $\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}$ after projection. Let us notice that $\mathcal{M}_{\mathcal{F}}(0) = \mathcal{M}_{\mathcal{F}}$ is non-empty, and that $\mathcal{M}_{\mathcal{F}}(1) = \mathcal{M}_{\mathcal{A}}$ is the set of shit-invariant measures on $\Omega_{\mathcal{A}}$.

We are now interested in the *stability* of noisy measures, *i.e.* in the fact that $\mathcal{M}_{\mathcal{F}}(\varepsilon)$ gets close to $\mathcal{M}_{\mathcal{F}}$ in some sense – for some topology – as ε goes to 0.

2.2 Weak-$*$ stability

A natural topology on measures to consider first is the weak-* topology, but we will see here that as $\varepsilon \to 0$, any adherence point of a sequence of noisy measures is in $\mathcal{M}_{\mathcal{F}}$.

Definition 2.10 (Weak-* Topology)**.** *We can define the weak-* topology on the space of probability measures on $\mathcal{A}^{\mathbb{Z}^d}$ as the smallest topology such that, for any finite pattern w, the evaluation $\mu \mapsto \mu([w])$ is continuous.*

Note that this topological space is Hausdorff and compact, and that the subset $\mathcal{M}_{\mathcal{A}}$ of σ-invariant measures is a closed subset.

Lemma 2.11. *Consider $\mu \in \mathcal{M}_{\mathcal{A}}$ a σ-invariant measure on $\mathcal{A}$. If, for any pattern $w \in \mathcal{F}$, we have $\mu([w]) = 0$, then $\mu \in \mathcal{M}_{\mathcal{F}}$.*

Proof. To show that $\mu \in \mathcal{M}_{\mathcal{F}}$, we need to show that the measure is supported on $\Omega_{\mathcal{F}}$. The complement of $\Omega_{\mathcal{F}}$ is the set $\bigcup_{k\in\mathbb{Z}^d} \bigcup_{w\in\mathcal{F}} \sigma_k([w])$. By σ-invariance, for any $w \in \mathcal{F}$ and $k \in \mathbb{Z}^d$, we have $\mu\left(\sigma_k([w])\right) = \mu([w]) = 0$, thus $\mu\left(\Omega_{\mathcal{F}}^c\right) = 0$, so that μ is indeed supported on $\Omega_{\mathcal{F}}$. □

Proposition 2.12. *Let $\mu_n \in \mathcal{M}_{\mathcal{F}}(\varepsilon_n)$ be a sequence of noisy measures, with $\varepsilon_n \underset{n\to\infty}{\longrightarrow} 0$. Then any adherence value of the sequence is in $\mathcal{M}_{\mathcal{F}}$.*

Proof. Consider a weakly converging subsequence $\mu_{\theta(n)} \to^* \mu$, with $\theta : \mathbb{N} \to \mathbb{N}$ an increasing map. Naturally, the limit μ is also σ-invariant.

Notice that, as ε-noises are defined by forcing the measure of the noise cylinder $[1]$ to belong to the closed set $[0, \varepsilon]$, the set $\widetilde{\mathcal{M}_{\mathcal{F}}}(\varepsilon)$ is naturally weakly closed. Hence, by monotonous inclusion, for any $\varepsilon > 0$ we have $\mu \in \mathcal{M}_{\mathcal{F}}(\varepsilon)$.

Consider $\lambda_\varepsilon \in \widetilde{\mathcal{M}_{\mathcal{F}}}(\varepsilon)$ that projects to μ, and a forbidden pattern $w \in \mathcal{F}$. Thus, $\mu([w]) = \lambda_\varepsilon\left(\left[w, \{0,1\}^{I(w)}\right]\right)$. In particular, because $\lambda_\varepsilon \in \mathcal{M}_{\widetilde{\mathcal{F}}}$ and $\left(w, 0^{I(w)}\right) \in \widetilde{\mathcal{F}}$, we have $\lambda_\varepsilon\left(\left[w, 0^{I(w)}\right]\right) = 0$. Hence, if $\mu([w]) > 0$, then at least one of the cells of the window $I(w)$ must be obscured for λ_ε, so $\mu([w]) \le |I(w)| \times \lambda_\varepsilon([\mathcal{A}, 1]) \le |I(w)|\varepsilon$ by union bound. As ε goes to 0, we conclude that $\mu([w]) = 0$. Using the previous lemma, $\mu \in \mathcal{M}_{\mathcal{F}}$. □

Hence, all SFTs are weakly "stable" in the sense that there is no sequence of measures without adherence values as $\varepsilon \to 0$ in this topology (see Proposition 3.2 for a more formal statement), so this property yields no interesting classification.

In the following section, we will introduce a general notion of metric stability and convergence speed – as there is no *canonical* metric associated to the weak topology, we did not try to quantify the speed of convergence in this case.

The main issue with the weak-$*$ topology is that it looks at things on a local scale, on finite patterns, without really forcing any kind of behaviour on $\mathbb{Z}^d$ as a whole. To better discriminate between SFTs, we will introduce the Besicovitch distance d_B on measures, that looks at configurations globally and quantifies the frequency of differences.

Thereafter, we will prove that the stability for d_B is conjugacy-invariant, in order to illustrate how this distance is manipulated.

3 Stability under Besicovitch topology and conjugacy invariance

Before diving into the Besicovitch world, let us briefly introduce our general notion of stability.

Definition 3.1 (Stability)**.** *Consider here a distance d on $\mathcal{M}_{\mathcal{A}}$, a noise class $\mathcal{N} \subset \mathcal{M}_{\{0,1\}}$, and a non-decreasing function $f : [0,1] \to \mathbb{R}^+$, right-continuous at 0 with $f(0) = 0$.*

The SFT induced by $\mathcal{F}$ is said to be f-stable for the distance d on the class $\mathcal{N}$ if:

$$\forall \varepsilon \in [0,1], \sup_{\mu \in \mathcal{M}_{\mathcal{F}}^{\mathcal{N}}(\varepsilon)} d\left(\mu, \mathcal{M}_{\mathcal{F}}\right) \leq f(\varepsilon).$$

The SFT is stable if it is f-stable for some function f. We say that $\Omega_{\mathcal{F}}$ is linearly stable (resp. polynomialy stable) if it is f-stable with $f(\varepsilon) = O(\varepsilon)$ (resp. $f(\varepsilon) = O\left(\varepsilon^{\alpha}\right)$ for some $0 < \alpha \leq 1$).

Proposition 3.2 (Weak-* Stability)**.** *Given a metric d_* which induces the weak-$*$ topology, any SFT is stable for the distance d_*.*

Proof. This result is a consequence of Proposition 2.12. Indeed, by contraposition, assume some SFT $\Omega_{\mathcal{F}}$ is not stable for the distance d_*.

Then, there exists a sequence $\varepsilon_n \underset{n\to\infty}{\longrightarrow} 0$ and measures $\mu_n \in \mathcal{M}_{\mathcal{F}}\left(\varepsilon_n\right)$ such that $\inf_{n\in\mathbb{N}} d_*\left(\mu_n, \mathcal{M}_{\mathcal{F}}\right) = d > 0$. By compactness, this sequence admits a weak-$*$ adherence value μ, and $d_*\left(\mu, \mathcal{M}_{\mathcal{F}}\right) \geq d$. In particular, $\mu \notin \mathcal{M}_{\mathcal{F}}$, which contradicts Proposition 2.12, hence stability. □

Note that this result gives no quantitative bound on the speed of convergence. In order to obtain such bounds, we may perhaps make a clever use of the reconstruction function, but as stated earlier, the lack of canonical choice for d_* discouraged this study.

In this section, we will first introduce the Besicovitch distance d_B, and then prove that stability for d_B is conjugacy-invariant on the class of all noises $\mathcal{M}_{\{0,1\}}$. At last, we mention the notion of domination, which will allow us to extend the conjugacy-invariant stability to $\mathcal{B}$.

3.1 Besicovitch topology

In order to compare *measures*, we need first to be able to compare *configurations*.

Definition 3.3 (Hamming-Besicovitch Distance)**.** *On a finite window $I \subset \mathbb{Z}^d$, we define the Hamming distance between two finite patterns $x, y \in \mathcal{A}^I$ as:*

$$d_I(x,y) = \frac{1}{|I|} \left|\{k \in I, x_k \neq y_k\}\right|.$$

For a given increasing sequence $\left(I_n\right)_{n\in\mathbb{N}}$ such that $\bigcup_{n\in\mathbb{N}} I_n = \mathbb{Z}^d$, we can define a pseudometric on $\Omega_{\mathcal{A}}$, such that for $x, y \in \Omega_{\mathcal{A}}$:

$$d_H(x,y) = \limsup_{n\to\infty} d_{I_n}\left(x|_{I_n}, y|_{I_n}\right).$$

This pseudometric d_H is usually called the Hamming-Besicovitch distance. Remark that the Hamming distances d_I are clearly measurable for the product topology, thus so is the limit d_H.

In order to use an extension of Birkhoff's pointwise ergodic theorem, we need (I_n) to be a sequence of *boxes* (products of intervals). A general statement of this theorem can be found in *Ergodic Theorems* [17, Chapter 6]. Let us more specifically use the boxes $B_n = [\![-n, n]\!]^d$ further on.

Definition 3.4 (Besicovitch Distance)**.** *A measure $\lambda \in \mathcal{M}_{\mathcal{A}\times\mathcal{A}}$ is said to be a coupling between the measures $\mu, \nu \in \mathcal{M}_{\mathcal{A}}$ if $\pi_1^*(\lambda) = \mu$ and $\pi_2^*(\lambda) = \nu$.*

For two measure $\mu, \nu \in \mathcal{M}_{\mathcal{A}}$ we define their Besicovitch distance as:

$$d_B(\mu,\nu) = \inf_{\lambda \text{ a coupling}} \int d_H(x,y) \mathrm{d}\lambda(x,y).$$

Note that we can always consider the independent coupling $\mu \otimes \nu$, so the set of couplings is non-empty. By the ergodic theorem, d_H is obtained by averaging $\mathbb{1}_{x_0 \neq y_0}$ over translations, so $\lambda\left(d_H\right) = \lambda\left([x_0 \neq y_0]\right)$ in the equation above.

For the coupling, we can more generally consider probability measures λ on some general space Ω and two measurable maps $\psi_1 : \Omega \to \Omega_{\mathcal{A}}$ (resp. ψ_2) such that $\psi_1^*(\lambda) = \mu$ (resp. $\psi_2^*(\lambda) = \nu$) and consider $d_H\left(\psi_1(\omega), \psi_2(\omega)\right) \mathrm{d}\lambda(\omega)$ in the integral instead, as long as we have $\left(\psi_1, \psi_2\right)^*(\lambda) \in \mathcal{M}_{\mathcal{A}\times\mathcal{A}}$. We will use this more general version to build couplings that use additional information, notably which cells should be obscured in the noisy framework, or the value of an additional independent random variable.

The Besicovitch distance d_B has been quite used in the recent research literature, but it was already introduced in earlier works, sometimes also named $\overline{d}$ as in the monograph by Glasner [11, Chapter 15]. The main interest of d_B, in the context of these works, is that the measure entropy is continuous for this topology. Even though d_B has been widely studied, let us prove here that d_B is indeed a distance, in order to get acquainted with the notion.

Lemma 3.5. *The function d_B is a distance on $\mathcal{M}_{\mathcal{A}}$, and $d_B(\mu, \nu)$ is always* reached *for some coupling between the measures.*

Proof. The function d_B is trivially symmetric, and $d_B(\mu, \mu) = 0$ for any measure $\mu \in \mathcal{M}_{\mathcal{A}}$.

To prove the triangle inequality, consider three measures $\mu_1, \mu_2, \mu_3 \in \mathcal{M}_{\mathcal{A}}$. Consider a coupling $\lambda_{1,2} \in \mathcal{M}_{\mathcal{A}\times\mathcal{A}}$ (resp. $\lambda_{2,3} \in \mathcal{M}_{\mathcal{A}\times\mathcal{A}}$) between μ_1 and μ_2 (resp. μ_2 and μ_3). The measures $\lambda_{1,2}$ and $\lambda_{2,3}$ are compatible in the sense that they share a common projection $\pi_2^*\left(\lambda_{1,2}\right) = \pi_1^*\left(\lambda_{2,3}\right) = \mu_2$. Thence, it is known [11, Chapter 6] that there exists a coupling $\lambda_{1,2,3}$ between them, such that $\left(\pi_1, \pi_2\right)^*\left(\lambda_{1,2,3}\right) = \lambda_{1,2}$ and likewise $\left(\pi_2, \pi_3\right)^*\left(\lambda_{1,2,3}\right) = \lambda_{2,3}$. In particular, $\lambda_{1,2,3}$ gives us a coupling between μ_1 and μ_3, and we have:

$$
\begin{array}{rcl}
d_B\left(\mu_1, \mu_3\right) & \leq & \int d_H(x, z) \mathrm{d}\lambda_{1,2,3}(x, y, z) \\
& \leq & \int d_H(x, y) + d_H(y, z) \mathrm{d}\lambda_{1,2,3}(x, y, z) \\
& = & \int d_H(x, y) \mathrm{d}\lambda_{1,2}(x, y) + \int d_H(y, z) \mathrm{d}\lambda_{2,3}(y, z).
\end{array}
$$

Now, by taking the infimum over all couplings $\lambda_{1,2}$ and $\lambda_{2,3}$ we finally obtain the upper bound $d_B\left(\mu_1, \mu_3\right) \leq d_B\left(\mu_1, \mu_2\right) + d_B\left(\mu_2, \mu_3\right)$, the triangle inequality.

Consider now some coupling $\lambda \in \mathcal{M}_{\mathcal{A}\times\mathcal{A}}$ between two measures $\mu, \nu \in \mathcal{M}_{\mathcal{A}}$. Using our pointwise ergodic theorem, it follows that $d_H(x, y)$ is an actual limit λ-a.s., and that $\int d_H(x, y) \mathrm{d}\lambda(x, y) = \int \mathbb{1}_{\{x_0 \neq y_0\}} \mathrm{d}\lambda(x, y)$. Hence, $\lambda \mapsto \int d_H(x, y) \mathrm{d}\lambda(x, y)$ is clearly a weakly continuous mapping on $\mathcal{M}_{\mathcal{A}\times\mathcal{A}}$, so by compactness the distance d_B is reached by some coupling λ.

Assume now that $d_B(\mu, \nu) = 0$ is reached for some coupling λ. Then, λ-a.s., we have $x_0 = y_0$. As λ is σ-invariant, it is more generally true for any cell $k \in \mathbb{Z}^d$ that $x_k = y_k$ almost surely. By taking the countable intersection of such events, $x = y$ almost surely, so λ is supported on the diagonal of $\Omega_{\mathcal{A}\times\mathcal{A}} = \Omega_{\mathcal{A}} \times \Omega_{\mathcal{A}}$. Thence, we have $\mu = \pi_1^*(\lambda) = \pi_2^*(\lambda) = \nu$. Conversely, distinct measures in $\mathcal{M}_{\mathcal{A}}$ are distinguishable. □

The important part of this lemma is that we need the σ-invariance of the measures to conclude that d_B is not a mere *pseudometric* but a distance, which shows how the Besicovitch distance is appropriate to our specific setting.

Remark 3.6 (Universal Linear Lower Bound)**.** If $\mu \in \mathcal{M}_{\mathcal{F}}(\varepsilon)$, then obscured cells have a frequency ε. For an arbitrary set of forbidden patterns $\mathcal{F}$, it is reasonable to assume that whenever a cell is obscured, the cell contains a "wrong" letter with positive probability, with respect to any globally admissible configuration $\omega \in \Omega_{\mathcal{F}}$. Hence, the best and fastest bounds we can reasonably obtain are linear, *i.e.* $f(\varepsilon) = \Omega(\varepsilon)$ using the Ω Landau notation.

Now that the Besicovitch distance d_B has been properly introduced, let us prove a that stability for d_B is conjugacy-invariant in some sense.

3.2 Conjugacy-invariant stability

Conjugate SFTs share the same dynamical properties, hence proving the invariance of stability under conjugacy would imply that stability is indeed a property of an SFT $\Omega_{\mathcal{F}}$ and not of the specific rules $\mathcal{F}$ used to define it. In order to study the invariance by conjugacy, let us first properly define what a conjugacy is.

Definition 3.7 (Morphism and Conjugacy)**.** *Consider two sets of forbidden patterns $\mathcal{F}$ on the alphabet $\mathcal{A}_1$ and $\mathcal{G}$ on $\mathcal{A}_2$, not necessarily on the same alphabet, but on the same grid $\mathbb{Z}^d$. A morphism from the SFT $\Omega_{\mathcal{F}}$ to $\Omega_{\mathcal{G}}$ is a continuous σ-invariant mapping $\theta : \Omega_{\mathcal{F}} \to \Omega_{\mathcal{G}}$.*

Equivalently [13], we can define $\theta : \mathcal{A}_1^J \to \mathcal{A}_2$ as a local map, on a finite window $J \subset \mathbb{Z}^d$, and extend it on configurations $x \in \Omega_{\mathcal{F}}$ so that we have $\theta(x)_k := \theta\left(x|_{J+k}\right)$ for any cell $k \in \mathbb{Z}^d$. Note that, when defining the local map θ, any value can be chosen for $\theta(w)$ when $w \in \mathcal{A}_1^J$ is not *globally admissible, we just need to fix a choice so that things are well-defined in the noisy case.*

When $\mathcal{A} = \mathcal{A}_1 = \mathcal{A}_2$ and $\mathcal{F} = \mathcal{G} = \emptyset$, $\theta : \Omega_{\mathcal{A}} \to \Omega_{\mathcal{A}}$ is also called a cellular automaton.

Two SFTs $\Omega_{\mathcal{F}}$ and $\Omega_{\mathcal{G}}$ are conjugate if there is a bijective morphism $\theta : \Omega_{\mathcal{F}} \to \Omega_{\mathcal{G}}$ (in which case θ^{-1} is consequently also *a morphism).*

Later on, we will always consider the *local* definition of morphisms as extensions of local mappings $\theta : \mathcal{A}_1^J \to \mathcal{A}_2$. An interest of this viewpoint is that any morphism from $\Omega_{\mathcal{F}}$ to $\Omega_{\mathcal{G}}$ is actually simply the restriction of a morphism on the full-shifts $\Omega_{\mathcal{A}_1}$ and $\Omega_{\mathcal{A}_2}$.

Definition 3.8 (Thickened Noise)**.** *Let $\gamma_n : \{0,1\}^{B_n} \to \{0,1\}$ be the cellular automaton defined by $\gamma_n(w) = \max_{k \in B_n} w_k$. We say that $\gamma_n(\omega)$ is n-thickened for $\omega \in \Omega_{\{0,1\}}$ in the sense that if the cell $c \in \mathbb{Z}^d$ is obscured in ω, then its n-neighbourhood $c + B_n$ is obscured in $\gamma_n(\omega)$.*

These specific morphisms will allow us to obscure the forbidden patterns that may appear when using a morphism or a measurable application on $\Omega_{\mathcal{A}}$ later on.

