

TOPOLOGICAL ENTROPY OF PSEUDO-ANOSOV MAPS FROM A TYPICAL THURSTON'S CONSTRUCTION

HYUNGRYUL BAIK, INHYEOK CHOI, AND DONGRYUL M. KIM

ABSTRACT. In this paper, we develop a way to extract information about a random walk associated with a typical Thurston's construction. We first observe that a typical Thurston's construction entails a free group of rank 2. We also present a proof of the spectral theorem for random walks associated with Thurston's construction that have finite second moment with respect to the Teichmüller metric. Its general case was remarked by Dahmani and Horbez. Finally, under a hypothesis not involving moment conditions, we prove that random walks eventually become pseudo-Anosov.

As an application, we first discuss a random analogy of Kojima and McShane's estimation of the hyperbolic volume of a mapping torus with pseudo-Anosov monodromy. As another application, we discuss non-probabilistic estimations of stretch factors from Thurston's construction and the powers for Salem numbers to become the stretch factors of pseudo-Anosovs from Thurston's construction.

Keywords. Topological entropy, Thurston's construction, pseudo-Anosov surface homeomorphism, Markov chains, hyperbolic volume of mapping torus, Salem numbers.

MSC classes: 37B40, 37D40, 37E30, 60B15, 60G50.

1. INTRODUCTION

For a closed orientable connected surface S (or S_g when g plays a role) of genus $g > 1$, its *mapping class group* $\text{Mod}(S)$ is defined as follows.

$$\text{Mod}(S) := \text{Homeo}^+(S)/\text{Isotopy}$$

The elements of $\text{Mod}(S)$ are called *mapping classes* and classified, by the Nielsen–Thurston classification, into three categories: periodic, reducible, and pseudo-Anosov elements. Among them, pseudo-Anosov mapping classes are considered the generic ones (cf. [EST20]) but it is not straightforward to come up with concrete examples.

In this regard, some constructions for pseudo-Anosovs have been developed, such as Thurston's construction [T⁺88] and Penner's construction [Pen88]. Both constructions use Dehn twists along a pair of multicurves which fills the surface. Penner originally conjectured that every pseudo-Anosov mapping class has some power arising from Penner's construction. However, Shin and Strenner showed in [SS16] that coronal pseudo-Anosov

Date: May 29, 2021.

mapping classes, which are characterized by their stretch factors, have no powers constructed from Penner's construction. This answers Penner's conjecture negatively.

We can then ask whether a typical pseudo-Anosov mapping class is obtained from Penner's construction. We can also ask an analogous question for Thurston's construction. Motivated by this, we study the behavior of pseudo-Anosov maps obtained from Thurston's construction, especially in a typical circumstance.

Let (A, B) be a filling pair of multicurves on S . Thurston's construction describes the subgroup $\langle T_A, T_B \rangle$ of $\text{Mod}(S)$ generated by multitwists T_A and T_B along A and B , respectively. In [Lei04], Leininger provided a necessary and sufficient condition for (A, B) to result in $\langle T_A, T_B \rangle \cong F_2$, the free group of rank 2. It asserts that $\langle T_A, T_B \rangle \not\cong F_2$ if and only if A and B have a certain tree-like position as in Figure 2. We summarize this as follows:

Observation 2.7. *For a typical filling pair of multicurves (A, B) , we have $\langle T_A, T_B \rangle \cong F_2$.*

From now on, we assume the conclusion of Observation 2.7. We then consider probability measures supported on the subgroup $\langle T_A, T_B \rangle$ and random mapping classes arising from them.

Let $\nu : \text{Mod}(S) \rightarrow [0, 1]$ be a probability measure with the support $\text{supp } \nu := \{g \in \text{Mod}(S) : \nu(g) > 0\}$ contained in $\langle T_A, T_B \rangle$. Let us also fix an initial mapping class $g_0 \in \text{Mod}(S)$, which is assumed to be the identity element unless otherwise stated. Then ν induces a probability measure \mathbb{P} on the space of sample paths $\text{Mod}(S)^{\mathbb{N}}$ as follows. Given elements $g_1, \dots, g_n \in \text{Mod}(S)$, we define a cylinder set

$$[g_1, \dots, g_n] := \{\omega = (\omega_i) \in \text{Mod}(S)^{\mathbb{N}} : \omega_i = g_i \text{ for } 1 \leq i \leq n\}$$

and assign

$$\mathbb{P}([g_1, \dots, g_n]) = \nu(g_0^{-1}g_1)\nu(g_1^{-1}g_2) \cdots \nu(g_{n-1}^{-1}g_n).$$

Throughout the paper, ω stands for a random walk or a sample path on $\text{Mod}(S)$, and ω_n for the n -th step mapping class of ω .

In this setting, we investigate how much and how often ω produces pseudo-Anosov maps. This topic was handled by Dahmani and Horbez in a more general setting, and we present one of their spectral theorems below. In the following, ν is said to be non-elementary if $\langle \text{supp } \nu \rangle$ contains a pair of pseudo-Anosov maps with disjoint sets of fixed points in $\mathcal{PMF}(S)$.

Theorem 3.9 (Dahmani–Horbez, [DH18]). *Let ν be a non-elementary probability measure on $\text{Mod}(S)$ with finite support, and \mathbb{P} be the induced measure on $\text{Mod}(S)^{\mathbb{N}}$. Then for \mathbb{P} -a.e. $\omega \in \text{Mod}(S)^{\mathbb{N}}$, ω_n is eventually pseudo-Anosov. Moreover, the stretch factor λ_{ω_n} of a pseudo-Anosov ω_n satisfies*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \lambda_{\omega_n} = L_{\mathcal{T}} > 0$$

where the constant $L_{\mathcal{T}} := \lim_{n \rightarrow \infty} \frac{1}{n} d_{\mathcal{T}}(\mathcal{X}, \omega_n \cdot \mathcal{X})$ is the drift of ω with respect to the Teichmüller metric $d_{\mathcal{T}}$.

According to Masai [Mas18], the drift $L_{\mathcal{T}}$ is almost surely equal to the topological entropy $h(\omega)$ of ω if ν is non-elementary and has finite first moment with respect to the Teichmüller metric. See Theorem 3.7. In addition, the topological entropy $h(\varphi)$ of a pseudo-Anosov map φ is equal to $\log \lambda_{\varphi}$, where λ_{φ} is the stretch factor of φ (cf. [FLP79, Exposé Ten]). Thus, Dahmani and Horbez's result can be interpreted as an asymptotic relation between the topological entropy of the entire random walk ω and that of each random mapping class ω_n . To keep this interpretation, we write spectral theorems in terms of $h(\omega)$ even though it does not depend on the choice of sample path ω by [Mas18].

In [DH18, Remark 2.7, 3.2], the authors assert that the finite support condition in Theorem 3.9 can be replaced with the finite second moment condition. For the sake of completeness, we present the proof of this assertion in our setting.

Theorem A. *Let (A, B) be a filling pair of multicurves on S and ν be a non-elementary probability measure on $\langle T_A, T_B \rangle$. Suppose that ν has finite second moment with respect to the Teichmüller metric, i.e., for some $\mathcal{X} \in \mathcal{T}(S)$,*

$$\sum_{g \in \text{Mod}(S)} d_{\mathcal{T}}(\mathcal{X}, g \cdot \mathcal{X})^2 \nu(g) < \infty.$$

Then for \mathbb{P} -a.e. $\omega \in \text{Mod}(S)^{\mathbb{N}}$, ω_n is eventually pseudo-Anosov and

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \lambda_{\omega_n} = h(\omega).$$

As noted in this theorem, random mapping classes eventually become pseudo-Anosov *almost surely* if the second moment is finite. In contrast, without moment condition, previously known results are in the sense of *phenomena in probability*, which is weaker than almost sure phenomena. For example, Rivin [Riv08] and Kowalski [Kow08] proved that nearest-neighbor random walks on $\text{Mod}(S)$ become pseudo-Anosov with asymptotic probability one. Note that their strategies are applicable in a broader setting, not only for mapping class groups. In addition, Maher proved in [M⁺11] the convergence $\mathbb{P}(\omega_n \text{ is pseudo-Anosov}) \rightarrow 1$ as $n \rightarrow \infty$ if the governing probability measure ν on $\text{Mod}(S)$ is non-elementary.

Hence, it is natural to ask whether an almost sure phenomenon can be obtained without any moment condition. We answer this question in the setting of Thurston's construction, utilizing facts about free groups.

Theorem B. *Let (A, B) be a filling pair of multicurves on S such that $\langle T_A, T_B \rangle \cong F_2$, and $\nu : \text{Mod}(S) \rightarrow [0, 1]$ be a non-elementary probability measure with $\text{supp } \nu \subseteq \langle T_A, T_B \rangle$ and $\text{rank}(\text{supp } \nu) < \infty$. Then for \mathbb{P} -a.e. $\omega \in \text{Mod}(S)^{\mathbb{N}}$, there exists $N \in \mathbb{N}$ such that*

$$\omega_n \text{ is pseudo-Anosov for all } n > N.$$

There are reasons why we focus on Thurston's construction. First, Thurston's construction makes use of a Teichmüller disk $f(\mathbb{H}^2)$ in $\mathcal{T}(S)$, an isometrically embedded copy of the hyperbolic disk \mathbb{H}^2 of curvature -4 . Mapping classes from Thurston's construction fix $f(\mathbb{H}^2)$ and act as isometries of $f(\mathbb{H}^2)$ ([HS07], [T⁺88], [Lei04]). Hence, the Teichmüller distance can be translated into the hyperbolic distance, allowing us an explicit computation.

Moreover, let $\rho : \langle T_A, T_B \rangle \rightarrow \mathrm{PSL}(2, \mathbb{R})$ be the representation obtained by Thurston's construction. Then there exists a point $\mathcal{X} \in f(\mathbb{H}^2)$ such that

$$\log \|\rho(w)\| = d_{\mathcal{T}}(\mathcal{X}, w \cdot \mathcal{X})$$

for each $w \in \langle T_A, T_B \rangle$, where $\|T\| := \sup_{|x|=1} |Tx|$ is the operator norm. If a measure ν has finite first moment with respect to the Teichmüller metric, then

$$\int \log \|\rho(\omega_1)\| d\nu < \infty.$$

In this case, using the following theorem of Furstenberg and Kesten, we can approximate the topological entropy $h(\omega)$ without referring to individual sample paths. We remark that an explicit formula for $h(\omega)$ can be obtained by the work [Pol10] of Pollicott when ν is supported on finitely many mapping classes that correspond to strictly positive matrices in $\mathrm{SL}(2, \mathbb{R})$.

Theorem 1.1 (Furstenberg–Kesten, [FK60]). *Let (X_n) be an independent and identically distributed stochastic process with values in the set of $k \times k$ matrices such that*

$$\mathbb{E} \max\{\log \|X_1\|, 0\} < \infty.$$

Then as $n \rightarrow \infty$,

$$\frac{1}{n} \log \|X_1 \cdots X_{n-1} X_n\| \rightarrow E \quad \text{almost surely,}$$

where $E = \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \log \|X_1 \cdots X_{n-1} X_n\|$.

Especially, the subadditivity of $\mathbb{E} \log \|X_1 \cdots X_{n-1} X_n\|$ implies that $E = \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \log \|X_1 \cdots X_{n-1} X_n\| = \inf_n \frac{1}{n} \mathbb{E} \log \|X_1 \cdots X_{n-1} X_n\|$. This allows us to approximate or provide an explicit upper bound for $h(\omega)$.

As an application of above results, we estimate the hyperbolic volumes of random mapping tori arising from a typical Thurston's construction. In [KM18], Kojima and McShane fixed a pseudo-Anosov map φ and investigated the hyperbolic volume of a mapping torus with monodromy φ . It represents the dynamics of the iterated sequence $\varphi, \varphi^2, \dots, \varphi^n, \dots$

In order to translate this into a random version, we replace the iterate φ^n with the element ω_n of a random walk associated with a typical Thurston's construction. Thanks to Theorem B, such elements are eventually pseudo-Anosov almost surely, and thus one can discuss the hyperbolic volume for the mapping torus with monodromy ω_n for large n . Consequently, our random analogy of Kojima and McShane's estimate follows:

Theorem C. *Let (A, B) be a filling pair of multicurves on S such that $\langle T_A, T_B \rangle \cong F_2$ and $\nu : \text{Mod}(S) \rightarrow [0, 1]$ be a non-elementary probability measure with $\text{supp } \nu \subseteq \langle T_A, T_B \rangle$ and $\text{rank}(\text{supp } \nu) < \infty$. Suppose also that ν has finite first moment with respect to the Teichmüller metric. Then*

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \text{vol}(M_{\omega_n}) \leq -3\pi\chi(S) \cdot h(\omega)$$

for \mathbb{P} -a.e. $\omega \in \text{Mod}(S)^\mathbb{N}$ where $\chi(S)$ is the Euler characteristic of S .

This estimation supports the point of view that the topological entropy of a random walk shares common characteristics with the entropy of a single pseudo-Anosov. Indeed, Kojima and McShane's estimate asserts that

$$\frac{1}{n} \text{vol}(M_{\varphi^n}) \leq -3\pi\chi(S) \cdot h(\varphi)$$

where $h(\varphi)$ is the topological entropy of a pseudo-Anosov φ , which can be regarded as an interchange of $h(\omega)$ and $h(\varphi)$. See Remark 4.4 for details.

As another application, we establish an estimation for stretch factors obtained by Thurston's construction; note that the estimation is non-probabilistic.

Theorem D. *Let (A, B) be a filling pair of multicurves on S . For a finite generating set Λ of $\langle T_A, T_B \rangle$, let $|\cdot|_\Lambda$ be a word norm with respect to Λ . Then there exist constants $L > 0$ and N such that*

$$\log \lambda_\varphi < L|\varphi|_\Lambda$$

for all pseudo-Anosov $\varphi \in \langle T_A, T_B \rangle$ with $|\varphi|_\Lambda > N$.

See Subsection 4.2 for the detailed nature of L . In particular, for some choice of Λ , L can be much smaller than the first moments of measures on Λ . In contrast, L is always greater than the drift $L_{\mathcal{T}}$ of a measure on Λ .

As stretch factors of pseudo-Anosovs from Thurston's construction are related to Salem number, an algebraic unit λ whose all Galois conjugates except $\lambda^{\pm 1}$ lie on the unit circle in \mathbb{C} , one can also obtain a consequence regarding Salem numbers. Pankau proved in [Pan17] that every Salem number has a power, which is the stretch factor of a pseudo-Anosov arising from Thurston's construction. Regarding this, one can ask how large the power should be. We provide one estimation here.

Theorem E. *Let λ be a Salem number and suppose that λ^k is the stretch factor of a pseudo-Anosov φ obtained from Thurston's construction with $\rho : \langle T_A, T_B \rangle \rightarrow \text{PSL}(2, \mathbb{R})$. Then*

$$k \leq \frac{K|\varphi|}{\log \lambda}$$

where $|\varphi|$ is the cyclically reduced word norm of φ in terms of T_A^\pm and T_B^\pm and $K = \log \frac{\sqrt{\mu} + \sqrt{4+\mu}}{2}$.

