TOPOLOGICAL MIXING OF RANDOM SUBSTITUTIONS

EDEN DELIGHT MIRO, DAN RUST, LORENZO SADUN, GWENDOLYN TADEO

ABSTRACT. We investigate topological mixing of compatible random substitutions. For primitive random substitutions on two letters whose second eigenvalue is greater than one in modulus, we identify a simple, computable criterion which is equivalent to topological mixing of the associated subshift. This generalises previous results on deterministic substitutions. In the case of recognisable, irreducible Pisot random substitutions, we show that the associated subshift is not topologically mixing. Without recognisability, we rely on more specialised methods for excluding mixing and we apply these methods to show that the random Fibonacci substitution subshift is not topologically mixing.

1. INTRODUCTION

The question of what it means for a dynamical system to be *disordered* is a subtle one with a rich history dating back to the birth of dynamics itself with Poincaré's study of chaotic orbits in celestial mechanics [24]. Mathematicians measure disorder using a variety of tools including entropy, the dynamical spectrum, Lyapunov exponents, recurrence phenomena, and mixing properties. While a system deemed to be disordered is ideally identified as so by each of these measures, this is not always the case. Indeed, systems exist which are disordered according to one measure, but well-ordered according to another. Subshifts arising from *random substitutions* are prototypical examples.

First studied in the context of mathematical quasicrystals by Godrèche and Luck [13], the random Fibonacci substitution has the curious property of giving rise to a dynamical system which is locally disordered but globally well-ordered, characterised by the simultaneous presence of positive topological entropy and a non-trivial pure-point component in its associated diffraction spectrum. Their results immediately encourage the further study of the dynamics of the random Fibonacci substitution and related systems. In this work, we consider one such dynamical property, namely topological mixing of the associated subshift and the associated tiling space.

Random substitutions are maps which send letters from a finite alphabet to finite collections of words over the same alphabet. They have received renewed attention in recent years, following the work of Rust and Spindeler [26], in which they initiated the study of random substitutions in the context of topological dynamics and ergodic theory and they posed numerous open questions and routes for further investigation. Since then, random substitution subshifts have been studied in terms of (among others) their frequency measures [18], diffraction spectrum [5,14], periodicity [25], automorphism groups [12], topological and measure theoretic entropy [15,16], and relations to shifts of finite type [17]. In their work, Rust and Spindeler determined by ad hoc methods that the subshift associated with the random period doubling substitution was not topologically mixing and they asked if a more general method could be established to determine

1

when an arbitrary random substitution gives rise to a topologically mixing subshift. We address that question here and give an answer for large families.

For deterministic substitutions, the classification of which substitution subshifts are topologically mixing is still incomplete. Dekking and Keane studied topological mixing for deterministic substitutions in the 70s [9] and provided the first example of a minimal system which is topologically strongly mixing of order two but not order three. Importantly, they also established that the presence of mixing depends not just on the abelianisation of the substitution, but also on the order of letters. The most recent attempt at providing a full classification is the work of Kenyon, Sadun and Solomyak [19], where they were able to determine the presence of mixing for large families of substitutions (their conditions do depend only on the abelianisation). While we do not address the same question (that is, in the deterministic setting), we use their results as a guide in our investigation of topological mixing of random substitutions, and as a tool in the study itself, capitalising on the close relationship between random substitutions and their deterministic counterparts.

In Section 2, we introduce the basic definitions of primitive and compatible random substitutions, as well as topological mixing in the context of subshifts over finite alphabets. We also outline our main results and compare them with the corresponding results on deterministic substitutions established in [19]. In Section 3, we prove our main results on determining topological mixing for primitive compatible random substitutions when the second largest eigenvalue of the substitution matrix is greater than 1 in modulus. In Section 4, we show that if a primitive compatible random substitution is irreducible Pisot, then its subshift is C-balanced and use this to prove that irreducible Pisot random substitutions are not topologically mixing as long as a mild recognisability condition is satisfied. In Section 5, we exhibit examples for each of our main theorems to illustrate their application. Significantly, we establish that the random Fibonacci substitution subshift is not topologically mixing; we weakened the recognisability condition enough in Section 4 for the main result to apply to this case (the random Fibonacci substitution does not satisfy full recognisability). Finally, in Section 6 we prove analogous theorems for $\mathbb{R}$-actions on random substitution tiling spaces.

2. Definitions and main results

2.1. **Random substitution subshifts.** Let $\mathcal{A} = \{a_1, a_2, \ldots, a_d\}$ be a finite alphabet. For a word $w = w_1 \cdots w_n$ over $\mathcal{A}$, let $|w| := n$ denote the length of w and let $\mathcal{A}^n$ denote the set of all words whose length is equal to n. Let $\mathcal{A}^+ = \bigcup_{n=1}^{\infty} \mathcal{A}^n$ be the set of all non-empty words over $\mathcal{A}$. Appending the empty word ϵ then yields the free monoid $\mathcal{A}^* = \mathcal{A}^+ \cup \{\epsilon\}$ of all finite words over $\mathcal{A}$ under concatenation. A *subword* of a word $w = w_1 \cdots w_n$ is a word $w_{[i,j]} = w_i \cdots w_j$ for some $1 \leq i \leq j \leq n$. The set $\mathcal{A}^{\mathbb{Z}}$ of all bi-infinite sequences over $\mathcal{A}$ forms a compact metrisable space under the product topology, where $\mathcal{A}$ is endowed with the discrete topology. The *shift map* $\sigma \colon \mathcal{A}^{\mathbb{Z}} \to \mathcal{A}^{\mathbb{Z}}$ given by $\sigma(x)_n = x_{n+1}$ is then a homeomorphism. A *subshift* $X \subseteq \mathcal{A}^{\mathbb{Z}}$ is a closed non-empty subspace of $\mathcal{A}^{\mathbb{Z}}$ that is invariant under the shift action σ. That is, $\sigma(X) = X$. The *language* of a subshift $\mathcal{L}(X)$ is the set of all subwords of elements of X,

$$\mathcal{L}(X) = \{w \in \mathcal{A}^* \mid w \text{ is a finite subword of some } x \in X\}.$$

A *deterministic substitution* θ on $\mathcal{A}$ is a map $\theta \colon \mathcal{A} \to \mathcal{A}^+$ and uniquely extends to a map $\theta \colon \mathcal{A}^+ \to \mathcal{A}^+$ by concatenation $\theta(w_1 \cdots w_n) := \theta(w_1) \cdots \theta(w_n)$. The action of a deterministic substitution on a bi-infinite element $x \in \mathcal{A}^{\mathbb{Z}}$ is defined analogously, where the only ambiguity

is where to place the origin. The usual convention is for $\theta(x)_0$ to be the first letter of $\theta(x_0)$. Deterministic substitutions are well-studied [4, 11]. We are interested in a generalisation which assigns a finite set of words, rather than only a single word, to each letter of the alphabet.

Definition 1. Let $\mathcal{A}$ be a finite alphabet, and let $\mathcal{P}(\mathcal{A}^+)$ denote the power set of $\mathcal{A}^+$. A *random substitution* is a map $\vartheta\colon \mathcal{A} \to \mathcal{P}(\mathcal{A}^+)\backslash\varnothing$ such that, for each $a \in \mathcal{A}$, $\vartheta(a)$ is a non-empty finite set. We can extend ϑ to a function $\mathcal{A}^+ \to \mathcal{P}(\mathcal{A}^+)\backslash\varnothing$ by concatenation

$$\vartheta(v_1\cdots v_n) = \vartheta(v_1)\cdots\vartheta(v_n) \coloneqq \{w_1\cdots w_n \mid w_i \in \vartheta(v_i),\ 1 \le i \le n\}.$$

We then extend ϑ to a function $\vartheta\colon \mathcal{P}(\mathcal{A}^+)\backslash\varnothing \to \mathcal{P}(\mathcal{A}^+)\backslash\varnothing$ by $\vartheta(B) \coloneqq \bigcup_{w \in B} \vartheta(w)$. Consequently, we can now take powers of ϑ by composition and so $\vartheta^k\colon \mathcal{P}(\mathcal{A}^+)\backslash\varnothing \to \mathcal{P}(\mathcal{A}^+)\backslash\varnothing$ is defined for any $k \ge 0$, where $\vartheta^0 \coloneqq \mathrm{Id}_{\mathcal{P}(\mathcal{A}^+)\backslash\varnothing}$ is the identity and $\vartheta^{k+1} \coloneqq \vartheta \circ \vartheta^k$. An element of $\vartheta^k(a)$, where $a \in \mathcal{A}$, is called a *super-word* of degree k.

Definition 2. A word $w \in \mathcal{A}^+$ is called a *realisation of* ϑ on a word $v \in \mathcal{A}^+$ if $w \in \vartheta(v)$. A word w that is a realisation of ϑ on a legal word $v \in \mathcal{L}_\vartheta$ is called an *inflation word*. Similarly, a bi-infinite sequence $x \in \mathcal{A}^{\mathbb{Z}}$ is a *realisation* of ϑ on a bi-infinite sequence $y \in \mathcal{A}^{\mathbb{Z}}$ if $x \in \vartheta(y)$.