Lemma 3.9. *Consider the SFTs $\Omega_{\mathcal{F}}$ and $\Omega_{\mathcal{G}}$, and a morphism from $\Omega_{\mathcal{F}}$ to $\Omega_{\mathcal{G}}$, locally defined as $\theta : \mathcal{A}_1^J \to \mathcal{A}_2$. Then for any $x, y \in \Omega_{\mathcal{A}_1}$, we have $d_H\left(\theta(x), \theta(y)\right) \leq D_\theta d_H(x,y)$ with the constant $D_\theta = |J|$. Consequently, for any measures $\mu, \nu \in \mathcal{M}_{\mathcal{A}_1}$, we have $d_B\left(\theta^*(\mu), \theta^*(\nu)\right) \leq D_\theta d_B(\mu,\nu)$.*

There exists a radius r_θ such that the morphism $\widetilde{\theta} := (\theta, \gamma_{r_\theta}) : \Omega_{\widetilde{\mathcal{A}_1}} \to \Omega_{\widetilde{\mathcal{A}_2}}$ satisfies $\widetilde{\theta}\left(\Omega_{\widetilde{\mathcal{F}}}\right) \subset \Omega_{\widetilde{\mathcal{G}}}$. Moreover, there is a constant C_θ such that, whenever $\gamma_{r_\theta}^*\left(\mathcal{N}\right) \subset \mathcal{N}'$, for any $\varepsilon > 0$:

$$\widetilde{\theta}^*\left(\widetilde{\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}}(\varepsilon)\right) \subset \widetilde{\mathcal{M}_{\mathcal{G}}^{\mathcal{N}'}}\left(C_\theta \times \varepsilon\right).$$

Proof. Assume that $\theta(x)_k \neq \theta(y)_k$. Then x and y *must* differ in at least one cell of the window $J + k$. Conversely, each cell of $\mathbb{Z}^d$ can appear into *at most* $|J|$ such windows, so that we naturally obtain the bound $d_H\left(\theta(x), \theta(y)\right) \leq |J| d_H(x,y)$. Now, assuming $d_B(\mu,\nu)$ is reached for a coupling $\lambda \in \mathcal{M}_{\mathcal{A}_1 \times \mathcal{A}_1}$, then $(\theta,\theta)^*(\lambda) \in \mathcal{M}_{\mathcal{A}_2 \times \mathcal{A}_2}$ is a coupling between $\theta^*(\mu)$ and $\theta^*(\nu)$, and we consequently obtain the analogous bound for d_B.

Just like $\theta : \mathcal{A}_1^J \to \mathcal{A}_2$ naturally sends $\Omega_{\mathcal{A}_1}$ onto $\Omega_{\mathcal{A}_2}$, it sends any finite pattern $v \in \mathcal{A}_1^{J+I}$ onto $\theta(v) \in \mathcal{A}_2^I$. The "local" property that characterises $\theta\left(\Omega_{\mathcal{F}}\right) \subset \Omega_{\mathcal{G}}$ is *not* that it preserves locally admissible patterns, but that it preserves globally admissible ones.

If a locally admissible pattern $v \in \mathcal{A}_1^{J+I}$ is not globally admissible, nothing forbids $\theta(v) \in \mathcal{A}_2^I$ from containing forbidden patterns of $\mathcal{G}$. In such a case, let us extend v into $\omega \in \Omega_{\mathcal{A}_1}$ by filling the empty cells outside of $I + J$ with any letter $a \in \mathcal{A}_1$, and consider the noise $b = \mathbb{1}_{(I+J)^c}$ that obscures all the cells outside of $I + J$. Then naturally $(\omega, b) \in \Omega_{\widetilde{\mathcal{F}}}$ is locally admissible but $(\theta(\omega), b) \notin \Omega_{\widetilde{\mathcal{G}}}$ is not. Thence, we cannot simply extend the morphism $\theta : \Omega_{\mathcal{A}_1} \to \Omega_{\mathcal{A}_2}$ as $\widetilde{\theta}$ by leaving the second coordinate unchanged.

More precisely, assume that $w = \theta(v) \in \mathcal{G}$ is a forbidden pattern, with $v \in \mathcal{A}_1^{J+I(w)}$. Then v must not be a *globally* admissible pattern itself. As explained in Remark 2.4, using the reconstruction function, we have $r(w) = \varphi_{\mathcal{F}}(J + I(w)) \in \mathbb{N}$ such that, if we can extend v into a locally admissible pattern $v \in \mathcal{A}_1^{J+I(w)+B_{r(w)}}$, then v itself must be globally admissible.

Let us define $r_\theta = \max_{w\in\mathcal{G}} r(w) + \max_{c\in J}\|c\|_\infty$. Consider $(\omega, b) \in \Omega_{\widetilde{\mathcal{F}}}$. If $\theta(\omega)$ contains a forbidden pattern w in the window $c + I(w)$, then it follows that the window $c + J + I(w)$ of ω is not globally admissible, so the window $c + I(w) + B_{r_\theta}$ of $\omega \in \Omega_{\mathcal{A}_1}$ must not be locally admissible. As (ω, b) is locally admissible, this implies that at least one cell in $c + I(w) + B_{r_\theta}$ must be obscured. We proved that, if $(\omega, b) \in \Omega_{\widetilde{\mathcal{F}}}$, then $(\theta(\omega), \gamma_{r_\theta}(b)) \in \Omega_{\widetilde{\mathcal{G}}}$, so $\widetilde{\theta} = (\theta, \gamma_{r_\theta})$ is the morphism we wanted.

Finally, we need to exhibit the constant C_θ. Consider a noisy measure $\lambda \in \widetilde{\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}}(\varepsilon)$, with an ε-noise $\nu = \pi_2^*(\lambda) \in \mathcal{N}$. Notice that $\pi_2 \circ \widetilde{\theta} = \gamma_{r_\theta}$, so the noise of $\widetilde{\theta}^*(\lambda)$ is actually $\gamma_{r_\theta}^*(\nu) \in \mathcal{N}'$. Remark that the clear configuration 0^∞ is a fixed point of γ_{r_θ}. As in the proof of Lemma 3.5, using a pointwise ergodic theorem, the amount of noise in ν is:

$$\nu([1]) = \int \mathbb{1}_{\{x_0\neq 0\}} \mathrm{d}\nu(x) = \int d_H\left(x, 0^\infty\right) \mathrm{d}\nu(x) = d_B\left(\nu, \delta_{0^\infty}\right).$$

Thus, if we apply the first part of the current lemma to the current morphism γ_{r_θ}, with the set $J = B_{r_\theta}$, we conclude that $d_B\left(\gamma_{r_\theta}^*(\nu), \delta_{0^\infty}\right) \leq C_\theta d_B\left(\nu, \delta_{0^\infty}\right) \leq C_\theta \times \varepsilon$ with the constant $C_\theta = |J| = (2r_\theta + 1)^d$. At last, $\gamma_{r_\theta}^*(\nu)$ is a $(C_\theta\varepsilon)$-noise, $\widetilde{\theta}^*(\lambda) \in \widetilde{\mathcal{M}_{\mathcal{G}}^{\mathcal{N}'}}\left(C_\theta\varepsilon\right)$. ☐

Assume that the SFT $\Omega_{\mathcal{F}}$ is f-stable, and that it is sent on $\Omega_{\mathcal{G}}$ by θ. The lemma suggests that the *subset* $\pi_1^*\left(\widetilde{\theta}^*\left(\widetilde{\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}}(\varepsilon)\right)\right) = \theta^*\left(\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}(\varepsilon)\right)$ of $\mathcal{M}_{\mathcal{G}}^{\gamma_{r_\theta}(\mathcal{N})}\left(C_\theta\varepsilon\right)$ is roughly $D_\theta \times f$-stable. However, this still does not give us enough information to obtain a full-fledged and well-defined stability property for $\mathcal{G}$. To obtain such a result, we will now assume that θ is not only a morphism but a conjugacy between $\Omega_{\mathcal{F}}$ and $\Omega_{\mathcal{G}}$.

Theorem 3.10 (Conjugacy-Invariant Stability)**.** *Consider a conjugacy $\theta : \Omega_{\mathcal{F}} \to \Omega_{\mathcal{G}}$, and assume that $\Omega_{\mathcal{F}}$ is f-stable for d_B on a class $\gamma_{r_{\theta^{-1}}}^*(\mathcal{N})$ of noises.*

Then there exists a constant E such that $\Omega_{\mathcal{G}}$ is g-stable on $\mathcal{N}$ with the speed

$$g : \varepsilon \mapsto D_\theta f\left(C_{\theta^{-1}}\varepsilon\right) + E\varepsilon.$$

Proof. We will use the result of Lemma 3.9 for both $\theta : \Omega_{\mathcal{F}} \to \Omega_{\mathcal{G}}$ *and* its inverse $\theta^{-1} : \Omega_{\mathcal{G}} \to \Omega_{\mathcal{F}}$. Note that, on the larger domain $\Omega_{\mathcal{A}_2}$, the cellular automaton $\theta \circ \theta^{-1}$ is still well-defined, but is not necessarily the identity function outside of the domain $\Omega_{\mathcal{G}}$. Now, if we consider two measures $\mu, \nu \in \mathcal{M}_{\mathcal{A}_2}$:

$$\begin{array}{rcl} d_B(\mu,\nu) & \leq & d_B\left(\mu, \left(\theta\circ\theta^{-1}\right)^*(\mu)\right) \\ & + & d_B\left(\left(\theta\circ\theta^{-1}\right)^*(\mu), \left(\theta\circ\theta^{-1}\right)^*(\nu)\right) \\ & + & d_B\left(\left(\theta\circ\theta^{-1}\right)^*(\nu), \nu\right). \end{array}$$

The idea behind this back-and-forth is that, by going from $\Omega_{\mathcal{G}}$ to $\Omega_{\mathcal{F}}$, we reach a stable SFT while still keeping the noise under control, and then going from $\Omega_{\mathcal{F}}$ to $\Omega_{\mathcal{G}}$ allows us to maintain this stability while comparing the new configuration to the old one. In particular, if $\nu \in \mathcal{M}_{\mathcal{G}}$, then it is supported on the domain $\Omega_{\mathcal{G}}$, where $\theta \circ \theta^{-1}$ is the identity function, so that $d_B\left(\left(\theta\circ\theta^{-1}\right)^*(\nu), \nu\right) = 0$.

Consider a measure $\mu \in \mathcal{M}_{\mathcal{G}}^{\mathcal{N}}(\varepsilon)$, and $\nu_{\mathcal{F}} \in \mathcal{M}_{\mathcal{F}}$ that achieves $d_B\left(\left(\theta^{-1}\right)^*(\mu), \mathcal{M}_{\mathcal{F}}\right)$. If we denote $\nu_{\mathcal{G}} = \theta^*(\nu_{\mathcal{F}}) \in \mathcal{M}_{\mathcal{G}}$, then:

$$d_B(\mu, \mathcal{M}_{\mathcal{G}}) \leq d_B(\mu, \nu_{\mathcal{G}}) \leq d_B\left(\mu, \left(\theta \circ \theta^{-1}\right)^*(\mu)\right) + d_B\left(\left(\theta \circ \theta^{-1}\right)^*(\mu), \nu_{\mathcal{G}}\right).$$

In particular, using Lemma 3.9 for θ^{-1}, we know that $\left(\theta^{-1}\right)^*(\mu) \in \mathcal{M}_{\mathcal{F}}^{\gamma^*_{r_{\theta^{-1}}}(\mathcal{N})}(C_{\theta^{-1}}\varepsilon)$. Thence, using Lemma 3.9 for θ, as $\Omega_{\mathcal{F}}$ is f-stable on $\gamma^*_{r_{\theta^{-1}}}(\mathcal{N})$, we get the bound:

$$\begin{array}{rcl} d_B\left(\left(\theta \circ \theta^{-1}\right)^*(\mu), \nu_{\mathcal{G}}\right) & \leq & D_\theta d_B\left(\left(\theta^{-1}\right)^*(\mu), \nu_{\mathcal{F}}\right) \\ & = & D_\theta d_B\left(\left(\theta^{-1}\right)^*(\mu), \mathcal{M}_{\mathcal{F}}\right) \\ & \leq & D_\theta f\left(C_{\theta^{-1}}\varepsilon\right). \end{array}$$

To conclude the proof, we just need to have a linear control on $d_B\left(\mu, \left(\theta \circ \theta^{-1}\right)^*(\mu)\right)$ as $\varepsilon \to 0$. To do so, we will study $d_H\left(x, \theta \circ \theta^{-1}(x)\right)$ for any $x \in \Omega_{\mathcal{A}_2}$. More precisely, whenever $(x, b) \in \Omega_{\widetilde{\mathcal{F}}}$, we want a bound $d_H\left(x, \theta \circ \theta^{-1}(x)\right) \leq E d_H\left(b, 0^\infty\right)$. Assuming such a bound holds, consider $\lambda \in \widetilde{\mathcal{M}_{\mathcal{G}}^{\mathcal{N}}}(\varepsilon)$ that projects to μ, which naturally gives a coupling between $\mu = \pi_1^*(\lambda)$ and $\left(\theta \circ \theta^{-1}\right)^*(\mu) = \left(\pi_1 \circ \widetilde{\theta} \circ \widetilde{\theta^{-1}}\right)^*(\lambda)$. Then we obtain:

$$d_B\left(\mu, \left(\theta \circ \theta^{-1}\right)^*(\mu)\right) \leq \int\limits_{\Omega_{\widetilde{\mathcal{G}}}} d_H\left(x, \theta \circ \theta^{-1}(x)\right) \mathrm{d}\lambda(x, b) \leq E \int\limits_{\Omega_{\widetilde{\mathcal{G}}}} d_H\left(b, 0^\infty\right) \mathrm{d}\lambda(x, b) \leq E\varepsilon.$$

The sketch of the proof from now on is pretty much the same as in Lemma 3.9. Let us suppose that $(x, b) \in \Omega_{\widetilde{\mathcal{G}}}$ and that $x_k \neq \theta \circ \theta^{-1}(x)_k$ for some cell $k \in \mathbb{Z}^d$. Consider the window $J = J_{\theta^{-1}} + J_\theta$ such that the value of $\theta \circ \theta^{-1}(x)_k$ only depends on the pattern $x|_{J+k}$. Let us assume that $0 \in J$ without loss of generality. If $x|_{J+k}$ was globally admissible, then we could extend it into a globally admissible configuration $y \in \Omega_{\mathcal{G}}$, such that $\theta \circ \theta^{-1}(x)_k = \theta \circ \theta^{-1}(y)_k = y_k = x_k$. This contradicts our hypothesis, so x_{J+k} is not globally admissible. This means that, using once again the reconstruction function φ from Remark 2.4 for the SFT $\Omega_{\mathcal{G}}$, $x|_{J+k+B_{\varphi(J)}}$ is not locally admissible, so the same windows in b contains at least one obscured cell. Hence, $d_H\left(x, \theta \circ \theta^{-1}(x)\right) \leq E d_H\left(b, 0^\infty\right)$ with the constant $E = \left|J + B_{\varphi(J)}\right|$, which concludes the proof. ☐

Corollary 3.11 (Conjugacy-Invariance for Stable Noise Classes)**.** *If $\mathcal{N}$ is stable under the action of any γ_n, then for any two conjugate SFTs $\Omega_{\mathcal{F}}$ and $\Omega_{\mathcal{G}}$, $\Omega_{\mathcal{F}}$ is stable (resp. linearly stable, polynomially stable) on the class $\mathcal{N}$ if and only if $\Omega_{\mathcal{G}}$ is.*

In particular, this corollary holds for the class of all noises $\mathcal{N} = \mathcal{M}_{\{0,1\}}$. This stability hypothesis is actually quite restrictive. For example, we naturally have the inclusion $\gamma_n(\mathcal{B}) \subset \mathcal{D}_{2n+1}$ but $\gamma_n(\mathcal{B}) \not\subset \mathcal{D}_{2n}$. Thence, $\gamma_n(\mathcal{B}) \not\subset \mathcal{B}$, the previous conjugacy-invariance corollary does not apply on the class $\mathcal{N} = \mathcal{B}$.

3.3 Stability and domination

We will now introduce the notion of domination between noises, which will allow us to send $\mathcal{D}_k$ back into $\mathcal{B}$, in order to obtain a conjugacy-invariant stability result for the class $\mathcal{B}$.

Definition 3.12 (Domination)**.** *A Borel set $B \subset \{0,1\}^{\mathbb{Z}^d}$ is said to be increasing if, for any $b \in B$ and $b' \geq b$ (on each coordinate), we have $b' \in B$.*

Consider $\nu_1, \nu_2 \in \mathcal{M}_{\{0,1\}}$. We say that ν_2 dominates ν_1, and we denote $\nu_2 \geq \nu_1$, if $\nu_2(B) \geq \nu_1(B)$ for any increasing Borel set B. Equivalently [19, Theorem 2.4], $\nu_2 \geq \nu_1$

if there exists some coupling ν_{dom} between $\nu_1 = \pi_1^*(\nu_{\mathrm{dom}})$ and $\nu_2 = \pi_2^*(\nu_{\mathrm{dom}})$ which is supported on $\Omega_{\leq} := \left\{(b_1, b_2) \in \Omega_{\{0,1\}}^2, b_1 \leq b_2\right\}$. Note that in the reference, this equivalence is stated in a general non-σ-invariant framework, so that ν_{dom} is not a priori *in $\mathcal{M}_{\{0,1\}^2}$, but as $\Omega_{\leq}$ is compact, we can replace ν_{dom} by any weak-* adherence value of $\left(\frac{1}{|I_n|}\sum_{k\in I_n}\sigma_k^*(\nu_{\mathrm{dom}})\right)_{n\in\mathbb{N}}$ to obtain a likewise σ-invariant coupling.*

We can extend this notion to classes of measures. Let $g : [0,1] \to [0,1]$ be a non-decreasing function, right-continuous in 0 with $g(0) = 0$. We say that the class $\mathcal{N}$ is g-dominated by $\mathcal{N}'$ if, for any $\varepsilon > 0$ and any ε-noise $\nu \in \mathcal{N}$, there exists a $g(\varepsilon)$-noise $\nu' \in \mathcal{N}'$ such that $\nu' \geq \nu$.

Lemma 3.13 (Disintegration Theorem [15, Theorem 8.5])**.** *Let λ be any probability measure on $\Omega_{\mathcal{A}} \times \Omega_{\mathcal{A}'}$, with $\mu = \pi_1^*(\lambda)$. We can factorise $\lambda(A \times B) = \int_A \nu_x(B)\mathrm{d}\mu(x)$, such that $x \mapsto \nu_x(B)$ is measurable for any measurable set B, and that $B \mapsto \nu_x(B)$ is a probability measure for μ-a.e. $x \in \Omega_{\mathcal{A}}$.*

Using this domination property along with the disintegration theorem, we can then prove the following result, that most notably does not depend on the distance d used for the stability.

Proposition 3.14. *If the SFT $\Omega_{\mathcal{F}}$ is f-stable on the class $\mathcal{N}'$ for the distance d, and $\mathcal{N}$ is g-dominated by $\mathcal{N}'$, then $\Omega_{\mathcal{F}}$ is $(f \circ g)$-stable on the class $\mathcal{N}$ for the distance d.*

Proof. Let us assume that $\Omega_{\mathcal{F}}$ is f-stable on the class $\mathcal{N}'$. Consider $\mu \in \mathcal{M}_{\mathcal{F}}^{\mathcal{N}}(\varepsilon)$. If we prove that $\mu \in \mathcal{M}_{\mathcal{F}}^{\mathcal{N}'}(g(\varepsilon))$, then $d_B(\mu, \mathcal{M}_{\mathcal{F}}) \leq f(g(\varepsilon))$ by f-stability.