Moreover, when $\langle T_A, T_B \rangle \cong F_2$, we further have

$$\frac{\log |\varphi|}{4 \log \lambda} \leq k \leq \frac{K|\varphi|}{\log \lambda}.$$

The paper is organized as follows. In Section 2, we review Thurston's construction of pseudo-Anosov mapping classes. We present Leininger's characterization in [Lei04] of filling multicurves A and B whose multitwists generate a free group F_2 of rank 2 and its typicality. We also classify elements in $\langle T_A, T_B \rangle \cong F_2$ using the representation obtained from Thurston's construction.

In Section 3, we investigate the dynamical properties arising from (typical) Thurston's construction. After reviewing the theory of topological entropy, we present a spectral theorem of Dahmani and Horbez. For the sake of completeness, we prove a version of the theorem asserted in [DH18] in the setting of Thurston's construction. We also investigate the eventually pseudo-Anosov behavior of random walks without moment condition.

Section 4 deals with applications of our results. Following Kojima and McShane, we consider the hyperbolic volume of a mapping torus with pseudo-Anosov monodromy and stretch factors from Thurston's construction. While our strategy relies on probabilistic tools, we successfully obtain a non-probabilistic estimate regarding the stretch factors.

Let us finish the introduction by suggesting further questions:

- (1) Is there a generalization of Thurston's construction involving n multicurves A_1, \dots, A_n so that $\langle T_{A_1}, \dots, T_{A_n} \rangle \cong F_n$ in a typical case?
- (2) On which subgroup of $\text{Mod}(S)$ almost every random walk associated with it is eventually pseudo-Anosov, as in Theorem B?
- (3) Does the spectral theorem hold under the finite p -th moment condition with respect to the Teichmüller metric, for $1 \leq p < 2$? Or, can we provide the threshold for p ?

Acknowledgments. We truly appreciate Mladen Bestvina, Ilya Gekhtman, Camille Horbez, Paul Jung, Wanmo Kang, Sang-hyun Kim, Hidetoshi Masai, Bram Petri, and Giulio Tiozzo for fruitful conversations. Finally we thank the anonymous referees for their valuable comments which improved the organization of the paper a lot. The first and second authors were partially supported by Samsung Science & Technology Foundation grant No. SSTF-BA1702-01. The third author was supported by KAIST Undergraduate Research Participation program.

2. THURSTON'S CONSTRUCTION

2.1. Quick review of Thurston's construction. In this section, we review Thurston's construction of pseudo-Anosov mapping classes using multitwists along filling multicurves. Throughout the paper, every surface is a closed orientable connected surface of genus $g > 1$; every simple closed curve is essential, i.e., not homotopic to a single point, unless otherwise stated.

Definition 2.1 (Filling multicurves). *A multicurve A on a surface S is a finite set of disjoint simple closed curves on S . A pair (A, B) of multicurves A and B is filling if $S \setminus \bigcup_{\gamma \in A \cup B} \gamma$ is a disjoint union of open disks.*

Note that in the above definition of a multicurve, a multicurve is allowed to have two isotopic simple closed curves as distinct elements.

Definition 2.2 (Multitwist). *For a multicurve A on a surface S , the multitwist T_A along A is the product*

$$T_A := \prod_{\alpha \in A} T_\alpha$$

where T_α is the Dehn twist along a simple closed curve $\alpha \in A$.

Given a filling pair (A, B) of multicurves, Thurston's construction provides a $\text{PSL}(2, \mathbb{R})$ -representation of $\langle T_A, T_B \rangle$. We can then apply the classification of elements in $\text{PSL}(2, \mathbb{R})$ to deduce that in $\langle T_A, T_B \rangle$.

We now explain Thurston's construction. The first ingredient is as follows:

Lemma 2.3 (Perron–Frobenius). *Let M be an $n \times n$ matrix with integer entries. Assume M is primitive, i.e., M is nonnegative and M^m is positive for some $m \in \mathbb{N}$. Then M has a unique nonnegative unit eigenvector v . Moreover, v is positive and has a positive eigenvalue, called the Perron–Frobenius eigenvalue, which is larger in absolute value than all other eigenvalues.*

Let $A = \{\alpha_1, \dots, \alpha_n\}$ and $B = \{\beta_1, \dots, \beta_m\}$. We define a matrix N by

$$N_{jk} = i(\alpha_j, \beta_k)$$

where $i(\cdot, \cdot)$ is the geometric intersection number. It can be shown that NN^t is primitive, and thus by Perron–Frobenius theorem, its Perron–Frobenius eigenvalue $\mu = \mu(A, B)$ is obtained. Together with this observation, Thurston constructed pseudo-Anosov mapping classes in the following theorem.

Theorem 2.4 ([T⁺88], [FM12, Chapter 14]). *Let (A, B) be a filling pair of multicurves on S . For the Perron–Frobenius eigenvalue $\mu = \mu(A, B)$ obtained by the above procedure, there is a representation $\rho : \langle T_A, T_B \rangle \rightarrow \text{PSL}(2, \mathbb{R})$ defined by*

$$T_A \mapsto \begin{bmatrix} 1 & -\sqrt{\mu} \\ 0 & 1 \end{bmatrix} \quad T_B \mapsto \begin{bmatrix} 1 & 0 \\ \sqrt{\mu} & 1 \end{bmatrix}.$$

The representation ρ has the following properties:

- (1) For $\varphi \in \langle T_A, T_B \rangle$, φ is periodic, reducible, or pseudo-Anosov if $\rho(\varphi)$ is elliptic, parabolic, or hyperbolic, respectively.
- (2) If $\rho(\varphi)$ is parabolic, then φ is a multitwist.
- (3) If $\rho(\varphi)$ is hyperbolic, the largest eigenvalue of $\rho(\varphi)$ is the stretch factor of the pseudo-Anosov mapping class φ .

2.2. Group-theoretic aspects of Thurston's construction. We now connect the group-theoretic aspects of $\langle T_A, T_B \rangle$ with the topological aspects of multicurves A and B .

The crucial observation on Thurston's construction is that $\langle T_A, T_B \rangle \cong F_2$, a free group of rank 2, for a typical configuration of A and B . To see the typicality, we first follow the work of Leininger [Lei04] that provides a necessary and sufficient condition for A and B to result in $\langle T_A, T_B \rangle \cong F_2$. We begin by defining a graph associated with multicurves A and B :

Definition 2.5 (Configuration graph, [Lei04]). *For a finite set C of simple closed curves on S , the configuration graph $G(C)$ for C is a graph defined by the following rule:*

- The vertex set of $G(C)$ is C and;
- There are $i(c_1, c_2)$ edges between vertices corresponding to $c_1, c_2 \in C$.

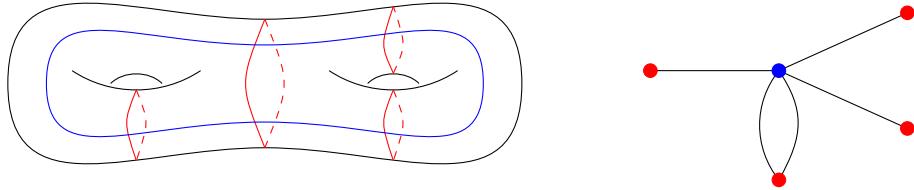


FIGURE 1. Filling multicurves A , B on S and the configuration graph $G(A \cup B)$.

With this definition, Leininger proved the following theorem.

Theorem 2.6 (Leininger, [Lei04]). *For a filling pair (A, B) of multicurves,*

$$\langle T_A, T_B \rangle \not\cong F_2 \Leftrightarrow \begin{array}{l} G(A \cup B) \text{ is one of} \\ \text{the graphs in Figure 2:} \end{array}$$

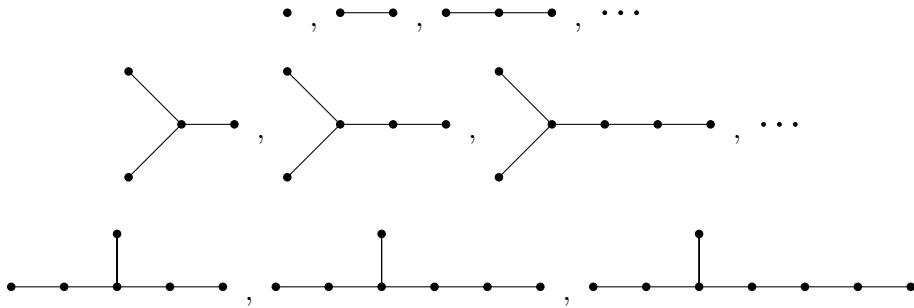


FIGURE 2. Configuration graphs that do not entail F_2 .

In other words, all configurations of A and B not as in Figure 2 result in $\langle T_A, T_B \rangle \cong F_2$. Hence we observe the following typicality of F_2 :

Observation 2.7. *For a typical filling pair of multicurves (A, B) , we have $\langle T_A, T_B \rangle \cong F_2$.*

We now consider a quantitative model that demonstrates this typicality. Let us introduce the convention that $[A] := A/\text{isotopy}$ ($[B] := B/\text{isotopy}$, resp.), and n_α (m_β , resp.) is the number of representatives in A (B , resp.) of an isotopy class $\alpha \in [A]$ ($\beta \in [B]$, resp.). We also note the result of Pankau [Pan17] that any nonsingular positive integer square matrix is realized as the intersection matrix of a filling pair of multicurves.

Proposition 2.8 (A quantitative reasoning for the typicality). *Consider the following random model for multicurves A and B :*

- $|[A]| = n$, $|[B]| = m$;
- $i(\alpha, \beta) \sim \text{Unif}(0, 1, \dots, k-1)$ for $\alpha \in [A]$, $\beta \in [B]$ and;
- $n_\alpha, m_\beta \sim \text{Unif}(1, 2, \dots, k)$ for $\alpha \in [A]$, $\beta \in [B]$

where the random variables are independent. Then,

$$\mathbb{P}(A, B \text{ are filling and } \langle T_A, T_B \rangle \not\cong F_2) \leq 2 - (1 - m/k^{m+1})^n - (1 - n/k^{n+1})^m.$$

Fixing a surface of genus g , we have a restriction $n, m \leq 3g - 3$ and the asymptotic probability $\mathbb{P}(A, B \text{ are filling and } \langle T_A, T_B \rangle \not\cong F_2) \rightarrow 0$ as $k \rightarrow \infty$. In other words, the probability gets closer to 0 as we allow more intersections among the given isotopy classes. Even in an extreme case that $n = m = 1$ and $k = 4$, the above bound is calculated as $1/8$.

In order to prove the proposition, we first provide a sufficient condition for $\langle T_A, T_B \rangle \cong F_2$ in terms of $i(\cdot, \cdot)$, n_α , and m_β . Observing that $\langle T_\alpha, T_\beta \rangle \cong F_2$ if $i(\alpha, \beta) \geq 2$ for simple closed curves α and β , we expect a similar phenomenon for multitwists, generalizing the case of single curves:

Lemma 2.9. *For multicurves A and B on S , suppose that there are subsets $\emptyset \neq A' \subseteq A$ and $\emptyset \neq B' \subseteq B$ satisfying the following:*

- (1) $n_\alpha i(\alpha, [B']) = n_\alpha \sum_{\beta \in [B']} i(\alpha, \beta) \geq 2$ for all $\alpha \in [A']$
- (2) $m_\beta i([A'], \beta) = m_\beta \sum_{\alpha \in [A']} i(\alpha, \beta) \geq 2$ for all $\beta \in [B']$

Then the group $\langle T_A, T_B \rangle$ generated by multitwists is isomorphic to F_2 .

This lemma is a generalization of [HT02, Theorem 3.2]. Proof of this lemma relies on the following ping-pong lemma:

Lemma 2.10 (Ping-pong lemma). *Let $\langle g_1, \dots, g_n \rangle$, $n \geq 2$, be a group acting on a set X . Suppose that there are disjoint nonempty subsets $X_1, \dots, X_n \subseteq X$ such that for each $i \neq j$, $g_i^k(X_j) \subseteq X_i$ for all $0 \neq k \in \mathbb{Z}$. Then*

$$\langle g_1, \dots, g_n \rangle \cong F_n.$$

Proof of Lemma 2.9. To apply the ping-pong lemma, let X be the set of isotopy classes of essential simple closed curves on S . Then for $x \in X$ and

$0 \neq n \in \mathbb{Z}$, [FLP79, Proposition A.1] implies

$$(2.1) \quad \left| i(T_A^n x, [B']) - |n| \sum_{\alpha \in [A]} n_\alpha i(x, \alpha) i(\alpha, [B']) \right| \leq i(x, [B']).$$

Now let

$$\begin{aligned} X_{A'} &= \{x \in X : i(x, [A']) < i(x, [B'])\}, \\ X_{B'} &= \{x \in X : i(x, [B']) < i(x, [A'])\}. \end{aligned}$$

They are nonempty since $\emptyset \neq [A'] \subseteq X_{A'}$ and $\emptyset \neq [B'] \subseteq X_{B'}$, and are disjoint. For $x \in X_{B'}$ and $0 \neq n \in \mathbb{Z}$, it follows from Equation 2.1 that

$$i(T_A^n x, [B']) \geq |n| \sum_{\alpha \in [A]} n_\alpha i(x, \alpha) i(\alpha, [B']) - i(x, [B']).$$

Since $x \in X_{B'}$ and $A' \subseteq A$,

$$\begin{aligned} |n| \sum_{\alpha \in [A]} n_\alpha i(x, \alpha) i(\alpha, [B']) - i(x, [B']) &> |n| \sum_{\alpha \in [A]} n_\alpha i(x, \alpha) i(\alpha, [B']) - i(x, [A']) \\ &\geq |n| \sum_{\alpha \in [A']} n_\alpha i(x, \alpha) i(\alpha, [B']) - i(x, [A']) \\ &= \sum_{\alpha \in [A']} i(x, \alpha) (|n| n_\alpha i(\alpha, [B']) - 1). \end{aligned}$$

From the condition (1) given in the statement,

$$\begin{aligned} \sum_{\alpha \in [A']} i(x, \alpha) (|n| n_\alpha i(\alpha, [B']) - 1) &\geq \sum_{\alpha \in [A']} i(x, \alpha) = i(x, [A']) \\ &= i(T_A^n x, T_A^n [A']) = i(T_A^n x, [A']) \end{aligned}$$

which implies

$$T_A^n X_{B'} \subseteq X_{A'}.$$

A similar argument after interchanging A and B shows that

$$T_B^n X_{A'} \subseteq X_{B'}.$$

Therefore, by the ping-pong lemma, we conclude that

$$\langle T_A, T_B \rangle \cong F_2. \quad \square$$

Based on this lemma, we now prove Proposition 2.8.

