Random walks on mapping class groups

Hyungryul Baik and Inhyeok Choi

Abstract. This survey is concerned with random walks on mapping class groups. We illustrate how the actions of mapping class groups on Teichmüller spaces or curve complexes reveal the nature of random walks and vice versa. Our emphasis is on the analogues of classical theorems such as laws of large numbers and central limit theorems and the properties of harmonic measures under optimal moment conditions. We also explain the geometric analogy between Gromov hyperbolic spaces and Teichmüller spaces that has been used to copy the properties of random walks from one to the other.

Dedicated to Dennis Sullivan on the occasion of his 80th birthday

1. Introduction

A random walk is an example of a Markov chain or more generally a stochastic process. Various models of random walks have been suggested for applications to physics, economics, finance, biology, ecology, and many other fields. Among them are random walks on groups that deal with products of random group elements chosen with the same probability distribution. Since Kesten's pioneering work [55], the connection between the group structure (e.g., solvability, amenability, etc.) and the asymptotic behavior of random walks on the group has been studied in depth (see also [33]). This turned out to be fruitful from the perspective of both probability theory and group theory, and even geometry, following the development of the relatively new field of geometric group theory.

In this survey, we study random walks on mapping class groups via their actions on Teichmüller spaces equipped with the Teichmüller metric. These actions are generalizations of the action of $SL(2, \mathbb{Z})$ on $\mathbb{H}^2$. Here, $SL(2, \mathbb{Z})$ can be viewed from two important perspectives in the study of non-commutative random walks: one as a discrete group acting on a negatively curved space, and the other one as a lattice in a Lie group acting on a homogeneous space. Kaimanovich and Masur relate random walks on mapping class groups with both perspectives, as we explain in Section 3.

The prototypes from the first perspective are random walks on free groups. In this case, random walks escape to infinity and reveal the harmonic properties of the group.

Explicit computations regarding this escape led to the analogues of classical limit laws, e.g., laws of large numbers, central limit theorems, and local limit theorems. For these computations, we refer the readers to [62, 91] and the references therein.

Appropriate notions that generalize free groups are hyperbolic groups introduced by M. Gromov. The corresponding geometric property is δ-hyperbolicity: hyperbolic groups admit a geometric action on a proper δ-hyperbolic space. Thanks to recent developments, one can deal with not only random walks on hyperbolic groups but also those on weakly hyperbolic groups acting on non-proper δ-hyperbolic spaces. We discuss Kaimanovich's theory of random walks on hyperbolic groups and Maher–Tiozzo's theory of random walks on weakly hyperbolic groups in Section 4.

However, Teichmüller spaces are not δ-hyperbolic and mapping class groups are not lattices in semi-simple Lie groups of higher rank. Hence, neither viewpoint applies directly to mapping class groups (as Kaimanovich and Masur explain). Despite this contrast, we will first pursue the former perspective and make a due modification. Namely, Masur–Minsky's theory guided the usage of curve complexes, non-proper δ-hyperbolic spaces, in the study of mapping class groups. We explain how to compare the actions of mapping class groups on Teichmüller spaces and curve complexes in Section 5.

Having established the relevant theories, Section 6 and Section 7 deal with limit theorems for random walks on mapping class groups. We finish this survey by explaining a counting problem in mapping class groups and suggesting future directions.

Our survey is certainly not exhaustive at all. Especially, we regret not to explain the ideas of Sisto [85], Arzhantseva–Cashen–Tao [3], and Yang [92, 93] that make use of contracting elements. This idea naturally covers the case of hyperbolic groups, CAT(0)-groups, mapping class groups and right-angled Artin groups. We also lack an explanation of the Martin boundary and its comparison with other boundaries. This topic dates back to Martin's paper [69] and Doob's paper [24], and is still being actively researched, e.g., [39].

