# Support Stability of Maximizing Measures for Shifts of Finite Type


Juliano S. Gonschorowski[1], Anthony Quas[2], and Jason Siefken[3]

[1]Universidade Tecnológica Federal do Paraná - Av. Prof. Laura Pacheco Bastos, 800 - CEP 85053-525 - Guarapuava - PR - Brasil (e-mail: julianod@utfpr.edu.br)
[2]Department of Mathematics and Statistics, University of Victoria, Victoria BC, Canada V8W 3R4 (e-mail: aquas@uvic.ca)
[3]Department of Mathematics, University of Toronto, Toronto Ontario, Canada M5S 2E4 (e-mail: siefkenj@math.utoronto.ca)


August 13, 2019


**Abstract**

This paper establishes a fundamental difference between $\mathbb{Z}$ subshifts of finite type and $\mathbb{Z}^2$ subshifts of finite type in the context of ergodic optimization. Specifically we consider a subshift of finite type $X$ as a subset of a full shift $F$. We then introduce a natural penalty function $f$, defined on $F$, which is 0 if the local configuration near the origin is legal and $-1$ otherwise.

We show that in the case of $\mathbb{Z}$ subshifts, for all sufficiently small perturbations, $g$, of $f$, the $g$-maximizing invariant probability measures are supported on $X$ (that is the set $X$ is *stably maximized* by $f$). However, in the two-dimensional case, we show that the well-known Robinson tiling fails to have this property: there exist arbitrarily small perturbations, $g$, of $f$, for which the $g$-maximizing invariant probability measures are supported on $F \setminus X$.

# 1 Introduction

Ergodic optimization is the study of the extreme points of the functional $P_f : \mathcal{M}_T(X) \to \mathbb{R}$ defined by

$$P_f(\mu) = \int f \, d\mu,$$

where $T : X \to X$ is a continuous dynamical system, $\mathcal{M}_T(X)$ is the set of $T$-invariant probability measures, and $f : X \to \mathbb{R}$ is a continuous function called a potential (see Jenkinson [5], [6], [7]).

Even in a simple setting—when $T$ is the full shift on a finite alphabet and $f$ is an exponentially decaying potential—the structure of $\mathcal{M}_T(X)$ can be quite complicated, giving rise to subtle issues (see Chazottes, Gambaudo, Hochman and Ugalde [4]).

Closely related to ergodic optimization is the study of Gibbs measures. In the study of Gibbs measures, there are fundamental differences between the one-dimensional and higher-dimensional cases. A well known example is the Ising model (a finite range model for interaction of particles). In one dimension, it is known that there is a unique Gibbs measure at all temperatures, whereas in higher dimensions, there is a unique Gibbs measure for high temperatures, but a phase transition gives multiple Gibbs measures at low temperatures.

By analogy, in this paper, we exhibit fundamental differences between the situation for ergodic optimization for one-dimensional dynamical systems and that for higher-dimensional dynamical systems. We show this difference by focusing on *stability* in the context of ergodic optimization.

Yuan and Hunt [12] showed that in one dimension, invariant measures supported on periodic orbits may be *stably maximizing* (small perturbations of the potential remain maximized by the same invariant measure), whereas aperiodic invariant measures (measures such that the collection of periodic points has measure 0) are never stably maximizing. This result was extended by Contreras [1] to show that for a generic Lipschitz function, there is a unique maximizing invariant measure, and that measure is supported on a periodic orbit. Further, this invariant measure is stably maximizing.

In this paper, we consider a shift of finite type as a subset of a full shift and start with a short-range potential function that penalizes local configurations that are in the set of forbidden words for the shift of finite type. Not surprisingly for these potentials, the maximizing measures are precisely those measures supported on the shift of finite type. We then considered what happens when the potential is perturbed. Here we establish a key difference between the one-dimensional and higher-dimensional cases: In one dimension, for sufficiently small perturbations of the original potential, all maximizing measures remain supported on the shift of finite type. However, in two (or higher) dimensions, there exist shifts of finite type such that arbitrarily small perturbations of the potential lead to maximizing measures whose support lies outside the shift of finite type.

We conisder shifts of finite type over a finite alphabet $\Sigma$. If $L$ is a finite collection of finite blocks (or words), $\mathrm{SFT}(L)$ will denote the subshift of finite type consisting of those points in $\Sigma^{\mathbb{Z}^d}$ containing no copy of a block in $L$.

Specifically, we obtain the following results.

**Theorem 1.** *Let $X = SFT(L) \subset \Sigma^{\mathbb{Z}}$ be a one-dimensional aperiodic irreducible subshift of finite type, where $L$ is a set of forbidden words of length $2$. Define the Lipschitz function*

$$f(x) = \begin{cases} -1 & \text{if } x_0 x_1 \in L \\ 0 & \text{otherwise} \end{cases}.$$

*Then there exists an $\varepsilon > 0$ such that for all $g$ with $\|f - g\|_{Lip} < \varepsilon$, every $g$-maximizing measure is supported on $X$.*

**Theorem 2.** *There exists a shift of finite type $X = SFT(L) \subset \Sigma^{\mathbb{Z}^2}$, where $L$ is a set of $2 \times 2$ forbidden blocks, such that $X$ has the following property. The Lipschitz function*

$$f(x) = \begin{cases} -1 & \text{if } \begin{matrix} x_{01} & x_{11} \\ x_{00} & x_{10} \end{matrix} \in L \\ 0 & \text{otherwise} \end{cases}$$

*is uniquely, but unstably, optimized by a measure supported on $X$. That is, for all $\varepsilon > 0$, there exists $g \in Lip(X)$, such that $\|f - g\|_{Lip} < \varepsilon$, and $\mu_g$, the $g$-maximizing measure, is supported on $X^c$.*

In both of these theorems, $f$ can be thought of as a *penalty* function. That is, $f$ penalizes measures which contain the forbidden blocks in their support.