In order to prove this, let us consider a measure $\lambda \in \widetilde{\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}}(\varepsilon)$ such that $\pi_1^*(\lambda) = \mu$, with $\pi_2^*(\lambda) = \nu \in \mathcal{N}$ an ε-noise. By domination, there exists a $g(\varepsilon)$-noise $\nu' \in \mathcal{N}'$ such that $\nu' \geq \nu$, with a coupling ν_{dom} between them.

Then, using the disintegration theorem, for ν-a.e. $b \in \Omega_{\{0,1\}}$, there is a measure μ_b on $\Omega_{\mathcal{A}}$ such that, for any two Borel sets $A \subset \Omega_{\mathcal{A}}$ and $B \subset \Omega_{\{0,1\}}$:

$$\lambda(A \times B) = \int_B \mu_b(A)\mathrm{d}\nu(b) = \int \mu_b(A)\mathbb{1}_B(b)\mathrm{d}\nu_{\mathrm{dom}}(b, b').$$

Now, we can naturally define the measure λ' on $\Omega_{\mathcal{A}} \times \Omega_{\{0,1\}}$ as:

$$\lambda'(A \times B) = \int \mu_b(A)\mathbb{1}_B(b')\,\mathrm{d}\nu_{\mathrm{dom}}(b, b').$$

By taking $B = \Omega_{\{0,1\}}$, it is clear that $\pi_1^*(\lambda') = \pi_1^*(\lambda) = \mu$. Now, by taking $A = \Omega_{\mathcal{A}}$, we conclude that $\pi_2^*(\lambda') = \pi_2^*(\nu_{\mathrm{dom}}) = \nu'$. Moreover, consider $\widetilde{w} = \left(w, 0^{I(w)}\right) \in \widetilde{\mathcal{F}}$ a forbidden pattern. Since ν_{dom} is supported by $\Omega_{\leq} = \{(b, b'), b \leq b'\}$:

$$\lambda'([\widetilde{w}]) = \int \mu_b([w])\mathbb{1}_{0^{I(w)}}(b')\,\mathrm{d}\nu_{\mathrm{dom}}(b, b') \leq \int \mu_b([w])\mathbb{1}_{0^{I(w)}}(b)\mathrm{d}\nu_{\mathrm{dom}}(b, b') = \lambda([\widetilde{w}]) = 0.$$

Thence, λ' is supported on $\Omega_{\widetilde{\mathcal{F}}}$. Without loss of generality, we can replace λ' by a weak-* adherence value of the averages $\left(\frac{1}{|B_n|}\sum_{k\in B_n}\sigma_k^*(\lambda')\right)_{n\in\mathbb{N}}$, which still projects to μ and ν', but is also σ-invariant, so that $\lambda' \in \widetilde{\mathcal{M}_{\mathcal{F}}^{\mathcal{N}'}}(g(\varepsilon))$. At last, we demonstrated that $\mu = \pi_1^*(\lambda') \in \mathcal{M}_{\mathcal{F}}^{\mathcal{N}'}(g(\varepsilon))$, which concludes the proof. ☐

Now, in order to use this result, we need to dominate $\gamma_n(\mathcal{B}) \subset \mathcal{D}_{2n+1}$ by $\mathcal{B}$. By adapting a classical result, we can obtain the following bound:

Proposition 3.15 ([20, Theorem 1.3])**.** *The k-dependent noise class $\mathcal{D}_k$ is polynomially g_k-dominated by $\mathcal{B}$, with $g_k(\varepsilon) \leq C\varepsilon^{1/(2k+1)^d}$ for some constant C that does* not *depend on k nor d.*

Corollary 3.16. *If the SFT $\Omega_{\mathcal{F}}$ is stable (resp. polynomially stable) on the Bernoulli class $\mathcal{B}$, then it is also stable (resp. polynomially stable) on any dependent class $\mathcal{D}_k$.*

Under the further assumption that there exists a conjugacy with some other SFT $\theta : \Omega_{\mathcal{F}} \to \Omega_{\mathcal{G}}$, then $\Omega_{\mathcal{G}}$ is also stable (resp. polynomially stable) on the class $\mathcal{B}$.

Proof. For the first part of the result, assume that $\Omega_{\mathcal{F}}$ is f-stable on $\mathcal{B}$. As the class $\mathcal{D}_k$ is g_k-dominated by $\mathcal{B}$, we may apply Proposition 3.14, so the SFT is $(f \circ g_k)$-stable on $\mathcal{D}_k$. In particular, for the polynomial case, if f is $O\left(\varepsilon^{\alpha}\right)$, then $f \circ g_k$ is $O\left(\varepsilon^{\alpha/(2k+1)^d}\right)$, still of polynomial order. For the second part of the result, we may use Theorem 3.10, as $\Omega_{\mathcal{F}}$ is now $\left(f \circ g_{2r_{\theta^{-1}}+1}\right)$-stable on $\gamma_{r_{\theta^{-1}}}(\mathcal{B}) \subset \mathcal{D}_{2r_{\theta^{-1}}+1}$. In particular, if $f \circ g_{2r_{\theta^{-1}}+1}$ is $O\left(\varepsilon^{\alpha/\left(4r_{\theta^{-1}}+3\right)^d}\right)$, then so is $D_\theta f \circ g_{2r_{\theta^{-1}}+1}\left(C_{\theta^{-1}}\varepsilon\right) + E\varepsilon$. □

Notice how, because of the domination, we are unable to preserve *linear* stability. Still, we have proven that stability on the class $\mathcal{B}$ is a conjugacy-invariant property. In particular, stability on the class $\mathcal{B}$ is an intrinsic property of an SFT $\Omega_{\mathcal{F}}$, which actually does not depend on the set of forbidden patterns $\mathcal{F}$ used to describe it.

As $g_k(\varepsilon) \approx \varepsilon^{1/(2k+1)^d} \underset{k\to\infty}{\longrightarrow} 1$ for any fixed value of ε, we conclude that even though an SFT stable on $\mathcal{B}$ is stable on all the classes $\mathcal{D}_k$, this stability does not reach the limit class $\mathcal{D} = \bigcup_{k\in\mathbb{N}} \mathcal{D}_k$, which would be the most natural generalisation of $\mathcal{B}$ stable under all the morphisms γ_k.

Remark 3.17 (Other Classes of Noise)**.** So far, we have only talked about noises in the class $\mathcal{D}$ of finite-range dependence, which we brought back to the independent case $\mathcal{B}$. This focus is purposeful, as pretty much all our further stability results will be proven on the class $\mathcal{B}$.

If we consider infinite-range dependencies, then we allow periodic noises, *i.e.* noises defined as uniform laws among the finite set of translations of a periodic configuration $b \in \Omega_{\{0,1\}}$. In most of the interesting cases, the rigid structure of such noises allows us to explicitly construct measures that do not converge to $\mathcal{M}_{\mathcal{F}}$ for d_B, as in Subsections 4.2 and 5.1.

The remaining in-between case would be that of infinite-range dependencies but with correlations that decrease and go to 0 as the distance goes to ∞. This case notably encompasses the Gaussian Free Field, as well as some Gibbs measures. This may be the most physically realistic case, but is also the harder to study, so we will set it aside for the rest of this exploratory work.

4 Classification of the one-dimensional stability

Now that we have proved a general conjugacy invariance of the stability, let us focus on a more specific framework, the one-dimensional (1D) case. This case has already been widely studied, and since the set of configurations of a 1D SFT can be seen as the set of bi-infinite paths in an word automaton, a lot of properties have been classified [21].

The section will be concluded by a discussion on how to transpose general SFTs from d to $d+1$ dimensions while preserving their (un)stable behaviour; this subsection is more technical and may be skipped without harming the reading of the rest of the article.

We will now briefly introduce the main tool allowing for such a classification, word automata, and then use it to classify stability as a consequence of the aperiodicity of an automaton. To put it shortly, stability of the SFT will be roughly equivalent to the uniqueness of a communication class in the automaton, which must be aperiodic.

4.1 1D SFTs and word automata

In the 1D case, patterns and configurations are also called *words*. Because of their linear structure, words exhibit some automatic properties not encountered in higher dimensions.

Definition 4.1 (Diameter of a Set of Forbidden Patterns)**.** *For a window of cells $I \subset \mathbb{Z}$, we denote $d(I) = \max(I) - \min(I)$ its diameter. For a word $w \in \mathcal{A}^I$, $d(w) = d(I)$. Finally, for a set of forbidden patterns $\mathcal{F}$, we denote $d(\mathcal{F}) = \max_{w\in\mathcal{F}} d(w)$ its maximal diameter.*

Consider an automaton $G^d_{\mathcal{A}}$, a directed graph with labelled edges, where states are words in $\mathcal{A}^d$, with transitions $au \xrightarrow{b} ub$ for any $u \in \mathcal{A}^{d-1}$ and $a, b \in \mathcal{A}$. There is a natural correspondence between bi-infinite words $w \in \mathcal{A}^{\mathbb{Z}}$ and bi-infinite sequences of transitions in this word automaton.

Note that this definition looks at words left-to-right, but we could likewise look at right-to-left $ub \xrightarrow{a} au$ transitions without changing any of the following (a)periodicity properties nor the (in)stability results they imply.

Definition 4.2 (Word Automaton)**.** *Consider a set of forbidden words $\mathcal{F}$. We define the automaton $G_{\mathcal{F}}$ induced by restricting $G^{d(\mathcal{F})}_{\mathcal{A}}$ to the states $w \in \mathcal{A}^d$ that contain no forbidden pattern, that are locally admissible.*

Note that a configuration $w \in \Omega_{\mathcal{A}}$ corresponds to a bi-infinite sequence of transitions of the automaton $G_{\mathcal{F}}$ if and only if $w \in \Omega_{\mathcal{F}}$ is a configuration of the SFT.

A SFT $\Omega_{\mathcal{F}}$ can be equivalently described by an automaton $G_{\mathcal{F}}$ instead of a set of forbidden patterns $\mathcal{F}$, and we can conversely compute $\mathcal{F}'$ out of $G_{\mathcal{F}}$ so that $\Omega_{\mathcal{F}'} = \Omega_{\mathcal{F}}$.

As the number of states is finite, an infinite path exists if and only if $G_{\mathcal{F}}$ contains a cycle, which allows us to algorithmically decide whether $\Omega_{\mathcal{F}} = \emptyset$ is empty or not in polynomial time.

Definition 4.3 (Weakly Irreducible Automaton)**.** *Two states $u, v \in \mathcal{A}^d$ of $G_{\mathcal{F}}$ communicate if there is a path from u to v and v to u in the directed graph induced by $G_{\mathcal{F}}$.*

This gives us a partial equivalence relation, whose classes are the communication classes. As long as $\Omega_{\mathcal{F}} \neq \emptyset$, there is a cycle in $G_{\mathcal{F}}$ so such a class always exists.

We say that $G_{\mathcal{F}}$ is weakly irreducible if this class is unique. Note that this does not imply that all *the states of $G_{\mathcal{F}}$ are in the class. For example, in the directed graph represented by $a \to b \circlearrowleft$, $\{b\}$ is the only communication class, because there is no path from b to a.*

Note that this definition differs from the usual (stronger) notion of irreducible directed graph [21, Definition 2.2.13], that requires *all* the vertices to be in the unique communication class. We use here a weaker notion because "purely" transient states (not part of any class) will not prevent stability in the irreducible aperiodic case in Theorem 4.8, and only affect the value of the constant in the $O(\varepsilon)$ convergence speed.

Definition 4.4 (Periodic Automaton)**.** *Consider $G_{\mathcal{F}}$ a weakly irreducible automaton. We say that it is p-periodic if p is the greatest common divisor of the lengths of all the cycles found inside $G_{\mathcal{F}}$. $G_{\mathcal{F}}$ is* aperiodic *if $p = 1$.*

In the p-periodic case, there exists a partition $C = \bigsqcup_{j\in\mathbb{Z}/p\mathbb{Z}} C_j$ of the communication class such that for any transition $u \to v$ of the automaton we must have $u \in C_j$ and $v \in C_{j+1}$ for some $j \in \mathbb{Z}/p\mathbb{Z}$.

4.2 A uniquely ergodic unstable example

For a stable SFT, as ε goes to 0, a generic noisy configuration has arbitrarily few differences with a generic clear configuration of the SFT. In the specific case of uniquely ergodic SFTs, since there is only one measure in $\mathcal{M}_{\mathcal{F}}$, a stronger structure is expected for generic clear configurations, hence a prior motivation to study this case in particular.

In the 1D case, uniquely ergodic SFTs are reduced to the finite orbit of a periodic configuration. Hence, consider the simplest non-trivial uniquely ergodic 1D SFT, whose only two configurations are $\omega_0 = (01)^{\mathbb{Z}}$ and $\omega_1 = (10)^{\mathbb{Z}}$ (such that $\omega_i(k) \equiv k + i[2]$). This system is induced by the forbidden patterns $\mathcal{F} = \{00, 11\}$, it admits a unique invariant measure (hence it is uniquely ergodic), and it is irreducible 2-periodic.

We define the p-periodic noise ν_p, uniform among the p translations of $\left(0^{p-1}1\right)^{\mathbb{Z}}$. With this noise, $\nu_p([1]) = \frac{1}{p}$ goes to 0 as $p \to \infty$.

Consider then $\lambda_p \in \widetilde{\mathcal{M}_{\mathcal{F}}}\left(\frac{1}{p}\right)$ such that $\pi_2^*(\lambda_p) = \nu_p$, and on each clear window of size $p-1$, we use alternatively the restriction of ω_0 or ω_1. Up to the values under obscured cells, which will bear no influence on the following proposition, we may assume without loss of generality that λ_p is supported on $2p$-periodic configurations.

Proposition 4.5. *Let $\Omega_{\mathcal{F}}$ be the 1D uniquely ergodic SFT defined in the previous paragraphs. We have $d_B\left(\pi_1^*(\lambda_p), \mathcal{M}_{\mathcal{F}}\right) = \frac{1}{2} - O\left(\frac{1}{p}\right)$.*

Proof. Consider (w, b) a $2p$-periodic configuration for λ_p, and an interval $I = [\![k, k+2p[\![$ of size $2p$. The restriction of w in the window must coincide with the restriction of ω_0 in (at least) $p-1$ cells, and *cannot* coincide with ω_0 on the $p-1$ cells specifically aligned with ω_1, thus $d_{2p}\left(w|_I, \omega_0|_I\right) \geq \frac{p-1}{2p} = \frac{1}{2} - \frac{1}{2p}$. More generally, for any choice of $n = 2pq + r$ with $0 \leq r < 2p$, and any interval I of size n, which contains q distinct intervals of size $2p$, $d_n\left(w|_I, \omega_0|_I\right) \geq \frac{q(p-1)}{n} = \frac{p-1}{2p+\frac{r}{q}} \underset{n\to\infty}{\longrightarrow} \frac{1}{2} - \frac{1}{2p}$, which naturally gives the limit bound $d_H\left(w, \omega_i\right) \geq \frac{1}{2} - \frac{1}{2p}$.

As the lower bound on d_H holds for any configuration (w, b) in the support of λ_p and both globally admissible configurations ω_0 and ω_1, it extends to $d_B\left(\pi_1^*(\lambda_p), \mathcal{M}_{\mathcal{F}}\right)$. ☐

We can generalise this result to all periodic SFTs without much effort, provided we use periodic noises of the form $\left(0^p 1^d\right)^{\mathbb{Z}}$ with $d \geq d(\mathcal{F})$ and $p \to \infty$. We will instead exhibit another d_B-instability, but with Bernoulli noises in a further subsection.

4.3 Stability for weakly irreducible aperiodic automata

In a 1D setup, as long as $\Omega_{\mathcal{F}} \neq \emptyset$, there is always a cycle in the word automaton, thus a periodic configuration. The aperiodicity of the automaton only implies the *existence* of aperiodic configurations in $\Omega_{\mathcal{F}}$ (unless the aperiodic class only has one state that self-loops, which corresponds to a 1-periodic configuration $a^{\mathbb{Z}}$ for some $a \in \mathcal{A}$), which will prove to be sufficient to obtain stability.

Let us denote $L\left(\Omega_{\mathcal{F}}\right)$ the language of the SFT, the set of words in $\mathcal{A}^*$ that are a restriction of a configuration of $\Omega_{\mathcal{F}}$.

Remark 4.6 (Aperiodic Automata and Mixing SFTs)**.** Consider a set of forbidden words $\mathcal{F}$ such that the automaton $G_{\mathcal{F}}$ has a unique communication class which is aperiodic. It easily follows from the aperiodicity of $G_{\mathcal{F}}$ that there exists a constant $n_0 \in \mathbb{N}$ such that, for any $u, v \in L\left(\Omega_{\mathcal{F}}\right)$ and $n \geq n_0$, there exists a word $w \in \mathcal{A}^n$ such that $uwv \in L\left(\Omega_{\mathcal{F}}\right)$. This constant can easily be computed from $G_{\mathcal{F}}$ in polynomial time – with respect to the size $\left(|\mathcal{A}| + \sum_{w\in\mathcal{F}} |w|\right) \in \mathbb{N}$ for example.

For more details on the basic properties of the 1D case, one may refer to the classic book by Lind and Marcus [21].

Assuming we can cut an obscured configuration from $\Omega_{\widetilde{\mathcal{F}}}$ into globally admissible words all distant by at least n_0, then we will be able to rewrite these gaps in order to obtain a globally admissible configuration. If we only exclude obscured cells, then we will obtain a sequence of locally admissible clear words instead, that may not be globally admissible, and the gaps between these words may be too small. By leaving out the

$\left\lceil \frac{n_0}{2} \right\rceil$-neighbourhood around each obscured cell, we make sure the gaps are big enough to be fillable.

The following proposition gives a stronger 1D version of the reconstruction function φ described in Remark 2.4.

Proposition 4.7. *Given a set of 1D forbidden patterns $\mathcal{F}$, there exists a constant $C(\mathcal{F})$ such that, for any locally admissible word $u \in \mathcal{A}^*$, by removing (at most) C letters on each end, we obtain instead a globally admissible word $v \in L(\Omega_{\mathcal{F}})$.*

Proof. Note that a path of length n in $G_{\mathcal{F}}$ visits $n+1$ vertices (each a word of length d), and represents a word of length $d(\mathcal{F})+n$. Thus, we may assume that $C \geq \frac{d(\mathcal{F})}{2}$, so that we only need to consider words long enough to represent a finite path in the automaton $G_{\mathcal{F}}$.

As long as we visit vertices in the communication class of $G_{\mathcal{F}}$, we can infinitely extend the path on both directions, thence the word we encode is globally admissible.

Issues arise when we visit other states, which explicitly correspond to vertices that never occur in a bi-infinite path, thus non-globally admissible words. As there is only one communication class, no path can cycle through such a state. Hence, if there are k states of $G_{\mathcal{F}}$ outside of the communication class, by removing k states on each end of the path, we make sure that the path only visits the communication class, thus corresponds to a globally admissible word. Hence, $C = \max\left(k, \left\lceil \frac{d(\mathcal{F})}{2} \right\rceil\right)$ is big-enough.