Let us briefly observe the connection between the theory of random walks and the Patterson–Sullivan theory. Given a negatively curved manifold $X = \widetilde{X}/\Gamma$ and its universal cover $\widetilde{X}$, both theories construct measures on $\partial \widetilde{X}$ using the deck transformation of $\widetilde{X}$. First, consider a random walk on Γ generated by the transition probability μ. By applying this random walk to a point $x \in \widetilde{X}$, one obtains a μ-harmonic measure ν_x on $\partial \widetilde{X}$ as the weak limit of the orbit distribution at step n. In the Patterson–Sullivan theory, each orbit point gx is assigned the mass $e^{-sd(x,gx)}$, where $s > 0$ is a parameter that determines how far the mass is spread toward infinity. When normalized, this assignment gives rise to a probability measure $\nu'_{x,s}$. As s approaches the growth exponent δ of G, the measure $\nu'_{x,s}$ escapes any bounded region. By taking the weak limit at $s = \delta$, we obtain the Patterson–Sullivan measure ν'_x on $\partial \widetilde{X}$ that is Γ-conformal with dimension δ. These measures entail rich information about the geometry of X (e.g., number of loops) and the dynamics on X (e.g., mixing geodesic flow). Furthermore, in certain circumstances, $\widetilde{X}$ is symmetric if and only if ν_x and ν'_x are proportional.

Given these observations, the next goal is to build a parallel theory for Teichmüller spaces. Harmonic measures and conformal densities will shed light on the geometry of

the moduli space, and it matters to clarify which measure serves which role. In particular, given the non-homogeneity of Teichmüller spaces, one can ask whether two measures on $\mathcal{PMF}$ differ or not.

Another important goal is to implement the Patterson–Sullivan theory on groups, which dates back to Coornaert's work [19]. In [40], Gekhtman–Taylor–Tiozzo utilized the automatic structure of the group to decompose the Patterson–Sullivan measure into countably many harmonic measures. This approach is intimately related to the following counting problem. Given a desirable property P of a group element or an orbit point (e.g., being loxodromic, two norms having a bounded ratio, and so on), we count the number of the elements satisfying P and compare it with the number of the entire elements. When we are discussing infinite groups such as mapping class groups, we truncate the group at a fixed word metric radius and see the asymptotic proportion of the elements satisfying P. One famous problem is Farb's question asking whether pseudo-Anosovs are generic in the sense of a counting problem. The Patterson–Sullivan theory on mapping class groups seems not developed enough to answer this question.

2. Preliminary

From the viewpoint of probability theory, Furstenberg writes that random walks on groups are non-commuting generalizations of random walks on $\mathbb{R}^n$ [33]. We are particularly interested in random walks on an infinite group with an interesting group-theoretic structure and actions on geometric spaces, mapping class groups being notable examples. Moreover, from the viewpoint of geometric group theory, random walks on mapping class groups arise as a natural way to depict typical mapping classes. Moreover, not only do random walks lead to rigidity theorems for mapping class groups, but they also give rise to boundary structures of Teichmüller spaces that fit into Thurston's theory. We briefly review the notion of random walks, mapping class groups, and the spaces related to mapping class groups.

Unless stated otherwise, G denotes a finitely generated group that acts on a metric space X by isometries. S denotes a finite generating set of G. All measures are probability measures, and μ always denotes a non-elementary measure on G. 'Almost every' and 'almost surely' are abbreviated to 'a.e.' and 'a.s.', respectively.

Given μ, we can consider the step space $(G^{\mathbb{Z}}, \mu^{\mathbb{Z}})$, the product space of G equipped with the product measure of μ. Each element $(g_n)_{n \in \mathbb{Z}}$ of the step space is called a *step path*. Each step path $(g_n)_n$ is associated with the sample path $(\omega_n)_n$ defined by

$$
\omega_n := \begin{cases} g_1 \cdots g_n, & n > 0, \\ \mathrm{id}, & n = 0, \\ g_0^{-1} \cdots g_{n+1}^{-1}, & n < 0. \end{cases}
$$

We also define the Bernoulli shift $T : G^{\mathbb{Z}} \to G^{\mathbb{Z}}$ by $T : (g_n)_n \mapsto (g_{n+1})_n$.