# 2 One-dimensional shifts of finite type

To prove Theorem 1, we make use of the technique of *coupling and splicing* (described in more detail in [9]).

**Lemma 2.1.** *Let $B \subset X$ a subset of a compact metric space $X$, $\varepsilon > 0$, and $f$ a Lipschitz function such that $f|_B = 0$. If a Lipschitz function $g$ satisfies $\|f - g\|_{Lip} < \varepsilon$, then $|g(x) - g(y)| < \varepsilon d(x, y)$ for all $x, y \in B$.*

*Proof.* Let $c_h$ denote the Lipschitz constant for the function $h$. Since $c_{f|_B} = 0$, we have

$$c_{f-g|_B} = c_{-g|_B} = c_{g|_B} < \varepsilon.$$

□

*Proof of Theorem 1.* The argument is based on the *coupling and splicing* technique where, starting from an invariant measure, $\mu$, supported on $X^c$, one obtains a related invariant measure, $\nu$, supported on $X$ such that $\int g\, d\nu > \int g\, d\mu$ whenever $g$ is a small perturbation of $f$. The measure $\nu$ is obtained from $\mu$ via a 'path-wise surgery', where realizations of $\mu$ are edited to obtain realizations of $\nu$.

Since $X$ is an aperiodic irreducible subshift of finite type, there exists $N > 0$ such that $A^N > 0$ where $A$ is the adjacency matrix of the subshift of finite type. In particular for each $\ell \geq N$ and each pair of symbols $(i, j)$, there is a path of length exactly $\ell$ connecting them containing no forbidden words.

Denote by $I$ the set where $f$ is zero ($I^c$ is where $f$ is $-1$), and note that $X = SFT(L)$ is a subset of $I$. Also note that $X$ and $X^c$ are shift-invariant, so that any ergodic invariant measure is supported either on $X$ or $X^c$. Let $\varepsilon = 1/(6N)$ and $g$ satisfy $\|f - g\|_{Lip} < \varepsilon$.

Let $\mu$ be an ergodic invariant measure supported on $X^c$. If $\mu(I^c) \geq \frac{1}{2N}$, using the fact that $|f(x) - g(x)| \leq \varepsilon$ for all $x$ and $\int f\, d\mu = -\mu(I^c) \leq -\frac{1}{2N}$ we get

$$\int g\, d\mu = \int f\, d\mu + \int (g - f)\, d\mu < -\frac{1}{2N} + \varepsilon = -\frac{1}{3N}.$$

On the other hand, if $\nu$ is an invariant measure supported on $X$, we have

$$\int g\, d\nu \geq -\varepsilon > -\frac{1}{3N}.$$

Hence $\mu$ cannot be $g$-maximizing.

Now if $\mu(I^c) < 1/(2N)$, using the Birkhoff ergodic theorem, we have for $\mu$-almost every $x$,

$$\lim_{n\to\infty} \frac{1}{n} S_n \mathbf{1}_{I^c}(x) = \int \mathbf{1}_{I^c}\, d\mu = \mu(I^c),$$

where $S_n f(x)$ denotes the Birkhoff sum $\sum_{j=0}^{n-1} f(T^j x)$.

Hence for all sufficiently large $n$, the number of $j$ in the range 0 to $n-1$ such that $x_j x_{j+1} \in L$ is smaller than $n/(2N)$. Hence $x$ must almost surely contain infinitely many subwords of length $N+1$ with no forbidden blocks.

Now we split into bad and good blocks. For any $x$ in $I \setminus X$, we define a relation $r$ on $S := \{n \in \mathbb{Z} \colon T^n(x) \in I^c\}$ where for $n, n' \in S$, $n\, r\, n'$ if $|n - n'| < N$. Taking the transitive closure, we obtain an equivalence relation $\sim$ on $S$, where $k \sim m$ if there exist $n_1, n_2, \ldots, n_p \in S$ such that $k r n_1; n_1 r n_2; \ldots; n_p r m$. For $n \in S$, define $[n] = \{m \in S : m \sim n\}$. For such a set we define the interval in $\mathbb{Z}$, $B = [\alpha, \beta]$, called a *bad block*, where $\alpha(n) = \min\{i : i \in [n]\}$ and $\beta(n) = \max\{\max\{i : i \in [n]\}, \alpha(n) + N\}$. Note that each bad block has size $|B| \geq N + 1$ and if $n < n'$ are such that $[n] \neq [n']$, then by the above definition, $\beta(n) < \alpha(n')$, so the $[\alpha(n), \beta(n)]$ are disjoint and contain all the forbidden words in $x$.

Now for each $x \in I \setminus X$ that has infinitely many blocks of length $N$ with no forbidden words, define a new sequence $\tilde{x} \in X$ to be equal to $x$ off $\bigcup B$, and on each bad block B to be equal to the lexicographically minimal word joining $x_{\alpha(B)}$ to $x_{\beta(B)+1}$ with no forbidden subwords (such a word exists by the aperiodicity and irreducibility of $X$).