We say that ϑ has *constant length* ℓ if all of the super-words of degree 1 have the same length $\ell \ge 1$. A deterministic substitution θ is called a *marginal* of a random substitution ϑ if $\theta(a) \in \vartheta(a)$ for all $a \in \mathcal{A}$. The random substitution ϑ is then said to be a *local mixture* of its marginals $\{\theta_i\}_{i \in I}$. We say that a word $w \in \mathcal{A}^+$ is ϑ-*legal* if there is a natural number k such that w is a subword of some super-word of degree k.

Example 3. The *random Fibonacci substitution* is given by $\vartheta\colon a \mapsto \{ab, ba\}, b \mapsto \{a\}$ and was first introduced by Godréche and Luck [13]. Note that aa is a subword of $baa \in \vartheta(ab) \subset \vartheta^2(a)$, while bb is a subword of $abba \in \vartheta(aa)$, and so is a subword of an element of $\vartheta^3(a)$. This makes both aa and bb ϑ-legal. However, we will see that bbb and $aaaaa$ are not ϑ-legal.

The marginals of the random Fibonacci substitution ϑ are given by the deterministic Fibonacci substitution $\theta_1\colon a \mapsto ab, b \mapsto a$ and its reflection $\theta_2\colon a \mapsto ba, b \mapsto a$. While the word bb is ϑ-legal, it is neither θ_1-legal nor θ_2-legal, since the only way for bb to be part of an inflation word is for the first b to be part of $\theta_1(a) = ab$ and for the second b to be part of $\theta_2(a) = ba$.

Let $\Phi\colon \mathcal{A}^* \to \mathbb{N}_0^d\colon w \mapsto (|w|_{a_1}, |w|_{a_2}, \ldots, |w|_{a_d})^T$ denote the *abelianisation function*, where $|w|_a$ denotes the number of occurrences of a in w. That is, Φ takes a word $w \in \mathcal{A}^*$ and enumerates the number of occurrences of each letter in w.

Definition 4. Let ϑ be a random substitution on $\mathcal{A} = \{a_1, a_2, \ldots, a_d\}$ and let $\{\theta_i\}_{i \in I}$ denote the set of marginals of ϑ. We say that ϑ is *compatible* if the abelianisations of its marginals coincide. That is, for each $a \in \mathcal{A}$, all words in $\vartheta(a)$ have the same abelianisation. In that case, the *substitution matrix* M_ϑ of ϑ is given by $(M_\vartheta)_{ij} \coloneqq |\vartheta(a_j)|_{a_i}$ for all $1 \le i, j \le d$. That is, the common abelianisation of $\vartheta(a_i)$ determines the ith column of M (not the ith row).

A matrix M is *primitive* if there exists a power p such that all entries of M^p are positive. We say that the compatible random substitution ϑ is primitive if M_ϑ is primitive. In that case, each super-word of degree p contains at least one copy of each letter. If ϑ is primitive, we let $\lambda_1 \coloneqq \lambda_{PF}$ denote the Perron–Frobenius (PF) eigenvalue of M_ϑ, the unique largest eigenvalue. A compatible random substitution is *Pisot* if λ_1 is a Pisot number—an algebraic integer greater than 1, all of whose algebraic conjugates lie in the open unit disk. If the characteristic polynomial of M_ϑ is irreducible, then we likewise call ϑ irreducible.

When a compatible primitive random substitution ϑ is defined on a two-letter alphabet, we refer to the PF-eigenvalue λ_1 as the *first eigenvalue* of M_ϑ and to the other eigenvalue λ_2 as the *second eigenvalue*.

Example 5. The random Fibonacci substitution $\vartheta \colon a \mapsto \{ab, ba\}, b \mapsto \{a\}$ is compatible as $\Phi(ab) = \Phi(ba) = (1,1)^T$ and $\Phi(b) = (1,0)^T$. The corresponding substitution matrix is $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix} \right)$ with irreducible characteristic polynomial $\lambda^2 - \lambda - 1$, first eigenvalue $\lambda_1 = \frac{1+\sqrt{5}}{2}$ (the golden ratio), and second eigenvalue $\lambda_2 = \frac{1-\sqrt{5}}{2}$. Since $|\lambda_2| < 1$, λ_1 is a Pisot number. Thus the random Fibonacci substitution is irreducible Pisot.

Definition 6. The *language of $\mathcal{L}_\vartheta$* of a random substitution ϑ on $\mathcal{A}$ is

$$\mathcal{L}_\vartheta = \{ w \in \mathcal{A}^* \mid w \text{ is } \vartheta\text{-legal} \}.$$

The set of length-n legal words for ϑ is denoted $\mathcal{L}_\vartheta^n := \mathcal{L}_\vartheta \cap \mathcal{A}^n$. The *random substitution subshift* of ϑ (RS-subshift) is

$$X_\vartheta := \left\{ x \in \mathcal{A}^{\mathbb{Z}} \mid w \text{ is a finite subword of } x \Rightarrow w \in \mathcal{L}_\vartheta \right\}.$$

It is easy to verify that X_ϑ is a closed, shift-invariant subspace of the full shift $\mathcal{A}^{\mathbb{Z}}$, so X_ϑ is a subshift. A primitive compatible random substitution gives rise to a non-empty RS-subshift, since each of its marginals already gives rise to a non-empty subshift.

A subspace $Y \subset X_\vartheta$ is called *substitutive* if there exists a primitive deterministic substitution θ such that $Y = X_\theta$. Many results relating to the dynamics and topology of primitive RS-subshifts were presented in the works of Gohlke, Rust and Spindeler [17, 26] including the following results that will be useful later.

Theorem 7 ([17, 26])**.** *Let ϑ be a primitive random substitution with a non-empty RS-subshift X_ϑ. Then:*

 (a) X_ϑ contains an element with a dense shift-orbit;
 (b) either X_ϑ is substitutive or there are infinitely many distinct substitutive subspaces of X_ϑ;
 (c) the union of the substitutive subspaces of X_ϑ is dense in X_ϑ.

2.2. **Mixing of substitution subshifts.** In the remainder of this section, we outline the main results of our work. First, we introduce the property in which we are principally interested.

Definition 8. A dynamical system (X, T) is said to be *topologically mixing* if for any two non-empty open subsets $U, V \subset X$, there exists a natural number N such that for all natural numbers $n \geq N$, $T^n(U) \cap V \neq \varnothing$.

If $v = v_1 \cdots v_m$ is a word of length m and if n is an integer, the *cylinder set* $[v]_n$ is the set of $x \in X_\vartheta$ such that $x_{i+n-1} = v_i$ for $i = 1, \ldots, m$. That is, it is the set of bi-infinite words that contain the word v starting at position n. These cylinder sets form a basis for the topology of a subshift X. If $\sigma \colon X \to X$ is the shift map, then for large n, the statement that $T^n([u]_{n_1}) \cap [v]_{n_2} \neq \varnothing$ is equivalent to the existence of a word $uwv \in \mathcal{L}(X)$ where uw has length $n + n_2 - n_1$. This allows us to recast topological mixing for subshifts in purely combinatorial terms.

Definition 9. A subshift (X, σ) is said to be *topologically mixing* if for any two words $u, v \in \mathcal{L}(X)$, there exists a natural number N such that for all natural numbers $n \geq N$, there exists a word w of length n such that $uwv \in \mathcal{L}(X)$.

Topological mixing is preserved under taking factor maps. That is, if $f\colon X \to Y$ is a factor map of dynamical systems and X is mixing, then Y is also mixing. In particular, topological mixing is preserved under topological conjugacy. It is well-known [20] that a shift of finite type X_A is topologically mixing if and only if A is a primitive matrix, where A is the adjacency matrix of a directed graph G and X_A is the vertex-shift on G. It has been shown by Gohlke, Rust and Spindeler [17] that every topologically transitive shift of finite type can be realised up to topological conjugacy as a primitive (though often non-compatible) RS-subshift. This then gives us a class of RS-subshifts that are topologically mixing, including the full shift $\mathcal{A}^{\mathbb{Z}}$ and the golden mean shift X_A with $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. There also exist RS-subshifts that are not topologically mixing such as the RS-subshift associated with the random period doubling substitution $\vartheta_{PD}\colon a \mapsto \{ab, ba\}, b \mapsto \{aa\}$ [26].

Recall that a subshift X is aperiodic if X contains no periodic orbits. We say that a random substitution ϑ is aperiodic if its associated RS-subshift X_ϑ is aperiodic (and similarly for deterministic substitutions). Our main results generalise those appearing in the work of Kenyon, Sadun and Solomyak [19] to the random setting.

Theorem 10 ([19, Prop. 1.1])**.** *Let θ be a primitive aperiodic deterministic substitution on an alphabet $\mathcal{A}$. If the subshift X_θ is topologically mixing, then $\gcd\{|\theta^n(a)| : a \in \mathcal{A}\} = 1$ for all $n \geq 1$.*

When $|\lambda_2| > 1$ and the alphabet only consists of two letters, they were able to show that this criterion is also sufficient for topological mixing.