Proof of Proposition 2.8. We fix $|[A]| = n$ and $|[B]| = m$ as in the statement, and then denote

$$\mathbb{P}(i(\alpha, \beta) = j) = p_j \quad \mathbb{P}(n_\alpha = j) = \mathbb{P}(m_\beta = j) = q_j.$$

Note that

$$\begin{aligned} \mathbb{P}(A, B \text{ are filling}, \langle T_A, T_B \rangle \not\cong F_2) &\leq \mathbb{P}(n_\alpha i(\alpha, [B]) = 1 \text{ for some } \alpha \in [A]) \\ &\quad + \mathbb{P}(m_\beta i([A], \beta) = 1 \text{ for some } \beta \in [B]). \end{aligned}$$

We first estimate $\mathbb{P}(n_\alpha i(\alpha, [B]) = 1 \text{ for some } \alpha \in [A])$. For a fixed $\alpha \in [A]$,

$$\mathbb{P}(n_\alpha i(\alpha, [B]) = 1) = \mathbb{P}(n_\alpha = 1) \mathbb{P}(i(\alpha, [B]) = 1) = q_1 m p_1 p_0^{m-1}$$

where the first equality comes from the independence. Hence, we have

$$\mathbb{P}(n_\alpha i(\alpha, [B]) \neq 1 \text{ for all } \alpha \in [A]) = (1 - q_1 m p_1 p_0^{m-1})^n$$

from the independence, and thus

$$\mathbb{P}(n_\alpha i(\alpha, [B]) = 1 \text{ for some } \alpha \in [A]) = 1 - (1 - q_1 m p_1 p_0^{m-1})^n.$$

Similarly,

$$\mathbb{P}(m_\beta i([A], \beta) = 1 \text{ for some } \beta \in [B]) = 1 - (1 - q_1 n p_1 p_0^{n-1})^m$$

which deduces the desired estimation

$$\mathbb{P}(A, B \text{ are filling}, \langle T_A, T_B \rangle \not\cong F_2) \leq 2 - (1 - q_1 m p_1 p_0^{m-1})^n - (1 - q_1 n p_1 p_0^{n-1})^m.$$

Now plugging in $p_j = q_j = 1/k$, we conclude

$$\mathbb{P}(A, B \text{ are filling}, \langle T_A, T_B \rangle \not\cong F_2) \leq 2 - (1 - m/k^{m+1})^n - (1 - n/k^{n+1})^m. \quad \square$$

Note that the estimation above is based on a criterion weaker than Lemma 2.9. Thus, the actual probability for not being F_2 will be even smaller than our estimation.

According to the classification of subgroups of $\text{Mod}(S)$, proven by McCarthy and Papadopoulos [MP89] and Kida [Kid08], every subgroup of $\text{Mod}(S)$ is classified into one of the below four types :

- *Non-elementary subgroup*, i.e., it contains a pair of pseudo-Anosov mapping classes with disjoint fixed point sets in $\mathcal{PMF}(S)$.
- *Virtually infinite cyclic subgroup*
- *Reducible subgroup*, i.e., there exists a multicurve on S fixed by an action of the subgroup
- *Finite subgroup*

This classification deduces the following proposition about F_2 subgroup of $\text{Mod}(S)$ generated by two multitwists along filling multicurves.

Proposition 2.11. *For a filling pair (A, B) of multicurves on S resulting in $\langle T_A, T_B \rangle \cong F_2$, $\langle T_A, T_B \rangle$ is a non-elementary subgroup of $\text{Mod}(S)$.*

Proof. $\langle T_A, T_B \rangle \cong F_2$ is clearly not a finite subgroup of $\text{Mod}(S)$. Now suppose that $\langle T_A, T_B \rangle$ fixes a multicurve C . Let M be the number of curves in C , and let γ be one such curve. Since T_A permutes the curves in C , we observe $T_A^{M!}(\gamma) = \gamma$. This implies that A and γ are disjoint, possibly after perturbing A a bit. Similarly, $T_B^{M!}(\gamma) = \gamma$, implying that B and γ are disjoint. Thus, a complementary region of $A \cup B$ contains the essential curve γ ; that region cannot be null-homotopic, which contradicts the filling condition of (A, B) . In conclusion, $\langle T_A, T_B \rangle$ cannot be reducible.

Now suppose that $\langle T_A, T_B \rangle$ is a virtually infinite cyclic subgroup, i.e.,

$$\langle T_A, T_B \rangle = \bigsqcup_{i=1}^n h_i \langle g \rangle$$

for some $g, h_i \in \langle T_A, T_B \rangle$ and $n \in \mathbb{N}$. Let $H, T \in \{T_A, T_B, T_A^{-1}, T_B^{-1}\}$ be the first and the last alphabet of the reduced word representing g , respectively. Furthermore, let $N = 1 + \max_i |h_i|$ where $|\cdot|$ denotes the word norm in terms of T_A^\pm and T_B^\pm . Finally, let $R \in \{T_A, T_B, T_A^{-1}, T_B^{-1}\} \setminus \{H^{-1}, T\}$. Then

$$R^N = h_j g^k = h_j (H \cdots T)^k$$

holds for some $j \in \{1, \dots, n\}$ and $k \in \mathbb{Z}$ since $R^N \in \langle T_A, T_B \rangle = \bigsqcup_{i=1}^n h_i \langle g \rangle$.

If $k = 0$, then $R^N = h_j$, deducing $|h_j| = N > |h_j|$, contradiction.

If $k > 0$, then

$$R^N T^{-1} \cdots H^{-1} = h_j$$

where the left-hand side is in its reduced form. Hence, it follows that $|h_j| = |R^N T^{-1} \cdots H^{-1}| \geq N > |h_j|$, a contradiction.

If $k < 0$, then

$$R^N H \cdots T = h_j$$

where the left-hand side is in its reduced form. Hence, $|h_j| = |R^N H \cdots T| \geq N > |h_j|$, a contradiction.

Therefore, $\langle T_A, T_B \rangle \cong F_2$ cannot be virtually infinite cyclic subgroup and we conclude that it is a non-elementary subgroup. \square

2.3. Elements of $\langle T_A, T_B \rangle \cong F_2$. We finish this section by classifying elements of $\langle T_A, T_B \rangle \cong F_2$ for a filling pair (A, B) . First, we observe the following property of the Perron–Frobenius eigenvalue $\mu = \mu(A, B)$:

Lemma 2.12. *For filling multicurves A and B on S , with $\langle T_A, T_B \rangle \cong F_2$,*

$$\mu = \mu(A, B) \geq 4.$$

Proof. Suppose to the contrary that $\mu < 4$. Then,

$$\text{tr}(\rho(T_A)\rho(T_B)) = 2 - \mu \in (-2, 2)$$

where ρ is the representation for Thurston's construction. Hence $\rho(T_A T_B)$ is elliptic in $\text{PSL}(2, \mathbb{R})$ and thus it follows from Theorem 2.4, Thurston's construction, that $T_A T_B$ is periodic, which contradicts to $\langle T_A, T_B \rangle \cong F_2$. \square

As an application of Lemma 2.12, we get the following classification of elements of $\langle T_A, T_B \rangle \cong F_2$ by Leininger [Lei04, Proposition 6.4].

Theorem 2.13. *Let A and B be filling multicurves on S with $\langle T_A, T_B \rangle \cong F_2$. If $\mu(A, B) > 4$, then*

$$\varphi \in \langle T_A, T_B \rangle \quad \Leftrightarrow \quad \begin{array}{l} \#\{g \in \langle T_A, T_B \rangle \text{ such that} \\ \text{is pseudo-Anosov}\} \end{array} \quad \begin{array}{l} \#\{g \in \langle T_A, T_B \rangle \text{ such that} \\ g\varphi g^{-1} \in \langle T_A \rangle \cup \langle T_B \rangle\} \end{array}.$$

Otherwise (i.e., $\mu(A, B) = 4$), we have

$$\varphi \in \langle T_A, T_B \rangle \Leftrightarrow \begin{array}{l} \#\text{g} \in \langle T_A, T_B \rangle \text{ such that} \\ \text{is pseudo-Anosov} \end{array} \quad g\varphi g^{-1} \in \langle T_A \rangle \cup \langle T_B \rangle \cup \langle T_A T_B \rangle.$$

Proof. According to Thurston's construction, it suffices to classify elements of $\rho(\langle T_A, T_B \rangle)$. To do this, let us consider its fundamental domain. Since $\mu \geq 4$ by Lemma 2.12, there are two cases:

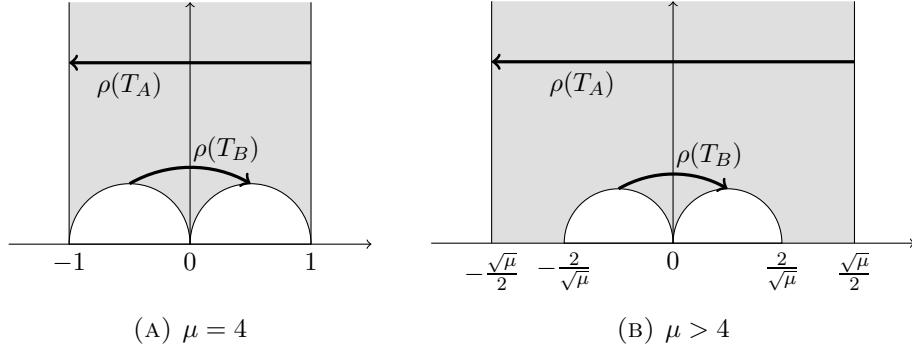


FIGURE 3. Fundamental domains of $\rho(\langle T_A, T_B \rangle)$

Case 1: $\mu = 4$. Its fundamental domain is as in Figure 3a. Accordingly, the quotient $\mathbb{H}^2/\rho(\langle T_A, T_B \rangle)$ is a thrice-punctured sphere. Since hyperbolic elements of $\rho(\langle T_A, T_B \rangle)$ correspond to loops in $\mathbb{H}^2/\rho(\langle T_A, T_B \rangle)$ which are freely homotopic to closed geodesics, the theorem follows in this case.

Case 2: $\mu > 4$. Fundamental domain for this case is as in Figure 3b, which implies that the quotient $\mathbb{H}^2/\rho(\langle T_A, T_B \rangle)$ is a twice-punctured disk. Again, by the same arguments, all elements of $\rho(\langle T_A, T_B \rangle)$ but conjugates of powers of T_A and T_B are hyperbolic. \square

We explore the properties of $\langle \rho(T_A), \rho(T_B) \rangle$ in detail. Given a pivot fundamental domain depicted in Figure 3a and 3b, we consider a tiling of \mathbb{H}^2 with its translations by $\langle \rho(T_A), \rho(T_B) \rangle$ as in Figure 4. This tiling has the Cayley graph \mathcal{T} of $\langle \rho(T_A), \rho(T_B) \rangle \cong F_2$ as a dual tree (blue and red lines in the figure). Each edge of \mathcal{T} is colored and directed, and each vertex is labelled with a reduced word of alphabets $\{a^\pm, b^\pm\}$. The labelling rule of vertices is as follows: if \overrightarrow{vw} is the blue (red, resp.) directed edge, w is represented by concatenating a to v from the right.

We now make a 1-1 correspondence among the set $V(\mathcal{T})$ of vertices of \mathcal{T} , the set $\text{Isom}(\mathcal{T})$ of isometries of \mathcal{T} , and the Deck transformation group $\langle \rho(T_A), \rho(T_B) \rangle \leq \text{Isom}^+(\mathbb{H}^2)$. For each reduced word w of $\{a^\pm, b^\pm\}$, there exists a cyclically reduced word v such that $w = cvc^{-1}$ for some word c . Then the translation $t(w)$ of \mathcal{T} along the axis connecting $\{cv^n\}_{n \in \mathbb{Z}}$ by v is the unique isometry of \mathcal{T} sending 0 to w . The corresponding word w' of $\{T_A^\pm, T_B^\pm\}$ determines the isometry $\rho(w')$ of \mathbb{H}^2 .

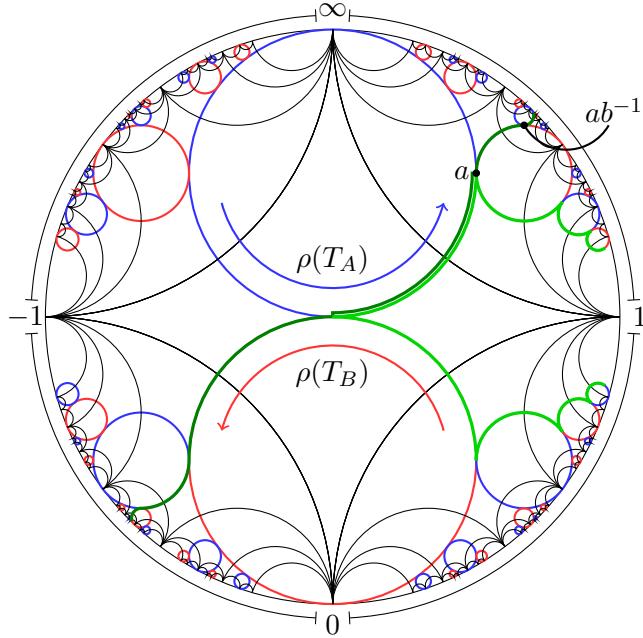


FIGURE 4. Fundamental domains for $\langle \rho(T_A), \rho(T_B) \rangle$ when $\mu = 4$ and the Cayley graph \mathcal{T} of F_2 as a dual tree. Blue lines correspond to the (right) cosets of $\rho(T_A)$ and red lines correspond to the (right) cosets of $\rho(T_B)$. The invariant axis of an element in $\text{Isom}(\mathcal{T})$ sending 0 to ab^{-1} (ab , resp.) is shown as the dark (light, resp.) green line.

The axes of $t(w)$ and $\rho(w')$ are homotopic relative to $\partial\mathbb{H}^2 = S^1$. When w is a conjugate of a power of a or b (a , b , or ab in the case of $\mu = 4$), the entire \mathcal{T} is on the left/right side of the axis of $t(w)$, and $\rho(w')$ is parabolic. If not, $\rho(w')$ is hyperbolic. In this case, the edges of fundamental domains intersecting the axis of $t(w)$ and those intersecting the axis of $\rho(w')$ coincide.

Let us now consider a hyperbolic word $w = cvc^{-1}$ with cyclically reduced form $v = \alpha_1\alpha_2 \cdots \alpha_n$ (α_i 's are alphabets of v). Here, n is called *cyclically reduced word norm* of w . For each $l \in \mathbb{Z}$ and $k \in \{1, \dots, n\}$, the invariant axis of $\rho(w')$ passes through the edge of the tiling that corresponds to $cv^l\alpha_1 \cdots \alpha_k$. Let us denote this edge by e_{ln+k} . We then define a collection \mathcal{J} of indices j such that:

- $e_{j-2} \rightarrow e_{j-1}$ was a left turn but $e_{j-1} \rightarrow e_j$ is a right turn, or
- $e_{j-2} \rightarrow e_{j-1}$ was a right turn but $e_{j-1} \rightarrow e_j$ is a left turn, or
- $e_{j-1} \rightarrow e_j$ is a diagonal move.

See Figure 5 for details. Since $\rho(w')$ is not parabolic, \mathcal{J} is not empty; it is periodic with a period n .

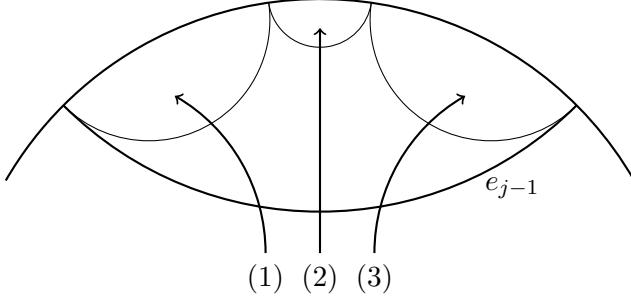


FIGURE 5. Consecutive edges passed through by the invariant axis. (1) left turn, (2) diagonal move, and (3) right turn.

We now pick arbitrary $j_0 \in \mathcal{J}$, and inductively collect $j_i \in \mathcal{J}$ by setting $j_{i+1} = \inf\{j \in \mathcal{J} : j \geq j_i + 2\}$ to form \mathcal{J}' . Then \mathcal{J}' is also nonempty and is eventually periodic with a period $2n$. Suppose that $j_1, \dots, j_m = j_1 + 2n$ form a period of length $2n$.