The idea now is to estimate the Birkhoff sums $S_n g(x)$ and $S_n g(\tilde{x})$ in order to show that we can construct another measure $\nu$ such that $\int g\,d\nu > \int g\,d\mu$, showing that $\mu$ is not maximizing. We also make use of the notation $S_a^b f(x) = \sum_{j=a}^{b} g(T^j x)$.

For any bad block $B = [\alpha, \beta]$, the number of $i \in B$ such that $T^i(x) \in I^c$ is at least $|B|/N$ and on this set $g(x) < -1 + \varepsilon$. For the other values of $i$ (at most $|B| - |B|/N$ of them), $T^i(x) \in I \setminus X$ and on this set $-\varepsilon < g(x) < \varepsilon$. In this way we have:

$$
\begin{aligned}
S_\alpha^\beta g(x) - S_\alpha^\beta g(\tilde{x}) &< \frac{|B|}{N}(-1+\varepsilon) + \left(|B| - \frac{|B|}{N}\right)\varepsilon + |B|\varepsilon \\
&= -\frac{|B|}{N} + 2|B|\varepsilon = -\frac{2|B|}{3N} \\
&\leq -\frac{2}{3N}\#\{i \in B : T^i(x) \in I^c\}.
\end{aligned}
$$

We enumerate the bad blocks from left to right as $(B_n)_{n\in\mathbb{Z}}$ with $B_0$ the leftmost bad block such that $\alpha(B_0) \geq 0$. Let $m_k = \alpha(B_k)$ and $n_k = \beta(B_k)$ so that $\cdots < m_{-1} < n_{-1} < m_0 < n_0 < \cdots$ and $m_{-1} < 0 \leq m_0$. Now we will estimate the Birkhoff sum between two bad blocks $B_k$ and $B_{k+1}$, that is the sum from $n_k + 1$ to $m_{k+1} - 1$.

When $i$ varies from $n_k + 1$ to $n_k + l$ where $l = \lfloor \frac{m_{k+1} - n_k}{2} \rfloor$, we have

$$d(T^i x, T^i \tilde{x}) = \frac{1}{2^{i-n_k}};$$

and when $i$ varies from $n_k + l + 1$ to $m_{k+1} - 1$,

$$d(T^i x, T^i \tilde{x}) = \frac{1}{2^{m_{k+1}-i}}.$$

By Lemma 2.1:

$$\sum_{i=1}^{l} g(T^{n_k+i}(x)) - g(T^{n_k+i}(\tilde{x})) < \sum_{i=1}^{l} \frac{1}{2^i}\varepsilon < \varepsilon$$

and

$$\sum_{i=l+1}^{m_{k+1}-n_k-1} g(T^{n_k+i}(x)) - g(T^{n_k+i}(\tilde{x})) < \sum_{i=l+1}^{m_{k+1}-n_k-1} \frac{1}{2^{m_{k+1}-n_k-i}}\varepsilon < \varepsilon.$$

In this way we get

$$S_{n_k+1}^{m_{k+1}-1} g(x) - S_{n_k+1}^{m_{k+1}-1} g(\tilde{x}) < 2\epsilon = \frac{1}{3N} \tag{2.1}$$

The same bound applies if the interval $[n_k+1, m_{k+1}-1]$ is replaced by any sub-interval.

Now let $k$ be any positive integer. Combining (2.1) and (2.2), we have

$$\begin{aligned} S_0^{n_k} g(x) - S_0^{n_k} g(\tilde{x}) &\leq -\frac{2}{3N} \#\{i \in [0, n_k) \colon T^i(x) \in I^c\} + \frac{(k+1)}{3N} \\ &\leq -\frac{1}{3N}(\#\{i \in [0, n_k) \colon T^i(x) \in I^c\} - 1). \end{aligned}$$

Hence we see

$$\liminf_{m\to\infty} \frac{1}{m} S_m g(x) + \frac{\mu_g(I^c)}{3N} < \liminf_{m\to\infty} \frac{1}{m} S_m g(\tilde{x}).$$

So by passing to a subsequence of the sequence of empirical measures for $\tilde{x}$, we see that there exists an invariant probability measure $\nu$ with support in $X$ such that

$$\int g\, d\mu < \int g\, d\nu,$$

establishing that $\mu$ is not $g$-maximizing.

□

# 3 Two-dimensional shifts of finite type

The higher-dimensional analogue of Theorem 1 is not true; in this section we prove Theorem 2.

### 3.1 The Robinson Tiles

The Robinson tiles are a finite set of tiles which tile the plane but only do so aperiodically. Though not the first tile set with this property, the Robinson tiles are well known and the tilings they produce exhibit a hierarchical structure. The Robinson tile set, among other things, has been extended to allow the embedding of Turing machines [10, 3] and provides a counterexample to Wang's conjecture on the decidability of whether a given subshift of finite type is non-empty [11].

In Figure 1, we show the 28 Robinson tiles annotated with arrows. We denote by $\Sigma$ the set of tiles, and $X \subset \Sigma^{\mathbb{Z}^2}$ the Robinson system consisting of tilings obeying the rules:

1. at each intersection of two tiles, all arrow heads must meet arrow tails and vice versa; and

2. there is a translate of the sub-lattice $2\mathbb{Z} \times 2\mathbb{Z}$ at which every tile is a cross (one of tiles 1–4).