Remark. It is worth making a distinction between left and right random walks. The above definition deals with right random walks: ω_{n+1} grows from ω_n by the multiplication of g_{n+1} on the right. Random walks on the ambient space X are then modeled by applying these random walks to a reference point $o \in X$. One advantage of this model is that a.e. orbit path converges to a boundary point.

Meanwhile, left random walks naturally arise when random isometries are successively applied to an object (see [51] and [7]). Although the asymptotic behavior of orbit paths differs in the right random walks and left random walks, the limit theorems for displacement and translation length can be copied from one setting to the other one via the inversion

$$\omega_n = g_1 \cdots g_n \leftrightarrow \omega_n' = g_n^{-1} \cdots g_1^{-1}$$

and the identity $d(o, \omega_n o) = d(o, w_n' o)$.

We now move on to discussing groups and spaces. Fixing a closed orientable surface Σ with genus at least 2, we define its mapping class group $\mathrm{Mod}(\Sigma)$ as the group of self-homeomorphisms (or self-diffeomorphisms) of Σ modulo homotopy. This group has been studied for decades due to its rich group-theoretic properties and connections to other topics in geometry and topology. Among such connection is Royden's theorem [82], which states that $\mathrm{Mod}(\Sigma)$ is an index-2 subgroup of the isometry group of the Teichmüller space $\mathcal{T}(\Sigma)$ of Σ. Here, $\mathcal{T}(\Sigma)$ is the space of marked complex structures (or equivalently, marked hyperbolic structures) on Σ and plays a significant role in the theory of Riemann surfaces and 3-manifolds.

We note that $\mathcal{T}(\Sigma)$ comes equipped with several metrics, the Teichmüller metric and the Weil–Petersson metric. These metrics arise from natural optimization problems among Riemann surfaces (i.e., maximal dilatation, energy, etc.) and induces a $\mathrm{Mod}(\Sigma)$-equivariant geodesic flow on the tangent bundle of $\mathcal{T}(\Sigma)$. Hence, it is worth studying the action of $\mathrm{Mod}(\Sigma)$ on $\mathcal{T}(\Sigma)$ in terms of these metrics.

In addition to the metric structures, $\mathcal{T}(\Sigma)$ is linked to several bordifications. We will be interested in Thurston's compactification among them. We first regard each point in $\mathcal{T}(\Sigma)$ as a projectivized functional of simple closed curves on Σ. In this setting, limit points of $\mathcal{T}(\Sigma)$ include the (projectivized) intersection number with a fixed curve γ, which is approximated by first fixing a point x in $\mathcal{T}(\Sigma)$ and then shrinking γ or applying powers of Dehn twist about γ on x. In general, limit points are projectivized measured foliations on Σ and we denote their collection (called the Thurston boundary) by $\mathcal{PMF}(\Sigma)$. $\mathcal{T}(\Sigma) \cup \mathcal{PMF}(\Sigma)$ is homeomorphic to a $(6g - 6)$-dimensional closed ball and has been used to reveal the dynamical properties of mapping classes, e.g., the Nielsen–Thurston classification. Our aim is to observe that this bordification also naturally arises in the study of random walks on $\mathrm{Mod}(\Sigma)$. For more details on measured foliations and Thurston's compactification, consult [30].

One obstacle to this plan is that the geometry of $\mathcal{T}(\Sigma)$ is not exactly the same with, say, negatively curved manifolds. $\mathcal{T}(\Sigma)$ does have an analogy with a simply-connected hyperbolic space and its quotient $\mathcal{T}(\Sigma)/\mathrm{Mod}(\Sigma)$ resembles a cusped negatively curved

manifold. Nevertheless, $\mathcal{T}(\Sigma)$ equipped with the Teichmüller metric is not Gromov hyperbolic ([52, 74]), which means that geodesic triangles in $\mathcal{T}(\Sigma)$ are not uniformly thin. Hence, it is tempting to invent a new Gromov hyperbolic space by collapsing all 'fat' regions of $\mathcal{T}(\Sigma)$.