Let us estimate $d_{\mathbb{H}^2}(e_{j_k}, e_{j_{k+1}})$ in the following two cases:

- (1) $j_{k+1} = j_k + 2$: then $e_{j_k} \rightarrow e_{j_{k+1}} \rightarrow e_{j_{k+1}}$ is either a left turn after a right turn, a right turn after a left turn, or a diagonal move after any move. In either case, $d_{\mathbb{H}^2}(e_{j_k}, e_{j_{k+1}}) \geq \frac{1}{2} \log(3 + 2\sqrt{2})$.
- (2) $j_{k+1} \geq j_k + 3$: then $j \notin \mathcal{J}$ for $j_k + 2 \leq j < j_{k+1}$. This implies that $e_{j_{k+1}} \rightarrow e_{j_{k+2}}$ is either left or right turn, and $e_{j_{k+1}} \rightarrow e_{j_{k+2}} \rightarrow \dots \rightarrow e_{j_{k+1}-1}$ are turns of the same type, but $e_{j_{k+1}-1} \rightarrow e_{j_{k+1}}$ is not in that type. Then the distance between e_{j_k} and $e_{j_{k+1}}$ is at least $\frac{1}{2} \log(l + \sqrt{l^2 - 1})$ where $l = 2(j_{k+1} - j_k - 2) + 1 \geq j_{k+1} - j_k$.

As a result, the distance between e_{j_k} and $e_{j_{k+1}}$ is greater than $1/2$ and $\frac{1}{2} \log(j_{k+1} - j_k)$. This gives the estimation

$$d_{\mathbb{H}^2}(e_{j_1}, e_{j_m}) \geq \sum_{k=1}^{m-1} d_{\mathbb{H}^2}(e_{j_k}, e_{j_{k+1}}) \geq \frac{1}{2} \sum_{k=1}^{m-1} \log(j_{k+1} - j_k).$$

Since $j_{k+1} - j_k \geq 2$ and $f(x, y) = \log(x+y) - \log x - \log y$ has negative partial derivatives with $f(2, 2) = 0$, $\sum_{k=1}^{m-1} \log(j_{k+1} - j_k) \geq \log \left(\sum_{k=1}^{m-1} (j_{k+1} - j_k) \right)$. From $j_m = j_1 + 2n$, we conclude that

$$d_{\mathbb{H}^2}(e_{j_1}, e_{j_m}) \geq \frac{1}{2} \log(2n).$$

Note that the translation length of $\rho(w')$ is equal to $\log \lambda_{\rho(w')}$ and is greater than $\frac{1}{2} d_{\mathbb{H}^2}(e_{j_1}, e_{j_1+2n})$. As such, we deduce the following proposition.

Proposition 2.14. *Let A and B be filling multicurves on S such that $\langle T_A, T_B \rangle \cong F_2$. Then for a pseudo-Anosov $\varphi \in \langle T_A, T_B \rangle$ and its stretch*

factor λ_φ , we have

$$\log \lambda_\varphi \geq \frac{1}{4} \log |\varphi|$$

where $|\cdot|$ is the cyclically reduced word norm in terms of T_A^\pm and T_B^\pm .

3. ASYMPTOTIC BEHAVIOR OF RANDOM WALKS OBTAINED FROM THURSTON'S CONSTRUCTION

3.1. Introduction to the theory of topological entropy. Before stating the main result, let us introduce the theory of topological entropy. The topological entropy, defined below, is a quantity representing the dynamics of a continuous map $f : X \rightarrow X$ on a compact Hausdorff space X .

Definition 3.1 (Topological entropy, [AKM65]). *Let X be a compact Hausdorff topological space, and $f : X \rightarrow X$ be a continuous map. For any open cover \mathcal{U} of X , let us denote the minimal size of subcover of \mathcal{U} by $N(\mathcal{U})$, which is a finite integer. Then, the following limit exists:*

$$H(f, \mathcal{U}) := \lim_{n \rightarrow \infty} \frac{1}{n} \log N(\mathcal{U} \vee f^{-1}(\mathcal{U}) \vee \cdots \vee f^{-(n-1)}(\mathcal{U}))$$

where $\mathcal{A} \vee \mathcal{B} := \{A \cap B : A \in \mathcal{A}, B \in \mathcal{B}\}$. The topological entropy of f is defined as

$$h(f) := \sup_{\mathcal{U}} H(f, \mathcal{U}).$$

In particular, one can consider the topological entropy of a self-homeomorphism φ of a surface S . Regarding this, Thurston proved the following.

Theorem 3.2 (Thurston, [FLP79, Exposé 10]). *For a pseudo-Anosov homeomorphism $\varphi : S \rightarrow S$ of a surface S ,*

$$h(\varphi) = \log \lambda_\varphi$$

where λ_φ is its stretch factor. Moreover, it minimizes the topological entropy in its isotopy class.

Furthermore, as $\text{Mod}(S)$ acts on the Teichmüller space $\mathcal{T}(S)$ of S , one can ask the dynamics of pseudo-Anosovs on S appear in $\mathcal{T}(S)$, in the perspective of the topological entropy. Bers' classification of mapping classes according to the action on $\mathcal{T}(S)$ helps to answer this question.

Theorem 3.3 (Bers, [B⁺78]). *Let $\varphi : S \rightarrow S$ be a pseudo-Anosov homeomorphism. Then the translation length of its action on $\mathcal{T}(S)$ is realized, which is equal to $\log \lambda_\varphi$. In other words, there exists $\mathcal{X}_\varphi \in \mathcal{T}(S)$ such that*

$$\log \lambda_\varphi = d_{\mathcal{T}}(\mathcal{X}_\varphi, \varphi \cdot \mathcal{X}_\varphi) = \inf_{\mathcal{Y} \in \mathcal{T}(S)} d_{\mathcal{T}}(\mathcal{Y}, \varphi \cdot \mathcal{Y})$$

where $d_{\mathcal{T}}$ is the Teichmüller distance on $\mathcal{T}(S)$.

Here \mathcal{X}_φ lies on the invariant geodesic Γ_φ of φ in $\mathcal{T}(S)$, on which φ acts as a translation. Combining Theorem 3.2 and Theorem 3.3 yields the following.

Corollary 3.4. *Let $\varphi : S \rightarrow S$ be a pseudo-Anosov homeomorphism. Then for any $\mathcal{X} \in \mathcal{T}(S)$, we have*

$$h(\varphi) = \log \lambda_\varphi = \lim_{n \rightarrow \infty} \frac{1}{n} d_{\mathcal{T}}(\mathcal{X}, \varphi^n \cdot \mathcal{X}).$$

Proof. It suffices to prove the second equality, since the first equality follows from Theorem 3.2. For any $\mathcal{X} \in \mathcal{T}(S)$ and $n \in \mathbb{N}$,

$$n \log \lambda_\varphi = \log \lambda_{\varphi^n} \leq d_{\mathcal{T}}(\mathcal{X}, \varphi^n \cdot \mathcal{X})$$

by Theorem 3.3. Accordingly, we obtain

$$\log \lambda_\varphi \leq \liminf_{n \rightarrow \infty} \frac{1}{n} d_{\mathcal{T}}(\mathcal{X}, \varphi^n \cdot \mathcal{X}).$$

To prove the opposite inequality, we observe

$$d_{\mathcal{T}}(\mathcal{X}, \varphi^n \cdot \mathcal{X}) \leq d_{\mathcal{T}}(\mathcal{X}, \mathcal{X}_\varphi) + d_{\mathcal{T}}(\mathcal{X}_\varphi, \varphi^n \cdot \mathcal{X}_\varphi) + d_{\mathcal{T}}(\varphi^n \cdot \mathcal{X}_\varphi, \varphi^n \cdot \mathcal{X})$$

where $\mathcal{X}_\varphi \in \Gamma_\varphi$ realizes the translation length as in Theorem 3.3. Since φ acts on $\mathcal{T}(S)$ as an isometry and acts on Γ_φ as a translation, we have

$$\begin{aligned} d_{\mathcal{T}}(\mathcal{X}, \varphi^n \cdot \mathcal{X}) &\leq 2d_{\mathcal{T}}(\mathcal{X}, \mathcal{X}_\varphi) + nd_{\mathcal{T}}(\mathcal{X}_\varphi, \varphi \cdot \mathcal{X}_\varphi) \\ &= 2d_{\mathcal{T}}(\mathcal{X}, \mathcal{X}_\varphi) + n \log \lambda_\varphi \end{aligned}$$

Therefore, we get

$$\limsup_{n \rightarrow \infty} \frac{1}{n} d_{\mathcal{T}}(\mathcal{X}, \varphi^n \cdot \mathcal{X}) \leq \log \lambda_\varphi$$

which completes the proof. \square

From Definition 3.1 and Corollary 3.4, the topological entropy of a homeomorphism φ indicates the dynamical behavior of the sequence of iterations $\varphi, \varphi^2, \dots, \varphi^n, \dots$. Recall that a pseudo-Anosov map φ minimizes the topological entropy in its isotopy class $[\varphi]$. From now on, we define the topological entropy of $[\varphi]$ by that of φ . In this regard, the notion of topological entropy can be generalized by considering the dynamical behavior of a sequence $\varphi_1, \varphi_2, \dots, \varphi_n, \dots$ in $\text{Mod}(S)$.

Definition 3.5 (Topological entropy of a sequence in $\text{Mod}(S)$, [Mas18]). *Let $\varphi = (\varphi_n) \in \text{Mod}(S)^{\mathbb{N}}$ be a sequence of mapping classes and $\phi_n \in \text{Diff}^+(S)$ be a representative of φ_n for each $n \in \mathbb{N}$. Let $\phi = (\phi_n)$ be the sequence of the chosen representatives.*

For an open cover \mathcal{U} , we define

$$H(\phi, \mathcal{U}) := \limsup_{n \rightarrow \infty} \frac{1}{n} \log N(\mathcal{U} \vee \phi_1(\mathcal{U}) \vee \dots \vee \phi_{n-1}(\mathcal{U}))$$

where $N(\cdot)$ and \vee are defined as in Definition 3.1. Then we define the topological entropy of ϕ by

$$h(\phi) := \sup_{\mathcal{U}} H(\phi, \mathcal{U}).$$

Finally, the topological entropy of φ is defined by

$$h(\varphi) := \inf_{\phi} h(\phi).$$

Remark 3.6. For a pseudo-Anosov mapping class, its iteration sequence has the same topological entropy as its pseudo-Anosov representative.

With this definition, Masai proved a random version of Corollary 3.4.

Theorem 3.7 (Masai, [Mas18]). Let ν be a probability measure on $\text{Mod}(S)$ such that

- ν has a finite first moment with respect to the Teichmüller metric on $\mathcal{T}(S)$, i.e. for any $\mathcal{X} \in \mathcal{T}(S)$,

$$\sum_{g \in \text{Mod}(S)} \nu(g) d_{\mathcal{T}}(\mathcal{X}, g \cdot \mathcal{X}) < \infty,$$

and

- ν is non-elementary, i.e. $\langle \text{supp } \nu \rangle$ is a non-elementary subgroup of $\text{Mod}(S)$.

Let \mathbb{P} be the probability measure on $\text{Mod}(S)^{\mathbb{N}}$ induced by ν . Then for \mathbb{P} -a.e. $\omega = (\omega_n) \in \text{Mod}(S)^{\mathbb{N}}$ and any $\mathcal{X} \in \mathcal{T}(S)$,

$$h(\omega) = \lim_{n \rightarrow \infty} \frac{1}{n} d_{\mathcal{T}}(\mathcal{X}, \omega_n \cdot \mathcal{X})$$

and it does not depend on $\omega \in \text{Mod}(S)^{\mathbb{N}}$. In other words, $h(\omega)$ is almost surely constant.

One way to investigate the dynamical property of a mapping class is to observe its action on each simple closed curve on the ambient surface. Let us equip the surface S with a Riemannian metric; $l(\gamma)$ denotes the length of the geodesic representative of a closed curve γ . Then for a pseudo-Anosov $\varphi \in \text{Mod}(S)$ and a simple closed curve α ,

$$h(\varphi) = \lim_{n \rightarrow \infty} \frac{1}{n} \log l(\varphi^{-n} \alpha).$$

For details, see [FM12, Theorem 14.23]. Note that this does not depend on the choice of α and the Riemannian metric.

The following theorem of Karlsson, together with Theorem 3.7, asserts that the topological entropy of a random walk has a similar property.

Theorem 3.8 (Karlsson, [Kar14]). There exists λ , called the Lyapunov exponent of the random walk, such that for \mathbb{P} -a.e. $\omega = (\omega_n) \in \text{Mod}(S)^{\mathbb{N}}$, for any simple closed curve α and a Riemannian metric on S ,

$$\log \lambda = \lim_{n \rightarrow \infty} \frac{1}{n} \log l(\omega_n^{-1} \alpha)$$

holds. Moreover, for any $\mathcal{X} \in \mathcal{T}(S)$,

$$\log \lambda = \lim_{n \rightarrow \infty} \frac{1}{n} d_{\mathcal{T}}(\mathcal{X}, \omega_n \cdot \mathcal{X}).$$

3.2. Spectral theorem for random walks on $\text{Mod}(S)$. So far, we dealt with the topological entropy of a single mapping class and that of a random walk. The former contains both topological information and dynamical information; for pseudo-Anosov $\varphi \in \text{Mod}(S)$, $h(\varphi)$ is the log of its stretch factor and $h(\varphi) = \lim_{n \rightarrow \infty} \frac{1}{n} d_{\mathcal{T}}(\mathcal{X}, \varphi^n \cdot \mathcal{X})$ for any $\mathcal{X} \in \mathcal{T}(S)$. The latter also describes the dynamics of a random walk in a similar way; $h(\omega) = \lim_{n \rightarrow \infty} \frac{1}{n} d_{\mathcal{T}}(\mathcal{X}, \omega_n \cdot \mathcal{X})$ for any $\mathcal{X} \in \mathcal{T}(S)$ under certain condition. With this similarity, it is natural to ask how both topological entropies are related. More precisely, for a random walk $\omega = (\omega_n) \in \text{Mod}(S)^{\mathbb{N}}$, how do $h(\omega_n)$ and $h(\omega)$ interact? Dahmani and Horbez answered this in the following spectral theorem when the underlying measure is finitely supported.

Theorem 3.9 (Dahmani–Horbez, [DH18]). *Let ν be a non-elementary probability measure on $\text{Mod}(S)$ with finite support, and \mathbb{P} be the induced measure on $\text{Mod}(S)^{\mathbb{N}}$. Then for \mathbb{P} -a.e. $\omega \in \text{Mod}(S)^{\mathbb{N}}$, ω_n is eventually pseudo-Anosov. Moreover, the stretch factor λ_{ω_n} of a pseudo-Anosov ω_n satisfies*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \lambda_{\omega_n} = L_{\mathcal{T}} > 0$$

where the constant $L_{\mathcal{T}} := \lim_{n \rightarrow \infty} \frac{1}{n} d_{\mathcal{T}}(\mathcal{X}, \omega_n \cdot \mathcal{X})$ is the drift of ω with respect to the Teichmüller metric $d_{\mathcal{T}}$.

Together with Theorem 3.7 and Theorem 3.2, the above theorem concludes that $\frac{1}{n} h(\omega_n) \rightarrow h(\omega)$ as $n \rightarrow \infty$. In other words, the linear growth rate of the topological entropies of random mapping classes equals the topological entropy of the random walk.