Note that in rule 2, although every tile on the sub-lattice contains a cross, these are not the only crosses that appear. Though this rule does not appear to be a shift of finite type rule, as described in [8] this rule can be imposed by expanding the tile set to a set of 56 tiles by adding additional markings to record parity with respect to the $2\mathbb{Z} \times 2\mathbb{Z}$ lattice. Using these 56 tiles in place of the 28 tiles with rule 2 does not affect the proof that follows.

In Figure 2, two configurations that belong to the language of $X$ are shown.

The configurations in $X$ are known to have a hierarchical structure. We now describe this structure, referring to [8] for more details. We let $y^{(1)}_{\mathrm{NW}}$, $y^{(1)}_{\mathrm{NE}}$, $y^{(1)}_{\mathrm{SW}}$ and $y^{(1)}_{\mathrm{SE}}$ denote the $1 \times 1$ blocks consisting of a cross tile of types 1 to 4 respectively (the 'directions' of the crosses refer to the sides with the double arrows).

We then inductively build a sequence of larger blocks $y^{(N)}_{\mathrm{NW}}$, $y^{(N)}_{\mathrm{NE}}$, $y^{(N)}_{\mathrm{SW}}$ and $y^{(N)}_{\mathrm{SE}}$, where each of the $y^{(N)}$ blocks is of size $(2^N - 1) \times (2^N - 1)$ and

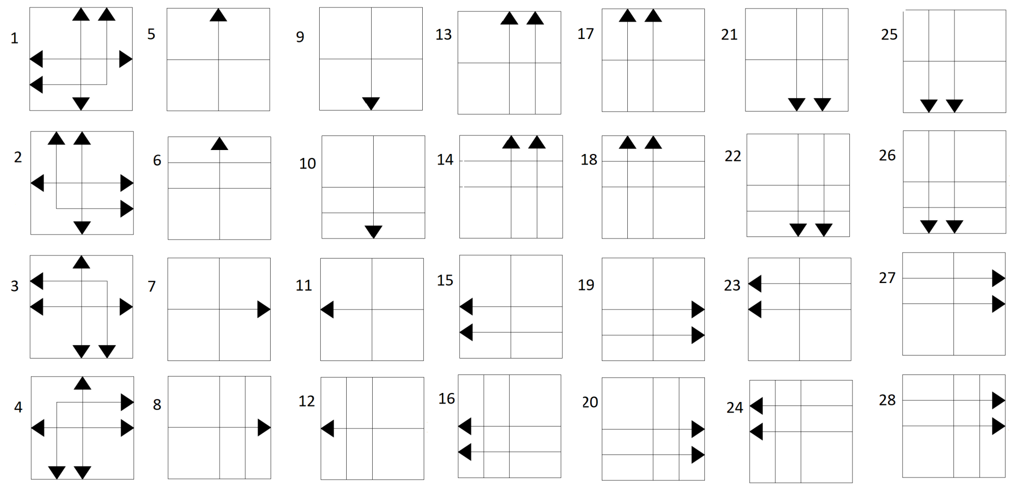


Figure 1: The Robinson tiles.

is obtained by placing copies of each of $y_{\mathrm{SE}}^{(N-1)}$, $y_{\mathrm{SW}}^{(N-1)}$, $y_{\mathrm{NE}}^{(N-1)}$ and $y_{\mathrm{NW}}^{(N-1)}$ in the four corners of a $(2^N - 1) \times (2^N - 1)$ grid respectively with gaps of size 1 in between. The central tile is then one of the four crosses ($y_{\mathrm{XX}}^{(N)}$ with XX$\in$ {SE,SW,NE,NW} has a cross of type XX in the middle) and the gaps are filled in as shown in Figure 3 (b). More precisely, in the particular case of $y_{\mathrm{SE}}^{(N)}$ illustrated, the tiles filling the gap above the central cross are of type 5 except for that the middle tile is of type 6; the tiles filling the gap to the left of the central cross are of type 11 except the middle tile is of type 12; the tiles filling the gap below the central cross are of type 25, except for the middle one which is of type 26; and the tiles filling the gap to the right of the central cross are of type 27, except for the middle one which is of type 28.

Note that by construction, these blocks satisfy the second rule defining the Robinson tiling: inductively, we see that the tiles of each $y^{(N-1)}$ occupying the "odd sub-lattice" (where both coordinates are odd) are all crosses. By the way that these tiles are assembled to form $y^{(N)}$, one sees that the same is true at the next level.