If we want a better constant, we can replace k by the maximum of the length of the longest path among vertices *outside of* yet *connected to* the communication class, and half of the longest path not connected to the class. □

Just like n_0, C can be computed from $G_{\mathcal{F}}$ in polynomial time. Now, if we remove a C-neighbourhood around each obscured cell, then we obtain a sequence of globally admissible words. Finally, by removing a D-neighbourhood with $D = \max\left(C, \left\lceil \frac{n_0}{2} \right\rceil\right)$, we make sure that we obtain alternately globally admissible words and fillable gaps. This is the key idea of the following theorem, whose proof mostly aims at properly explaining why the transformation we perform is a σ-invariant morphism that returns a clear globally admissible configuration.

Theorem 4.8. *Let $\Omega_{\mathcal{F}}$ be a 1D SFT with a weakly irreducible aperiodic automaton $G_{\mathcal{F}}$. Then $\Omega_{\mathcal{F}}$ is linearly stable, with an explicit constant in the $O(\varepsilon)$.*

Proof. In order to obtain linear stability, we will consider a measure $\lambda \in \widetilde{\mathcal{M}_{\mathcal{F}}}(\varepsilon)$, and build a measurable mapping $\psi : \Omega_{\widetilde{\mathcal{F}}} \to \Omega_{\mathcal{F}}$, so that $d_H(\pi_1(\omega), \psi(\omega))$ is small. Let us notice that ψ does not need to be defined on $\Omega_{\widetilde{\mathcal{F}}}$, but only on a high-probability support $S \subset \Omega_{\widetilde{\mathcal{F}}}$. In such a case, we may add a third *independent* coordinate to λ that follows some given law in $\mathcal{M}_{\mathcal{F}}$, and project onto this coordinate with ψ outside of the event S, to simply use the trivial upper bound $d_H \leq 1$ in this low-probability case. This way:

$$d_B\left(\pi_1^*(\lambda), \mathcal{M}_{\mathcal{F}}\right) \leq \int_S d_H\left(\pi_1(\omega), \psi(\omega)\right) \mathrm{d}\lambda(\omega) + \lambda\left(S^c\right).$$

Consider the cellular automaton γ_D on $\Omega_{\{0,1\}}$, as defined in Definition 3.8. This morphism obscures the cells in the D-neighbourhood, as described in Definition 3.8. This process is clearly measurable, and naturally extends as a morphism on $\Omega_{\widetilde{\mathcal{F}}}$. We now need to map this subset of $\Omega_{\widetilde{\mathcal{F}}}$ into $\Omega_{\mathcal{F}}$ in a measurable way.

The issue now is that, while we can manually fill each gap, issues may arise with the order of the operations. Indeed, assume we decide on a word $w(u_1, u_2, n)$ for any words $u_1, u_2 \in L(\Omega_{\mathcal{F}})$ and any gap of size $n \geq n_0$, as in Remark 4.6. Naturally, if we have *three* globally admissible words u_1, u_2 and u_3 as well as two gaps i and j, then we can fill the leftmost gap first, with $v = u_1 w\left(u_1, u_2, i\right) u_2$ and then the second one with $vw\left(v, u_3, j\right) u_3$.

There are several ways to proceed, but we chose here to be able to fill those gaps simultaneously, so that the σ-invariance of the morphism directly follows. To ensure we can fill the gaps simultaneously, we simply need to know the leftmost and rightmost states of $G_{\mathcal{F}}$ corresponding to u_2, which requires in turn $|u_2| \geq d(\mathcal{F})$. By looking at the $\left\lceil \frac{d(\mathcal{F})}{2} \right\rceil$-neighbourhood of a clear cell in a configuration of $\gamma_D\left(\Omega_{\{0,1\}}\right)$, we can see whether it belongs to a long-enough globally admissible clear word, and obscure it if it does not. Let us name θ the cellular automaton on $\Omega_{\{0,1\}}$ obtained by applying γ_D and then this new measurable process. We identify θ with the morphism on $\Omega_{\widetilde{\mathcal{F}}}$ that leaves the first coordinate unchanged.

For a configuration $\omega \in \Omega_{\widetilde{\mathcal{F}}}$, the obscured cells in $\theta(\omega)$ are all in an E-neighbourhood (with $E = D + \left\lceil \frac{d(\mathcal{F})}{2} \right\rceil$) of the original obscured cells, so we still have a linear control on the frequency of obscured cells.

Now, all clear words of an obscured configuration $\theta(\omega)$ are of length at least $d(\mathcal{F})$. For such words, we can define $w\left(u_1, u_2, n\right)$ using only the d rightmost letters of u_1 and the d leftmost letters of u_2, which won't change if we change letters on the other end of u_1 or u_2. Several choices may be possible, what matters is to choose one. Then, we can simultaneously replace all the entirely obscured windows of $\theta(\omega)$ of length n_0 by the corresponding clear words, using a cellular automaton, hence in a measurable σ-invariant way. We can iterate the process for each length $n \geq n_0$, to obtain a configuration $\psi(\omega)$ at the limit, still in a measurable σ-invariant way.

There is one last issue to deal with, *i.e.* the fact that $\psi(\omega)$ consists of one big globally admissible clear word, but that it may have an infinite obscured window on the left or the right. Let us name S the set of configurations where this phenomenon does not happen. So far, we obtained a set S and defined a morphism $\psi : S \to \Omega_{\mathcal{F}}$, as stated in the first paragraph of the proof, so let us now study the two terms of the bound.

First, inside of S, $d_H\left(\pi_1(\omega), \psi(\omega)\right) \leq d_H\left(\pi_2(\theta(\omega)), 0^\infty\right) \leq (2E+1) d_H\left(\pi_2(\omega), 0^\infty\right)$. Thus, $\int_S d_H\left(\pi_1(\omega), \psi(\omega)\right) \mathrm{d}\lambda(\omega) \leq (2E+1) \int d_H\left(\pi_2(\omega), 0^\infty\right) \mathrm{d}\lambda(\omega)$. So far, the bound holds for any configuration in $\Omega_{\widetilde{\mathcal{F}}}$, any measure $\lambda \in \mathcal{M}_{\widetilde{\mathcal{F}}}$.

Assume now that $\lambda \in \widetilde{\mathcal{M}_{\mathcal{F}}}(\varepsilon)$. Then, using Birkhoff's pointwise ergodic theorem, $\int d_H\left(\pi_2(\omega), 0^\infty\right) \mathrm{d}\lambda(\omega) = \pi_2^*(\lambda)([1]) \leq \varepsilon$. We just need to study $\lambda\left(S^c\right)$ to conclude. Now, $\lambda\left(S^c\right) \leq \lambda\left(T_L\right) + \lambda\left(T_R\right)$ where T_L (resp. T_R) is the event where there is a infinite obscured window on the left (resp. right) in the configuration $\theta(\omega)$. If $\omega \in T_*$, then 1 must at least have a $\frac{1}{2D+1}$ density in the configuration $\pi_2(\omega)$ to begin with, by averaging in the appropriate direction, so that $\lambda\left(T_*\right) \times \frac{1}{2D+1} + \lambda\left(T_*^c\right) \times 0 \leq \varepsilon$ and $\lambda\left(T_*\right) \leq (2D+1)\varepsilon$.

At last, we obtain the explicit bound $d_B\left(\pi_1^*(\lambda), \mathcal{M}_{\mathcal{F}}\right) \leq 3(2E+1)\varepsilon$, with E an explicit constant, computable in polynomial time. □

Remark that, when using independent ε-Bernoulli noises, as $\lambda(S) = 1$, then we lose the factor 3 in this upper bound, but the constant is still in the same general order of magnitude.

4.4 Instability for weakly irreducible periodic automata

In the previous subsection, we proved stability in the aperiodic case. The proof made full use of aperiodicity, in the sense that the obscured cells can induce a gap of arbitrary size into any globally admissible configuration, and aperiodicity is needed to guarantee such a gap is fillable. We also clearly saw how this approach failed in the introductory 2-periodic example, using a periodic noise to precisely quantify the amount of differences on finite windows in order to obtain d_B-instability. Our objective is now to prove a broader periodic d_B-instability result, but for seemingly more natural noisy configurations, using ε-Bernoulli noises.

Theorem 4.9 (Periodic Instability)**.** *Consider a 1D SFT $\Omega_{\mathcal{F}}$ such that $G_{\mathcal{F}}$ is weakly irreducible p-periodic ($p \geq 2$). Then for any $\varepsilon > 0$ there exists $\mu_\varepsilon \in \mathcal{M}^{\mathcal{B}}_{\mathcal{F}}(\varepsilon)$ such that $d_B(\mu_\varepsilon, \mathcal{M}_{\mathcal{F}}) \geq \frac{p-1}{pd(\mathcal{F})} - \varepsilon$.*

Proof. Let us begin by considering the partition of $\Omega_{\mathcal{F}}$ into p sets $(\Omega_j)_{j\in\mathbb{Z}/p\mathbb{Z}}$ induced by the states of $G_{\mathcal{F}}$, so that if $\omega \in \Omega_i$, then $\sigma_k(\omega) \in \Omega_{i+k}$.

Consider also once and for all a periodic word $\omega_0 \in \Omega_{\mathcal{F}}$, that corresponds to an infinite cycle of $G_{\mathcal{F}}$. Note that this cycle may not be of length p but a multiple of it – *e.g.* if $G_{\mathcal{F}}$ is made of a 6-cycle and a 10-cycle joined in a vertex, it is 2-periodic but has no 2-cycle. What matters is that ω_0 has a finite orbit under translations. What is more, by looking at a window of size $d(\mathcal{F})$ of a translation of ω_0, we can identify to which state of $G_{\mathcal{F}}$ it corresponds and thus deduce to which class Ω_j the translated configuration belongs to. To construct $\mu_\varepsilon \in \mathcal{M}^{\mathcal{B}}_{\mathcal{F}}(\varepsilon)$, consider the measure λ_ε obtained by:

1. taking the independent Bernoulli noise $\mathcal{B}(\varepsilon)^{\otimes\mathbb{Z}}$ first,
2. identifying intervals of consecutive obscured cells, of length at least $d(\mathcal{F})$, and writing down letters of $\mathcal{A}$ uniformly at random under each such block,
3. in-between two such intervals, in a window that must have a clear cell on each end and may contain some short obscured blocks in the middle, we choose uniformly at random a translation of ω_0 to write it down on the cells, whether clear or obscured.

It is apparent that this measure has an ε-Bernoulli noise, and that it is σ-invariant by construction. The measure λ_ε is also strongly mixing, thus ergodic. Indeed, consider two finite windows $I, J \subset \mathbb{Z}$ such that $\min(J) - \max(I) = n > d(\mathcal{F})$. Conditionally to the fact that the window $[\![\max(I)+1, \min(J)-1]\!]$ contains an obscured window of size $d(\mathcal{F})$, the windows I and J behave independently from each other. As the probability of having such an obscured window goes to 1 as $n \to \infty$, we deduce the mixing property on cylinders, so that λ_ε itself is strongly mixing.

Hence, if we cut down $\mathbb{Z}$ into consecutive windows of length $d(\mathcal{F})$, we obtain a measure on $\left(\mathcal{A}^d \times \{0,1\}^d\right)^{\mathbb{Z}}$. This induced measure is *also* σ-invariant and strongly mixing (thus ergodic), so that Birkhoff's pointwise ergodic theorem applies. Hence, the frequency of a $d(\mathcal{F})$-interval in a configuration is λ_ε-a.s. equal to its probability under λ_ε.

A clear $d(\mathcal{F})$-interval has probability $(1-\varepsilon)^{d(\mathcal{F})}$ of happening, which we bound below by $1 - d(\mathcal{F})\varepsilon$. Under such an event, by construction, we can identify the state of $G_{\mathcal{F}}$ it represents, thus to which class Ω_i it comes from. Note that on such a clear window, if ω and ω' belong to different classes Ω_i and Ω_j, then in particular they correspond to different states of $G_{\mathcal{F}}$ thus must differ in *at least* one cell.

Thus, for *any* globally admissible state $\omega \in \Omega_{\mathcal{F}}$ and λ_ε-a.e. locally admissible state $(\omega', b) \in \Omega_{\widetilde{\mathcal{F}}}$, we have:

$$d_H(\omega, \omega') \geq (1 - d(\mathcal{F})\varepsilon) \times \frac{p-1}{p} \times \frac{1}{d(\mathcal{F})}.$$

The first factor comes from the frequency of clear windows, the second one from the probability of ω' not being in the same class as ω conditionally to some clear window, and the third one from the minimal number of differences in such a window of size $d(\mathcal{F})$ under the previous event.

It immediately follows that $d_B(\mu_\varepsilon, \mathcal{M}_{\mathcal{F}}) \geq \frac{p-1}{pd(\mathcal{F})} - \frac{p-1}{p}\varepsilon$, which concludes the proof. □

Note that this proof can be adapted from the periodic case to the non-irreducible case where there are several communication classes, by using finite trajectories evolving inside distinct communication classes instead of words "aligned" along different periods of the system, even if all the classes are aperiodic. We thus obtain the following theorem, the proof of which we will omit for the sake of brevity, as it offers no further insight on the topic.

Theorem 4.10 (Non-Irreducible Instability). *Consider an SFT $\Omega_{\mathcal{F}}$ such that $G_{\mathcal{F}}$ is a non-irreducible automaton, with p communication classes ($p \geq 2$). Then for any $\varepsilon > 0$ there exists a measure $\mu_\varepsilon \in \mathcal{M}_{\mathcal{F}}^{\mathcal{B}}(\varepsilon)$ such that $d_B(\mu_\varepsilon, \mathcal{M}_{\mathcal{F}}) \geq \frac{p-1}{pd(\mathcal{F})} - \varepsilon$.*

4.5 Extension to higher dimensions

This subsection dives deeper into the intricacies of couplings from a measure theory viewpoint, which offers a different insight on the objects we are working on, but can also be skipped by an unfamiliar reader as it is independent of everything that will follow.

Our goal here is to translate both stable and unstable SFTs into higher dimensions, from the one-dimensional case in particular. First of all, let us quickly characterise σ-invariant couplings, which will be useful for the main result of this subsection.

Lemma 4.11 (σ-Invariant Disintegration). *Consider λ a probability measure on $\Omega_{\mathcal{A}} \times \Omega_{\mathcal{A}}$, with $\mu = \pi_1^*(\lambda)$, and its disintegration $\mathrm{d}\lambda(x,y) = \mathrm{d}\nu_x(y)\mathrm{d}\mu(x)$ as in Lemma 3.13.*

Assume now that μ is σ-invariant. Then λ is also σ-invariant if and only if the equality $\nu_x(B) = \nu_{\sigma_k(x)}(\sigma_k(B))$ holds for any $k \in \mathbb{Z}^d$, any measurable cylinder B, and μ-a.e. $x \in \Omega_{\mathcal{A}}$.

Proof. Consider two cylinders A and B, as well as $k \in \mathbb{Z}^d$. As stated, we have the equality $\lambda(A \times B) = \int_A \nu_x(B)\mathrm{d}\mu(x)$. Likewise, as μ itself is σ-invariant:

$$\lambda(\sigma_k(A \times B)) = \lambda(\sigma_k(A) \times \sigma_k(B)) = \int_{\sigma_k(A)} \nu_x(\sigma_k(B))\,\mathrm{d}\mu(x) = \int_A \nu_{\sigma_k(y)}(\sigma_k(B))\,\mathrm{d}\mu(y).$$

Now, the measure λ is σ-invariant if and only if, for any cylinder B and $k \in \mathbb{Z}^d$, we have $\lambda(A \times B) = \lambda(\sigma_k(A \times B))$ for any cylinder A. Using the integral expressions, $\int_A \nu_x(B)\mathrm{d}\mu(x) = \int_A \nu_{\sigma_k(x)}(\sigma_k(B))\,\mathrm{d}\mu(x)$. It is equivalent for this equality to hold for any A and for the functions to be μ-a.s. equal, which concludes the proof. $\square$

Note how the measures ν_x are not necessarily σ-invariant. In particular, using the Dirac measures $\nu_x = \delta_x$ – which are obviously not σ-invariant – gives us a *diagonal* coupling between μ and itself, such that $\pi_2^*(\lambda) = \mu$ too, which is σ-invariant.

Now, given a d-dimensional SFT $\Omega_{\mathcal{F}}$, it is possible to extend $\mathcal{F}$ into $\mathcal{F}'$ in $d+1$ dimensions, by replacing every forbidden pattern $w \in \mathcal{A}^{I(w)}$ on the window $I(w) \subset \mathbb{Z}^d$ by $w' \in \mathcal{A}^{I(w)\times\{0\}}$ with $I(w) \times \{0\} \subset \mathbb{Z}^{d+1}$. This way, $\Omega_{\mathcal{F}'} = \left\{(\omega_i)_{i\in\mathbb{Z}}, \forall i \in \mathbb{Z}, \omega_i \in \Omega_{\mathcal{F}}\right\}$. In other words, each slice (with a fixed last coordinate) represents a copy of the original SFT, with no constraints on how to align the slices. In particular, if $\mu \in \mathcal{M}_{\mathcal{F}}(\varepsilon)$, then by coupling all these layers independently, we obtain $\mu^{\otimes\mathbb{Z}} \in \mathcal{M}_{\mathcal{F}'}(\varepsilon)$.

Let us now prove that (in)stability of an SFT is in some sense preserved through this transformation. Thus, as we exhibited (un)stable 1D examples earlier in this section, this will imply the existence of (un)stable systems in *any* dimension.

Consider the projection $\zeta : b \in \{0,1\}^{\mathbb{Z}^{d+1}} \mapsto b|_{\mathbb{Z}^d\times\{0\}} \in \{0,1\}^{\mathbb{Z}^d}$, that commutes with translations in $\mathbb{Z}^d$. More generally, we will use ζ as a multipurpose projector for any alphabet $\mathcal{A}$ instead of $\{0,1\}$. For a given class of $(d+1)$-dimensional noises $\mathcal{N}'$, we obtain the d-dimensional class $\mathcal{N} = \zeta^*(\mathcal{N}')$. In particular, if $\mathcal{N}'$ is the class of $(d+1)$-dimensional Bernoulli noises, then $\mathcal{N}$ is the class of d-dimensional Bernoulli noises.

To make things easier to read, we will distinguish the Besicovitch distances d_B in d dimensions and d'_B in $d+1$ dimensions (resp. d_H and d'_H).

Proposition 4.12. *Using the projection ζ introduced in the previous paragraph, by extending the d-dimensional forbidden patterns $\mathcal{F}$ as $\mathcal{F}'$ and assuming that $\mathcal{N} = \zeta^* (\mathcal{N}')$:*

1. *For any $\mu' \in \mathcal{M}_{\mathcal{F}'}^{\mathcal{N}'}(\varepsilon)$, we have $\mu = \zeta^* (\mu') \in \mathcal{M}_{\mathcal{F}}^{\mathcal{N}}(\varepsilon)$.*
2. *For any $\mu \in \mathcal{M}_{\mathcal{F}}^{\mathcal{N}}(\varepsilon)$, there exists $\mu' \in \mathcal{M}_{\mathcal{F}'}^{\mathcal{N}'}(\varepsilon)$ such that $\mu = \zeta^* (\mu')$.*
3. *In both cases, $d_B (\mu, \mathcal{M}_{\mathcal{F}}) = d'_B (\mu', \mathcal{M}_{\mathcal{F}'})$.*

Proof. In all cases, going from dimension $d+1$ to dimension d is a mere matter of projection through ζ, whereas going from dimension d to $d+1$ is a bit trickier and will require us to make use of Lemma 4.11. Note that the n in the couplings λ_{n} used below, to prove the first two items, stands for the noise on its second coordinate. This is to better distinguish it from the coupling λ with the same alphabet $\mathcal{A}$ for the two coordinates, in the proof of the last item.