Not only for finitely supported measure, Dahmani and Horbez also asserted in [DH18, Remark 2.7, 3.2] that the same result holds under the finite second moment condition. For the sake of completeness, we provide a detailed proof of this assertion in the setting of Thurston's construction.

Theorem A. *Let (A, B) be a filling pair of multicurves on S and ν be a non-elementary probability measure on $\langle T_A, T_B \rangle$. Suppose that ν has finite second moment with respect to the Teichmüller metric, i.e., for some $\mathcal{X} \in \mathcal{T}(S)$,*

$$\sum_{g \in \text{Mod}(S)} d_{\mathcal{T}}(\mathcal{X}, g \cdot \mathcal{X})^2 \nu(g) < \infty.$$

Then for \mathbb{P} -a.e. $\omega \in \text{Mod}(S)^{\mathbb{N}}$, ω_n is eventually pseudo-Anosov and

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \lambda_{\omega_n} = h(\omega).$$

Before proving this theorem, we briefly review the Busemann boundary of hyperbolic spaces: see [BQ16b] for details. For a proper quasiconvex Gromov hyperbolic space (M, d) and a reference point $\mathcal{X} \in M$, the *Busemann compactification* M_B consists of functions $h_x : M \rightarrow \mathbb{R}$ of the form

$$h_x(m) := \lim_{n \rightarrow \infty} d(m, x_n) - d(\mathcal{X}, x_n)$$

that are well-defined on M , where $x = (x_n)$ is a sequence in M . From the definition, we observe that $h_x(m) = d(m, x) - d(\mathcal{X}, x)$ for a convergent sequence $x_n \rightarrow x \in M$ and $h_x = h_y$ implies $x = y$ for $x, y \in M$. As such, it is possible to regard M as a subspace of M_B ; indeed, M_B is a compactification of M under a suitable topology. Then we define the *Busemann boundary* of M as $\partial_B M := M_B \setminus M$.

We now compare the Busemann compactification M_B of M with another compactification M_G , called the *Gromov compactification*. Since a bounded sequence in M is convergent if it defines a point in the Busemann compactification, there exists a natural map $\pi : M_B \rightarrow M_G$ which is the identity on M , and this is indeed surjective.

We can extend the Gromov product on M to M_B , by defining

$$(m|x)_{\mathcal{X}} = (x|m)_{\mathcal{X}} := \frac{1}{2} (d(m, \mathcal{X}) - h_x(m)) \text{ and}$$

$$(x|x')_{\mathcal{X}} := - \inf_{m \in M} \frac{1}{2} (h_x(m) + h_{x'}(m))$$

for $m \in M$ and $x, x' \in \partial_B M$. For $x = (x_n)$, the first formula reads

$$(m|x)_{\mathcal{X}} = \lim_{n \rightarrow \infty} (m|x_n)_{\mathcal{X}}.$$

Lemma 3.10 (Lemma 5.10 of [MT18]). *Let M be a Gromov hyperbolic space and $\mathcal{X} \in M$. Then there exist $K, K' > 0$ satisfying the following. For $a, b, c, d \in M$, if $(a|b)_{\mathcal{X}}, (c|d)_{\mathcal{X}} \geq A$ and $(a|c)_{\mathcal{X}} \leq A - K$ holds for some A , then $|(a|c)_{\mathcal{X}} - (b|d)_{\mathcal{X}}| \leq K'$.*

We are now ready to prove Theorem A:

Proof of the theorem. Thurston's construction in [T⁺88] describes how a Teichmüller disk $f : \mathbb{H}^2 \rightarrow \mathcal{T}(S)$ is determined from the pair (A, B) (cf. [Lei04], [HS07]). Elements in $\langle T_A, T_B \rangle$ stabilize $f(\mathbb{H}^2)$ and act on $f(\mathbb{H}^2)$ as isometries. Thus, the isometric embedding f induces a representation $\rho : \langle T_A, T_B \rangle \rightarrow \text{PSL}(2, \mathbb{R}) = \text{Isom}^+(\mathbb{H}^2)$. Throughout, $O(\delta)$ denotes a constant solely depending on the hyperbolicity δ of \mathbb{H}^2 . $d(\cdot, \cdot)$ denotes the hyperbolic distance of \mathbb{H}^2 , which induces a Gromov product $(\cdot|\cdot)$.

We now estimate the translation length of a random mapping class ω_n viewed as an isometry of \mathbb{H}^2 . From the work [MT18] of Maher and Tiozzo, the translation length can be estimated as follows:

Lemma 3.11 (Maher–Tiozzo). *There exists a constant $C_0 > 0$ such that for any isometry g of \mathbb{H}^2 , if $d(\mathcal{X}, g \cdot \mathcal{X}) \geq 2(g \cdot \mathcal{X} | g^{-1} \cdot \mathcal{X})_{\mathcal{X}} + C_0$, then*

$$|(\text{translation length of } g) - (d(\mathcal{X}, g \cdot \mathcal{X}) - 2(g \cdot \mathcal{X} | g^{-1} \cdot \mathcal{X})_{\mathcal{X}})| \leq O(\delta).$$

In view of Lemma 3.11, we should estimate $d(\mathcal{X}, \omega_n \cdot \mathcal{X})$ and control the term $(\omega_n \cdot \mathcal{X} | \omega_n^{-1} \cdot \mathcal{X})_{\mathcal{X}}$ in order to obtain the translation length of ω_n . We do this by referring to [BQ16b, Proposition 4.1].

Lemma 3.12 (Benoist–Quint). *Let $p > 1$, (M, d) be a proper quasiconvex Gromov hyperbolic space, and $\mathcal{X} \in M$. Let ν be a non-elementary Borel probability measure on $G \leq \text{Isom}^+(M)$ such that $\int_G d(\mathcal{X}, g \cdot \mathcal{X})^p d\nu(g) < \infty$.*

Then $\lambda := \lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} d(\mathcal{X}, \omega_n \cdot \mathcal{X}) > 0$. Moreover, for every $\varepsilon > 0$, there exist constants C_n such that $\sum_{n \geq 1} n^{p-2} C_n < \infty$ and for any $x \in \partial_B M$,

$$\mathbb{P}(|h_x(\omega_n^{-1} \cdot \mathcal{X}) - n\lambda| > \varepsilon n) \leq C_n, \text{ and}$$

$$\mathbb{P}(|d(\mathcal{X}, \omega_n \cdot \mathcal{X}) - n\lambda| > \varepsilon n) \leq C_n.$$

Since $\langle T_A, T_B \rangle$ is discrete ([T⁺88], [Lei04]), ν and $\hat{\nu} := \nu(\cdot^{-1})$ are Borel measures and we can apply Lemma 3.12. Note that λ is the drift of ω .

For each $\varepsilon > 0$, we now obtain a summable sequence $C_{n;\varepsilon}$ such that

$$\mathbb{P}(|h_x(\omega_n^{-1} \cdot \mathcal{X}) - n\lambda| > \varepsilon n) \leq C_{n;\varepsilon}, \text{ and}$$

$$\mathbb{P}(|d(\mathcal{X}, \omega_n^{-1} \cdot \mathcal{X}) - n\lambda| > \varepsilon n) \leq C_{n;\varepsilon}.$$

Since $(\omega_n^{-1} \cdot \mathcal{X} | x)_{\mathcal{X}} = \frac{1}{2} (d(\mathcal{X}, \omega_n^{-1} \cdot \mathcal{X}) - h_x(\omega_n^{-1} \cdot \mathcal{X}))$, we have

$$\begin{aligned} \mathbb{P}((\omega_n^{-1} \cdot \mathcal{X} | x)_{\mathcal{X}} > \varepsilon n) &\leq \mathbb{P}(|d(\mathcal{X}, \omega_n^{-1} \cdot \mathcal{X}) - n\lambda| > \varepsilon n) \\ &\quad + \mathbb{P}(|h_x(\omega_n^{-1} \cdot \mathcal{X}) - n\lambda| > \varepsilon n) \\ &\leq 2C_{n;\varepsilon}. \end{aligned}$$

A similar argument shows that $\mathbb{P}((\omega_n \cdot \mathcal{X} | x)_{\mathcal{X}} > \varepsilon n) < 2C_{n;\varepsilon}$.

Now let $0 < \varepsilon < \lambda/10$. We set $m = m(n) = \lceil n/2 \rceil$ and $u_{m(n)} := \omega_m^{-1} \omega_n = g_{m+1} \cdots g_n$. Note that $\omega_{m(n)}$ and $u_{m(n)}$ are independent for each $n \in \mathbb{N}$. Note also that $u_{m(n)}$ and $\omega_{n-m(n)}$ have the same distribution.

Using the independence, we deduce

$$\mathbb{P}((u_m^{-1} \cdot \mathcal{X} | \omega_m \cdot \mathcal{X})_{\mathcal{X}} > \varepsilon n) = \sum_{g \in \text{Mod}(S)} \mathbb{P}(u_m^{-1} = g) \mathbb{P}((g \cdot \mathcal{X} | \omega_m \cdot \mathcal{X})_{\mathcal{X}} > \varepsilon n).$$

If $g = id$, then $(g \cdot \mathcal{X} | \omega_m \cdot \mathcal{X})_{\mathcal{X}} = 0$. For $g \neq id$, we consider the half-geodesic $\Gamma = \overrightarrow{\mathcal{X} x}$ from \mathcal{X} to some $x \in \partial_B \mathbb{H}^2$ such that $g \mathcal{X} \in \Gamma$. Let $\{\mathcal{X}_i\}$ be a sequence on Γ tending to x , where $\mathcal{X}_1 = g \mathcal{X}$. For each i , we obtain

$$\begin{aligned} (\mathcal{X}_i | \omega_m \cdot \mathcal{X})_{\mathcal{X}} &= \frac{1}{2} [d(\mathcal{X}, \mathcal{X}_i) + d(\mathcal{X}, \omega_m \cdot \mathcal{X}) - d(\mathcal{X}_i, \omega_m \cdot \mathcal{X})] \\ &\geq \frac{1}{2} [(d(\mathcal{X}, \mathcal{X}_1) + d(\mathcal{X}_1, \mathcal{X}_i)) + d(\mathcal{X}, \omega_m \cdot \mathcal{X}) - (d(\mathcal{X}_1, \omega_m \cdot \mathcal{X}) + d(\mathcal{X}_1, \mathcal{X}_i))] \\ &= \frac{1}{2} [d(\mathcal{X}, \mathcal{X}_1) + d(\mathcal{X}, \omega_m \cdot \mathcal{X}) - d(\mathcal{X}_1, \omega_m \cdot \mathcal{X})] \\ &= (g \cdot \mathcal{X} | \omega_m \cdot \mathcal{X})_{\mathcal{X}}. \end{aligned}$$

This implies that

$$\mathbb{P}((g \cdot \mathcal{X} | \omega_m \cdot \mathcal{X})_{\mathcal{X}} > \varepsilon n) \leq \mathbb{P}((x | \omega_m \cdot \mathcal{X})_{\mathcal{X}} > \varepsilon n) \leq 2C_{m;\varepsilon}.$$

From this and the fact that $\sum_{g \in \text{Mod}(S)} \mathbb{P}(u_m^{-1} = g) = 1$, we conclude that

$$(3.1) \quad \mathbb{P}((u_m^{-1} \cdot \mathcal{X} | \omega_m \cdot \mathcal{X})_{\mathcal{X}} > \varepsilon n) \leq 1 \cdot 2C_{m;\varepsilon}.$$

Next, we calculate

$$\begin{aligned}
& \mathbb{P}((\omega_m \cdot \mathcal{X} | \omega_n \cdot \mathcal{X})_{\mathcal{X}} < (\lambda - \varepsilon)m) \\
&= \sum_{g \in \text{Mod}(S)} \mathbb{P}(\omega_m = g) \mathbb{P}((g \cdot \mathcal{X} | g\omega_m \cdot \mathcal{X})_{\mathcal{X}} < (\lambda - \varepsilon)m) \\
&\leq \sum_{\substack{d(\mathcal{X}, g \cdot \mathcal{X}) \\ \leq (\lambda - \varepsilon/2)m}} \mathbb{P}(\omega_m = g) + \sum_{\substack{d(\mathcal{X}, g \cdot \mathcal{X}) \\ > (\lambda - \varepsilon/2)m}} \mathbb{P}(\omega_m = g) \mathbb{P}((g \cdot \mathcal{X} | g\omega_m \cdot \mathcal{X})_{\mathcal{X}} < (\lambda - \varepsilon)m) \\
&= \mathbb{P}\left(d(\mathcal{X}, \omega_m \cdot \mathcal{X}) \leq \left(\lambda - \frac{\varepsilon}{2}\right)m\right) \\
&\quad + \sum_{\substack{d(\mathcal{X}, g \cdot \mathcal{X}) \\ > (\lambda - \varepsilon/2)m}} \mathbb{P}(\omega_m = g) \mathbb{P}((\mathcal{X} | \omega_{n-m} \cdot \mathcal{X})_{g^{-1}\mathcal{X}} < (\lambda - \varepsilon)m).
\end{aligned}$$

The first term is bounded by $C_{m; \varepsilon/2}$. To deal with the second term, we fix g with $d(\mathcal{X}, g \cdot \mathcal{X}) > (\lambda - \varepsilon/2)m$. Let $\Gamma = \overrightarrow{\mathcal{X}x}$ be the half-geodesic from \mathcal{X} to some $x \in \partial_B \mathbb{H}^2$ such that $g^{-1}\mathcal{X} \in \Gamma$. Let $\{\mathcal{X}_i\}$ be a sequence on Γ tending to x , where $\mathcal{X}_1 = g^{-1}\mathcal{X}$. If $(\mathcal{X} | \omega_{n-m} \cdot \mathcal{X})_{g^{-1}\mathcal{X}} < (\lambda - \varepsilon/2)m$, then

$$\begin{aligned}
(\mathcal{X}_i | \omega_{n-m} \cdot \mathcal{X})_{\mathcal{X}} &\geq (g^{-1} \cdot \mathcal{X} | \omega_{n-m} \cdot \mathcal{X})_{\mathcal{X}} \\
&= d(g^{-1} \cdot \mathcal{X}, \mathcal{X}) - (\mathcal{X} | \omega_{n-m} \cdot \mathcal{X})_{g^{-1}\mathcal{X}} \\
&\geq \left(\lambda - \frac{\varepsilon}{2}\right)m - (\lambda - \varepsilon)m \\
&= \frac{\varepsilon}{2}m.
\end{aligned}$$

Thus, we deduce

$$\mathbb{P}((\mathcal{X} | \omega_{n-m} \cdot \mathcal{X})_{g^{-1}\mathcal{X}} < (\lambda - \varepsilon)m) \leq \mathbb{P}((\mathcal{X} | \omega_{n-m} \cdot \mathcal{X})_{\mathcal{X}} \geq \varepsilon m/2) \leq 2C_{m; \varepsilon/2}.$$

From this we conclude that

$$(3.2) \quad \mathbb{P}((\omega_m \cdot \mathcal{X} | \omega_n \cdot \mathcal{X})_{\mathcal{X}} < (\lambda - \varepsilon)m) \leq C_{m; \varepsilon/2} + 1 \cdot 2C_{m; \varepsilon/2}.$$

A similar argument implies

$$(3.3) \quad \mathbb{P}((u_m^{-1} \cdot \mathcal{X} | \omega_n^{-1} \cdot \mathcal{X})_{\mathcal{X}} < (\lambda - \varepsilon)m) \leq 3C_{m; \varepsilon/2}.$$

The probability of the events in Equation 3.1, 3.2 and 3.3 is summable. Thus, by the Borel–Cantelli lemma, \mathbb{P} –a.e. ω satisfies

$$\begin{aligned}
(u_m^{-1} \cdot \mathcal{X} | \omega_m \cdot \mathcal{X})_{\mathcal{X}} &\leq \varepsilon n, \\
(\omega_m \cdot \mathcal{X} | \omega_n \cdot \mathcal{X})_{\mathcal{X}} &\geq (\lambda - \varepsilon)m, \text{ and} \\
(u_m^{-1} \cdot \mathcal{X} | \omega_n^{-1} \cdot \mathcal{X})_{\mathcal{X}} &\geq (\lambda - \varepsilon)m
\end{aligned}$$

for sufficiently large n . Then Lemma 3.10 asserts that $(\omega_n^{-1} \cdot \mathcal{X} | \omega_n \cdot \mathcal{X})_{\mathcal{X}} = (u_m^{-1} \cdot \mathcal{X} | \omega_m \cdot \mathcal{X})_{\mathcal{X}} + O(\delta) \leq 2\varepsilon n$ for sufficiently large n .