We now turn to rule 1. The key observation, proved inductively, is that

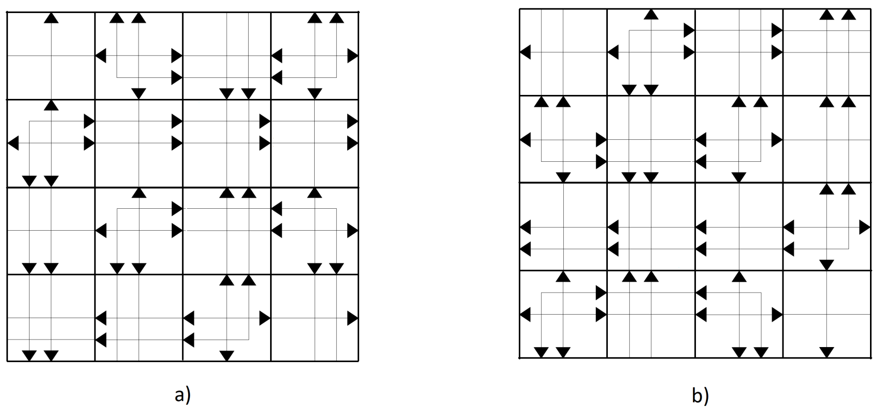


Figure 2: Configurations in $X$

in $y^{(N)}_{\mathrm{XX}}$, there is a single outward-pointing arrow coming from each external edge of each tile forming an edge of the block except for the central tile of each edge where the outward-pointing arrows match those of $y^{(1)}_{\mathrm{XX}}$. One can then check that this hypothesis is preserved by the inductive construction of $y^{(N)}$ from $y^{(N-1)}$ described above,

Now assume that the $y^{(N-1)}$'s satisfy rule 1. By the above observation, one can check that rule 1 is satisfied by the tiles which are used to fill in the gaps between the $y^{(N-1)}$'s, so that the $y^{(N)}$'s also satisfy rule 1. Hence we can show inductively that the $y^{(N)}$'s satisfy the rules for the Robinson tiling.

The remarkable feature of the tiles is that this is essentially the only way that they fit together. For any element $x \in X$, either every finite block is contained in a single copy of $y^{(N)}_{\mathrm{XX}}$ for some $N$ and XX$\in$ {SE,SW,NE,NW} or there are exceptional tilings with a single horizontal 'fault line' of height 1, a single vertical fault line of width 1, or there is a fault line of one type, and one or both of the two half planes created by removing this fault line has a fault half-line in the orthogonal direction; in the exceptional tilings, any block not intersecting a fault line is contained in a single copy of $y^{(N)}_{\mathrm{XX}}$ for some $N$.

We record this as follows:

**Theorem 3.1** (Johnson-Madden [8]). *Each point $x$ of $X$ at most two hori-*

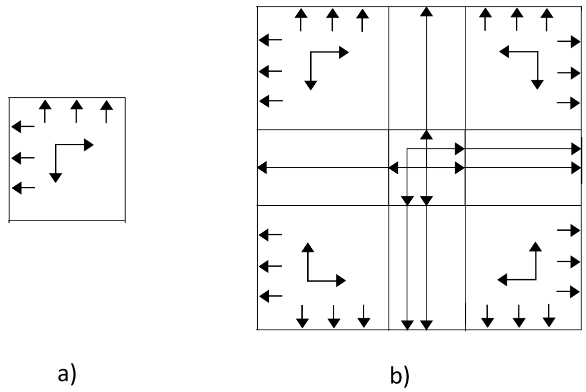


Figure 3: Induction used in [8]

*zontal and vertical fault half-lines. In each remaining region $x$ has the hierarchical structure depicted in Figure 4, where the squares represent $y_i^{(N)}$ blocks with the orientations depicted. The gaps (of size 1) between the $y_i^{(N)}$ blocks are then partially filled in to obtain a similar periodic configuration of $y_i^{(N+1)}$ blocks etc.*

We shall also require the statement above describing the outer boundaries of the $y^{(N)}$ tiles.

**Proposition 3.2.** *For all $N$, the boundaries of each of $y_{SE}^{(N)}$, $y_{SW}^{(N)}$, $y_{NE}^{(N)}$ and $y_{NW}^{(N)}$ each consist of single out-pointing arrows except for the middles of the two sides corresponding to the direction named in the block. For example, $y_{SE}^{(N)}$ has a boundary consisting of single out-pointing arrows except for the middles of the bottom ('south') and right ('east') sides, which have double out-pointing arrows.*

From now on we work with only one of $y_{\mathrm{XX}}^{(N)}$'s, namely $y_{SE}^{(N)}$, and we will denote it just by $y^{(N)}$. We denote by $x^{(N)}$ the $2^N \times 2^N$ square which is $y^{(N)}$ but padded on the top and left sides as in Figure 5(b).

In order to prove Theorem 2, we will create a periodic configuration $x \notin X$ obtained by periodically repeating the square $x^{(N)}$. This periodic point will have some tiles that don't obey the matching rules. We will show that the

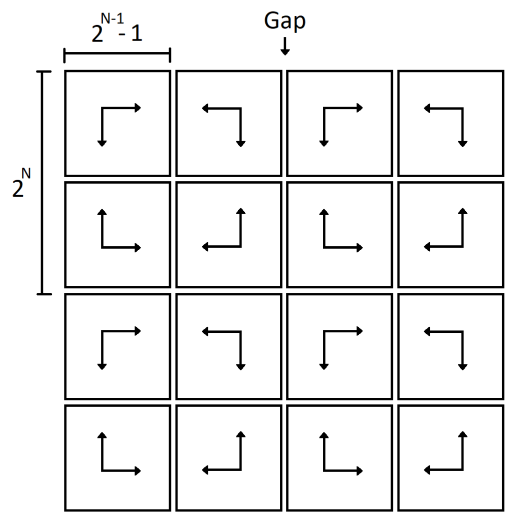


Figure 4: General configuration in $X$.

number of these errors per period remains bounded as $N$ varies (where we say that $x$ has an error at position $\vec{v}$ if $x_{\vec{v}+\{0,1\}^2}$ contains violates the tile matching rules described above).