First, assume there is $\lambda'_{\mathrm{n}} \in \widetilde{\mathcal{M}_{\mathcal{F}'}^{\mathcal{N}'}}(\varepsilon)$ such that $\mu' = \pi_1^* (\lambda'_{\mathrm{n}}) \in \mathcal{M}_{\mathcal{F}}$. Thus, we have $\lambda_{\mathrm{n}} = \zeta^* (\lambda'_{\mathrm{n}}) \in \widetilde{\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}}(\varepsilon)$, so that:

$$\mu = \zeta^* (\mu') = \zeta^* (\pi_1^* (\lambda'_{\mathrm{n}})) = \pi_1^* (\lambda_{\mathrm{n}}) \in \mathcal{M}_{\mathcal{F}}^{\mathcal{N}}(\varepsilon).$$

This proves the first item.

Conversely, consider $\lambda_{\mathrm{n}} \in \widetilde{\mathcal{M}_{\mathcal{F}}^{\mathcal{N}}}(\varepsilon)$ such that $\mu = \pi_1^* (\lambda_{\mathrm{n}})$, and let us build the desired measure μ'. Using Lemma 4.11, we have $\mathrm{d}\lambda_{\mathrm{n}}(\omega, b) = \mathrm{d}\mu_b(\omega)\mathrm{d}\nu(b)$ with $\nu = \pi_2^* (\lambda_{\mathrm{n}}) \in \mathcal{N}$ an ε-noise, and $\mathrm{d}\mu_{\sigma_k(b)} (\sigma_k(\omega)) = \mathrm{d}\mu_b(\omega)$ for ν-a.e. $b \in \Omega_{\{0,1\}}$. As $\nu \in \mathcal{N} = \zeta^* (\mathcal{N}')$, there is $\nu' \in \mathcal{N}'$ such that $\nu = \zeta^*(\nu')$. In particular, ν' is σ-invariant and must be an ε-noise too. Now, for any families of d-dimensional layers $\omega' = (\omega_i)_{i\in\mathbb{Z}} \in \Omega_{\mathcal{F}'}$ and $b' = (b_i)$, we define the measures $\mathrm{d}\mu'_{b'} (\omega') = \prod_{i\in\mathbb{Z}} \mathrm{d}\mu_{b_i} (\omega_i)$ and then $\mathrm{d}\lambda'_{\mathrm{n}} (\omega', b') = \mathrm{d}\mu_{b'} (\omega') \,\mathrm{d}\nu' (b')$. Naturally, the measures $\mu'_{b'}$ are $\sigma_{e_{d+1}}$-invariant – invariant by translations on the last coordinate – by construction, and satisfy the criterion of Lemma 4.11 because the measures μ_b did. Thence, λ'_{n} is σ-inviariant, so that $\lambda'_{\mathrm{n}} \in \widetilde{\mathcal{M}_{\mathcal{F}'}^{\mathcal{N}'}}(\varepsilon)$. At last, $\mu' = \pi_1^* (\lambda'_{\mathrm{n}})$ is such that $\zeta^* (\mu') = \mu$, which proves the second item.

Finally, consider any two σ-invariant measures μ' and $\mu = \zeta^*(\mu)$. We begin with the easier inequality, by considering λ' a coupling between μ' and $\nu' \in \mathcal{M}_{\mathcal{F}'}$ such that $d'_B (\mu', \mathcal{M}_{\mathcal{F}'}) = d'_B (\mu', \nu') = \int d'_H (x', y') \,\mathrm{d}\lambda' (x', y')$. Using Birkhoff's pointwise ergodic theorem, this is equal to $d'_B (\mu', \nu') = \int \mathbb{1}_{x'_0 \neq y'_0} (x', y') \,\mathrm{d}\lambda' (x', y')$. Likewise, $\lambda = \zeta^* (\lambda')$ is a coupling between μ and $\nu = \zeta^* (\nu') \in \mathcal{M}_{\mathcal{F}}$, not necessarily such that d_B is reached, but:

$$d_B (\mu, \mathcal{M}_{\mathcal{F}}) \leq d_B (\mu, \nu) \leq \int \mathbb{1}_{x_0 \neq y_0} \mathrm{d}\lambda(x, y) = \int \mathbb{1}_{x'_0 \neq y'_0} \mathrm{d}\lambda' (x', y') = d'_B (\mu', \mathcal{M}_{\mathcal{F}'}) .$$

For the reverse inequality, consider λ a coupling between μ and $\nu \in \mathcal{M}_{\mathcal{F}}$ such that $d_B (\mu, \mathcal{M}_{\mathcal{F}}) = \int \mathbb{1}_{x_0 \neq y_0} \mathrm{d}\lambda$. As in Lemma 4.11, we can factorise $\mathrm{d}\lambda(x, y) = \mathrm{d}\nu_x(y)\mathrm{d}\mu(x)$. With $x', y' \in \Omega_{\mathcal{F}'} = \Omega_{\mathcal{F}}^{\mathbb{Z}}$, we define the family of measures $\mathrm{d}\nu'_{x'} (y') = \prod_{i\in\mathbb{Z}} \mathrm{d}\nu_{x_i} (y_i)$, and then $\mathrm{d}\lambda' (x', y') = \mathrm{d}\nu'_{x'} (y') \,\mathrm{d}\mu' (x')$. The measures $\nu'_{x'}$ satisfy the criterion of Lemma 4.11, so that λ' is σ-invariant. This implies that it is a coupling between the measures μ' and $\nu' = \pi_2^* (\lambda') \in \mathcal{M}_{\mathcal{F}'}$, once again not necessarily optimal, such that:

$$d'_B (\mu', \mathcal{M}_{\mathcal{F}'}) \leq d'_B (\mu', \nu') \leq \int \mathbb{1}_{x'_0 \neq y'_0} \mathrm{d}\lambda' (x', y') = \int \mathbb{1}_{x_0 \neq y_0} \mathrm{d}\lambda(x, y) = d_B (\mu, \mathcal{M}_{\mathcal{F}}) .$$

This concludes the proof of the last item. □

Corollary 4.13. *The d-dimensional SFT $\Omega_{\mathcal{F}}$ is f-stable (resp. unstable) on the class $\mathcal{N}$ if and only if the $(d+1)$-dimensional SFT $\Omega_{\mathcal{F}'}$ is f-stable (resp. unstable) on the class $\mathcal{N}'$.*

Proof. Going from (un)stability on $\Omega_{\mathcal{F}'}$ to $\Omega_{\mathcal{F}}$ is once again a simple matter of projecting measures with ζ so we won't insist further on these implications.

If $\Omega_{\mathcal{F}}$ is f-stable on $\mathcal{N}$, and $\mu' \in \mathcal{M}_{\mathcal{F}'}^{\mathcal{N}'}(\varepsilon)$, then $\mu = \zeta^*(\mu') \in \mathcal{M}_{\mathcal{F}}^{\mathcal{N}}(\varepsilon)$ using Item 1 of the previous proposition, so that $d'_B(\mu', \mathcal{M}_{\mathcal{F}'}) = d_B(\mu, \mathcal{M}_{\mathcal{F}}) \leq f(\varepsilon)$ using Item 3. Thus, $\Omega_{\mathcal{F}'}$ is f-stable.

Now, if $\Omega_{\mathcal{F}}$ is unstable, we have a sequence of measures $\mu_n \in \mathcal{M}_{\mathcal{F}}(\varepsilon_n)$ with $\varepsilon_n \underset{n\to\infty}{\longrightarrow} 0$ such that $\inf_{n\in\mathbb{N}} d_B(\mu_n, \mathcal{M}_{\mathcal{F}}) = d > 0$. Then, with the measures $\mu'_n \in \mathcal{M}_{\mathcal{F}'}^{\mathcal{N}'}(\varepsilon_n)$ given by Item 2, we conclude that $\inf_{n\in\mathbb{N}} d'_B(\mu'_n, \mathcal{M}_{\mathcal{F}'}) \geq d$ too with Item 3, thence $\Omega_{\mathcal{F}'}$ also is unstable. $\square$

Using this corollary, we can in particular extend the (un)stable 1D SFTs we exhibited earlier in order to obtain (un)stable SFTs in *any* dimension. Of course, these examples are not really satisfactory and we will now strive for other higher-dimensional examples in the following sections of this paper.

Remark 4.14 (Alphabet Extension)**.** Another way to extend SFTs, not in dimension but in alphabet size, is the direct product. If we consider two d-dimensional SFTs $\Omega_{\mathcal{F}}$ on the alphabet $\mathcal{A}$ and $\Omega_{\mathcal{F}'}$ on the alphabet $\mathcal{A}'$, then we can build the SFT $\Omega_{\mathcal{F}} \times \Omega_{\mathcal{F}'}$ on the alphabet $\mathcal{A} \times \mathcal{A}'$.

Let us note that, with $\omega_1, \omega_2 \in \Omega_{\mathcal{A}}$ and $\omega'_1, \omega'_2 \in \Omega_{\mathcal{A}'}$, we have the inequalities:

$$d_H(\omega_1, \omega_2) \leq d_H([\omega_1, \omega_2], [\omega'_1, \omega'_2]) \leq d_H(\omega_1, \omega_2) + d_H(\omega'_1, \omega'_2).$$

Thence, if $\Omega_{\mathcal{F}}$ (resp. $\Omega_{\mathcal{F}'}$) is f-stable (resp. f'-stable) on the same class $\mathcal{N}$, then the product $\Omega_{\mathcal{F}} \times \Omega_{\mathcal{F}'}$ is $(f + f')$-stable on the class $\mathcal{N}$. If one of the SFTs is unstable, then the product is unstable with the same lower bound.

5 Stability of 2D+ periodic SFTs with Bernoulli noise

In this section, we will explore the notion of stability for higher-dimensional (*2D+*) periodic SFTs. Here, by periodicity of the SFT we really mean periodicity of all its configurations, not of some associated structure like the word automaton of Section 4. In particular, a SFT $\Omega_{\mathcal{F}}$ is periodic *iff* it is a finite set [4, Theorem 3.8]. First, we will show how to obtain instability using a grid noise, like we did in Subsection 4.2 for the 1D case. We will then focus on Bernoulli noises and prove that, using a percolation argument for 2D+, we have linear stability in this framework.

There are several non-equivalent notions of periodicity in the 2D+ case. We will in this case consider the strongest notion of periodicity, *i.e.* the existence of $\mathbb{Z}$-independent vectors $x_1, \ldots, x_d \in \mathbb{Z}^d$, such that *any* configuration $\omega \in \Omega_{\mathcal{F}}$ is invariant under any translation among those ($\sigma_{x_i}(\omega) = \omega$ for any $1 \leq i \leq d$). Equivalently, we can always assume that these d vectors align with the d axes of $\mathbb{Z}^d$, so that we can actually simply repeat a base pattern defined on a hyper-rectangle along those d base directions.

Up to an added redundancy along some of those axes, we may even go one step further and assume the base pattern is defined on a hypercube whose edge-length is the smallest common multiple of those of the hyper-rectangle. This added hypothesis will worsen the constants obtained in the following proofs, but will make notations a bit lighter as a trade-off.

5.1 Instability for grid noises

Definition 5.1 (Grid Noise)**.** *Consider $k, n \in \mathbb{N}^*$ two positive integers. We define the base pattern $b_{k,n}$ on a $(k+n)$-hypercube, such that for $x \in [\![0, k+n-1]\!]^d$, we have $b(x) = 1$ iff $\min_{1\leq i\leq d} x_i < k$. We then identify $b_{k,n}$ to the configuration obtained by extending this* *base pattern in all directions. We finally define the σ-invariant noise:*

$$\nu_{k,n} = \frac{1}{(k+n)^d} \sum_{x \in [\![0,k+n-1]\!]^d} \delta_{\sigma_x(b_{k,n})}.$$

The probability of an obscured cell in this noise is $1 - \left(\frac{k}{k+n}\right)^d$.

Assuming k is greater than the maximal diameter of the forbidden patterns of $\mathcal{F}$, then two distinct clear hypercubes (both translations of $[\![0, n-1]\!]^d$) are insulated from each other, and can be tiled independently, as no forbidden pattern could have cells in both windows. We will work under this assumption from now on.

Proposition 5.2. *For any non-trivial 2D+ periodic SFT ($|\Omega_\mathcal{F}| \geq 2$), there exists a constant $\delta(\mathcal{F}) > 0$ such that, for any $\varepsilon > 0$, there is a measure $\mu \in \mathcal{M}_\mathcal{F}(\varepsilon)$ (with a grid noise) at distance at least δ from $\mathcal{M}_\mathcal{F}$.*

Proof. Let us assume that $\Omega_\mathcal{F}$ is a periodic SFT, on the base hypercube $[\![0, N-1]\!]^d$. Then two distinct configurations $\omega \neq \omega' \in \Omega_\mathcal{F}$ differ on *at least* one cell in any translation of the N-hypercube.

By monotonicity, we only need to prove it for arbitrarily small values of ε. We will prove the result for the noises $\nu_{k,nN}$ as $n \to \infty$, for which the frequency of obscured cells is equal to $\varepsilon_n = \left(\frac{k}{k+nN}\right)^d$, hence $\varepsilon_n \underset{n\to\infty}{\longrightarrow} 0$.

What we mean here by non-trivial is that there exists a non-constant configuration $\omega_0 \in \Omega_\mathcal{F}$ such that $\omega_0 \neq \sigma_{e_j}(\omega_0)$ for some $1 \leq j \leq d$, thus $\Omega_0 := \left\{\sigma_x(\omega_0), x \in \mathbb{Z}^d\right\}$ has between 2 and N^d elements.

We define the noisy measure $\lambda \in \widetilde{\mathcal{M}_\mathcal{F}}(\varepsilon_n)$ as follows:

- first, pick a noise grid at random following the measure $\nu_{k,nN}$,
- under any obscured cell pick a letter uniformly at random,
- then, independently from the noise, and independently on each clear nN-hypercube, pick a configuration $\omega \in \Omega_0$ uniformly at random, and finally restrict it to the corresponding hypercube.

Fix a configuration $\omega \in \Omega_\mathcal{F}$. For λ-a.e. noisy tiling (ω', b), on the first coordinate $\omega' \in \Omega_\mathcal{A}$, a proportion $\frac{1}{|\Omega_0|} \geq \frac{1}{N^d}$ of the clear hypercubes from b contains each translation of ω_0. Hence, in a proportion greater or equal to $\frac{|\Omega_0|-1}{|\Omega_0|} \geq \frac{1}{2}$, the configuration chosen for this hypercube is *not* ω. For such a clear window, as a translation of $[\![0, nN-1]\!]^d$, contains n^d distinct translations of $[\![0, N-1]\!]^d$. On each such sub-hypercube, ω and ω' differ on at least one cell. Finally:

$$d_H(\omega, \omega') \geq \frac{1}{2} \times \left(\frac{n}{k+nN}\right)^d.$$

This inequality holds λ-a.s. for *any* configuration $\omega \in \Omega_\mathcal{F}$, so for big enough values of $n \geq k$, we obtain the lower bound $d_B\left(\pi_1^*(\lambda), \mathcal{M}_\mathcal{F}\right) \geq \frac{1}{2(N+1)^d}$. ☐

5.2 From noisy SFTs to percolations

In the 1D case, under a Bernoulli noise, having room for aperiodicity was what helped us correct defects in the noisy configurations from $\Omega_{\widetilde{\mathcal{F}}}$ in order to couple them with globally admissible configurations in $\Omega_\mathcal{F}$, while intrinsic periodicity of the SFT was precisely what prevented stability. Yet, in the 2D+ case, we will see that periodicity *helps* stability as long as most of the clear cells are connected to each other in an induced percolation process.

Once again, let us consider a variant of the reconstruction function described in Remark 2.4. Here, $\varphi_{\mathcal{F}} : \mathbb{N} \to \mathbb{N}$ is a non-decreasing function such that, for any integer $n \in \mathbb{N}^*$, $\varphi(n) \geq n$ and whenever $\omega \in \mathcal{A}^{B_{\varphi(n)}}$ is a locally admissible pattern, its restriction $\omega|_{B_n}$ is globally admissible. For the 1D case, we proved in Proposition 4.7 that this function can always be chosen as $\varphi(n) = n + c$ for some $c \in \mathbb{N}$. This property allowed us to convert a locally admissible configuration into a globally admissible clear one, up to some "peeling" around obscured cells, in the case of aperiodic word automata. What we now want is to transpose this argument into the 2D+ case, using purposely the redundancy induced by the periodicity.

Lemma 5.3 (Connected Reconstruction)**.** *Consider a 2D+ periodic SFT $\Omega_{\mathcal{F}}$. There exists a constant $c(\mathcal{F}) \in \mathbb{N}$ such that, for any* connected *cell window $I \subset \mathbb{Z}^d$, if $w \in \mathcal{A}^{I+B_c}$ is locally admissible, then $w|_I$ is globally admissible.*

Proof. As the SFT is periodic, like before, consider N the size of a base hypercube such that any configuration of $\omega \in \Omega_{\mathcal{F}}$ is invariant under any $k \in N\mathbb{Z}^d$, *i.e.* $\sigma_k(\omega) = \omega$.

Let us begin with the case where $I = \{e\}$ is made of a single cell. Assuming a pattern u on the window $e + B_{\lceil \frac{N}{2} \rceil}$ is *globally* admissible, then u actually is the restriction of a configuration $\omega_e \in \Omega_{\mathcal{F}}$ that coincides with u on the window, and in particular $\omega_e|_I = u|_I$. Thus, it is sufficient to consider $c = \varphi\left(\left\lceil \frac{N}{2} \right\rceil\right)$, such that whenever u is locally admissible on $e + B_c$, it is globally admissible on $e + B_{\lceil \frac{N}{2} \rceil}$ so the previous paragraph applies.

More generally, consider any connected window of cells I, and $u \in \mathcal{A}^{I+B_c}$ a locally admissible pattern. For any cell $e \in I$, we can likewise obtain a configuration $\omega^e \in \Omega_{\mathcal{F}}$ such that, on the domain $e + B_{\lceil \frac{N}{2} \rceil}$, the pattern u and the configuration ω^e coincide.

Consider now two neighbouring cells $e, f \in I$. As we left a bit of margin to begin with, the intersection $\left(e + B_{\lceil \frac{N}{2} \rceil}\right) \cap \left(f + B_{\lceil \frac{N}{2} \rceil}\right)$ contains an N-hypercube (that contains both e and f), which contains the same base pattern for both ω^e and ω^f, so that we actually have equality $\omega^e = \omega^f$.