Theorem 11 ([19, Thm. 1.2])**.** *Let θ be a primitive aperiodic deterministic substitution defined on a two-letter alphabet $\mathcal{A}$ with second eigenvalue λ_2 greater than 1 in modulus. The subshift X_θ is topologically mixing if and only if $\gcd\{|\theta^n(a)| : a \in \mathcal{A}\} = 1$ for all $n \geq 1$.*

Our first goal is to extend Theorem 10 to the random case. With this in mind, we introduce an additional condition in Section 3.1 referred to as *local recognisability*, which means that legal words of the RS-subshift can be uniquely *desubstituted* given a sufficiently large legal neighbourhood of the word. This condition is an analogue of local recognisability for deterministic substitutions which is equivalent to aperiodicity in the primitive deterministic setting, thanks to a celebrated result of Mossé [23]. Unfortunately, no such result is known for random substitutions, and the situation must necessarily be more complicated as, in contrast to the deterministic case, recognisability does not follow from aperiodicity. A classic example is given by the random Fibonacci substitution whose RS-subshift is aperiodic because its PF-eigenvalue $\lambda_1 = \frac{1+\sqrt{5}}{2}$ is not an integer, but it is not recognisable [25]. It is known that the set of periodic points of a primitive RS-subshift is either empty (hence, aperiodic) or dense [26].

By assuming local recognisability, which holds automatically for aperiodic deterministic substitutions, we are able to generalise Theorem 10 to the random setting.

Theorem 12. *Let ϑ be a primitive compatible random substitution on an alphabet $\mathcal{A}$ that is locally recognisable. If the RS-subshift X_ϑ is topologically mixing, then $\gcd\{|\vartheta^n(a)| : a \in \mathcal{A}\} = 1$ for all $n \geq 1$.*

Our second main result is an extension of Theorem 11 to the random case. To prove the necessary direction, we consider the substitutive subspaces of a given RS-subshift associated with the marginals of its powers. In order to apply Theorem 11 directly, we need to show that

each of these substitutive subspaces is aperiodic. We do this by providing a total classification of periodicity for primitive substitutions on two letters in Proposition 26. (To the best of our knowledge, this classification is new.)

Theorem 13. *Let ϑ be a primitive compatible random substitution on a two-letter alphabet $\mathcal{A}$ that is locally recognisable with second eigenvalue λ_2 greater than 1 in modulus. The RS-subshift X_ϑ is topologically mixing if and only if $\gcd\{|\vartheta^n(a)| : a \in \mathcal{A}\} = 1$ for all $n \geq 1$.*

For a primitive compatible random substitution ϑ of constant length $\ell > 1$, by compatibility we have $\gcd\{|\vartheta^n(a)| : a \in \mathcal{A}\} = \ell^n$ for all $n \geq 1$. The following then follows from Theorem 12.

Corollary 14. *Let ϑ be a primitive compatible constant-length random substitution on an alphabet $\mathcal{A}$ that is locally recognisable. The RS-subshift X_ϑ is not topologically mixing.*

Note however that even if the PF-eigenvalue of a random substitution is an integer, the substitution may not be constant length (see Example 37). In the case that $|\lambda_2| > 1$, we can extend the above corollary to the case of integral PF-eigenvalue using the following short argument.

Proposition 15. *Let ϑ be a primitive compatible random substitution on a two-letter alphabet that is locally recognisable with integer first eigenvalue λ_1 and second eigenvalue λ_2 greater than 1 in modulus. Then the RS-subshift X_ϑ is not topologically mixing.*

Proof. Let θ be a marginal of ϑ. By compatibility, λ_1 is also the first eigenvalue of θ. By a classical argument [8, Theorem 1], there exists a primitive deterministic substitution θ_{cl} of constant length $\ell = \lambda_1$ such that X_θ and $X_{\theta_{cl}}$ are topologically conjugate. Since ϑ is locally recognisable, the marginal θ is also locally recognisable, and so is aperiodic by Mossé's theorem [23]. Theorem 10 then states that $X_{\theta_{cl}}$ is not topologically mixing, insofar as θ_{cl} is constant-length. Thus, the subshift X_θ is also not topologically mixing. By Theorem 11, there must exist some $n \geq 1$ such that $\gcd\{|\theta^n(a)| : a \in \mathcal{A}\} > 1$. But then $\gcd\{|\vartheta^n(a)| : a \in \mathcal{A}\} = \gcd\{|\theta^n(a)| : a \in \mathcal{A}\} > 1$, so by Theorem 13, X_ϑ is not topologically mixing. $\square$

Our next main result concerns mixing properties of Pisot random substitutions. First, we show in Proposition 33 that the irreducible Pisot property implies C-balancedness. That is, the number of occurrences of a letter in a legal word does not differ from the statistically expected proportion by more than a uniformly bounded value.

Some well-known Pisot random substitutions such as the random Fibonacci substitution fail to be locally recognisable. However, they often satisfy a weaker condition which we call "admitting recognisable words at all levels". For a Pisot random substitution ϑ, this is enough to imply that X_ϑ is not topologically mixing via the following theorem:

Theorem 16. *Let ϑ be a C-balanced random substitution. There exists a constant N such that if ϑ admits a level-n recognisable word for some $n \geq N$, then the RS-subshift X_ϑ is not topologically mixing.*

In particular, we are able to show that the random Fibonacci substitution ϑ admits recognisable words at all levels (Example 38), so X_ϑ is not topologically mixing. It should be noted that in the deterministic setting, Theorem 16 is automatic in the Pisot case thanks to the fact that, for minimal system, topological mixing implies topological weak mixing. This is not necessarily

the case for non-minimal systems such as RS-subshifts, and so we require a direct proof of non-mixing.

Finally, we turn our attention to tilings associated with our subshifts. To each letter a_i we associate a labelled interval (tile) of length $\ell(a_i) = \ell_i$ and to each (possibly infinite) word we associate a concatenation of the tiles corresponding to the letters. Instead of considering a $\mathbb{Z}$-action on a space of bi-infinite words, we consider an $\mathbb{R}$-action on a space of tilings. Remarkably, the theorems for topological mixing of random substitution tilings are nearly identical to those of random substitution subshifts, and so are not repeated here, only with the condition that $\gcd\{|\vartheta^n(a)| : a \in \mathcal{A}\} = 1$ for all $n \geq 1$ replaced by the existence of two (or more) tiles whose lengths have an irrational ratio.

3. Topological mixing for general primitive random substitutions

Let ϑ be a primitive random substitution on a finite alphabet $\mathcal{A}$ and let u be a ϑ-legal word. By the definition of ϑ-legality, there is a natural number k such that u is a subword of a particular realisation $w_a \in \vartheta^k(a)$ for some letter $a \in \mathcal{A}$. For each $b \neq a$, pick such a word $w_b \in \vartheta^k(b)$ arbitrarily. Now, define the deterministic substitution $\theta_u \colon \mathcal{A} \to \mathcal{A}^+$ by $\theta_u(a) = w_a$ and $\theta_u(b) = w_b$ for each $b \in \mathcal{A}\backslash\{a\}$. Clearly, θ_u is a primitive deterministic substitution containing u as a θ_u-legal word whose associated subshift X_{θ_u} is contained in X_ϑ. Furthermore, the substitution matrix of θ_u is M_ϑ^k, which is primitive.

This construction is taken from [17] where it was used to prove Theorem 7(c). If a substitution θ is a marginal of a power of ϑ, we call the subspace $X_\theta \subset X_\vartheta$ a *basic subspace* of ϑ. All basic subspaces are substitutive by construction and each X_{θ_u} is a basic subspace. The union of the basic subspaces of ϑ is then dense in X_ϑ since, by the construction of the previous paragraph, every ϑ-legal word u appears in the language of the basic subspace X_{θ_u}. The basic subspaces of X_ϑ will play an important role in allowing us to convert mixing properties of deterministic substitutions to random substitutions.

3.1. **A sufficient condition.** To prove Theorem 12, we need to introduce the concept of local recognisability. Roughly speaking, this means that if x and y are bi-infinite words and $x \in \vartheta(y)$, then we can deduce the value of y_0 from the restriction of x to a region of fixed size around the origin. To make this precise, we must first introduce the notions of inflation word decompositions and induced inflation word decompositions. We then prove intermediate results concerning these constructions that will be used in the proof of Theorem 12.

Definition 17 ([12])**.** Let ϑ be a random substitution and let $u \in \mathcal{L}_\vartheta$ be a legal word. Let $n \geq 1$ be a natural number. For words $u_i \in \mathcal{A}^+$, the tuple $[u_1, \ldots, u_\ell]$ is called a ϑ^n-*cutting* of u if $u_1 \cdots u_\ell = u$ and there exists a ϑ-legal word $v = v_1 \cdots v_\ell$ such that

- For $i = 2, \ldots, \ell - 1$, u_i is a super-word of level n built from the letter v_i. That is, $u_i \in \vartheta^n(v_i)$. Note that u_i is a word, while v_i is a single letter.
- u_1 is the suffix of a super-word of level n built from v_1, and
- u_ℓ is the prefix of a super-word of level n built from v_ℓ.

That is, u is contained in a realisation of $\vartheta^n(v)$, which is a concatenation of n-super-words, with each of the interior u_i's being one of those n-super-words.

We call v a *root* of the ϑ^n-cutting and we call $([u_1, \ldots, u_\ell], v)$ the corresponding *level-n inflation word decomposition* of u. If u actually is a realisation of $\vartheta^n(v)$, so $u_1 \in \vartheta^n(v_1)$ and

$u_\ell \in \vartheta^n(v_\ell)$, then we say that u is an *exact* level-n inflation word. Finally, we let $D_{\vartheta^n}(u)$ denote the set of all level-n decompositions of the ϑ-legal word u.