**Proposition 3.3.** *The periodic extension of $x^{(N)}$ in which $x^{(N)}$ is repeated with vertical and horizontal periods $2^N$ has 7 errors per period for every $N \geq 3$.*

*Proof.* By Proposition 3.2, the boundary of $y^{(N)}$ consists of single outward-pointing arrows, except for the middle tile of the bottom and the middle tile of the right side. By the construction of $x^{(N)}$, it has only tails in the top and left sides as in Figure 6.

In this way, when we repeat $x^{(N)}$ with horizontal and vertical periods $2^N$ to form a periodic configuration, it will have errors in the middle of the four sides and errors caused by the mismatch between the top left cross and its neighbours to the top and the left.

As the point $x^{(N)}$ is shifted, the translates on which $f$ 'detects an error' (that is for which $f$ takes the value $-1$) are illustrated in Figure 7(b) $\square$

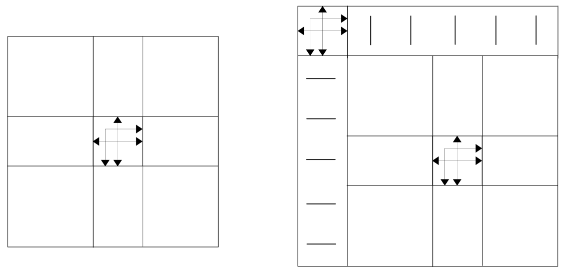

Figure 5: (a) $y^{(N)}$ and (b) $x^{(N)}$

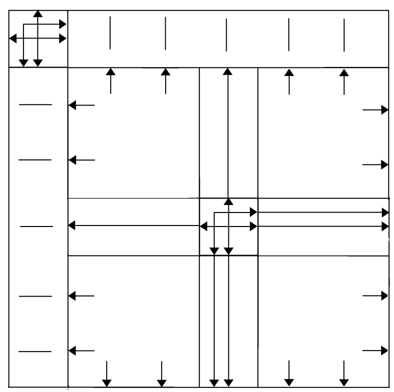

Figure 6: The block $x^{(N)}$

**Proposition 3.4.** $(T, X)$ *is uniquely ergodic.*

This proposition can be obtained as a corollary of Proposition 8 in the paper [2] of Cortez once we establish that the Robinson tiling is a regular Toeplitz system. Rather than justifying the hypotheses and referring to [2], we sketch a version of Cortez's proof in our situation.

*Proof.* Let $\mu$ be an invariant probability measure on $X$. By the Poincaré recurrence theorem, $\mu$-almost every configuration has no fault lines (as these are non-recurrent sets: horizontal fault lines do not recur under the vertical action; vertical fault lines do not recur under the horizontal action). Hence by Theorem 3.1, and for each $N > 0$, $\mu$-a.e. configuration looks like Figure

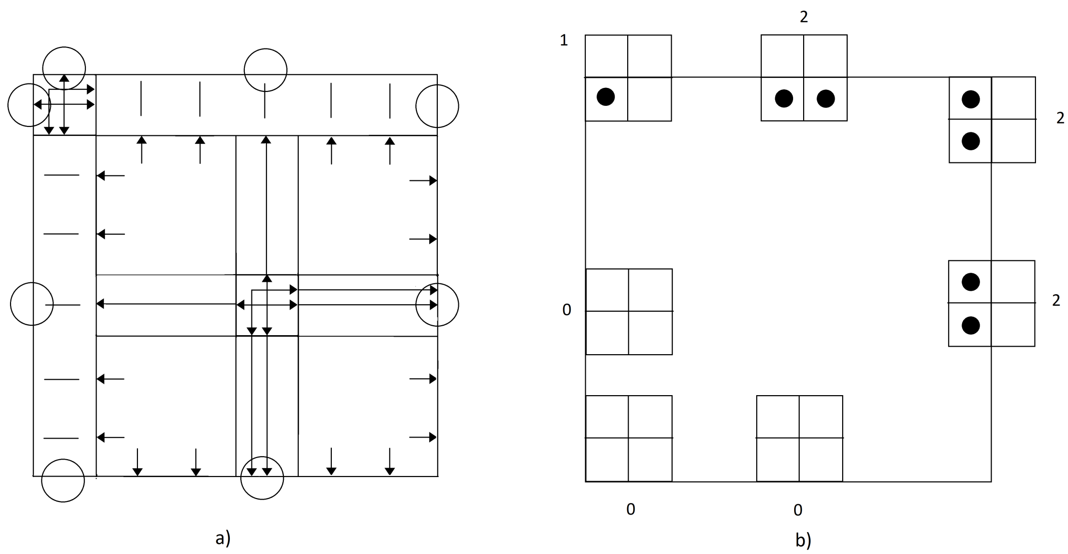


Figure 7: a) Errors in the periodic configuration and b) The dots indicate those translations of the configuration such that the function $f$ "sees an error" that is such that $f$ takes the value $-1$.

4, where the gaps are of width 1, and may be filled in in various ways, but the $y^{(N-1)}$'s repeat periodically (with period $2^N$).