As I is connected, by induction, the pattern $u|_I$ is actually a restriction of ω^e, hence globally admissible. □

For a noisy configuration $(\omega, b) \in \Omega_{\widetilde{\mathcal{F}}}$ to be close to a globally admissible one, we need a high-density connected window I such that all cells in $I + B_c$ are clear. If such a window occurs with high probability, then we will be able to control the distance of a noisy measure to $\mathcal{M}_{\mathcal{F}}$. Notice that this behaviour can be characterised by looking solely at the noise b, by studying a site percolation process on $\mathbb{Z}^d$. This is what we will do in the next subsection.

5.3 Study of the thickened percolation

We consider here the site percolation process on $\mathbb{Z}^d$, with configurations $b \in \Omega_{\{0,1\}}$. In our framework, the *open* cells will be the clear ones, with value 0, and the *closed* ones will be the obscured ones, with value 1.

As we want specific properties on the "thickness" of the infinite component of this percolation (the connected window I such that $I + B_c$ is open), we will induce an auxiliary percolation process. If $\nu \in \mathcal{M}_{\{0,1\}}$ defines a random percolation, then the percolation process on $\mathbb{Z}^d$ induced by the measure $\gamma_n^*(\nu)$ will be called the n-thickened ν-percolation, with the cellular automaton γ_n from Definition 3.8.

In the article *Density and Uniqueness in Percolation* [8, Theorem 2], it is shown that under a condition of *finite energy* on the measure μ, defined below, the percolation process almost surely has *at most* one infinite connected component. This property holds true for any Bernoulli noise in particular.

Definition 5.4 (Finite Energy)**.** *Consider $w \in \mathcal{A}^I$ a finite pattern. For a measurable set B, we define $B_w = \{\omega \in \Omega_\mathcal{A}, \exists \omega' \in B, \omega|_{I^c} = \omega'|_{I^c}, \omega|_I = w\}$ which is also measurable.*

A measure μ has finite energy if, for any finite pattern w and any measurable set B, we have $\mu(B_w) > 0$ whenever $\mu(B) > 0$.

Note that thickened measures cannot have the finite energy property. Indeed, a consequence of finite energy is that any cylinder has a positive measure. However, for an n-thickened percolation, we cannot have three adjacent cells with the pattern 010 in a configuration $\gamma_n(b)$, as the presence of a 1 in an n-hypercube of b implies its presence in the left-translated or right-translated hypercube. The result can nonetheless be effortlessly adapted to the case of thickened measures, and we will sketch its proof here for completeness.

Lemma 5.5. *When ν has the finite energy property, any thickened ν-percolation has* at most *one infinite connected component.*

Proof. The finite energy property still holds for the measures obtained through the ergodic decomposition theorem, hence we can assume ν is ergodic. As γ_n is σ-invariant, by definition of ergodicity, if ν is ergodic, then so is the n-thickened ν-percolation.

As a σ-invariant measurable function, the number $N(b)$ of infinite components in the percolation b is $\gamma_n^*(\nu)$-a.s. constant.

If N was infinite, then for a big-enough hypercube B, the probability of encountering three different infinite components in $\gamma_n(b)$ inside of it would be positive.

In the context of site percolation processes, a trifurcation of a configuration b is an open cell that is part of an infinite component, with exactly three open neighbours such that if the cell was closed then these neighbours would each be in a different infinite component.

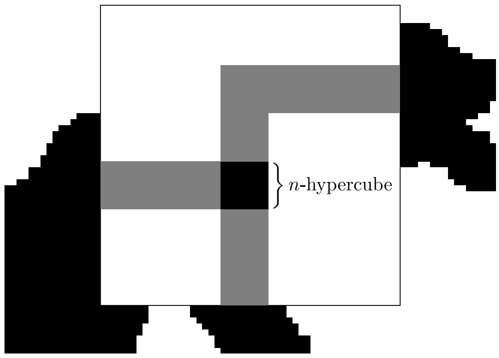

Figure 1: Schematic representation of a trifurcation in the n-thickened case.

Using the finite energy property to change the configuration b inside of B when it encounters three infinite thick components, as illustrated on Figure 1, there is a positive probability of observing a trifurcation inside of B for $\gamma_n(b)$.

The rest of the proof follows as in the original theorem: if the probability that a cell is a trifurcation is positive, then so is the frequency of trifurcations by Birkhoff's ergodic theorem on $\mathbb{Z}^d$, thus it must be of order n^d in a big hypercube. However, a theoretical $O\left(n^{d-1}\right)$ bound can be obtained on the number of trifurcations, thus a contradiction. The number N cannot be infinite.

With a similar but much simpler finite energy argument, N cannot be constant greater or equal to 2, as the probability of having at most $N-1$ components would be positive, by opening an entire hypercube encountering several components. □

Thanks to this result, we can from now on talk about *the* infinite component of the percolation process, whenever it exists. We now need to actually control the frequency of cells belonging to it. Further analyses will be done on a Bernoulli noise, but we still hope for a more general result to come from percolation theory.

Proposition 5.6 (Frequency of the Infinite Component). *Consider $I(b) \subset \mathbb{Z}^d$ the random infinite component of the n-thickened percolation $\gamma_n(b)$, with respect to the original ε-Bernoulli percolation process $\mathbb{P} = \mathcal{B}(\varepsilon)^{\otimes \mathbb{Z}^d}$. When such an infinite component does not exist, we use the convention $I(b) = \emptyset$.*

Then the constant $C_n^d = 48(2n+1)^d$ is such that $\mathbb{P}(0 \notin I) \leq C_n^d \times \varepsilon$.

Proof. Let us describe first what the event $\{0 \notin I\}$ represents. Either the cell 0 is closed in $\gamma_n(b)$ (*i.e.* $\gamma_n(b)_0 = 1$) so that it belongs to no component, or it is open, but its component is finite. The first scenario happens with probability $\left(1 - (1-\varepsilon)^{(2n+1)^d}\right)$.

In the second scenario, this implies that the component of 0 in the percolation process induced by $\gamma_n(x)$ on the lattice $\mathbb{Z}^2 \times \{0\}^{d-2}$ is also finite. Consider the sub-lattice $[(2n+1)\mathbb{Z}]^2 \times \{0\}^{d-2}$, where two cells are adjacent whenever one coordinate differs by $2n+1$. If two neighbouring cells e and f of this sub-lattice are open in $\gamma_n(b)$, then all the cells in $(e + B_n) \cup (f + B_n)$ must be open. Hence, if e and f are open, connected in the sub-lattice, then all the cells that lie in-between in $\mathbb{Z}^2$ are also open, so that e and f are in the same connected component of $\gamma_n(b)$. The interest of this trick is that, as those windows $e + B_n$ and $f + B_n$ are disjoint, the value of the cells e and f in $\gamma_n(b)$ are actually independent. To put it short, in this second scenario, the component of 0 in the sub-lattice $[(2n+1)\mathbb{Z}]^2$ must be finite too.

The percolation process on the sub-lattice is just a plane $\left(1 - (1-\varepsilon)^{(2n+1)^d}\right)$-Bernoulli independent site percolation. In this case, if the component of 0 is finite, then the outer boundary of this component must be a cycle of closed cells, where two neighbouring cells may be diagonally adjacent, so we just need an upper bound on the probability of *this* event.

We can easily start with the upper bound $1 - (1-\varepsilon)^{(2n+1)^d} \leq (2n+1)^d \varepsilon$ on the probability of a cell being closed. Now we need to count the number of cycles of a given length l. Such a cycle must necessarily intersect the half-line $\mathbb{N}^* \times \{0\}$, let's say at coordinates $(k, 0)$, and each of the columns $\{j\} \times \mathbb{Z}$ with $0 \leq j < k$ must cross the cycle at least twice, thus $l \geq 2k$ gives us an upper bound on the coordinate k. Note also that a cycle is in particular a self-avoiding path, so that, for a fixed value of k, we can upper bound the number of cycles by $9 \times 8^{l-1}$. Whenever $\varepsilon < \frac{1}{8(2n+1)^d}$, we have:

$$
\begin{aligned}
\mathbb{P}(0 \notin I) &\leq (2n+1)^d \varepsilon + \sum_{l \geq 4} \tfrac{l}{2} \times 9 \times 8^{l-1} \times \left((2n+1)^d \varepsilon\right)^l \\
&\leq \tfrac{9}{16} \varepsilon \times \sum_{l \geq 1} 8(2n+1)^d \times l \left(8(2n+1)^d \varepsilon\right)^{l-1} \\
&= \tfrac{9}{16} \varepsilon \times \tfrac{\mathrm{d}}{\mathrm{d}\varepsilon} \left[\sum_{l \geq 0} \left(8(2n+1)^d \varepsilon\right)^l\right] \\
&= \tfrac{9}{16} \varepsilon \times \tfrac{\mathrm{d}}{\mathrm{d}\varepsilon} \left[\tfrac{1}{1 - 8(2n+1)^d \varepsilon}\right] = \tfrac{9}{16} \varepsilon \times \tfrac{8(2n+1)^d}{(1 - 8(2n+1)^d \varepsilon)^2}.
\end{aligned}
$$

So far, this upper-bound is of the form $\varepsilon f(\varepsilon)$ for some function f that is positive increasing on the interval $\left[0, \frac{1}{8(2n+1)^d}\right[$ and goes to infinity on the right. If we find ε_0 in this interval such that $\varepsilon_0 f(\varepsilon_0) = 1$, then the upper bound by $f(\varepsilon_0) \varepsilon$ will hold on this interval as f is increasing, and the upper bound will hold for $\varepsilon_0 \leq \varepsilon \leq 1$ as $\mathbb{P}(0 \notin I) \leq 1 \leq f(\varepsilon_0) \varepsilon$ on this interval.

Let us denote $a = \frac{9}{16}$ and $b = 8(2n+1)^d$. Solving $\varepsilon f(\varepsilon) = 1$ equates finding the root of $b^2\varepsilon^2 - b(a+2)\varepsilon + 1$ on the interval $\left[0, \frac{1}{b}\right]$. The roots are $\varepsilon_\pm = \frac{\sqrt{a+2}}{b}\left(\frac{\sqrt{a}+\sqrt{a+2}}{2}\right)$ and only ε_- is in the desired interval. A direct computation then yields $f(\varepsilon_-) = \frac{2b}{1-a\left(\sqrt{1+\frac{2}{a}}-1\right)}$. Replacing a by its value, we obtain $1 - a\left(\sqrt{1+\frac{2}{a}}-1\right) = \frac{25-3\sqrt{41}}{16} > \frac{1}{3}$, thus finally $f(\varepsilon_-) < 6b$. At last, the constant $C_n^d = 48(2n+1)^d$ provides the desired upper bound. $\square$

This proof depends on the specific properties of the independent percolation process, but is quite elementary in exchange. In order to adapt the following periodic stability theorem to a more general class of noises $\mathcal{N}$, one would first need to obtain a similar lower bound on the frequency of cells in the (unique) infinite connected component, so that $\sup_{\nu\in\mathcal{N},\,\nu([1])\leq\varepsilon} \mathbb{P}_\nu(0 \notin I) \underset{\varepsilon\to 0}{\longrightarrow} 0$.

5.4 Stability theorem

Theorem 5.7 (2D+ Periodic Stability)**.** *Consider $\Omega_\mathcal{F}$ a 2D+ periodic SFT. Then $\Omega_\mathcal{F}$ is f-stable for d_B on the class $\mathcal{B}$ of Bernoulli noises, with linear speed $f(\varepsilon) = 2C^d_{c(\mathcal{F})}\varepsilon$.*

Proof. In order to obtain linear stability, we will consider a measure $\lambda \in \widetilde{\mathcal{M}_\mathcal{F}^\mathcal{B}}(\varepsilon)$, and build a measurable mapping $\psi : \Omega_{\widetilde{\mathcal{F}}} \to \Omega_\mathcal{F}$, so that $d_H(\omega, \psi(\omega,b))$ is small for a λ-typical configuration $(\omega,b) \in \Omega_{\widetilde{\mathcal{F}}}$.

Consider N the size of a base hypercube for the periodic SFT $\Omega_\mathcal{F}$, and c the constant obtained in Lemma 5.3. As $\mathcal{A}^{[\![0,N-1]\!]^d}$ is finite, then so is $\Omega_\mathcal{F}$. Thus, it makes sense to consider $\Omega_\mathcal{F}$ as a finite alphabet and to define the full-shift $\Omega_{\Omega_\mathcal{F}}$.

Let us define the morphism $\rho : \Omega_{\widetilde{\mathcal{F}}} \to \Omega_{\Omega_\mathcal{F}}$ such that, whenever the window B_c is clear in $\sigma_e(\omega,b) \in \Omega_{\widetilde{\mathcal{F}}}$, then $\rho(\omega,b)_e = \omega_e$ as in Lemma 5.3, but specifically for the window B_c of $\sigma_e(\omega,b)$ centred on 0. If the window is obscured, then we may default to some configuration $\omega' \in \Omega_\mathcal{F}$. The interest of "forgetting" the role of the coordinate e, of acting as if each cell was the centre of the lattice $0 \in \mathbb{Z}^d$, is that this way ρ is σ-invariant, we have a local characterisation of the morphism $\rho : \widetilde{\mathcal{A}}^{B_c} \to \Omega_\mathcal{F}$.

Without loss of generality, assume the finite set $(\Omega_\mathcal{F}, <)$ is strictly ordered. We may now define the adjusted majority rule cellular automaton $\theta_n : \Omega_\mathcal{F}^{B_n} \to \Omega_\mathcal{F}$ as follows. First, map each configuration of the pattern $(\omega_e)_{e\in B_n}$ onto the configuration $\sigma_{-e}(\omega_e)$, so that we locally undo the offset introduced by ρ by aligning all the configurations on a "common" centre. Only then we may apply a regular majority rule, on the family $(\sigma_{-e}(\omega_e))_{e\in B_n}$, by picking the maximal configuration for the arbitrarily introduced order in case of a tie.

Consider now the morphisms $\psi_n = \theta_n \circ \rho$ obtained by applying an adjusted majority rule over ρ. Using once again the order on $\Omega_\mathcal{F}$, we may define the pointwise limit $\psi = \overline{\lim}_{n\to\infty} \psi_n$, which is still σ-invariant and measurable. Note that the value of $\psi(\omega,b)$ in some cell may now depend on arbitrarily far values, so ψ is *not* a morphism.

Consider the configuration $(\omega,b) \in \Omega_{\widetilde{\mathcal{F}}}$, and let $I \subset \mathbb{Z}^d$ be the infinite component of the c-thickened percolation in b. As $\omega|_{I+B_c}$ is locally admissible, $\omega|_I$ is a globally admissible pattern, the restriction of some configuration $\omega_0 \in \Omega_\mathcal{F}$. For any cell $e \in I$, we have $\rho(\omega,b)_e = \sigma_e(\omega_0)$.

Assume now that $\varepsilon < \frac{1}{2C_c^d}$, so that in the Bernoulli percolation process, I has a density greater than $\frac{1}{2}$ according to Proposition 5.6. This means that, λ-a.s., after some rank n_0, strictly more than half of the cells $f \in e + B_n$ of (ω,b) are inside of I, thus are mapped by ρ onto translations $\sigma_f(\omega_0)$. Thence, after the very same rank n_0, $\psi_n(\omega,b)_e = \sigma_e(\omega_0)$. Consequently, by taking the limit $n \to \infty$, λ-a.s., $\psi(\omega,b)_e = \sigma_e(\omega_0)$ for any cell $e \in \mathbb{Z}^d$.

To sum it up, $(\omega, b) \mapsto \psi(\omega, b)_0 = \omega_0$ is a measurable mapping $\Omega_{\widetilde{\mathcal{F}}} \to \Omega_{\mathcal{F}}$, such that $d_H(\omega, \omega_0) \le C_c^d \varepsilon$ whenever $\varepsilon \le \frac{1}{2C_c^d}$. More generally, the bound $d_H(\omega, \psi(\omega, b)_0) \le 2C_c^d \varepsilon$ holds λ-a.s. for any choice of ε, which finally gives us the linear bound we wanted:

$$d_B\left(\pi_1^*(\lambda), \mathcal{M}_{\mathcal{F}}\right) \le d_B\left(\pi_1^*(\lambda), [\psi(\cdot)_0]^*(\lambda)\right) \le 2C_c^d \varepsilon. \qquad \square$$

This concludes our analysis of periodic SFTs in the 2D+ case. The explicit constant C_n^d could doubtlessly be improved, but such matters would require much more work without improving on the *linear* aspect of the bound.

A further track of reflection, as already mentioned earlier, may be to extend this theorem to a more general class of noises, using stronger percolation results, while leaving much of the actual proof of the theorem unchanged.

What we got interested in instead is the study of stability for aperiodic SFTs. We chose the well-known Robinson tiling, as it is already *almost* periodic, in order to adapt the previous scheme of proof as much as possible. This will be the topic of the last section of the paper.

6 The case of 2D (c_1, c_2)-robust tilesets

Before diving into the Robinson tiling, let us now digress a bit to contextualise our study. The aim of this section is to provide an informal analysis of an already existing Besicovitch stability result in our current framework. More precisely, we are interested in the notion of stability described by Durand, Romaschenko and Shen [9], which was then used to prove periodic stability in the 2D case in a further article by Ballier, Durand and Jeandel [5].

Here, we will provide a rough and qualitative estimate of the convergence speed obtained with their method. Yet, for this article to be as self-contained as possible, we will still introduce the essential definitions to understand the cited results.

The estimates provided here bear no influence on the following aperiodic stability result, so this section can be easily skipped in a first reading of the current article.

6.1 Robust tilesets and sparse sets

To obtain stability, instead of using a notion of percolation – which is best seen as a clear connected component that spans the whole obscured space – they introduce the notion of islands of errors – which is best seen as small clumps of obscured cells isolated in the whole clear space.

Definition 6.1 ((α, β)-Island of Errors). *Consider a noise configuration $b \in \{0,1\}^{\mathbb{Z}^2}$ which we identify with $E \subset \mathbb{Z}^2$ the set of obscured cells.*

A set $F \subset E$ is an (α, β)-island of E if F can be included in some α-square and its β-neighbourhood does not meet any other obscured cell of E, i.e. *$(F + B_\beta) \cap (E \backslash F) = \emptyset$.*

In this framework, the "right" way to obtain stability is to remove the islands of obscured cells, by changing the values of the tiles underneath on a small neighbourhood. This is well-encapsulated by the following notion of robustness.

Definition 6.2 ((c_1, c_2)-Robustness). *Let us denote by $R_{i,j} := B_{\frac{j-1}{2}} \backslash B_{\frac{i-1}{2}}$ (with $i < j$ two odd integers) the ring-shaped window obtained by removing the i-square at the centre of a j-square.*

Let $0 < c_1 \le c_2$ be two positive integers. A tileset $\mathcal{F}$ is (c_1, c_2)-robust if, for any $n \in \mathbb{N}$ and any locally admissible pattern $u \in \mathcal{A}^{R_{n,c_2 n}}$, there exists a locally admissible pattern $v \in \mathcal{A}^{S_{c_2 n}}$ such that u and v coincide on $R_{c_1 n, c_2 n}$ – which is a strict subset of the ring $R_{n, c_2 n}$ as long as $c_1 \ge 2$.

An explicit example of robust tileset is any one inducing a periodic SFT [5], roughly for the same reason we could obtain a globally admissible configuration by peeling a constant width of the border of any pattern in the previous section. However, this notion is much more general, and strongly aperiodic robust SFTs are proven to exist [9].