We now return to Lemma 3.11. There exists an event Ω_k for $\varepsilon = \lambda/(100k)$ such that $\mathbb{P}(\Omega_k) = 1$ and $\omega \in \Omega_k$ satisfies

$$|d(\mathcal{X}, \omega_n \cdot \mathcal{X}) - \lambda n| < \varepsilon n \text{ and } (\omega_n^{-1} \cdot \mathcal{X} | \omega_n \cdot \mathcal{X})_{\mathcal{X}} \leq 2\varepsilon n$$

for sufficiently large n . Thus, by Lemma 3.11,

$$\begin{aligned} \left| \frac{1}{n}(\text{translation length of } \omega_n) - \lambda \right| &\leq \frac{1}{n} |d(\mathcal{X}, g \cdot \mathcal{X}) + 2(\omega_n^{-1} \cdot \mathcal{X} | \omega_n \cdot \mathcal{X})_{\mathcal{X}} - n\lambda| \\ &\quad + \frac{1}{n} O(\delta) \\ &\leq 5\varepsilon + \frac{1}{n} O(\delta) \end{aligned}$$

for sufficiently large n . Then $\lim_n \frac{1}{n}(\text{translation length of } \omega_n) = \lambda > 0$ for $\omega \in \Omega := \cap_{k \in \mathbb{N}} \Omega_k$ and $\mathbb{P}(\Omega) = 1$. In particular, the translation length of ω_n becomes positive so ω_n is eventually pseudo-Anosov. \square

It is still open whether Theorem A holds under the finite first moment condition. However, a weaker assertion does hold true as follows.

Theorem 3.13. *Let (A, B) be a filling pair of multicurves on S and $\nu : \text{Mod}(S) \rightarrow [0, 1]$ be a non-elementary probability measure with $\text{supp } \nu \subseteq \langle T_A, T_B \rangle$. If the first moment with respect to the Teichmüller metric is finite, i.e. for any $\mathcal{X} \in \mathcal{T}(S)$,*

$$\sum_{g \in \text{Mod}(S)} d_{\mathcal{T}}(\mathcal{X}, g \cdot \mathcal{X}) \nu(g) < \infty,$$

then as $n \rightarrow \infty$ we have

$$\frac{1}{n} \log \lambda_{\omega_n} \rightarrow h(\omega) \quad \text{in probability.}$$

Proof. Since ν has finite first moment, $\lim_{n \rightarrow \infty} \frac{1}{n} d_{\mathcal{T}}(\mathcal{X}, \omega_n \cdot \mathcal{X}) = h(\omega)$ is almost surely equal to the drift $L_{\mathcal{T}} > 0$ (cf. [Mas18], [MT18, Theorem 1.2], [BQ16b, Proposition 3.3]).

We again use the notation $m = m(n) = \lceil n/2 \rceil$ and $u_m := \omega_m^{-1} \omega_n$. The following lemma of Maher and Tiozzo holds without any moment condition.

Lemma 3.14 ([MT18, Lemma 5.9, 5.11]). *Let G be a countable group of isometries of a separable Gromov hyperbolic space, and let ν be a non-elementary probability measure on G . Then for sufficiently small η ,*

$$\begin{aligned} \mathbb{P}((u_m^{-1} \cdot \mathcal{X} | \omega_m \cdot \mathcal{X})_{\mathcal{X}} > \eta n/2) &\rightarrow 0, \\ \mathbb{P}((\omega_m \cdot \mathcal{X} | \omega_n \cdot \mathcal{X})_{\mathcal{X}} < \eta n) &\rightarrow 0, \text{ and} \\ \mathbb{P}((u_m^{-1} \cdot \mathcal{X} | \omega_n^{-1} \cdot \mathcal{X})_{\mathcal{X}} < \eta n) &\rightarrow 0 \end{aligned}$$

as $n \rightarrow \infty$.

Suppose now that $0 < \varepsilon < L_{\mathcal{T}}/10$ is given. We then define

$$E_n = \left\{ \omega : \left| \frac{1}{k} d(\mathcal{X}, \omega_k \cdot \mathcal{X}) - L_{\mathcal{T}} \right| < \varepsilon/10 \text{ for all } k \geq n \right\}.$$

Then $E_n \nearrow E = \{ \omega : \limsup | \frac{1}{k} d(\mathcal{X}, \omega_k \cdot \mathcal{X}) - L_{\mathcal{T}} | < \varepsilon/10 \}$. Since $\frac{1}{k} d(\mathcal{X}, \omega_k \cdot \mathcal{X})$ tends to $L_{\mathcal{T}}$ almost surely, $P(E_n) \nearrow 1$.

Next, let $\eta = \varepsilon/100$ and define

$$\Omega_n = \left\{ \omega : \begin{array}{l} (u_m^{-1} \cdot \mathcal{X} | \omega_m \cdot \mathcal{X})_{\mathcal{X}} \leq \eta n/2, \quad (\omega_m \cdot \mathcal{X} | \omega_n \cdot \mathcal{X})_{\mathcal{X}} \geq \eta n, \\ \text{and } (u_m^{-1} \cdot \mathcal{X} | \omega_n^{-1} \cdot \mathcal{X})_{\mathcal{X}} \geq \eta n \end{array} \right\}.$$

Note that by Lemma 3.14, $\mathbb{P}(\Omega_n^c)$ tends to 0. We now let $N = 1000O(\delta)/\varepsilon$. Then for $n > N$ and $\omega \in E_n \cap \Omega_n$, Lemma 3.10 implies

$$2(\omega_n \cdot \mathcal{X} | \omega_n^{-1} \cdot \mathcal{X})_{\mathcal{X}} \leq \eta n + 2O(\delta) \leq 2\eta n < (L_{\mathcal{T}} - \varepsilon)n \leq d(\mathcal{X}, \omega_n \cdot \mathcal{X}).$$

Thus, we can apply Lemma 3.11 and deduce that

$$\begin{aligned} \left| \frac{1}{n} \log \lambda_{\omega_n} - L_{\mathcal{T}} \right| &\leq \frac{1}{n} |nL_{\mathcal{T}} - d(\mathcal{X}, \omega_n \cdot \mathcal{X}) - 2(\omega_n \cdot \mathcal{X} | \omega_n^{-1} \cdot \mathcal{X})_{\mathcal{X}}| + \frac{1}{n} O(\delta) \\ &\leq (\varepsilon/10) + 2\eta + (\varepsilon/1000) < \varepsilon. \end{aligned}$$

In conclusion, for $n > N$,

$$\mathbb{P} \left(\left| \frac{1}{n} \log \lambda_{\omega_n} - L_{\mathcal{T}} \right| \geq \varepsilon \right) \leq \mathbb{P}(E_n^c \cup \Omega_n^c).$$

Since $\mathbb{P}(E_n^c \cup \Omega_n^c) \rightarrow 0$ as $n \rightarrow \infty$, the desired convergence is proved. \square

This in particular implies that $\mathbb{P}(\omega_n \text{ is pseudo-Anosov}) \rightarrow 1$ as $n \rightarrow \infty$. This result in a general setting is due to Maher [M⁺11].

As a corollary of Theorem 3.13, we can deduce the almost sure equality between the topological entropy and limsup of the linear growth rate of translation lengths. This was also observed by Dahmani and Horbez in [DH18, Remark 2.7, 3.2]. For the sake of completeness, we include a proof of the result below.

Corollary 3.15. *Let (A, B) be a filling pair of multicurves on S and $\nu : \text{Mod}(S) \rightarrow [0, 1]$ be a non-elementary probability measure with $\text{supp } \nu \subseteq \langle T_A, T_B \rangle$. If the first moment with respect to the Teichmüller metric is finite, then for \mathbb{P} -a.e. $\omega \in \text{Mod}(S)^{\mathbb{N}}$, we have*

$$\limsup_{n \rightarrow \infty} \frac{1}{n} (\text{translation length of } \omega_n) = h(\omega)$$

where $h(\omega)$ is the topological entropy of a random walk.

Proof. For each k , we take n_k such that

$$\mathbb{P} \left(\left| \frac{1}{n_k} \log \lambda_{\omega_{n_k}} - h(\omega) \right| > 2^{-k} \right) < 2^{-k}.$$

The Borel–Cantelli lemma implies that \mathbb{P} -a.e. $\omega \in \text{Mod}(S)^{\mathbb{N}}$ has a subsequence $(\omega_{n_k})_{k > K(\omega)}$ such that $|\frac{1}{n_k} \log \lambda_{\omega_{n_k}} - h(\omega)| \leq 2^{-k}$. Thus, it makes sense to consider the quantity $\limsup_n \frac{1}{n} \log \lambda_{\omega_n} \geq h(\omega)$ for such paths.

The reverse inequality follows from Theorem 3.3: for pseudo-Anosov ω_n , we have $\log \lambda_{\omega_n} \leq d_{\mathcal{T}}(\mathcal{X}, \omega_n \cdot \mathcal{X})$ so

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log \lambda_{\omega_n} \leq h(\omega)$$

for \mathbb{P} -a.e. $\omega \in \text{Mod}(S)^\mathbb{N}$, as a result of Theorem 3.7. \square

3.3. Eventually pseudo-Anosov behavior of random walks. One consequence of Theorem A is that almost every random walk associated with Thurston's construction eventually becomes pseudo-Anosov. Our purpose is to obtain this phenomenon without any moment condition.

Theorem B. *Let (A, B) be a filling pair of multicurves on S such that $\langle T_A, T_B \rangle \cong F_2$, and $\nu : \text{Mod}(S) \rightarrow [0, 1]$ be a non-elementary probability measure with $\text{supp } \nu \subseteq \langle T_A, T_B \rangle$ and $\text{rank}(\text{supp } \nu) < \infty$. Then for \mathbb{P} -a.e. $\omega \in \text{Mod}(S)^\mathbb{N}$, there exists $N \in \mathbb{N}$ such that*

$$\omega_n \text{ is pseudo-Anosov for all } n > N.$$

Proof. Recall Theorem 2.13: the only non-pseudo-Anosov elements that appear in $\langle T_A, T_B \rangle$ are the conjugates of powers of T_A , T_B and $T_A T_B$.

We first prove that the conjugates of powers of T_A do not appear infinitely often. Let $\{v_1, \dots, v_r\}$ be a set of free generators of $\langle \text{supp } \nu \rangle \leq \langle T_A, T_B \rangle$. By [Hal49] and [Bur69], there exist $v_{r+1}, \dots, v_k \in \langle T_A, T_B \rangle$ so that v_1, \dots, v_k freely generate a finite index subgroup L of $\langle T_A, T_B \rangle$. Let $\{Lg_1, \dots, Lg_l\}$ be the collection of the right cosets of L in $\langle T_A, T_B \rangle$.

We define the abelianization map $Ab : L \rightarrow \mathbb{Z}^k$ sending each v_i to \vec{e}_i and a projection map $Q : \mathbb{Z}^k \rightarrow \mathbb{Z}^r$ discarding the last $k - r$ coordinates. Since $\langle \text{supp } \nu \rangle \leq L$, we obtain a new random walk $Q \circ Ab \circ \omega$ on $\mathbb{Z}^r \subseteq \mathbb{R}^r$. For $i = 1, \dots, l$, we pick $w_i \in \mathbb{Z}^r \setminus \{0\}$ such that

$$(Q \circ Ab)(\langle g_i T_A g_i^{-1} \rangle \cap L) \subseteq \langle w_i \rangle.$$

Suppose now that $\omega_n = g T_A^j g^{-1} \in L$ for some ω and $g \in Lg_i$. Since ω_n and $g_i g^{-1}$ are in L , we deduce

$$(Q \circ Ab)(\omega_n) = (Q \circ Ab)((g_i g^{-1}) \omega_n (g g_i^{-1})) = (Q \circ Ab)(g_i T_A^j g_i^{-1}) \in \langle w_i \rangle.$$

Thus, if ω visits the conjugates of powers of T_A infinitely often, then $(Q \circ Ab \circ \omega)$ visits $\langle w_i \rangle$ infinitely often for some i .

For each i , we consider the projection $P_i : \mathbb{R}^r \rightarrow \mathbb{R}^{r-1}$ onto the orthogonal complement of w_i . Suppose that $(P_i \circ Q \circ Ab)(\text{supp } \nu) \cdot \theta = 0$ for some $\theta \in \mathbb{R}^{r-1} \setminus \{0\}$. This implies that $(P_i \circ Q \circ Ab)(\langle \text{supp } \nu \rangle) \cdot \theta = P_i(\mathbb{Z}^r) \cdot \theta = 0$, which is a contradiction. Hence $P_i \circ Q \circ Ab \circ \omega$ is truly $(r-1)$ -dimensional; this implies the transience if $r \geq 4$ is assumed further ([Dur19], [CF08]). Consequently, we deduce

$$\mathbb{P} \left(\begin{array}{c} \omega_n \text{ visits the conjugates of} \\ \text{powers of } T_A \text{ i.o.} \end{array} \right) \leq \sum_{i=1}^l \mathbb{P}((Q \circ Ab)(\omega_n) \in \langle w_i \rangle \text{ i.o.}) = 0.$$

when $r \geq 4$ is assumed¹.

¹“i.o.” is an abbreviation of “infinitely often”.

It remains to deal with the cases $r = 2, 3$. Recall that v_1, \dots, v_r freely generate $\langle \text{supp } \nu \rangle$. Let $Ab = (Ab_1, \dots, Ab_r) : \langle \text{supp } \nu \rangle \rightarrow \mathbb{Z}^r$ be the abelianization map sending each v_i to \vec{e}_i . For $k = 0, 1, 2, 3$, we define

$$A_k := \{w \in \langle \text{supp } \nu \rangle : Ab_1(w) + kAb_2(w) \in 4\mathbb{Z}\},$$

$$B_k := \{w \in \langle \text{supp } \nu \rangle : kAb_1(w) + Ab_2(w) \in 4\mathbb{Z}\}.$$

Note that $(\cup_k A_k) \cup (\cup_k B_k) = \langle \text{supp } \nu \rangle$. Further, $\langle \text{supp } \nu \rangle / A_k$ and $\langle \text{supp } \nu \rangle / B_k$ are isomorphic to $\mathbb{Z}/4\mathbb{Z}$. Thus, A_k and B_k are index 4 subgroups of $\langle \text{supp } \nu \rangle$; the Nielsen–Schreirer theorem asserts that A_k, B_k are of rank $4r - 3$. See Figure 6 for the example A_0 .