Let $C$ be any block of size $k \times l$, say, and let $N$ be a large number such that $k, l < 2^{N-1}$. Let $n_{\mathrm{XX}}^{(N-1)}$ be the number of times that $C$ occurs in $y_{\mathrm{XX}}^{(N-1)}$ for XX$\in \{\mathrm{SE}, \mathrm{SW}, \mathrm{NE}, \mathrm{NW}\}$. We then notice that for $\mu$-almost every $x \in X$, the frequency of $C$'s is at least $(n_{\mathrm{SE}}^{(N-1)} + n_{\mathrm{SW}}^{(N-1)} + n_{\mathrm{NE}}^{(N-1)} + n_{\mathrm{NW}}^{(N-1)})/2^{2N}$ and at most $(n_{\mathrm{SE}}^{(N-1)} + n_{\mathrm{SW}}^{(N-1)} + n_{\mathrm{NE}}^{(N-1)} + n_{\mathrm{NW}}^{(N-1)})/2^{2N} + (k+l)/2^N$ (the second term in the upper estimate comes from counting the proportion of translates of a $k \times l$ window that intersect the non-periodic part of Figure 4). Since the upper and lower bounds differ by a quantity that approaches 0 as $N$ increases, we see that $\mu$-almost every $x \in X$ has the frequency of $C$'s given by

$$\lim_{N\to\infty} \frac{n_{\mathrm{SE}}^{(N-1)} + n_{\mathrm{SW}}^{(N-1)} + n_{\mathrm{NE}}^{(N-1)} + n_{\mathrm{NW}}^{(N-1)}}{2^{2N}},$$

and hence this gives the value of $\mu(C)$. In particular, if $\mu'$ is another invariant probability measure, then $\mu'(C) = \mu(C)$. Since this is true for all blocks, we deduce $\mu' = \mu$ as required. $\square$

Let $\mu$ be the unique invariant measure (Proposition 3.4) supported on $X$. We will show next that $\mu$ is invariant under rotation by 180 degrees.

For any configuration $x \in \Sigma^{\mathbb{Z}^2}$, we define the function $r_\pi : \Sigma^{\mathbb{Z}^2} \to \Sigma^{\mathbb{Z}^2}$ by: $r_\pi(x)_{\vec{u}} = R(x_{-\vec{u}})$ for all $\vec{u} \in \mathbb{Z}^2$, where $R$ is the map from the tiles to the tiles that rotates them by 180 degrees (e.g $R$(tile 1)=tile 4; $R$(tile 13)=tile 26 etc). An example of a configuration and its rotation under $r_\pi$ is shown in figure 2 (a) and (b). Note that since $r_\pi$ preserves the matching rules we have $r_\pi(X) = X$.

Now let $\nu = \mu \circ r_\pi^{-1}$ be the push-forward of $\mu$ under $r_\pi$. Notice that $r_\pi(X) = X$, so that $\operatorname{supp}\nu = \operatorname{supp}\mu$. The measure $\nu$ is invariant because $r_\pi \circ T^{\vec{u}} = T^{-\vec{u}} \circ r_\pi$ for all $\vec{u} \in \mathbb{Z}^2$, so that $\nu(T^{-\vec{u}}B) = \mu \circ r_\pi^{-1} \circ T^{-\vec{u}}(B) = \mu(T^{\vec{u}} r_\pi^{-1} B) = \mu(r_\pi^{-1} B) = \nu(B)$. By the uniqueness of $\mu$ we have that $\nu = \mu$, that is, $\mu$ is invariant under rotation (by 180°).

Now we are ready to prove Theorem 2.

*Proof of Theorem 2.* Given $\varepsilon > 0$ and $N > \log_2\left(\frac{7}{\varepsilon}\right)$, consider $\bar{x}^{(N)}$ be the periodic point in $\Sigma^{\mathbb{Z}^2}$ formed by repeating the block $x^{(N)}$. Now we define the function $g : \Sigma^{\mathbb{Z}^2} \to \mathbb{R}$ by $g = f + \varepsilon h$, where $h$ is the following function:

$$h = 1_{\tau_{25} \cup \tau_{27}} - 1_{\tau_{13} \cup \tau_{15}}.$$

Here, $\tau_{**}$ means Tile $**$ in the set $\Sigma$ (Figure 1). In Figure 8(a), we show the left, right, top and bottom sides of the central cross in $x^{(N)}$ with their respective tiles.

Let $\mu_{per}$ be the invariant probability measure supported on the orbit of $\bar{x}^{(N)}$ and $\mu$ the invariant probability measure supported on $X$. Then

$$\int f \, d\mu_{per} = \frac{-7}{(2^N)^2} \quad \text{and} \quad \int f \, d\mu = 0,$$

where the first equality follows from Proposition 3.4. On the other hand, as $\mu(\tau) = \nu(R(\tau))$ for all $\tau \in \Sigma$, we have $\mu(\tau_{15}) = \nu(\tau_{15}) = \mu(\tau_{27})$ and $\mu(\tau_{13}) = \nu(\tau_{13}) = \mu(\tau_{25})$. Hence $\int h \, d\mu = 0$. It follows that $\int g \, d\mu = 0$.