Example 18. Let $\vartheta \colon a \mapsto \{ab, ba\}, b \mapsto \{a\}$ be the random Fibonacci substitution. The word aab has two possible ϑ-cuttings $[a, ab]$ and $[a, a, b]$ with each having two distinct associated roots. The set of level-1 inflation word decompositions of aab is

$$D_\vartheta(aab) = \{([a, ab], ba), ([a, ab], aa), ([a, a, b], bba), ([a, a, b], aba)\}.$$

In this example, $abba$ is an exact level-1 inflation word and $abbaaaabbaaaabba$ is an exact level-1, -2, -3, and -4 inflation word. However, bb is not an exact (level-1) inflation word, since any concatenation of 1-super-words containing bb must also contain some a's.

Note that having a unique ϑ-cutting does not lead to the uniqueness of roots. For example, consider $\vartheta^2 \colon a \mapsto \{aba, baa, aab\}, b \mapsto \{ab, ba\}$. The word bb has a unique ϑ^2-cutting $[b, b]$, but this ϑ^2-cutting can come from four distinct inflation word decompositions:

$$D_{\vartheta^2}(bb) = \{([b, b], aa), ([b, b], ab), ([b, b], ba), ([b, b], bb)\}.$$

Similarly, having a unique root does not mean that one has a unique ϑ-cutting. For example, under the random period doubling substitution $\varrho \colon a \mapsto \{ab, ba\}, b \mapsto \{aa\}$, the word bab can only come from the legal word aa but it has two possible ϑ-cuttings; specifically, $D_\varrho(bab) = \{([ba, b], aa), ([b, ab], aa)\}$.

Let u be a ϑ-legal word and let $u_{[i,j]}$ be a subword. An inflation word decomposition of u restricts to an inflation word decomposition of $u_{[i,j]}$, which we call an *induced inflation word decomposition*. The idea is simple, but the precise definition is somewhat technical:

Definition 19 ([12]). Let ϑ be a random substitution and let $d = ([u_1, \ldots, u_\ell], v) \in D_\vartheta(u)$ be an inflation word decomposition of u. For $1 \le i \le j \le |u| - 1$, we write $d_{[i,j]}$ for the *induced inflation word decomposition on the subword* $u_{[i,j]}$, defined by

$$d_{[i,j]} = ([u_1, \ldots, u_\ell], v)_{[i,j]} = \left([\hat{u}_{k(i)}, u_{k(i)+1}, \ldots, u_{k(j)-1}, \hat{u}_{k(j)}], v_{[k(i),k(j)]} \right),$$

where $1 \le k(i) \le k(j) \le \ell$ are natural numbers such that

$$\left| u_1 \cdots u_{k(i)-1} \right| < i \le \left| u_1 \cdots u_{k(i)} \right| \text{ and } \left| u_1 \cdots u_{k(j)} \right| \le j < \left| u_1 \cdots u_{k(j)+1} \right|,$$

$\hat{u}_{k(i)}$ is a suffix of $u_{k(i)}$ and $\hat{u}_{k(j)}$ is a prefix of $u_{k(j)}$ such that

$$\hat{u}_{k(i)} u_{k(i)+1} \cdots u_{k(j)-1} \hat{u}_{k(j)} = u_{[i,j]}.$$

Example 20. Let $\vartheta \colon a \mapsto \{ab, ba\}, b \mapsto \{a\}$ be the random Fibonacci substitution. The legal word $u = ababa \in \mathcal{L}_\vartheta$ has exactly five level-1 inflation word decompositions given by

$$D_\vartheta(u) = \{([a, ba, ba], baa), ([a, ba, ba], aaa), ([ab, ab, a], aab), ([ab, ab, a], aaa), ([ab, a, ba], aba)\}.$$

For the subword $u_{[2,4]} = bab$ of u, the first two elements of $D_\vartheta(u)$ yield the induced decomposition $d_{[2,4]}^{(1)} = ([ba, b], aa)$, the next two elements yield $d_{[2,4]}^{(2)} = ([b, ab], aa)$, while the last element yields $d_{[2,4]}^{(3)} = ([b, a, b], aba)$. Thus, $\# \left\{ d_{[2,4]} \mid d \in D_\vartheta(u) \right\} = 3$. In this example, all three possible ϑ-cuttings of bab are induced from cuttings of u.

By contrast, the word $u' = bbaba$ has a unique inflation word decomposition $d' = ([b, ba, ba], aaa)$ which yields a unique induced inflation word decomposition on the subword $u'_{[2,4]} = bab$ given by $d'_{[2,4]} = ([ba, b], aa)$. That is, when bab sits inside u, the embedding tells us nothing about the

8

ϑ-cutting of bab, but when bab sits inside u', the embedding uniquely defines the ϑ-cutting of bab.

Definition 21 ([12]). Let ϑ be a random substitution and let $u \in \mathcal{L}_\vartheta$ be a legal word. We say that u is *recognisable* if there exists a natural number N such that for each legal word of the form $w = u^{(l)}uu^{(r)}$ with $\left|u^{(l)}\right| = \left|u^{(r)}\right| = N$, all inflation word decompositions of w induce the same inflation word decomposition of u. That is, knowing the N letters to the left of u and the N letters to the right of u determines a unique induced inflation word decomposition of u. We call the minimum such N the *radius of recognisability* for u. If u is recognisable with respect to the nth power ϑ^n of the random substitution, then we say that u is *level-n recognisable* with respect to ϑ.

Example 22. Let $\vartheta \colon a \mapsto \{ab, ba\}, b \mapsto \{a\}$ be the random Fibonacci substitution with marginals θ_1 and θ_2 such that $\theta_1(a) = ab$ and $\theta_2(a) = ba$. The word $w = abbaaaabbaaaabba$ is a level-1, -2, -3, and -4 recognisable word with radius 0 at each level, as it has the following sets of unique inflation word decompositions

$$
\begin{aligned}
D_\vartheta(w) &= \{([ab, ba, a, a, ab, ba, a, a, ab, ba], aabbaabbaa)\}, \\
D_{\vartheta^2}(w) &= \{([ab, baa, aab, baa, aab, ba], baaaab)\}, \\
D_{\vartheta^3}(w) &= \{([abbaa, aab, baa, aabba], abba)\}, \\
D_{\vartheta^4}(w) &= \{([abbaaaab, baaaabba], aa)\}.
\end{aligned}
$$

By contrast, the word bab turns out not to be recognisable at level 1 (or at any level, actually), as there exist arbitrarily long legal extensions of bab that admit both an inflation word decomposition that induces the decomposition $([ba, b], aa)$ of bab, and another decomposition that induces $([b, ab], aa)$. Specifically, if n is an odd natural number, let $u = \theta_1^n(aa)$ with the first letter removed, which is the same as $\theta_2^n(aa)$ with the last letter removed. The middle three letters of u are bab. The inflation word decomposition of u that comes from applying θ_2 to $\theta_2^{n-1}(aa)$ induces the decomposition $([ba, b], aa)$, while the decomposition of u that comes from applying θ_1 to $\theta_1^{n-1}(aa)$ induces $([b, ab], aa)$.

Although the random Fibonacci substitution has ϑ-legal words like bab that are not recognisable, for each natural number n there also exist ϑ-legal words that are recognisable at level n. This fact, which is sufficient to establish many mixing properties of X_ϑ, will be proven in Section 5.

Definition 23. We call a random substitution ϑ *locally recognisable* if there exists a natural number N such that every ϑ-legal word is recognisable with radius at most N. The minimum such N is called the *radius of recognisability for ϑ*. Similarly, for any natural number $n \geq 1$, if ϑ^n is locally recognisable, we denote by N_{ϑ^n} its radius of recognisability.

Remark 24. As in the deterministic setting, there is a notion of global recognisability for random substitutions, first appearing in the work of Rust [25]. A random substitution ϑ is globally recognisable if for each bi-infinite word $x \in X_\vartheta$ there is a unique preimage $y \in X_\vartheta$ and a unique integer n in a fixed bounded range such that $\sigma^n(x) \in \vartheta(y)$. Global and local recognisability are in fact equivalent, but in practice it is usually much easier to check local recognisability than global.

By a routine inductive argument, or by invoking the equivalence with global recognisability, one can show that local recognisability for compatible random substitutions is preserved under taking powers.

Proposition 25. *A compatible random substitution ϑ is locally recognisable if and only if ϑ^n is locally recognisable for all $n \geq 1$.* $\qquad\square$

Unfortunately, local recognisability of a random substitution is a rather strong condition; many examples, such as the random Fibonacci substitution, are not locally recognisable. However, for many proofs it is enough that a substitution admits recognisable words at all levels. We will return to this concept when we address the proof of Theorem 16.

We are now in a position to prove the first of our main results, Theorem 12. For a random substitution ϑ, we write $|\vartheta| := \max\{|u| \mid u \in \vartheta(\mathcal{A})\}$.

Proof of Theorem 12. We prove the contrapositive. Suppose that $\gcd\{|\vartheta^n(a)| : a \in \mathcal{A}\} = p > 1$ for some natural number $n \geq 1$. As ϑ is locally recognisable, ϑ^n is also locally recognisable by Proposition 25. So, without loss of generality, we can assume that $n = 1$. Let N be the radius of recognisability for ϑ, let u be a ϑ-legal word of length $|u| > 2N + 2|\vartheta|$, and let u' be the subword of u obtained by deleting the first N and last N letters of u. By local recognisability, every inflation word decomposition of u induces the same inflation word decomposition of u'. Since $|u'| > 2|\vartheta|$, this unique decomposition contains a complete 1-super-word u'' in a fixed position relative to u' (and hence relative to u).