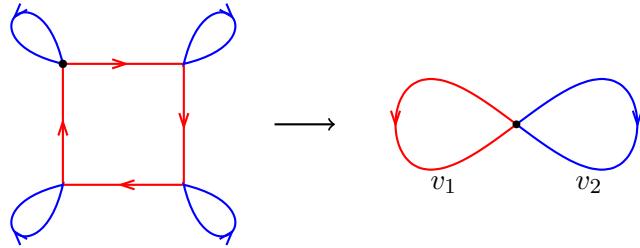


FIGURE 6. Covering space for A_0 in the case $r = 2$.

Let us focus on A_0 from now on. Note that $v_1 \in \langle \text{supp } \nu \rangle$ corresponds to a generator of $\langle \text{supp } \nu \rangle / A_0 \cong \mathbb{Z}/4\mathbb{Z}$. If each element of $\text{supp } \nu$ corresponds to one of $\bar{0}, \bar{2} \in \mathbb{Z}/4\mathbb{Z}$, so does each element of $\langle \text{supp } \nu \rangle$. This contradicts the existence of v_1 ; thus, some element f of $\text{supp } \nu$ should correspond to a generator of $\mathbb{Z}/4\mathbb{Z}$. In particular, $f^4 \in \langle \langle \text{supp } \nu \rangle \rangle \cap A_0$ holds².

We define a stopping time $\{n_m\}_{m \in \mathbb{Z}_{\geq 0}}$ as

$$n_0 = 0, \quad n_m = \inf\{n > n_{m-1} : \omega_n \in A_0\}.$$

Then $(\omega'_m) = (\omega_{n_m})$ becomes a new random walk on A_0 with transition probability ν' . Note that $\langle \langle \text{supp } \nu' \rangle \rangle = \langle \langle \text{supp } \nu \rangle \rangle \cap A_0$. Note also that ω_n visits the conjugates of powers of T_A inside A_0 i.o. if and only if ω'_m visits the conjugates of powers of T_A i.o..

Let $\{v'_1, \dots, v'_{4r-3}\}$ be a set of free generators of A_0 . By [Hal49] and [Bur69], there exist $v'_{4r-2}, \dots, v'_k \in \langle T_A, T_B \rangle$ so that v'_1, \dots, v'_k freely generate a finite index subgroup L of $\langle T_A, T_B \rangle$. Let $\{Lg_1, \dots, Lg_l\}$ be the collection of the right cosets of L .

As before, we define the abelianization map $Ab' : L \rightarrow \mathbb{Z}^k$ sending each v'_i to \vec{e}_i and a projection map $Q : \mathbb{Z}^k \rightarrow \mathbb{Z}^{4r-3}$ discarding the last

² $\langle \langle \cdot \rangle \rangle$ denotes the semigroup generated by $\{\cdot\}$.

$k - 4r + 3$ coordinates. For $i = 1, \dots, l$, we pick $w_i \in \mathbb{Z}^{4r-3} \setminus \{0\}$ such that $(Q \circ Ab')(\langle g_i T_A g_i^{-1} \rangle \cap L) \subseteq \langle w_i \rangle$. We then similarly deduce that

$$\mathbb{P} \left(\begin{array}{c} \omega'_m \text{ visits the conjugates of} \\ \text{powers of } T_A \text{ i.o.} \end{array} \right) \leq \sum_{i=1}^l \mathbb{P} ((Q \circ Ab')(\omega'_m) \in \langle w_i \rangle \text{ i.o.}).$$

For each i , we consider the projection $P_i : \mathbb{R}^{4r-3} \rightarrow \mathbb{R}^{4r-4}$ onto the orthogonal complement of w_i . Suppose that $(P_i \circ Q \circ Ab')(\text{supp } \nu') \cdot \theta = 0$ for some $\theta \in \mathbb{R}^{4r-4} \setminus \{0\}$. In particular, $f^4 \in \langle \langle \text{supp } \nu \rangle \rangle \cap A_0 = \langle \langle \text{supp } \nu' \rangle \rangle$ also satisfies $(P_i \circ Q \circ Ab')(f^4) \cdot \theta = 0$.

Since $(P_i \circ Q \circ Ab')(\langle \text{supp } \nu \rangle \cap A_0) = P_i(\mathbb{Z}^r) \cdot \theta \neq 0$, $(P_i \circ Q \circ Ab')(w) \cdot \theta \neq 0$ for some $w \in \langle \text{supp } \nu \rangle \cap A_0$. Here, $w = b_1^{k_1} \dots b_m^{k_m}$ for some alphabets $b_j \in \text{supp } \nu$ and $k_j \in \mathbb{Z}$. Using f , we express w as a concatenation of elements in A_0 :

$$w = (b_1^{k_1} f^{-t_1})(f^{t_1} b_2^{k_2} f^{-t_2}) \dots (f^{t_{m-1}} b_m^{k_m}),$$

where f^{t_j} and $b_1^{k_1} \dots b_j^{k_j}$ belong to the same right coset of A_0 in $\langle \text{supp } \nu \rangle$.

Since $(P_i \circ Q \circ Ab')(w) \cdot \theta \neq 0$, we deduce that $(P_i \circ Q \circ Ab')(f^t b^k f^s) \cdot \theta \neq 0$ for some $b \in \text{supp } \nu$ and $t, k, s \in \mathbb{Z}$ such that $f^t b^k f^s \in A_0$. By taking inverse if necessary, we may suppose $k \geq 0$. Recall again that $f^4 \in A_0$ and $(P_i \circ Q \circ Ab')(f^4) \cdot \theta = 0$; by multiplying f^{-4t} or f^{-4s} if necessary, we may also assume that $t, s \geq 0$. However, $f^t b^k f^s$ then belongs to $\langle \langle \text{supp } \nu \rangle \rangle \cap A_0 = \langle \langle \text{supp } \nu' \rangle \rangle$, which contradicts the assumption.

Thus, $(P_i \circ Q \circ Ab')(\text{supp } \nu') \cdot \theta \neq 0$ for each $\theta \in \mathbb{R}^{4r-4} \setminus \{0\}$; $P_i \circ Q \circ Ab' \circ \omega'$ is truly $(4r - 4)$ -dimensional and thus transient. Consequently, we deduce

$$\mathbb{P} \left(\begin{array}{c} \omega_n \text{ visits the conjugates of} \\ \text{powers of } T_A \text{ in } A_0 \text{ i.o.} \end{array} \right) \leq \sum_{i=1}^l \mathbb{P} ((Q \circ Ab')(\omega'_m) \in \langle w_i \rangle \text{ i.o.}) = 0.$$

Similarly, almost every ω_n eventually avoids the conjugates of powers of T_A in each of A_k and B_k . Put together, almost every ω_n eventually avoids any conjugates of powers of T_A . Similar argument applies to the conjugates of powers of T_B or $T_A T_B$. Then Theorem 2.13 completes the proof. \square

4. APPLICATIONS

4.1. Volumes of random mapping tori from Thurston's construction. As we explained earlier, the topological entropy of a random walk is closely related to the dynamics of random mapping classes. In this section, we discuss applications of the eventually pseudo-Anosov behavior of random walks and the spectral theorem. As the first application, we estimate the hyperbolic volumes of random mapping tori with monodromies obtained from Thurston's construction.

Definition 4.1 (Mapping torus). *For a surface S and a mapping class $\varphi \in \text{Mod}(S)$, the mapping torus M_φ with monodromy φ is the closed 3-manifold defined by*

$$M_\varphi := S \times [0, 1]/(x, 0) \sim (\varphi(x), 1).$$

Thurston proved that the geometry of a mapping torus M_φ reads off the nature of the mapping class φ . For instance, according to [Ota96], $\varphi \in \text{Mod}(S)$ is pseudo-Anosov if and only if M_φ admits a hyperbolic structure, which is unique by the Mostow rigidity theorem [Mos68]. In this case, the *hyperbolic volume* $\text{vol}(M_\varphi)$ of M_φ serves as an invariant of the mapping class.

Once we have a probability measure ν on $\text{Mod}(S)$ and the induced one \mathbb{P} on $\text{Mod}(S)^\mathbb{N}$, we get a sequence of random mapping tori M_{ω_n} corresponding to a random walk $\omega = (\omega_n) \in \text{Mod}(S)^\mathbb{N}$. For an eventually pseudo-Anosov sample path ω_n , it makes sense to consider the asymptotes of the hyperbolic volume of M_{ω_n} . For instance, Viaggi proved the following.

Theorem 4.2 (Viaggi, [Via19]). *Suppose that the probability measure ν on $\text{Mod}(S)$ satisfies $\langle \text{supp } \nu \rangle = \text{Mod}(S)$ and it has a finite symmetric support. Then for the induced probability measure \mathbb{P} on $\text{Mod}(S)^\mathbb{N}$,*

$$\lim_{n \rightarrow \infty} \frac{1}{n} \text{vol}(M_{\omega_n}) = V \quad \text{almost surely}$$

for some $V > 0$.

According to Viaggi [Via19], the condition in Theorem 4.2 can be weakened so that ν is non-elementary, i.e., $\langle \text{supp } \nu \rangle$ is not necessarily the whole $\text{Mod}(S)$ but a non-elementary subgroup. However, Viaggi's argument relies on the boundedness of $\text{supp } \nu$, that is, $\sup_{g \in \text{supp } \nu} d_{\mathcal{T}}(\mathcal{X}, g \cdot \mathcal{X}) < \infty$ for some $\mathcal{X} \in \mathcal{T}(S)$, in order to estimate $\text{vol}(M_{\omega_n})$ by applying triangle inequalities. Therefore, one cannot affirm that $\frac{1}{n} \text{vol}(M_{\omega_n})$ converges if the probability measure is not finitely supported.

We now consider this question in the setting of Thurston's construction. As before, let (A, B) be a filling pair of multicurves such that $\langle T_A, T_B \rangle \cong F_2$ and ν be a non-elementary probability measure on $\langle T_A, T_B \rangle$. This induces a measure \mathbb{P} on the path space $\text{Mod}(S)^\mathbb{N}$. If $\text{rank}(\text{supp } \nu) < \infty$, then Theorem B asserts that \mathbb{P} -a.e. random walk ω has a pseudo-Anosov tail. In this circumstance, we can discuss the long-term behavior of $\text{vol}(M_{\omega_n})$.

Regarding the hyperbolic volume of a mapping torus, Kojima and McShane provided the following estimation in terms of the topological entropy of the monodromy. We note that the work of Brock [Bro03] and Linch [Lin74] leads to a similar estimation with a different constant.

Theorem 4.3 (Kojima–McShane, [KM18]). *For any pseudo-Anosov $\varphi \in \text{Mod}(S)$ and its topological entropy $h(\varphi) = \log \lambda_\varphi$, we have*

$$\text{vol}(M_\varphi) \leq -3\pi\chi(S) \cdot h(\varphi)$$

where $\chi(S)$ is the Euler characteristic of S .

Combining this estimation with Theorem B gives the following statement.

Theorem C. *Let (A, B) be a filling pair of multicurves on S such that $\langle T_A, T_B \rangle \cong F_2$ and $\nu : \text{Mod}(S) \rightarrow [0, 1]$ be a non-elementary probability measure with $\text{supp } \nu \subseteq \langle T_A, T_B \rangle$ and $\text{rank}(\text{supp } \nu) < \infty$. Suppose also that ν has finite first moment with respect to the Teichmüller metric. Then*

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \text{vol}(M_{\omega_n}) \leq -3\pi\chi(S) \cdot h(\omega)$$

for \mathbb{P} -a.e. $\omega \in \text{Mod}(S)^\mathbb{N}$ where $\chi(S)$ is the Euler characteristic of S .

Proof. As we mentioned above, Theorem B ensures that ω_n is eventually pseudo-Anosov for \mathbb{P} -a.e. ω . By Theorem 4.3, we deduce

$$\text{vol}(M_{\omega_n}) \leq -3\pi\chi(S) \cdot \log \lambda_{\omega_n}$$

for large n , where λ_{ω_n} is the stretch factor of ω_n . By Theorem 3.3,

$$\frac{1}{n} \text{vol}(M_{\omega_n}) \leq -3\pi\chi(S) \cdot \frac{1}{n} \log \lambda_{\omega_n} \leq -3\pi\chi(S) \cdot \frac{1}{n} d_{\mathcal{T}}(\mathcal{Y}, \omega_n \cdot \mathcal{Y})$$

for any $\mathcal{Y} \in \mathcal{T}(S)$. Then applying Theorem 3.7 completes the proof. \square

Remark 4.4. *This estimation suggests a similarity between the topological entropy of a random walk and that of a single pseudo-Anosov. Indeed,*

$$\text{vol}(M_{\varphi^n}) = n \cdot \text{vol}(M_\varphi), \quad h(\varphi^n) = n \cdot h(\varphi)$$

and thus

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \text{vol}(M_{\varphi^n}) \leq -3\pi\chi(S) \cdot h(\varphi)$$

for any pseudo-Anosov $\varphi \in \text{Mod}(S)$.

4.2. The distribution of the stretch factors from Thurston's construction. We now discuss the distribution of the stretch factors. Note that for a non-elementary probability measure ν on $\langle T_A, T_B \rangle$ with finite first moment, Theorem 3.13 implies that for each $\varepsilon > 0$,

$$\mathbb{P}\left(e^{h(\omega)-\varepsilon} < \lambda_{\omega_n} < e^{h(\omega)+\varepsilon}\right) \rightarrow 1 \quad \text{as } n \rightarrow \infty.$$

We now discuss the case of finitely supported measures in detail. Let ν be a finitely supported non-elementary measure on $\langle T_A, T_B \rangle$ and denote $p := \min\{\nu(w) : w \in \text{supp } \nu\} > 0$. Let $\mathcal{X}_0 \in \mathcal{T}(S)$ be the point corresponding to the origin of the Poincaré disk \mathbb{H}^2 . Note that

$$\bigcup_{N=0}^{\infty} \{\omega_N : \omega = (\omega_n) \in \Omega\} = \langle\langle \text{supp } \nu \rangle\rangle$$

for any $\Omega \subseteq \text{Mod}(S)^\mathbb{N}$ with $\mathbb{P}(\Omega) = 1$. Hence, observing \mathbb{P} -a.e. ω is sufficient to deal with the stretch factors of all pseudo-Anosovs in $\langle\langle \text{supp } \nu \rangle\rangle$.

Since ν is finitely supported in addition, even stronger estimate is available. Using the estimate of Azuma [Azu67] for martingales with bounded

increments, the theory of Benoist and Quint in [BQ16a] yields the following: there exist constants C and σ such that for any $\varepsilon > 0$ and $x \in \partial_B M$,

$$(4.1) \quad \mathbb{P}(|h_x(\omega_n^{-1} \cdot \mathcal{X}_0) - nL_{\mathcal{T}}| > \varepsilon n) \leq Ce^{-(n\varepsilon^2)/2\sigma^2} \text{ and}$$

$$(4.2) \quad \mathbb{P}(|d_{\mathbb{H}^2}(\mathcal{X}_0, \omega_n \cdot \mathcal{X}_0) - nL_{\mathcal{T}}| > \varepsilon n) \leq Ce^{-(n\varepsilon^2)/2\sigma^2}.$$

Here $L_{\mathcal{T}}$ is the drift of the random walk. Taking $\varepsilon > 4\sigma\sqrt{-\log p}$, the proof of Theorem A asserts that

$$\mathbb{P}(|\log \lambda_{\omega_n} - nL_{\mathcal{T}}| > 6\varepsilon n) < p^n$$

for sufficiently large n . Since all atoms at step n has probability at least p^n , this implies that $n(L_{\mathcal{T}} - 6\varepsilon) < \log \lambda_{\omega_n} < n(L_{\mathcal{T}} + 6\varepsilon)$ for sufficiently large n . In other words, we conclude the following.