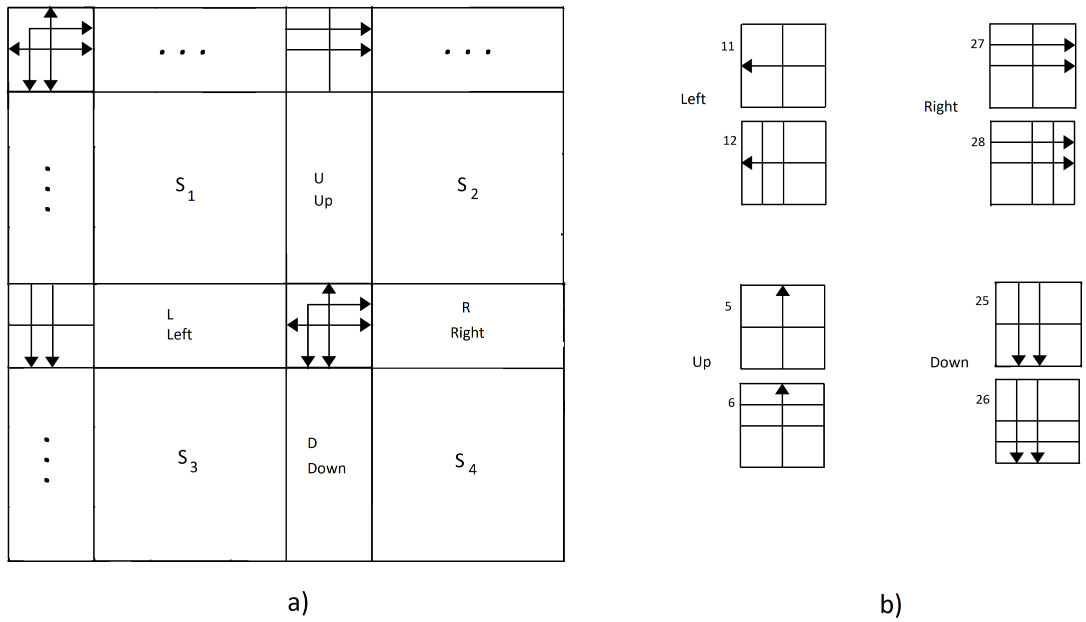


Figure 8: Tiles in $x^{(N)}$

Now in order to calculate $\int h\, d\mu_{per}$, we observe that inside the squares $S_1, \ldots, S_4$ in Figure 8, the number of tiles of type $\tau_{25}$ and $\tau_{27}$ in $S_1$ is equal to the number of tiles of type $\tau_{13}$ and $\tau_{15}$ in $S_4$. Similarly, the number of tiles of type $\tau_{13}$ and $\tau_{15}$ in $S_1$ is equal to the number of tiles of type $\tau_{25}$ and $\tau_{27}$ in $S_4$. The same holds for $S_2$ and $S_3$. Hence the contribution to $\int h\, d\mu_{\mathrm{per}}$ coming from the regions $S_1$, $S_2$, $S_3$, and $S_4$ is 0 since the terms cancel.

From this, we see

$$\int h\, d\mu_{per} = \frac{\#(x \setminus S)_{\tau_{25}} + \#(x \setminus S)_{\tau_{27}}}{(2^N)^2} = \frac{2[(2^N - 1) + (2^{N-1} - 1)]}{2^{2N}} > \frac{1}{2^N} > 0,$$

where $\#(x \setminus S)_{\tau_{25}}$ represents the number of tiles of type 25 in the square $x^{(N)}$ lying outside $S = \bigcup_{i=1}^4 S_i$. Finally, for $g$ we have

$$\int g\, d\mu_{per} > \frac{-7}{(2^N)^2} + \varepsilon \frac{1}{2^N} > 0 = \int g\, d\mu \quad \text{whenever} \quad N > \log_2\left(\frac{7}{\varepsilon}\right).$$

As the measure $\mu_{per}$ is not supported on $X$, the result is proved. $\square$

# References


[1] G. Contreras. Ground states are generically a periodic orbit. *Invent. Math.*, 205:383–412, 2016.

[2] M.-I. Cortez. $\mathbb{Z}^d$ Toeplitz arrays. *Discrete Contin. Dyn. Syst.*, 15:859–881, 2006.

[3] M. Hochman. On the dynamics and recursive properties of multidimensional symbolic systems. *Invent. Math.*, 176(1):131, 2009.

[4] M. Hochman, J. R. Chazottes, J. M. Gambaudo, and E. Ugalde. On the finite-dimensional marginals of shift-invariant measures. *Ergodic Theory Dynam. Systems*, 32:1485–1500, 2012.

[5] O. Jenkinson. Ergodic optimization. *Discrete Contin. Dyn. Syst.*, 15:197–224, 2006.

[6] O. Jenkinson. Every ergodic measure is uniquely maximizing. *Discrete Contin. Dyn. Syst.*, 16:383–392, 2006.

[7] O. Jenkinson. A partial order on $\times 2$-invariant measures. *Math. Res. Lett.*, 15:893–900, 2008.

[8] A. Johnson and K. Madden. Putting the pieces together: Understanding Robinson's nonperiodic tilings. *College Math. J.*, 28:172–181, 1997.

[9] A. Quas. Coupling and splicing. Lecture notes available at http://www.math.uvic.ca/faculty/∼aquas.

[10] R. M. Robinson. Undecidability and nonperiodicity for tilings of the plane. *Invent. Math.*, 12:177–209, 1971.

[11] H. Wang. Proving theorems by pattern recognition, part 2. *Bell Systems Technical Journal*, 40:1–41, 1961.

[12] G. Yuan and B. R. Hunt. Optimal orbits of hyperbolic systems. *Nonlinearity*, 12:1207–1224, 1999.