Now suppose that uwu is a ϑ-legal word and consider an inflation word decomposition of uwu. This induces inflation word decompositions of the first and second u, giving rise to appearances of u'' spaced exactly $|u| + |w|$ apart. However, the length of every 1-super-word is a multiple of p, so $|u| + |w|$ must be a multiple of p. Since $p > 1$, there are arbitrarily large values of $|w|$ that are impossible, so X_ϑ is not topologically mixing. $\qquad\square$

3.2. **Substitutions on two letters.** In order to apply Theorem 10, it will be helpful to classify all two-letter primitive substitutions whose subshifts are periodic. A fully detailed proof appears in the thesis of the fourth author [28] and so we only present a sketch of the proof here. The constant length version of this result appears in work of Baake, Coons and Mañibo [3]. Our proof for the general case is similar in spirit to theirs. In particular, we capitalise on the absence of asymptotic pairs for periodic subshifts.

Theorem 26. *Let θ be a primitive deterministic substitution on $\mathcal{A} = \{a, b\}$. X_θ is periodic if and only if θ takes one of the following forms $\{a \mapsto u^k, b \mapsto u^\ell\}$, $\{a \mapsto (ab)^k a, b \mapsto (ba)^\ell b\}$, or $\{a \mapsto (ba)^k b, b \mapsto (ab)^\ell a\}$ where u is a non-empty finite word and $k, \ell \geq 0$.*

Proof. It is an easy exercise to show that the three above forms lead to periodic subshifts. The non-trivial step is in showing that no other form of substitution on two-letters gives rise to a periodic subshift.

Let θ be a primitive periodic deterministic substitution. Suppose that θ is not of the form $a \mapsto (ab)^k a, b \mapsto (ba)^\ell b$ or $a \mapsto (ba)^k b, b \mapsto (ab)^l a$. It follows then that either aa or bb is a θ-legal word. Without loss of generality, assume that aa is legal. By periodicity and primitivity of θ, we must then have $aa, ab, ba \in \mathcal{L}_\theta$.

We first put the substitution into a standard form such that the leftmost letters of $\theta(a)$ and $\theta(b)$ are different. To do this, we take 'conjugates' of the substitution. That is, if $\theta(a) = xv$ and $\theta(b) = xw$ for $x \in \{a, b\}$, then we replace the substitution with the conjugate substitution $\theta' \colon a \mapsto vx, b \mapsto wx$. It is easy to see that conjugating a primitive substitution leaves the associated subshifts fixed, so $X_\theta = X_{\theta'}$. We may repeat the conjugation process. Either the iteration never stops, in which case θ is of the form $a \mapsto u^k, b \mapsto u^\ell$, or we eventually reach a

conjugate $\theta^{(n)}\colon a \mapsto xv_a, b \mapsto yv_b$ with $x \neq y$. By squaring if necessary, we can assume that $x = a$ and $y = b$. This new substitution $\theta\colon a \mapsto au_a s, b \mapsto bu_b t$ for $s, t \in \{a, b\}$ is then the standard form we may assume our substitution to be in.

Recall that $x, y \in X$ are called an *asymptotic pair* if $x \neq y$ and either $d(\sigma^n(x), \sigma^n(y)) \to 0$ or $d(\sigma^{-n}(x), \sigma^{-n}(y)) \to 0$ as $n \to \infty$. Periodic subshifts necessarily admit no asymptotic pairs. We now consider the possible cases for the letters s, t. In each case, the existence of an asymptotic pair $x, y \in X_\theta$ will allow us to conclude that X_θ is aperiodic and hence reach a contradiction. For the case $s = a, t = b$, the existence of $a.a$ as a legal seed for producing an asymptotic pair of fixed points for the substitution is essential. This is why we had to conclude that $aa, ab, ba \in \mathcal{L}_\theta$ earlier.

We will show this case. The other cases are similar (possibly with different seeds).

Case: $s = a$, $t = b$. The substitution admits two (distinct) fixed points (because $a.a$ and $a.b$ are both legal seeds)

$$\cdots \theta^2(au_a)\, \theta(au_a)\, au_a\, a \, . \, a\, u_a a\, \theta(u_a a)\, \theta^2(u_a a) \cdots$$
$$\cdots \theta^2(au_a)\, \theta(au_a)\, au_a\, a \, . \, b\, u_b b\, \theta(u_b b)\, \theta^2(u_b b) \cdots$$

which are elements in X_θ that are right-asymptotic to one another. $\qquad \square$

Corollary 27. *Let θ be a primitive periodic deterministic substitution on $\mathcal{A} = \{a, b\}$. The second eigenvalue λ_2 of the substitution matrix M_θ is either $0, 1,$ or -1.*

Proof. By Theorem 26, we only need to consider the three possible forms that θ can take. If θ is of the form $a \mapsto u^k, b \mapsto u^\ell$, then M_θ has rank one, with $\lambda_2 = 0$. For the other forms, the substitution matrix has $(1, -1)$ as a left-eigenvector with eigenvalue ± 1. $\qquad \square$

Lemma 28. *Let ϑ be a primitive compatible random substitution defined on $\mathcal{A} = \{a, b\}$ with $|\lambda_2| > 1$. If $\gcd\{|\vartheta^n(a)|, |\vartheta^n(b)|\} = 1$ for all $n \geq 1$ then the basic subspaces of ϑ are topologically mixing.*

Proof. Let X_θ be a basic subspace of ϑ. So, $\theta(a) \in \vartheta^k(a)$ and $\theta(b) \in \vartheta^k(b)$ for some natural number $k \geq 1$. The second eigenvalue of θ then has modulus $|\lambda_2^k| > 1$. In particular, this eigenvalue is not 0 or ± 1, so X_θ is non-periodic. By the compatibility of ϑ, we have

$$\gcd\left\{|\theta^n(a)|, |\theta^n(b)|\right\} = \gcd\left\{|\vartheta^{nk}(a)|, |\vartheta^{nk}(b)|\right\} = 1$$

for all $n \geq 1$. By Theorem 10, X_θ is then topologically mixing. $\qquad \square$

We now prove Theorem 13, the generalisation of Theorem 11 to the random setting.

Proof of Theorem 13. One direction is immediate from Theorem 12. So, it remains to show that if $\gcd\{|\vartheta^n(a)|, |\vartheta^n(b)|\} = 1$ for all $n \geq 1$, then X_ϑ is topologically mixing.

Let u and v be arbitrary legal words. There then exist natural numbers k_u and k_v such that u is a subword of a legal k_u-super-word and v is a subword of a legal k_v-super-word. Pick k bigger than the larger of k_u and k_v. If $a_i a_j$ is a legal 2-letter word then, by primitivity, there is an element of $\vartheta^k(a_i a_j)$ that contains u in the first k-super-word and v in the second. Thus there is a ϑ-legal word of the form uwv.

Let $\theta := \theta_{uwv}$ be a marginal of a power of ϑ such that uwv is θ-legal. By Lemma 28, X_θ is topologically mixing. Thus there exists a natural number $N > 0$ such that for all $n \geq N$, there is a word w_n of length n such that $uw_n v$ is θ-legal, and hence ϑ-legal. As u and v were chosen arbitrarily, it follows that X_ϑ is topologically mixing. $\qquad \square$

4. Pisot random substitutions

4.1. *C*-Balancedness. The concept of 'balancedness' was first introduced by Morse and Hedlund [22] to provide one of the many classifications of Sturmian sequences [11]. Here, we consider the more general notion of *C-balanced* sequences, also called sequences of *bounded discrepancy*. For primitive substitutions, *C*-balancedness was classified by Adamczewski [1] and for the most part, the story changes little in the study of *C*-balancedness for compatible random substitutions. Note, however, that we do not attempt a full classification of *C*-balancedness in this setting. We only consider the Pisot case, which is sufficient for our needs. An approachable account of *C*-balancedness is presented in the work of Berthé and Bernales [6], from which we adopt the following definitions and results.

Definition 29. A bi-infinite sequence $x \in \mathcal{A}^{\mathbb{Z}}$ is said to be *C-balanced* if there exists a constant C such that for every $a \in \mathcal{A}$ and all pairs w, w' of subwords of x with $|w| = |w'|$,

$$\big||w|_a - |w'|_a\big| \leq C.$$

A subshift $X \subseteq \mathcal{A}^{\mathbb{Z}}$ is said to be *C-balanced* if there exists a constant C such that for every $a \in \mathcal{A}$ and all pairs $w, w' \in \mathcal{L}(X)$ with $|w| = |w'|$,

$$\big||w|_a - |w'|_a\big| \leq C.$$

Remark 30. Note that we only consider *C*-balancedness with respect to letters, and not the more general notion of *C*-balancedness with respect to subwords, which is a uniform bound on the deviation of the number of appearances of any legal subword within pairs of words of the same length. Indeed, for our setting, this would be asking for too much from random substitutions, where such a property necessarily requires the corresponding subshift to be uniquely ergodic. This is never the case for non-degenerate RS-subshifts.