Note that, while the constants may change in the process, this notion of robustness is invariant under conjugacy (using the local viewpoint as in Definition 3.7), so that we cannot prove stability of a non-robust SFT by looking for a suitable robust conjugated.

Whenever $\beta \geq c_2\alpha$, we can "repair" an (α, β)-island of errors by changing the tiles in a $c_1\alpha$-square. Hence, we need some guarantees that E is entirely made out of islands we can correct.

Definition 6.3 ((α, β)-Sparse Set)**.** *A set $E = E_0$ is said to be sparse, given a sequence $(\alpha_k, \beta_k)_{k\in\mathbb{N}^*}$, if we can step by step remove all the (α_k, β_k)-islands from E_{k-1} to obtain a set E_k, in such a way that the decreasing limit set $E_\infty = \bigcap E_k$ is empty.*

Up to now, the definitions introduced were formal. For the rest of this section, we will provide a qualitative and quite handwavy analysis of the convergence speed we can obtain in this framework.

6.2 Qualitative analysis of the convergence speed

By the Borel-Cantelli theorem, any ε-Bernoulli noise will certainly contain islands for any pair (α, β), which may *a priori* be hard to correct. However, it is proven [9, Lemma 3] that, assuming $8\sum_{k=1}^{n-1} \beta_k < \alpha_n \leq \beta_n$ for any $n \in \mathbb{N}^*$ and $\sum_n \frac{\ln(\beta_n)}{2^n} < \infty$, then for ε small enough the random set E is almost surely (α, β)-sparse. Unfortunately, general bounds on ε would be quite hard to obtain, but we will provide rough estimates for our choice of (α, β).

It is also proven [9, Lemma 4] that in any (α, β)-sparse set E, the density of obscured cells is *at most* $\sum_n (\alpha_n/\beta_n)^2$ – the main argument is that each (α_n, β_n)-island contains *at most* α_n^2 obscured cells, among *at least* β_n^2 cells in a neighbourhood of the island disjoint of the other islands and their neighbourhoods. To properly quantify the convergence speed, we would need to take into account the density not of the islands of errors but of the $c_1\alpha$-square around them, but this approximation will suffice for the present qualitative analysis.

Consider $\alpha_n = 8^n(n-1)!n!$ and $\beta_n = 8^n(n!)^2$. It is clear that any k-*shift* of this sequence (starting at some rank $k+1$ instead of 1) will satisfy the previously stated hypotheses. For a given sparse set E for the k-shifted sequence, the density of errors is $\sum_{n=k+1}^{\infty} \frac{1}{n^2} \leq \int_k^\infty \frac{1}{t^2}\mathrm{d}t = \frac{1}{k}$.

To obtain the convergence speed, we now need to estimate the maximal value of k such that E is sparse for the k-shifted sequence for a given ε. Looking at the proof of the result [9, Lemma 3], it appears that the key property to obtain sparsity is that $\sum_n \frac{\ln(\beta_n)}{2^n} < \ln\left(\frac{1}{\varepsilon}\right)$. As $\ln(\beta_n) = 8\ln(n) + 2\ln(n!) \leq n^2$ after some rank, for the k-shifted sequence, we can bound the left term by $k^2 + 4k + 6$. Asymptotically, the best choice for k is thus $k(\varepsilon) \approx \sqrt{\ln(1/\varepsilon)}$, so that $f(\varepsilon) \approx \frac{1}{\sqrt{\ln(1/\varepsilon)}}$.

Note that, as $\frac{\alpha_n}{\beta_n} \to 0$, for *any* pair (c_1, c_2) (and any accordingly (c_1, c_2)-robust tileset), we will satisfy the $\beta_n \geq c_2\alpha_n$ condition after a rank n. It follows that this bound on the speed of convergence holds for any robust tileset, but on an interval $[0, \varepsilon_{\max}(c_2)]$ that depends on the constant c_2, with $k(\varepsilon_{\max})$ such that $\beta_k \geq c_2\alpha_k$. As a rule of thumb, the bigger c_2 gets, the smaller $\varepsilon_{\max}$ gets. In any case, this doesn't affect the asymptotic order of the bound outside of a multiplicative factor.

Considering all the small approximations we did on the way, what matters here is not the value of the bound but its order of magnitude. Indeed, $\frac{1}{\sqrt{\ln(1/\varepsilon)}}$ is much *much* slower than any polynomial speed, which legitimises our efforts to obtain a linear convergence speed in the periodic case.

The notion of islands and sparsity can be used as a black box to obtain percolation results [9, Section 9.3], hence as a tool it is in some ways more powerful than the percolation theory we used in the previous section. However, as we have seen here, this versatility comes at the cost of the precision and simplicity of the bounds we can obtain.

7 The Robinson tiling: an almost periodic stable example

The first aperiodic tiling defined by local rules was proposed by R. Berger [6], who used 20426 *Wang tiles* (with forbidden patterns only between neighbouring tiles that share an edge) to encode a hierarchical structure, and thus aperiodicity. The construction was strongly simplified by R. Robinson [22] who proposed a Wang tileset with 56 tiles, which once again forces a hierarchical structure. In fact, if we allow diagonal interactions between tiles, the number of tiles can be brought down to 6 tiles and their rotations and symmetries [22]. The simplicity of the tileset and its hierarchical structure, with arbitrary large squares which permits the embedding of space-time diagrams of Turing machines into it, explains why the Robinson tiling is certainly the most studied aperiodic tiling.

The Robinson tiling is *not* (c_1, c_2)-robust in the previous sense: it can have an infinite central cross in $\mathbb{Z}^2$ with a black arm in each direction, with only one obscured cell at the centre as in Figure 4, that no amount of local correction may turn into a globally admissible pattern. However, the hierarchical structure implies that for a given scale, the corresponding squares form a periodic structure, except for a small fraction of tiles that corresponds to the squares higher in the hierarchy. A similar technique that in Section 5 yields some stability at this scale, and allows us to deduce the stability of the Robinson tiling with a polynomial speed (Therorem 7.10).

7.1 The classic Robinson tiling

Our first attempt at 2D+ aperiodic stability used the folkloric Robinson tiles shown in Figure 2, and their rotations and symmetries – so that the total number of tiles is actually 32.

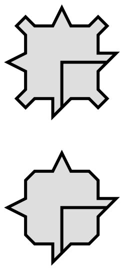

Figure 2: The six base Robinson tiles.

With this tileset, the forbidden patterns are self-evident: two laterally adjacent tiles must have matching borders, including the black lines drawn on them, and any square made of four tiles must use exactly one rotation of the top-left tile in Figure 2 with bumpy corners, so that the small diamond in the centre of the square is filled-in. Any non-matching pair or square of adjacent tiles is then a forbidden pattern in $\mathcal{F}$.

Note that forbidden patterns can occur with left-right and top-bottom neighbours, but *also* on diagonally adjacent tiles, unlike the tileset originally introduced by Robinson in the context of Wang tiles. The two associated SFTs can nonetheless be shown to be conjugate.

Definition 7.1 (Macro-Tiles)**.** *We define macro-tiles inductively. First, the* 1*-macro-tile is just the top-left tile of Figure 2, with bumpy corners.*

Then, the $(N+1)$*-macro-tile is obtained by sticking four* N*-macro-tiles in order to draw a square around a central cross, as shown in Figure 3.*

Definition 7.2 (Orientation Symbols)**.** *Let us use the symbol to denote the* default *orientation of an* N*-macro-tile, with the black arms of the central cross pointing on the bottom and on the right, as seen in Figure 3. Likewise we denote , and for the other orientations.*

By induction, we have that an N-macro-tile is a $2^N - 1$ tiles long square. One can prove that two N-macro-tiles cannot overlap. These fundamental properties can be found in Robinson's seminal article *Undecidability and nonperiodicity for tilings of the plane* [22], and are nicely condensed into seminar notes [23].

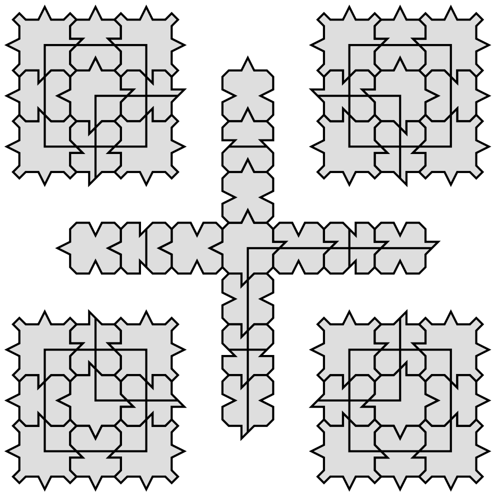

Figure 3: Four 2-macro-tiles around a central cross form a 3-macro-tile.

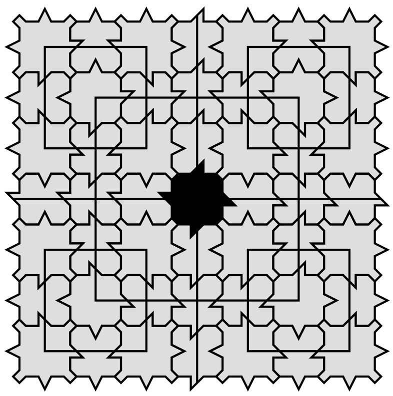

Figure 4: Four (arbitrarily long) lines around one obscure cell in a locally admissible square.

A Robinson tiling is almost periodic, in the sense that any given window of a tiling occurs periodically in the tiling, but not always with the same periodicity. Most notably, if you keep only the N-macro-tiles and forget about the thin grid sticking all of them together, you obtain a 2^{N+1}-periodic pattern, which has density $\left(1 - \frac{1}{2^N}\right)^2$.

The issue with this tileset is that the alignment of macro-tiles on such a grid is a consequence of the global structure of a Robinson tiling, and is not enforced by the local rules. This is illustrated by the two misaligned macro-tiles in Figure 5, and such a phenomenon can arise at *any* scale. This implies that we would not be able to ensure stability using a percolation argument as we did for the periodic case.

By pushing this phenomenon to the limit, we can obtain "pathological" Robinson tilings that exhibit a *cut*, an infinite horizontal or vertical line, with a misalignment on both sides.

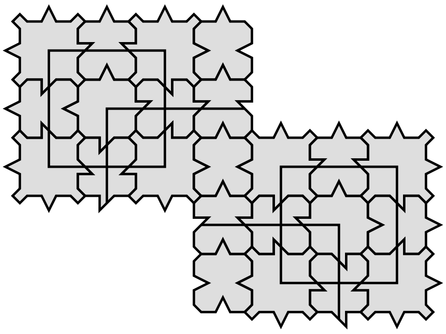

Figure 5: Two loosely aligned 2-macro-tiles, with one tile in common and a tiled gap.

7.2 An enhanced Robinson tiling

To work around the aforementioned issue, let us now introduce a variant tileset by *adding* information over the already existing tiles.

To force this alignment in a local way, we want for each macro-tile to send a "signal" from its central cross, which will force the correct alignment between neighbouring macro-tiles at any scale. The idea originates in Sylvère Gangloff's phd thesis [10] on another variant of the Robinson tiles, and we transpose it on our current tileset.

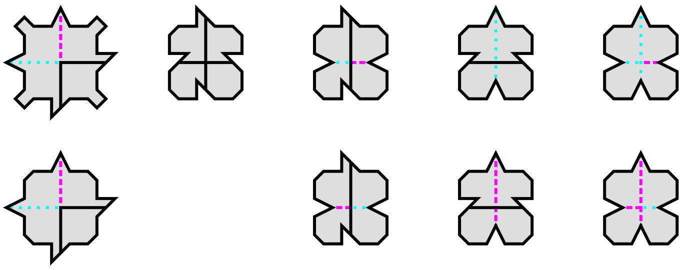

Figure 6: The nine enhanced Robinson tiles.

More precisely, consider the tiles on Figure 6, roughly grouped according to which of the previous tiles they come from. Now, all of the tiles have a cross-like pattern drawn upon them. In order to preserve their specific orientation, the two leftmost tiles must never undergo a symmetry, so that a tile always has a blue dotted line pointing left and a red dashed one pointing up. Up to symmetry of the other tiles and rotation, this brings the total number of tiles to 56.

We define the set of forbidden patterns as before, now in accordance with the crosses drawn upon the tiles. For the rest of the section, we will use $\mathcal{F}$ to denote this specific set of forbidden patterns. Using the same process as before, starting from the base 1-macro-tiles, there is a unique way to build macro-tiles inductively. For a given macro-tile of the initial Robinson tileset, we can without ambiguity deduce where the red dashed lines and blue dotted lines of the enhanced macro-tile are.

As there is a direct local projection (thus a morphism) of this enhanced tileset on the previous Robinson tiles, any configuration is still aperiodic. However, this morphism is not a bijection. On one hand, this morphism is not surjective, as we cannot reach tilings with a misaligned cut. On the other hand, this morphism is not injective, as we may have an aligned cut with an infinite red dashed *or* blue dotted line that gets projected onto the same configuration. The main interest of this added structure, as we will prove, is that it indeed *locally* enforces the alignment we lacked before.

Remark 7.3 (Limits of the Notion of Stability)**.** One of the initial motivations of studying the Besicovitch-stability of (aperiodic) tilings was the analogy with the structural stability of quasicrystals. In particular, we ultimately would like to obtain a stable aperiodic structure.

However, for the Robinson tilings here, this aperiodic structure is *not* preserved, and we will now give a rough idea of why. Fix here a globally admissible grid of N-macro-tiles, with a one-tile thick empty grid around them as in Figure 9, and then add the ε-Bernoulli noise. We can see this empty grid as a bond percolation process on $\mathbb{Z}^2$, with the sites being the crossings of the grid, and the bonds being the $\left(2^N - 1\right)$-long lines and columns between neighbouring sites. A bond is open *iff* all the cells inside are clear, with probability $(1-\varepsilon)^{2^N-1} \underset{N\to\infty}{\longrightarrow} 0$ (for a fixed ε). In particular, for a big-enough scale N, this percolation process has no infinite connected component. Each connected component can then be independently filled in a locally admissible way for the Robinson tileset. In this random noisy tiling, there is no well-identified aperiodic structure.

Thus, the best we can hope for (and what we will indeed obtain in Theorem 7.10) is structural stability up to a finite scale $N(\varepsilon)$, in such a way that $N(\varepsilon) \to \infty$ as $\varepsilon \to 0$.

7.3 Local alignment properties

As we already said, we want to study the almost periodicity obtained by looking only at N-macro-tiles.

Definition 7.4 (Well-Aligned and Well-Oriented Pairs)**.** *A pair of N-macro-tiles (both having an actual position in $\mathbb{Z}^2$) is said to be well-aligned if both of their centres have one coordinate in common, and the other differs by exactly 2^N so that there is a gap of precisely one line/column between them.*

More generally, we say the two N-macro-tiles are loosely *aligned (with $0 < k \leq 2^N - 1$ tiles in common) when one of the coordinates of their centres differs by exactly 2^N and the other by $2^N - k - 1$,* i.e. *we start with a well-aligned pair (with $2^N - 1$ tiles in common) and we translate one of them of k units in the direction of the gap in-between, as in Figure 5.*

A pair of well-aligned macro-tiles is said to be well-oriented if their central crosses form a pattern or (or a rotation of these), which can actually be filled by a central cross in the process of making a larger macro-tile.

Definition 7.5 (Edge Words of Macro-Tiles)**.** *We define the words l_N and t_N, obtained by reading the colours on the left and top edges of the N-macro-tile in a clockwise motion, with blue dotted lines encoded as a 0 and red dashed lines as a 1.*

For a binary word, we define $\overline{b} = 1 - b$ the binary complement of a letter, extended to binary words by a direct induction. We also define the mirror function on words such that $\mathrm{mirror}(uv) = \mathrm{mirror}(v)\mathrm{mirror}(u)$, that returns the same word but backwards. Both of these mappings are involutions and they commute with each other.

Lemma 7.6. *Let l_N and t_N be the previously defined edge words of an enhanced N-macro-tile. For any $N \in \mathbb{N}^*$, we have $t_N = \overline{\text{mirror}(l_N)}$.*

What is more, $|l_N| = |t_N| = 2^N - 1$ is odd, and these words actually differ of exactly one letter in their middle.

Proof. For $N = 1$, we simply have $l_1 = 0$ and $t_1 = 1$.

By induction, as seen in Figure 3, when building an $(N+1)$-macro-tile, on the left half from bottom to top, we first have a N-macro-tile that reads as t_N, then we read the 0 given by the blue dotted arm of the central cross, and finally l_N on the , so that $l_{N+1} = t_N 0 l_N$. Likewise, $t_{N+1} = t_N 1 l_N$. Hence:

$$\overline{\text{mirror}(l_{N+1})} = \overline{\text{mirror}(t_N 0 l_N)} = \overline{\text{mirror}(l_N)} 1 \overline{\text{mirror}(t_N)} = t_N 1 l_N = t_{N+1},$$

which concludes the proof by induction. □

In Figure 7, for example, we observe that $l_3 = 1100100$ and $t_3 = 1101100$.

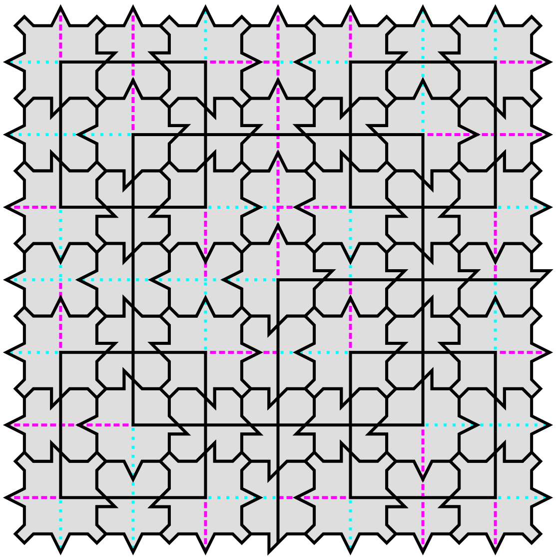

Figure 7: The 3-macro-tile obtained using the enhanced tileset.

Proposition 7.7 (Local Alignment of Macro-Tiles)**.** *Consider the enhanced Robinson tiling. For any scale $N \in \mathbb{N}^*$, a pair of loosely aligned N-macro-tiles with a tileable gap in-between (that can be filled in a* locally admissible *way as in Figure 5)* must *be well-aligned and well-oriented.*

Proof. Assume first that two well-aligned macro-tiles are not well-oriented. If only one of these tiles has a black arm that falls into the gap (*e.g.* a pattern), then this gap cannot be tiled. Up to a rotation, the remaining cases are the and patterns. In these cases, the right arm of the left cross and the left arm of the right cross have the same colour, thus no tile can fill the gap in-between. In other words, by contraposition, a well-aligned pair with a tileable gap must be well-oriented.

At the scale 1, if two tiles are loosely aligned they are actually well-aligned, thus if the gap is tileable they are well-oriented. This allows us to initialise the induction.

Assume the result holds up to scale $N \in \mathbb{N}^*$ and consider a pair of $(N+1)$-macro-tiles, once again loosely aligned with a tileable gap. The macro-tiles cannot have exactly one tile in common, which would imply that we have two 1-macro-tiles well-aligned with a tileable gap but ill-oriented, hence $k \geq 2$.