Theorem D. *Let (A, B) be a filling pair of multicurves on S . For a finite generating set Λ of $\langle T_A, T_B \rangle$, let $|\cdot|_{\Lambda}$ be a word norm with respect to Λ . Then there exist constants $L > 0$ and N such that*

$$\log \lambda_{\varphi} < L|\varphi|_{\Lambda}$$

for all pseudo-Anosov $\varphi \in \langle T_A, T_B \rangle$ with $|\varphi|_{\Lambda} > N$.

It remains to clarify σ in Inequality 4.1 and Inequality 4.2. For example, we can choose σ to be $M_1 := \max\{d_{\mathbb{H}^2}(\mathcal{X}_0, w \cdot \mathcal{X}_0) : w \in \text{supp } \nu\}$. Since $p \leq 1/2$, we obtain $n(L_{\mathcal{T}} + 6\varepsilon) > (24\sqrt{\log 2})nM_1$. This estimate is no better than nM_1 , which is obtained as follows. Let $w = a_1 \cdots a_n$ be a pseudo-Anosov with $a_i \in \text{supp } \nu$, and let $w_i = a_1 \cdots a_i$ for $i = 1, \dots, n$. We then observe

$$\log \lambda_w \leq d_{\mathbb{H}^2}(\mathcal{X}_0, w \cdot \mathcal{X}_0) \leq d_{\mathbb{H}^2}(\mathcal{X}_0, w_1 \cdot \mathcal{X}_0) + \sum_{i=2}^n d_{\mathbb{H}^2}(w_{i-1} \cdot \mathcal{X}_0, w_i \cdot \mathcal{X}_0) \leq nM_1.$$

Note also that this estimate applies to every step n .

However, M_1 may be exaggerated in some cases, whereas the drift $L_{\mathcal{T}}$ reads off the cancellation among elements and provides a better bound for the stretch factor. For example, we can fix $s \in \langle T_A, T_B \rangle$ and set $\nu(sT_A^{\pm 1}s^{-1}) = \nu(sT_B^{\pm 1}s^{-1}) = 1/4$. Then there exists a constant $C > 0$ such that $d_{\mathbb{H}^2}(\mathcal{X}_0, a \cdot \mathcal{X}_0) \geq 2d_{\mathbb{H}^2}(\mathcal{X}_0, s \cdot \mathcal{X}_0) - C$ for each $a \in \text{supp } \nu$. Thus, both M_1 and the first moment is at least $2d_{\mathbb{H}^2}(\mathcal{X}_0, s \cdot \mathcal{X}_0) - C$.

Now let w be a word with k alphabets in $\text{supp } \nu$. Then we obtain

$$d_{\mathbb{H}^2}(\mathcal{X}_0, w \cdot \mathcal{X}_0) \leq 2d_{\mathbb{H}^2}(\mathcal{X}_0, s \cdot \mathcal{X}_0) + kK$$

where

$$K = d_{\mathbb{H}^2}(\mathcal{X}_0, T_A^{\pm 1} \cdot \mathcal{X}_0) = d_{\mathbb{H}^2}(\mathcal{X}_0, T_B^{\pm 1} \cdot \mathcal{X}_0) = \log \frac{\sqrt{\mu} + \sqrt{4 + \mu}}{2}.$$

Thus, for any $K' > K$ and sufficiently large k ,

$$\max\{d_{\mathbb{H}^2}(\mathcal{X}_0, w \cdot \mathcal{X}_0) : w = a_1 \cdots a_k, a_i \in \text{supp } \nu\} < kK'.$$

In this situation, Inequality 4.1 and Inequality 4.2 with $\sigma = K'$ and $\varepsilon > 4K'\sqrt{\log 4}$ hold for sufficiently large n . Note also $L_{\mathcal{T}} \leq K'$. Thus, when $d_{\mathbb{H}^2}(\mathcal{X}_0, s \cdot \mathcal{X}_0) > 100K' + C$, $n(L_{\mathcal{T}} + 6\varepsilon)$ becomes a better bound than nM_1 .

Here, using the drift amounts to using $\mathcal{X} = s \cdot \mathcal{X}_0$. With this different pivot point, one can also obtain the bound $\log \lambda_{\omega_n} \leq nK$ for all n .

The lower bound $n(L_{\mathcal{T}} - 6\varepsilon)$ may not be informative when $\varepsilon > L_{\mathcal{T}}/6$. This happens, for example, if $\text{supp } \nu$ is symmetric. Nonetheless, Proposition 2.14 gives another lower bound for λ_{ω_n} in terms of the cyclically reduced word norm of ω_n . Combining them yields the following theorem.

Theorem 4.5. *Let (A, B) be a filling pair of multicurves on S . Then*

$$\lambda_{\varphi} \leq e^{K|\varphi|}$$

for any pseudo-Anosov $\varphi \in \langle T_A, T_B \rangle$ where $|\cdot|$ is the cyclically reduced word norm with respect to T_A^{\pm} and T_B^{\pm} and $K = \log \frac{\sqrt{\mu} + \sqrt{4 + \mu}}{2}$.

Moreover, if $\langle T_A, T_B \rangle \cong F_2$, each pseudo-Anosov $\varphi \in \langle T_A, T_B \rangle$ satisfies

$$|\varphi|e^{1/4} \leq \lambda_{\varphi} \leq e^{K|\varphi|}.$$

Remark 4.6. *This estimation is non-probabilistic.*

It is also interesting to understand the algebraic properties of the stretch factors of pseudo-Anosovs obtained from Thurston's construction. Many known examples are Salem numbers, which are defined as follows.

Definition 4.7 (Salem number). *A real algebraic unit $\lambda > 1$ is a Salem number if all Galois conjugates except $\lambda^{\pm 1}$ lie on the unit circle in \mathbb{C} .*

One can naturally ask whether every Salem number can be realized as the stretch factor of some pseudo-Anosov from Thurston's construction, which is a special case of Fried's conjecture. Pankau partially answered this question by proving that every Salem number has a power which is the stretch factor arising from Thurston's construction.

Theorem 4.8 (Pankau, [Pan17]). *For any Salem number λ , there are $k, g \in \mathbb{N}$ such that λ^k is the stretch factor of a pseudo-Anosov homeomorphism of the surface S_g genus g arising from Thurston's construction.*

In fact, Pankau's proof concludes that the pseudo-Anosov for a given Salem number can be obtained as a word of two positive multitwists T_A and T_B . Together with Theorem 4.5, we get the following corollary.

Corollary 4.9. *Let λ be a Salem number and suppose that λ^k is the stretch factor of a pseudo-Anosov φ from Thurston's construction. Then we have*

$$\lambda < \lambda^k \leq e^{K|\varphi|}.$$

Proof. The existence of k and a configuration for Thurston's construction is guaranteed by the work of Pankau. Then by Theorem 4.5, we conclude that $\lambda^k = \lambda_{\varphi} \leq e^{K|\varphi|}$. \square

One interesting consequence of above inequality, together with Proposition 2.14, is that we can estimate the power k .

Theorem E. *Let λ be a Salem number and suppose that λ^k is the stretch factor of a pseudo-Anosov φ obtained from Thurston's construction with $\rho : \langle T_A, T_B \rangle \rightarrow \mathrm{PSL}(2, \mathbb{R})$. Then*

$$k \leq \frac{K|\varphi|}{\log \lambda}$$

where $|\varphi|$ is the cyclically reduced word norm of φ in terms of T_A^\pm and T_B^\pm and $K = \log \frac{\sqrt{\mu} + \sqrt{4+\mu}}{2}$.

Moreover, when $\langle T_A, T_B \rangle \cong F_2$, we further have

$$\frac{\log |\varphi|}{4 \log \lambda} \leq k \leq \frac{K|\varphi|}{\log \lambda}.$$

In fact, Pankau constructed a pseudo-Anosov of word norm two whose stretch factor is some power of a given Salem number λ . Accordingly, one can rewrite the above inequalities if one focuses only on the specific configuration for Thurston's construction considered by Pankau in [Pan17].

REFERENCES

- [AKM65] Roy L Adler, Alan G Konheim, and M Harry McAndrew. Topological entropy. *Trans. Amer. Math. Soc.*, 114(2):309–319, 1965.
- [Azu67] Kazuoki Azuma. Weighted sums of certain dependent random variables. *Tohoku Math. J.*, 19(3):357–367, 1967.
- [B⁺78] Lipman Bers et al. An extremal problem for quasiconformal mappings and a theorem by thurston. *Acta Math.*, 141:73–98, 1978.
- [BQ16a] Yves Benoist and Jean-François Quint. Central limit theorem for linear groups. *Ann. Probab.*, 44(2):1308–1340, 2016.
- [BQ16b] Yves Benoist and Jean-François Quint. Central limit theorem on hyperbolic groups. *Izu. Math.*, 80(1):3, 2016.
- [Bro03] Jeffrey F Brock. Weil-petersson translation distance and volumes of mapping tori. *Comm. Anal. Geom.*, 11(5):987–1000, 2003.
- [Bur69] Roger G. Burns. A note on free groups. *Proc. Amer. Math. Soc.*, 23(1):14–17, 1969.
- [CF08] Kai Lai Chung and Wolfgang Heinrich Johannes Fuchs. On the distribution of values of sums of random variables. In *Selected Works Of Kai Lai Chung*, pages 157–168. World Scientific, 2008.
- [DH18] François Dahmani and Camille Horbez. Spectral theorems for random walks on mapping class groups and $\mathrm{Out}(F_N)$. *Int. Math. Res. Not. IMRN*, 2018(9):2693–2744, 2018.
- [Dur19] Rick Durrett. *Probability: theory and examples*, volume 49. Cambridge university press, 2019.
- [EST20] Viveka Erlandsson, Juan Souto, and Jing Tao. Genericity of pseudo-anosov mapping classes, when seen as mapping classes. *arXiv preprint arXiv:2001.03576*, 2020.
- [FK60] Harry Furstenberg and Harry Kesten. Products of random matrices. *Ann. Math. Stat.*, 31(2):457–469, 1960.

- [FLP79] Albert Fathi, François Laudenbach, and Valentin Poénaru. *Travaux de Thurston sur les surfaces*, volume 66 of *Astérisque*. Société Mathématique de France, Paris, 1979. Séminaire Orsay, With an English summary.
- [FM12] Benson Farb and Dan Margalit. *A primer on mapping class groups*, volume 49 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 2012.
- [Hal49] Marshall Jr. Hall. Coset representations in free groups. *Trans. Amer. Math. Soc.*, 67(2):421–432, 1949.
- [HS07] Frank Herrlich and Gabriela Schmithüsen. On the boundary of teichmüller disks in teichmüller and in schottky space. *Handbook of Teichmüller theory*, 1:293–349, 2007.
- [HT02] Hessam Hamidi-Tehrani. Groups generated by positive multi-twists and the fake lantern problem. *Algebr. Geom. Topol.*, 2(2):1155–1178, 2002.
- [Kar14] Anders Karlsson. Two extensions of thurston's spectral theorem for surface diffeomorphisms. *Bull. Lond. Math. Soc.*, 46(2):217–226, 2014.
- [Kid08] Yoshikata Kida. The mapping class group from the viewpoint of measure equivalence theory. *Mem. Amer. Math. Soc.*, no.916, 2008.
- [KM18] Sadayoshi Kojima and Greg McShane. Normalized entropy versus volume for pseudo-anosovs. *Geom. Topol.*, 22(4):2403–2426, 2018.
- [Kow08] Emmanuel Kowalski. *The large sieve and its applications: Arithmetic Geometry, Random Walks and Discrete Groups*, volume 175 of *Cambridge Tracts in Mathematics*. Cambridge University Press, 2008.
- [Lei04] Christopher J Leininger. On groups generated by two positive multi-twists: Teichmüller curves and lehmer's number. *Geom. Topol.*, 8(3):1301–1359, 2004.
- [Lin74] Michele Linch. A comparison of metrics on teichmüller space. *Proc. Amer. Math. Soc.*, 43(2):349–352, 1974.
- [M⁺11] Joseph Maher et al. Random walks on the mapping class group. *Duke Math. J.*, 156(3):429–468, 2011.
- [Mas18] Hidetoshi Masai. Topological entropy of random walks on mapping class groups. *Int. Math. Res. Not. IMRN*, 2018(3):739–761, 2018.
- [Mos68] George D Mostow. Quasi-conformal mappings in n -space and the rigidity of hyperbolic space forms. *Publ. Math. Inst. Hautes Études Sci.*, 34:53–104, 1968.
- [MP89] John McCarthy and Athanase Papadopoulos. Dynamics on thurston's sphere of projective measured foliations. *Comment. Math. Helv.*, 64(1):133–166, 1989.
- [MT18] Joseph Maher and Giulio Tiozzo. Random walks on weakly hyperbolic groups. *J. Reine Angew. Math.*, 2018(742):187–239, 2018.
- [Ota96] Jean-Pierre Otal. *Le théorème d'hyperbolisation pour les variétés fibrées de dimension 3*, volume 235. Société mathématique de France Paris, 1996.
- [Pan17] Joshua Pankau. Salem number stretch factors and totally real fields arising from thurston's construction. *arXiv preprint arXiv:1711.06374*, 2017.
- [Pen88] Robert C Penner. A construction of pseudo-anosov homeomorphisms. *Trans. Amer. Math. Soc.*, 310(1):179–197, 1988.
- [Pol10] Mark Pollicott. Maximal lyapunov exponents for random matrix products. *Invent. Math.*, 181(1):209–226, 2010.
- [Riv08] Igor Rivin. Walks on groups, counting reducible matrices, polynomials, and surface and free group automorphisms. *Duke Math. J.*, 142(2):353–379, 2008.
- [SS16] Hyunshik Shin and Balázs Strenner. Pseudo-anosov mapping classes not arising from penner's construction. *Geom. Topol.*, 19(6):3645–3656, 2016.
- [T⁺88] William P Thurston et al. On the geometry and dynamics of diffeomorphisms of surfaces. *Bull. Amer. Math. Soc. (N.S.)*, 19(2):417–431, 1988.
- [Via19] Gabriele Viaggi. Volumes of random 3-manifolds. *arXiv preprint arXiv:1905.04935*, 2019.

DEPARTMENT OF MATHEMATICAL SCIENCES, SEOUL NATIONAL UNIVERSITY, REPUBLIC OF KOREA

Email address: hrbaik85@gmail.com

DEPARTMENT OF MATHEMATICAL SCIENCES, KAIST, 291 DAEHAK-RO YUSEONG-GU, DAEJEON, 34141, SOUTH KOREA

Email address: inhyeokchoi@kaist.ac.kr

DEPARTMENT OF MATHEMATICS, YALE UNIVERSITY, 10 HILLHOUSE AVENUE, NEW HAVEN, CT 06520, USA

Email address: dongryul.kim@yale.edu