Recall that the *frequency* of a letter $a \in \mathcal{A}$ in a bi-infinite sequence $x \in \mathcal{A}^{\mathbb{Z}}$ is the limit

$$f_a := \lim_{n \to \infty} \frac{\big|x_{[-n,n]}\big|_a}{2n+1},$$

if it exists, and x is said to have *uniform letter-frequencies* if, for every letter $a \in \mathcal{A}$, the convergence of $\frac{|x_k \cdots x_{k+2n}|_a}{2n+1}$ towards f_a is uniform in k, when n tends to infinity. The amount $\big||w|_a - f_a|w|\big|$ that a subword w of x deviates from the expected number of occurrences of the letter a, according to the letter-frequency of a in x is called the *discrepancy* of the letter a in the word w.

Theorem 31 ([7, Prop. 2.4]). *A bi-infinite sequence $x \in \mathcal{A}^{\mathbb{Z}}$ is C-balanced for some C if and only if it has uniform letter frequencies and there exists another constant B such that for any finite subword w of x and any $a \in \mathcal{A}$,*

$$\big||w|_a - f_a|w|\big| \leq B,$$

where f_a is the frequency of a.

Theorem 32 ([6, Thm. 2.5]). *Let θ be a primitive deterministic irreducible Pisot substitution. Then the subshift X_θ is C-balanced for some C.*

We now extend this result to the random setting. The proof is a simple adaptation of the proof in the deterministic setting [7], generalising the usual Dumont–Thomas prefix-suffix decomposition [10] to the setting of random substitutions.

Theorem 33. *Let ϑ be a compatible primitive irreducible Pisot random substitution. Then the RS-subshift X_ϑ is C-balanced for some C.*

Proof. As ϑ is irreducible Pisot, we know that the second-largest eigenvalue λ_2 has modulus strictly bounded above by 1 and that the substitution matrix is diagonalisable. Let $|\vartheta| = k$ denote the longest inflation word for ϑ. By diagonalising M_ϑ, a routine calculation shows that for any letters $a, b \in \mathcal{A}$, the discrepancy of a in a level-n inflation word in $\vartheta^n(b)$ is bounded by $D|\lambda_2|^n$ for some fixed constant D.

Now consider an arbitrary ϑ-legal word w and an inflation word decomposition that includes a complete n-super-word with n as large as possible. That is, w fully contains between 1 and $2(k-1)$ complete n-super-words, followed by a (possibly empty) suffix s_n and preceded by a (possibly empty) prefix p_n such that s_n and p_n are both strictly contained in n-super-words. The suffix s_n then consists of at most $k-1$ $(n-1)$-super-words followed by a further suffix s_{n-1} that is strictly contained in an $(n-1)$-super-word. Continuing in this way, s_{n-r} comprises at most $(k-1)$ $(n-r-1)$-super-words and a further suffix s_{n-r-1}. Iterating, we get that s_n is the concatenation of up to $k-1$ $(n-1)$-super-words, up to $k-1$ $(n-2)$-super-words, etc., ending with up to $k-1$ individual letters. Likewise, the prefix p_n is the concatenation of up to $k-1$ letters, up to $k-1$ 1-super-words, up to $k-1$ 2-super-words, continuing through $(n-1)$-super words.

Since the discrepancy of w is bounded by the sum of the discrepancies of its pieces, and since the discrepancy of each i-super-word is bounded by $D|\lambda_2|^i$, the total discrepancy $\big||w|_a - f_a|w|\big|$ of a in w is bounded above by

$$\sum_{i=0}^{n} 2(k-1)D|\lambda_2|^i < \sum_{i=0}^{\infty} 2(k-1)D|\lambda_2|^i = \frac{2(k-1)D}{1 - |\lambda_2|}.$$

$\square$

Remark 34. This proof can be modified to show that any compatible random substitution with $|\lambda_2| < 1$ must be C-balanced. This class includes examples, like the random Thue-Morse substitution, that have 0 as an eigenvalue. The complication is that the substitution matrix need not be diagonalisable, in that the eigenvalue 0 (and only the eigenvalue 0!) can be associated with a non-trivial Jordan block. To show C-balancedness, the case $\lambda_2 = 0$ has to be treated separately.

We remark that, as a consequence of Theorem 33, bounded Rauzy fractals exist for irreducible Pisot random substitutions (and are almost surely independent of the choice of element chosen from the subshift). This potentially offers a new tool in the effort to solve the long-standing Pisot conjecture [2]. We therefore invite further exploration of these random substitution Rauzy fractals.

4.2. **Non-mixing of Pisot random substitution subshifts.** We are now in a position to prove Theorem 16. Since irreducible Pisot substitutions are C-balanced by Theorem 33, it is sufficient to prove the following.

Theorem 35. *Let ϑ be a primitive compatible random substitution such that X_ϑ is C-balanced. There exists a constant N such that if ϑ admits a level-n recognisable word for some $n \geq N$, then the RS-subshift X_ϑ is not topologically mixing.*

Proof. This proof has several steps:

(1) Showing that the existence of a level-n recognisable word implies the existence of a ϑ-legal word u such that, if uwu is ϑ-legal, then the length of uw is the length of an exact level-n inflation word. This step does not require C-balancedness.

(2) Showing that for each length L and natural number n, the set of abelianisations of words of the form $\vartheta^n(v)$, where v is a ϑ-legal word of length L, is bounded by a constant independent of L and n. This is where C-balancedness comes in.

(3) Showing that for sufficiently large n, the set of possible lengths of level-n inflation words has natural density strictly less than 1. This step uses step 2. Combined with step 1, this shows that X_ϑ is not topologically mixing.

Step 1: Suppose that there exists a level-n recognisable word $\hat{u}$ with recognisability radius R. Let u_l and u_r be words of length R such that $u = u_l \hat{u} u_r$ is ϑ-legal. By recognisability, all level-n inflation word decompositions of u result in the same decomposition of $\hat{u}$.

Now consider a level-n decomposition of a ϑ-legal word uwu. Since the decompositions of the two copies of $\hat{u}$ are identical, there are corresponding points in the two copies that are spaced by an exact level-n inflation word. However, the spacing between corresponding points in the two copies of $\hat{u}$ is precisely the length of uw.

Step 2: Suppose that our alphabet $\mathcal{A}$ consists of d letters. The number of a's in a word of fixed length L can take on at most $C + 1$ values, since the largest such value is at most C plus the smallest. Repeating for all but one letter and noting that the sum of the entries of the ablianisation is fixed, we see that the set of abelianisations of words of length exactly L has cardinality at most $(C + 1)^{d-1}$. Since the abelianisation of each element of $\vartheta^n(v)$ is just M_ϑ^n times the abelianisation of v, there are at most $(C + 1)^{d-1}$ abelianisations of exact level-n inflation words whose roots have length exactly L.

Step 3: The growth in the lengths of n-super-words is controlled by λ_1, the PF-eigenvalue of M_ϑ. Specifically, there exist positive constants c_1 and c_2 such that every n-super-word has length at least $c_1 \lambda_1^n$ and at most $c_2 \lambda_1^n$. If w is a ϑ-legal exact level-n inflation word with root v and if $|w| < L$, then $|v| < L\lambda_1^{-n}/c_1$. Since each possible length of $|v|$ gives rise to at most $(C + 1)^{d-1}$ possible abelianisations of w, and since each abelianisation determines a unique length, the number of possible lengths of w is bounded by $(C + 1)^{d-1}L\lambda_1^{-n}/c_1$. As L goes to ∞, the density of possible lengths of w is bounded by $(C + 1)^{d-1}/(c_1\lambda_1^n)$. If we pick N such that $\lambda_1^N > (C + 1)^{d-1}/c_1$, then for all $n \geq N$ this density is less than 1. $\qquad\square$

5. Additional Examples

Example 36. Consider the compatible random substitution $\vartheta \colon a \mapsto \{ababab a, bbbaaaa\}, b \mapsto \{babb, bbab\}$ whose substitution matrix is $\left(\begin{smallmatrix} 4 & 1 \\ 3 & 3 \end{smallmatrix}\right)$ with eigenvalues $\lambda_1 = \frac{1}{2}(7 + \sqrt{13})$ and $\lambda_2 = \frac{1}{2}(7 - \sqrt{13}) > 1$. It can be verified that $\gcd\{|\vartheta^n(a_i)| : a_i \in \mathcal{A}\} = 1$ for every $n \geq 1$. Thus, by Theorem 13, X_ϑ is topologically mixing.

Example 37. Consider the random substitution $\vartheta \colon a \mapsto \{bb\}, b \mapsto \{abaaba, ababaa\}$. Its substitution matrix is $\left(\begin{smallmatrix} 0 & 4 \\ 2 & 2 \end{smallmatrix}\right)$ with eigenvalues $\lambda_1 = 4$ and $|\lambda_2| = 2 > 1$. Here $\gcd\{|\vartheta(a_i)| : a_i \in \mathcal{A}\} = \gcd\{2, 6\} = 2 > 1$. Using the fact that the word bb can only come from a substituted a, one can easily show that ϑ is locally recognisable. Thus, as a consequence of Theorem 12, the subshift X_ϑ is not topologically mixing. This can also be seen directly, since the spacing between any two appearances of bb must be a 1-inflation word and so must have even length.

As previously noted, the random Fibonacci substitution is not locally recognisable. Likewise, neither are any of the random metallic means substitutions [21]. However, we do not need local recognisability to prove a lack of topological mixing. By Theorem 16, we just need to show that there exist recognisable words at a sufficiently high level. For the random Fibonacci substitution, we will do more, showing that recognisable words exist at *all* levels.