What is more, k cannot be even. Assuming k is even, this pair of $(N+1)$-macro-tiles contains a pair of 2-macro-tiles with a tileable gap and 2 tiles in common. It is clear that this cannot happen, by an exhaustion of cases. For example, looking at a well-aligned pair, if we move the right tile of one unit upwards, then the right arm of the left tile and the bottom-left corner of the right tile face a tileable gap with a red dashed line, which is impossible.

This concludes the case $N+1=2$, as $k \geq 3$ must then be equal to 3, maximal, so that the 2-macro-tiles are well-aligned. Likewise, when $N+1>2$, the N-macro-tiles must be well-aligned with k odd, so either the $(N+1)$-tiles are well-aligned, or only half of the N-macro-tiles actually face the gap and are well-aligned. In the second scenario, we are once again in a tileable ill-oriented case, impossible. Finally, the $(N+1)$-macro-tiles must be well-aligned thus well-oriented, which concludes the induction. ☐

Proposition 7.8 (Almost Reconstruction)**.** *For any scale $N \geq 2$, let $C_N = 2^N - 1$. This constant is such that for any $n \in \mathbb{N}$ and any clear locally admissible pattern ω on B_{n+C_N}, its restriction $\omega|_{B_n}$ is almost globally admissible, in the sense that up to a one-tile thick grid, $\omega|_{B_n}$ is the restriction of an enhanced Robinson tiling, with well-aligned and well-oriented N-macro-tiles as in Figure 9.*

Proof. We will demonstrate a slightly stronger result here, *i.e.* that by removing *at most* C_N layers of tiles on the top, bottom, left and right sides of any locally admissible square, and not necessarily the same amount of layers on each side, we obtain an actual family of well-aligned and well-oriented N-macro-tiles with respect to their neighbours. Thence, by actually peeling C_N layers on each side, we obtain the stated result.

To do so, we need to proceed inductively, as before. We cannot initialise the result at $N=1$ without proving the case $N=2$ at the same time, which is why we only give C_N for $N \geq 2$ in this proposition (though C_2 will work for the case $N=1$).

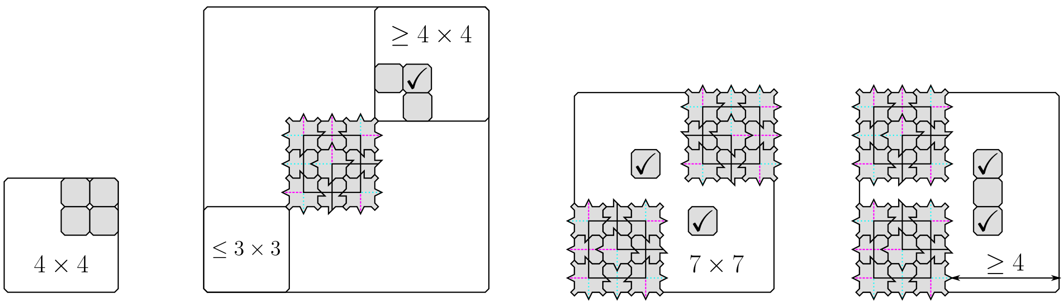

Figure 8: From left to right, key steps (a), (b), (c) and (d) of the case $N=2$.

Hence, let us now prove the case $N=2$, by peeling 3 layers of any admissible k-square B to obtain the announced restriction (remember that B_n is a *ball* of *radius* n, hence a $(2n+1)$-square). First, it is known that with the initial Robinson tileset, if a 3-square is tiled with a in the bottom-left corner, then it is tiled by a 2-macro-tile. This property still holds for the enhanced tileset, and can be easily checked by enumerating all the cases.

We will inductively build a rectangle of well-aligned 2-macro-tiles in the k-square B, assuming that $k \geq 10$ for now. If we look at the 4-square in the bottom-left corner, one of the four cells highlighted in the step (a) of Figure 8 must contain a 1-macro-tile, with bumpy corners. Then this bumpy corner is actually part of a 2-macro-tile in B. So far, we have a 1×1 rectangle of 2-macro-tiles.

As illustrated in step (b), considering where our first bumpy corner was, there is at *most* a 3×3 rectangle in the bottom-left corner (diagonally adjacent to the 2-macro-tile), and $k \geq 10$, so the top-right corner is at *least* a 4×4 rectangle of tiles. One of the three highlighted tiles in step (b) must be a bumpy corner too. If there was a corner in one of the unchecked tiles, it would be part of a 2-macro-tile, that should either intersect the one drawn on Figure 8 – which is impossible even for regular Robinson tileset – or be loosely aligned with it – which is impossible according to Proposition 7.7. Hence the checked cell must contain a tile with bumpy corners, and more precisely a for the same reasons. This tile can then be completed into a 2-macro-tile, which brings us to step (c). There, the two checked cells must contain a 1-macro-tile too, and each can be completed into its own 2-macro-tile, so that we obtain at last a 2×2 rectangle of 2-macro-tiles.

Just like the two diagonally adjacent 2-macro-tiles present in step (c) imply a square of 2-macro-tiles, the presence of two laterally adjacent 2-macro-tiles in step (d) implies a square of 2-macro-tiles. Thus, now that we have a rectangle with *at least* 2 macro-tiles on each side, we can repeat step (d) in each direction as long as 4 tiles or more remain. Hence, as long as $k \geq 10$, $C_2 = 3$ works well.

More generally, we can trivially peel a 9-square into one single tile of grid if we remove 4 layers on each side, so that $C_2 = 4$ works in this case. However, a more careful study of the cases $k \in \{7, 8, 9\}$ allows us to conclude that $C_2 = 3$ works for these cases and is optimal (to do so consider a 9-square *centred* on a 2-macro-tile, so that all the adjacent ones will be missing a layer). When $k \leq 6$, $C_2 = 3$ trivially works too, which concludes our study of $N = 2$.

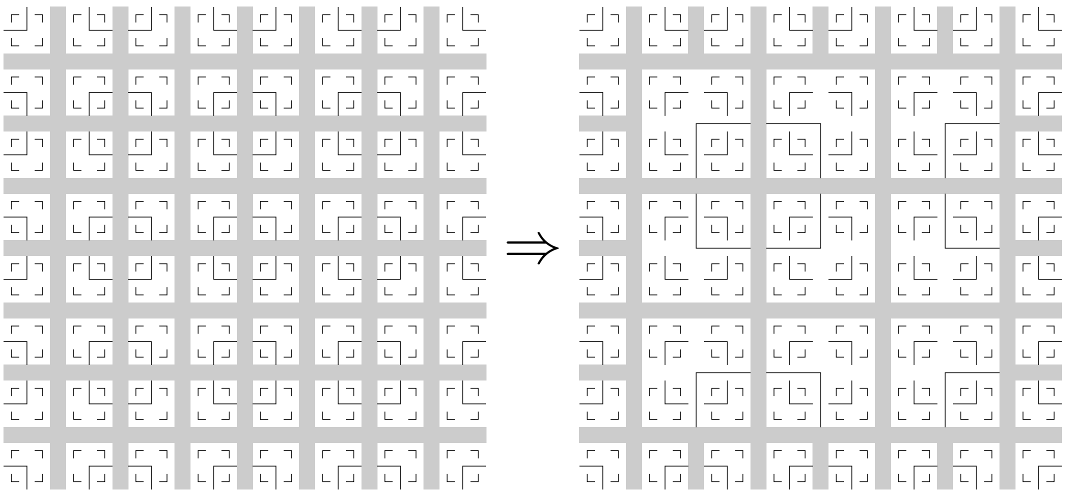

Figure 9: By filling the grid around N-macro-tiles, we obtain $(N+1)$-macro-tiles, up to one outer layer of N-macro-tiles.

Assume now that the result holds at rank N with the constant C_N and let us prove it at rank $N+1$. We can start by peeling away at most C_N tiles, using our induction hypothesis, to obtain a grid of well-aligned and well-oriented N-macro-tiles. A square of well-aligned N-macro-tiles can either form one $(N+1)$-macro-tile, represent the lateral interface between two $(N+1)$-macro-tiles or represent the central corner between four $(N+1)$-macro-tiles. Thus, by peeling at most one layer of N-macro-tiles on each border – an N-macro-tile not part of an $(N+1)$-macro-tile and the following grid, so 2^N tiles in total – we remove the incomplete interfaces and corners to obtain a grid of well-aligned $(N+1)$-macro-tiles (hence well-oriented by the previous proposition) as seen in Figure 9. In conclusion, the result holds at rank $N+1$ with the constant $C_{N+1} = C_N + 2^N$, hence $C_N = 2^N - 1$ by a direct induction. □

7.4 Almost stability at a fixed scale and stability

Proposition 7.9 (Almost Stability). *Let $\Omega_{\mathcal{F}}$ be the enhanced Robinson tiling. For any choice of $\varepsilon > 0$, any scale $N \geq 2$, and any measure $\mu \in \mathcal{M}_{\mathcal{F}}^{\mathcal{B}}(\varepsilon)$, we have a coupling s.t.:*

$$d_B(\mu, \mathcal{M}_{\mathcal{F}}) \leq 96\left(2^{N+2}+1\right)^2 \varepsilon + \frac{1}{2^{N-1}}.$$

Proof. For a given scale N, we want to apply the percolation argument as if we were looking at a $\left(2 \times 2^N\right)$-periodic SFT. This added factor 2 comes from the fact that, for any globally admissible configuration, the $(N+1)$-macro-tiles are well-aligned on a grid, and indistinguishable if we ignore their central cross, hence the N-macro-tiles form a unique 2^{N+1}-periodic pattern up to translation.

By looking at a *globally admissible* 2^N-square, we can always identify one, two or four partial N-macro-tiles arranged in a square pattern around a central cross. Thus, we can actually identify to which translation of the 2^{N+1}-periodic pattern this window corresponds. Note that unlike in the general k-periodic case, where we needed to look at k-squares to identify the translation, we only need to look at a window of size $\frac{k}{2}$ here because the Robinson tiling has a lot of intrinsic redundancy.

Just like in the periodic case, we can then look at the c-thickened percolation, with $c = \left\lceil \frac{2^{N+1}+1}{2} \right\rceil + C_N = 2^N + 1 + 2^N - 1 = 2^{N+1}$, as explained in Lemma 5.3 but using the almost reconstruction property from Proposition 7.8. As stated in Proposition 5.6, the infinite component of the c-thickened percolation has density at least $1 - 48\left(2c+1\right)^2 \varepsilon$.

Let us add a blank symbol $\square \notin \mathcal{A}$ to the original alphabet. Then, following the proof of Theorem 5.7, we can measurably map a noisy configuration (ω, b) onto a globally admissible configuration $\psi(\omega, b) \in \Omega_{\mathcal{F}}$ but on the extended alphabet $\mathcal{A} \sqcup \{\square\}$, such that almost surely:

$$d_H\left(\omega, \psi(\omega, b)\right) \leq 96\left(2^{N+2}+1\right)^2 \varepsilon + \frac{2^{N+1}-1}{2^{2N}}.$$

The second term comes from the density of the symbols $\square$ in $\psi(\omega, b)$, of the one-tile thick grid itself, which is equal to $1 - \left(\frac{2^N-1}{2^N}\right)^2$.

In order to conclude, we need to explain how to measurably project $\psi(\omega, b)$ back onto the original alphabet $\mathcal{A}$, how to fill-in the grid, so that we obtain an actual globally admissible enhanced Robinson tiling. To do so, we can simply consider some measure $\widetilde{\mu} \in \mathcal{M}_{\mathcal{F}}$, take a configuration $y \in \Omega_{\mathcal{F}}$ at random independently of the rest following $\widetilde{\mu}$, and then replace $\psi(\omega, b)$ by $\psi'(\omega, b, y)$ which is the unique translation of y by a vector $k \in \left[\!\left[0, 2^{N+1}-1\right]\!\right]^2$ such that the N-macro-tiles of $\psi'(\omega, b, y)$ and $\psi(\omega, b)$ are aligned. This whole process is measurable, σ-invariant, and only changes the values of $\psi(\omega, b)$ on the $\square$ tiles which were already taken into account in the upper bound, so that the same bound holds for $d_H\left(\omega, \psi'(\omega, b, y)\right)$.

Thence, we have a coupling such that $d_B(\mu, \mathcal{M}_{\mathcal{F}}) \leq 96\left(2^{N+2}+1\right)^2 \varepsilon + \frac{1}{2^{N-1}}$, which proves the bound. □

By taking N arbitrarily large, and then $\varepsilon \to 0$, we directly deduce the *stability* of our enhanced Robinson tiling for the Besicovitch distance. By optimising over N for a given value of ε, we will now conclude this analysis with an explicit non-linear upper bound on this speed.

Theorem 7.10 (Enhanced Robinson Stability)**.** *Let $\Omega_{\mathcal{F}}$ be the enhanced Robinson tiling. Then $\Omega_{\mathcal{F}}$ is f-stable for d_B on the class of Bernoulli noises $\mathcal{B}$, with $f(\varepsilon) = 48\sqrt[3]{6\varepsilon}$ for small-enough values of ε. In particular, $\Omega_{\mathcal{F}}$ is polynomially stable.*

Proof. To simplify things, we start by bounding $\left(2^{N+2}+1\right)^2 \leq 2^{2N+5}$, so that we are now trying to minimise $2\left(4^N \times 2^9 3\varepsilon + \frac{1}{2^N}\right)$. If we denote $c(\varepsilon) = \sqrt[3]{2^9 3\varepsilon} = 8\sqrt[3]{3\varepsilon}$, then the upper-bound can be rewritten as $2c\left(\left(2^N c\right)^2 + \frac{1}{2^N c}\right)$.

If we treat $x = 2^N c$ as a real-valued parameter, then $x^2 + \frac{1}{x}$ is minimal at $x = \sqrt[3]{\frac{1}{2}}$, equal to $\frac{3}{\sqrt[3]{4}}$. This gives us a $24\sqrt[3]{6\varepsilon}$ bound. As N must be integer, we cannot have $N = \log_2\left(\frac{x}{c}\right)$, but by replacing it with the nearest integer (at distance at most $\frac{1}{2}$), we obtain the previous bound up to a factor $4^{\frac{1}{2}} = 2$, thus the announced bound.

In order for this rounding argument to give a valid scale $N \geq 2$, we need $N(\varepsilon) \geq \frac{3}{2}$ to begin with, hence $\varepsilon \leq \frac{1}{49152\sqrt{2}}$. □

References

[1] David Aristoff and Charles Radin, *First order phase transition in a model of quasicrystals*, Journal of Physics A: Mathematical and Theoretical **44** (2011), no. 25, 11. MR2800867

[2] Eugene Asarin and Pieter Collins, *Noisy Turing machines*, ICALP 2005: Automata, Languages and Programming, 32nd International Colloquium, Lecture Notes in Computer Science, vol. 3580, Springer, 2005, pp. 1031–1042. MR2184698

[3] Nathalie Aubrun and Mathieu Sablik, *Simulation of effective subshifts by two-dimensional subshifts of finite type*, Acta Applicandae Mathematicae. An International Survey Journal on Applying Mathematics and Mathematical Applications **126** (2013), 35–63. MR3077943

[4] Alexis Ballier, Bruno Durand, and Emmanuel Jeandel, *Structural aspects of tilings*, STACS 2008: 25th International Symposium on Theoretical Aspects of Computer Science, vol. 1, 2008, pp. 61–72. MR2873724

[5] Alexis Ballier, Bruno Durand, and Emmanuel Jeandel, *Tilings robust to errors*, LATIN 2010: Theoretical Informatics, Springer Berlin Heidelberg, 2010, pp. 480–491. MR2673286

[6] Robert Berger, *The undecidability of the domino problem*, Memoirs of the American Mathematical Society, no. 66, AMS, 1966. MR0216954

[7] François Blanchard, Enrico Formenti, and Petr Kůrka, *Cellular automata in the Cantor, Besicovitch, and Weyl topological spaces*, Complex Systems **11** (1997), no. 2, 107–123. MR1673733

[8] Richard Max Burton and Michael Keane, *Density and uniqueness in percolation*, Communications in Mathematical Physics **121** (1989), 501–505. MR0990777

[9] Bruno Durand, Andrei Romashchenko, and Alexander Shen, *Fixed-point tile sets and their applications*, Journal of Computer and System Sciences **78** (2012), no. 3, 731–764. MR2900032

[10] Silvère Gangloff, *Algorithmic complexity of growth-type invariants of SFT under dynamical constraints*, Ph.D. thesis, Université d'Aix-Marseille, 2018, theses.fr/2018AIXM0231.

[11] Eli Glasner, *Ergodic theory via joinings*, AMS, 2003. MR1958753

[12] Peter Gács, *Reliable cellular automata with self-organization*, Journal of Statistical Physics **103** (2001), no. 1-2, 45–267. MR1828729

[13] Gustav Hedlund, *Endomorphisms and automorphisms of the shift dynamical system*, Mathematical Systems Theory **3** (1969), 320–375. MR0259881

[14] Michael Hochman, *On the dynamics and recursive properties of multidimensional symbolic systems*, Inventiones Mathematicae **176** (2009), no. 1, 131–167. MR2485881

[15] Olav Kallenberg, *Foundations of modern probability*, 3rd ed., Springer, 2002. MR4226142

[16] Jarkko Kari, *A small aperiodic set of Wang tiles*, Discrete Mathematics **160** (1996), 259–264. MR1417578

[17] Ulrich Krengel and Antoine Brunel, *Ergodic theorems*, De Gruyter, 1985. MR0797411

[18] Dov Levine and Paul Steinhardt, *Quasicrystals: A new class of ordered structures*, Physical Review Letters **53** (1984), 2477–2480. MR0831879

[19] Thomas Liggett, *Interacting particle systems*, Classics in mathematics, Springer, 2005. MR2108619

[20] Thomas Liggett, Roberto Schonmann, and A. M. Stacey, *Domination by product measures*, Annals of Probability **25** (1997), no. 1, 71–95. MR1428500

[21] Douglas Lind and Brian Marcus, *An introduction to symbolic dynamics and coding*, 2 ed., Cambridge Mathematical Library, Cambridge University Press, 2021. MR1369092

[22] Raphael Robinson, *Undecidability and nonperiodicity for tilings of the plane*, Inventiones mathematicae **12** (1971), 177–209. MR0297572

[23] Richard Schwartz, *The Robinson tiles*, www.math.brown.edu/reschwar/MFS/handout9.pdf, 2007.

[24] Dan Shechtman, Ilan Blech, Denis Gratias, and John Cahn, *Metallic phase with long-range orientational symmetry and no translational symmetry*, Physical Review Letters **53** (1984), 1951–1953, journals.aps.org/prl/abstract/10.1103/PhysRevLett.53.1951.

[25] Siamak Taati, *A finite-range lattice gas model with quasicrystal phases at positive temperatures*, Work in progress.

[26] Aernout van Enter, Jacek Miękisz, and Miloš Zahradník, *Nonperiodic long-range order for fast-decaying interactions at positive temperatures*, Journal of Statistical Physics **90** (1998), no. 5-6, 1441–1447. MR1628261

[27] Anatoly Vershik, *The Kantorovich metric: the initial history and little-known applications*, Journal of Mathematical Sciences (New York) **133** (2006), no. 4, 1410–1417. MR2117883