Example 38. Let ϑ be the random Fibonacci substitution given by

$$\vartheta \colon a \mapsto \{ab, ba\}, b \mapsto \{a\}.$$

Note that all four 2-letter words are ϑ-legal. We begin with the seeds $F_0 = a|a$ and $F_1 = ab|ba$. We successively construct words $F_n = L_n|R_n \in \vartheta(F_{n-1})$ as follows, where the bar divides F_n into level-n super-words. Let ϕ be a map that

- Replaces each b with an a.
- Replaces each a that is followed by a b with ba.
- Replaces each a that is preceded by a b with ab.
- Replaces each a that is preceded by an a with ab if the preceding a was replaced with ba, and with ba if the preceding a was replaced with ab.

These rules are consistent as long as the b's come in pairs (except for a possible lone b at the beginning or end of a word), separated by an even number of a's. This is true for F_0 and F_1, and follows inductively from the fact that $ba^{2k}b$ becomes $a(abba)^k a$.

We claim that all of the words F_n are recognisable inflation words of level 1 with radius 0, with F_{n-1} being the unique root of F_n. By induction, F_n is then a recognisable inflation word of level n, again with radius 0, with unique root aa. The first few iterates and their inflation word decompositions and corresponding roots are given in the following table.

n	$F_n := L_n \mid R_n$	$D_{\vartheta^n}(F_n)$
1	$ab \mid ba$	$\{([ab, ba], aa)\}$
2	$baa \mid aab$	$\{([baa, aab], aa)\}$
3	$aabba \mid abbaa$	$\{([aabba, abbaa], aa)\}$
4	$abbaaaab \mid baaaabba$	$\{([abbaaaab, baaaabba], aa)\}$
5	$baaaabbaabbaa \mid aabbaabbaaaab$	$\{([baaaabbaabbaa, aabbaabbaaaab], aa)\}$

We will show by hand that F_1, F_2 and F_3 have unique 1-decompositions and then generalise to F_n.

Since b's come only from substituting a's, the only way that bb can ever appear in a ϑ-legal word is as part of $abba$, decomposed as $([ab, ba], aa)$. That is, F_1 only decomposes as $\phi(F_0)$. Note that this also shows that bbb is not ϑ-legal.

F_2 contains $aaaa$. Since the middle two a's are not adjacent to b's, they must come from substituting b's. Since bbb is not ϑ-legal, the outer two letters of the root must be a's, so our root is $abba = F_1$. This argument also shows that $aaaaa$ is not ϑ-legal, since the middle three a's would have to all come from b's and bbb is not ϑ-legal.

F_3 contains $abbaabba$. As previously noted, each $abba$ can only come from substituting two a's. Since $aaaaa$ is not ϑ-legal, the outer letters of the root must be b's, so the root is $baaaab = F_2$.

When $n > 3$, we simply repeat these arguments on each portion $baab$ or $baaaab$ of F_n. Since $baab$ is preceded and followed by b's, it can only decompose as $([ba, ab], aa)$, and since $baaaab$ contains two central a's that can only come from b's, this piece can only decompose as $([ba, a, a, ab], abba)$. The decomposition of any remaining prefixes and suffixes is determined by

the fact that the root cannot contain three successive b's or five successive a's. Thus the only decomposition of F_n is as $\phi(F_{n-1})$. As ϑ is irreducible Pisot and admits recognisable words at all levels, it follows from Theorem 16 that the random Fibonacci subshift is not topologically mixing.

Incidentally, we can also build recognisable bi-infinite words as limits of the F_n's. These words form a (deterministic) substitution subshift in their own right, a morphic copy of the Fibonacci subshift, with a replaced by $A = baaaab$ and b replaced by $B = baab$. This is because the map ϕ sends $baaaab$ to $aabbaabbaa$ and sends $baab$ to $aabbaa$. After applying a conjugation that removes baa from the end of each substituted word and places it at the beginning, this is equivalent to $A \to AB$, $B \to A$. The fact that such words are recognisable under ϑ, and not merely under ϕ, follows from the same arguments as with the finite words F_n. Specifically, each $abba$ can only come from aa and all of the remaining a's can only come from b's.

6. RANDOM SUBSTITUTION TILING SPACES

Subshifts are one way to build a dynamical system from a random substitution. An equally valid way is to build a space of tilings on which $\mathbb{R}$ acts by translation [4, 27]. To each letter $a_i \in \mathcal{A}$ we associate an interval of length $\ell(a_i)$, which we sometimes abbreviate as ℓ_i. We call such an interval a *tile*. To each word $u = u_1 \cdots u_n$ we associate an interval of length $\sum_{i=1}^{n} \ell(u_i)$ obtained by laying tiles of type $u_1, u_2, \ldots, u_n$ end to end. We call such a concatenation of tiles a *patch*. A patch associated to an n-super-word is called an *n-supertile*. A patch is called ϑ-legal if it is contained in a supertile of some order. A *tiling* is a bi-infinite patch. The *tiling space* Ω_ϑ associated to the random substitution ϑ is the set of tilings with the property that every finite patch is ϑ-legal.

The group $\mathbb{R}$ acts on Ω_ϑ by translation. Specifically, if $\mathcal{T}$ is a tiling and $t \in \mathbb{R}$, we define $T^t(\mathcal{T})$ to be the tiling obtained by moving all tiles in $\mathcal{T}$ a distance t to the left. (From the point of view of the tiles, this is equivalent to moving the origin a distance t.) Instead of explicitly invoking the group action T, we usually just call the result $\mathcal{T} - t$.

Let P be a finite patch, understood to lie at a specific location, and let U be an open subset of $\mathbb{R}$. Then

$$\mathcal{Z}_{P,U} = \{\mathcal{S} \in \Omega_\vartheta \mid P \subset \mathcal{S} - y \text{ for some } y \in U\}$$

is called a *cylinder set*. By shifting the open set U, we can assume, without loss of generality, that the left endpoint of P is at the origin. These cylinder sets form a basis for the topology of Ω_ϑ. This topology is also the metric topology coming from the *big box metric*, in which two tilings are considered ϵ-close if, after translation by up to ϵ, they agree on the interval $[-\epsilon^{-1}, \epsilon^{-1}]$.

The definition of topological mixing for $\mathbb{R}$ actions is similar to that for subshifts.

Definition 39. The dynamical system (Ω_ϑ, T) is topologically mixing if, for any two open sets $\mathcal{Z}_1$ and $\mathcal{Z}_2$, there is a number R such that, for all $t > R$, $T^t(\mathcal{Z}_1) \cap \mathcal{Z}_2 \neq \emptyset$.

Although this definition is phrased only in terms of large positive values of t, topological mixing is also a property of T^t for t large and negative, since switching the roles of $\mathcal{Z}_1$ and $\mathcal{Z}_2$ is equivalent to changing the sign of t.

For subshifts, topological mixing is related to the number of letters in possible return words w between ϑ-legal words u and v. That is, the possible lengths of words w such that uwv is ϑ-legal. Topological mixing for tiling spaces boils down to studying the geometric lengths of the

patches associated to those return words. Specifically, let $S_{u,v}$ be the set of lengths of all patches associated to words uw, where uwv is ϑ-legal.

Definition 40. A discrete subset $S \subset \mathbb{R}$ is said to be *asymptotically dense* if, for each $\epsilon > 0$ there exists a number R_ϵ such that every $t > R_\epsilon$ is within ϵ of an element of S.

Proposition 41. *The tiling space Ω_ϑ is topologically mixing if and only if, for all ϑ-legal words u and v, $S_{u,v}$ is asymptotically dense.*

Proof. Without loss of generality, we can check topological mixing on sets of the form $\mathcal{Z}_{P,(-\epsilon,\epsilon)}$, since the definition is invariant under translation of the sets and since intervals of size 2ϵ form a basis for the topology of $\mathbb{R}$. We therefore consider sets $\mathcal{Z}_1 = \mathcal{Z}_{P_1,(-\epsilon_1,\epsilon_1)}$ and $\mathcal{Z}_2 = \mathcal{Z}_{P_2,(-\epsilon_2,\epsilon_2)}$, where P_1 is a patch starting at the origin based on the ϑ-legal word u and P_2 is a patch starting at the origin based on the ϑ-legal word v. The open set $\mathcal{Z}_1 - t$ consists of all tilings that exhibit a copy of P_1 starting within ϵ_1 of $-t$, while $\mathcal{Z}_2$ consists of all tilings that exhibit a copy of P_2 starting within ϵ_2 of the origin. These sets intersect if and only if there is a tiling in which the left endpoints of P_1 and P_2 are spaced between $t - (\epsilon_1 + \epsilon_2)$ and $t + (\epsilon_1 + \epsilon_2)$ apart. In other words, if and only if t is within $\epsilon_1 + \epsilon_2$ of an element of $S_{u,v}$. The condition that $T^t((Z)_1) \cap (Z)_2 \neq \emptyset$ for all sufficiently large t for all choices of ϵ_1 and ϵ_2 is equivalent to $S_{u,v}$ being asymptotically dense. $\square$

Our results for subshifts, and those of [19], carry over to tiling spaces with one key adjustment. The condition that $\gcd\{|\theta^n(a)| : a \in \mathcal{A}\} = 1$ for all $n \geq 1$ is replaced with the existence of letters a and b such that the ratio $\ell(a)/\ell(b)$ is irrational.

Theorem 42 ([19, Prop. 1.1])**.** *Let θ be a primitive aperiodic deterministic substitution on an alphabet $\mathcal{A}$. If the tiling space Ω_θ is topologically mixing, then there exist letters $a, b \in \mathcal{A}$ such that $l(a)/l(b)$ is irrational.*

Proof. If all the tile lengths are rational multiples of one another, then there is a basic length ℓ_0 such that all tile lengths are multiples of ℓ_0. But then every element of $S_{u,v}$ (where u and v are arbitrary ϑ-legal words) is a multiple of ℓ_0, so $S_{u,v}$ is not asymptotically dense. $\square$

Kenyon, Sadun and Solomyak [19] determined when a tiling space with a 2-letter deterministic substitution, with second eigenvalue either larger or smaller than 1, is topologically mixing. (The borderline case where $|\lambda_2| = 1$ and $\ell(a)/\ell(b)$ is irrational is subtle. There are examples where the tiling space is topologically mixing and examples where it isn't.)

Theorem 43 ([19, Thm. 1.2])**.** *Let θ be a primitive aperiodic deterministic substitution defined on a two-letter alphabet $\mathcal{A} = \{a, b\}$. If $|\lambda_2| < 1$ or if the ratio $\ell(a)/\ell(b)$ of the two tile lengths is rational, then the tiling space Ω_θ is not topologically mixing. If $|\lambda_2| > 1$ and the ratio is irrational, then the tiling space is topologically mixing.*

We now generalise these results to the random setting. Many of the arguments are the same as for subshifts and as such will be abbreviated.

Theorem 44. *Let Ω be a tiling space whose allowed sequences of tiles are given by a C-balanced subshift X. Then Ω is not topologically mixing.*

Proof. As in the proof of Theorem 35, there are at most $(C+1)^{d-1}$ abelianisations of words with a fixed number of letters, where d is the size of our alphabet. If ℓ_1 is the length of the shortest

tile, then any patch of length L has at most L/ℓ_1 letters. Thus there are at most $L(C+1)^{d-1}/\ell_1$ possible patch lengths of size L or less. In particular, the number of elements of $S_{u,v}$ of size L or less grows at most linearly with L, so $S_{u,v}$ cannot be asymptotically dense. $\qquad\square$

Since random Pisot substitutions are C-balanced, we have an immediate corollary that generalises half of Theorem 43. We highlight that this result is independent of any recognisability assumptions on the random substitution.

Theorem 45. *Let ϑ be a primitive compatible random substitution on an alphabet $\mathcal{A} = \{a_1, \ldots, a_d\}$ with second eigenvalue λ_2 less than 1 in modulus. Then the associated RS-tiling space Ω_ϑ is not topologically mixing.*

Finally, we generalise the remainder of Theorem 43:

Theorem 46. *Let ϑ be a primitive compatible random substitution on a two-letter alphabet with second eigenvalue λ_2 greater than 1 in modulus, such that the ratio of the two tile lengths is irrational. Then Ω_ϑ is topologically mixing.*

Proof. Let u and v be ϑ-legal words and let θ be a marginal of ϑ such that u and v are θ-legal. The second eigenvalue of M_θ is a positive power of λ_2 and so is larger than 1 in modulus. Applying Theorem 43 to Ω_θ, we see that Ω_θ is topologically mixing, which implies that the set of lengths of the return patches from u to v in Ω_θ is asymptotically dense. But that set is a subset of $S_{u,v}$, so $S_{u,v}$ is asymptotically dense. Since this is true for all pairs (u,v), Ω_ϑ is topologically mixing. $\qquad\square$

References

[1] B. Adamczewski, Balances for fixed points of primitive substitutions, *Theoret. Comput. Sci.* **307**(1) (2003), 47–75.

[2] S. Akiyama, M. Barge, V. Berthé, J.-Y. Lee, and A. Siegel, On the Pisot substitution conjecture. In *Mathematics of Aperiodic Order*, Vol. 309 of *Progr. Math.*, pp. 33–72. Birkhäuser/Springer, 2015.

[3] M. Baake, M. Coons, and M. Mañibo, Binary constant-length substitutions and Mahler measures of Borwein polynomials. In *From Analysis to Visualization*, Vol. 313 of *Springer Proceedings in Mathematics & Statistics*, pp. 303–322. Springer, Cham, 2020. Proceedings of the Jonathan Borwein Commemorative Conference.

[4] M. Baake and U. Grimm, *Aperiodic Order. Volume 1: A Mathematical Invitation*, Vol. 149 of *Encyclopedia Math. Appl.*, Cambridge Univ. Press, 2013.

[5] M. Baake, T. Spindeler, and N. Strungaru, Diffraction of compatible random substitutions in one dimension, *Indag. Math. (N.S.)* **29**(4) (2018), 1031–1071.

[6] V. Berthé and P. C. Bernales, Balancedness and coboundaries in symbolic systems, *Theoret. Comput. Sci.* **777** (2019), 93–110.

[7] V. Berthé and V. Delecroix, Beyond substitutive dynamical systems: S-adic expansions. In *Numeration and substitution 2012*, RIMS Kôkyûroku Bessatsu, B46, pp. 81–123. Res. Inst. Math. Sci. (RIMS), Kyoto, 2014.

[8] F. M. Dekking, The spectrum of dynamical systems arising from substitutions of constant length, *Z. Wahrscheinlichkeitsth. Verw. Geb.* **41**(3) (1978), 221–239.

[9] F. M. Dekking and M. Keane, Mixing properties of substitutions, *Z. Wahrscheinlichkeitsth. Verw. Geb.* **42**(1) (1978), 23–33.

[10] J. Dumont and A. Thomas, Systèmes de numération et fonctions fractales relatifs aux substitutions, *Theoret. Comput. Sci.* **65**(2) (1989), 153–169.

[11] N. P. Fogg, *Substitutions in Dynamics, Arithmetics and Combinatorics*, Vol. 1794 of *Lecture Notes in Math.*, Springer-Verlag, 2002.

[12] R. Fokkink, D. Rust, and V. Salo, Automorphism groups of random substitution subshifts, *in preparation* (2021).

[13] C. Godrèche and J. M. Luck, Quasiperiodicity and randomness in tilings of the plane, *J. Statist. Phys.* **55** (1989), 1–28.

[14] P. Gohlke, On a family of semi-compatible random substitutions, Master's thesis, Bielefeld University, 2017.

[15] P. Gohlke, Inflation word entropy for semi-compatible random substitutions, *Monatsh. Math.* **192**(1) (2020), 93–110.

[16] P. Gohlke, A. Mitchell, D. Rust, and T. Samuel, Measure theoretic entropy of random substitution subshifts, *in preparation* (2021).

[17] P. Gohlke, D. Rust, and T. Spindeler, Shifts of finite type and random substitutions, *Discrete Contin. Dyn. Syst.* **39** (2019), 5085–5103.

[18] P. Gohlke and T. Spindeler, Ergodic frequency measures for random substitutions, *Studia Math.* **255**(3) (2020), 265–301.

[19] R. Kenyon, L. Sadun, and B. Solomyak, Topological mixing for substitutions on two letters, *Ergodic Th. & Dynam. Syst.* **25**(6) (2005), 1919–1934.

[20] D. Lind and B. Marcus, *An Introduction to Symbolic Dynamics and Coding*, Cambridge Univ. Press, 1995.

[21] N. Manibo, E. D. P. Miro, D. Rust, and G. S. Tadeo, Zeckendorf representations and mixing properties of sequences, *Tsukuba Journal of Mathematics* (2021). in print.

[22] M. Morse and G. A. Hedlund, Symbolic dynamics II. Sturmian trajectories, *Amer. J. Math.* **62** (1940), 1–42.

[23] B. Mossé, Puissance de mots et reconnaissabilité des points fixes d'une substitution, *Theoret. Comput. Sci.* **99**(2) (1992), 327–334.

[24] H. J. Poincaré, *Les méthodes nouvelles de la mécanique céleste, Vols 1–3*, Gauthiers-Villars, Paris, 1892, 1893, 1899.

[25] D. Rust, Periodic points in random substitution subshifts, *Monatsh. Math.* **193**(3) (2020), 683–704.

[26] D. Rust and T. Spindeler, Dynamical systems arising from random substitutions, *Indag. Math. (N.S.)* **29**(4) (2018), 1131–1155.

[27] L. Sadun, *Topology of Tiling Spaces*, Vol. 46 of *Univ. Lecture Ser.*, Amer. Math. Soc., 2008.

[28] G. S. Tadeo, *Mixing of Random Substitutions*, PhD thesis, Ateneo de Manila University, 2019.

Department of Mathematics, Ateneo de Manila University, Quezon City 1108, Philippines

Email address: `eprovido@ateneo.edu`

School of Mathematics and Statistics, The Open University, Walton Hall, Milton Keynes, MK7 6AA, UK

Email address: `dan.rust@open.ac.uk`

Department of Mathematics, University of Texas at Austin, Austin TX 78712, USA

Email address: `sadun@math.utexas.edu`

Department of Mathematics, Saint Louis University, Baguio City, Philippines

Email address: `gstadeo@slu.edu.ph`