This was furnished by Harvey's construction of the curve complex $\mathcal{C}(\Sigma)$ of Σ, whose vertices are simple closed curves on Σ and edges are drawn between disjoint simple closed curves. $\mathcal{C}(\Sigma)$ is Gromov hyperbolic, and two spaces are related by the projection map $\pi : \mathcal{T}(\Sigma) \to \mathcal{C}(\Sigma)$ that picks the shortest curve on the surface. Moreover, this projection induces a quasi-isometry between $\mathcal{C}(\Sigma)$ and a modification of $\mathcal{T}(\Sigma)$. Here, the modification is as follows: for each simple closed curve γ on Σ, we collapse all points of $\mathcal{T}(\Sigma)$ with short γ. Hence, it is impossible to escape to infinity by shrinking a curve in this modified space. This implies that the boundary structure of $\mathcal{C}(\Sigma)$ does not include simple closed curves.

Nonetheless, $\mathcal{T}(\sigma)$ and $\mathcal{C}(\Sigma)$ partially share the boundary structure. Let $\mathcal{MIN}(\Sigma) \subseteq \mathcal{PMF}(\Sigma)$ be the set of projective measured foliations that correspond to minimal foliations, the foliations that do not have closed leaves. Let $\widetilde{\mathcal{MIN}}(\Sigma)$ be its quotient by the equivalence relation of having trivial intersection. Finally, $\mathcal{UE}(\Sigma)$ denotes the set of uniquely ergodic foliations on Σ, which can be regarded as a subset of both $\mathcal{MIN}(\Sigma)$ and $\widetilde{\mathcal{MIN}}(\Sigma)$. Klarreich's theorem in [59] states that the Gromov boundary of $\mathcal{C}(\Sigma)$ is the set of minimal foliations $\widetilde{\mathcal{MIN}}(\Sigma)$. For details on the Gromov products, Gromov hyperbolic spaces and the Gromov boundary, see [13, 23, 42, 88].

One dynamical quantity of an isometry g of X is the (asymptotic) translation length of g defined by

$$\tau(g) := \lim_{n \to \infty} \frac{1}{n} d(o, g^n o).$$

Recall that an isometry g of a Gromov hyperbolic space X falls into exactly one of the following categories (see [20, 45] for instance):

- g has a bounded orbit (elliptic);
- g is not elliptic and has a unique fixed point in ∂X (parabolic);
- g has two fixed points, an attractor and a repeller, in ∂X (loxodromic).

In particular, g is loxodromic if and only if $n \mapsto g^n x$ is a quasi-isometry for one (hence all) $x \in X$, if and only if $\tau(g) > 0$.

A similar classification of mapping classes, called the Nielsen–Thurston classification, was proven by referring to the mapping class group action on Teichmüller space and its boundary. Thurston observed that each $g \in \mathrm{Mod}(\Sigma)$ acts on $\mathcal{T}(\Sigma) \cup \mathcal{PMF}(\Sigma) \simeq \overline{B^{6g-6}}$ as a self-homeomorphism and has a fixed point. We have (not mutually exclusive) four cases:

(1) if g has a fixed point in $\mathcal{T}(\Sigma)$, then g is of finite order (periodic);

(2) if g fixes the projective class of a rational foliation, then g fixes a multicurve (reducible);

(3) if g fixes an arational measured foliation without scaling, then g is of finite order (periodic);

(4) if g scales up an arational measured foliation (hence fixes its projective class), then g is pseudo-Anosov.

Here, the last case and the other cases are mutually exclusive. In view of their actions on $\mathcal{C}(\Sigma)$, periodic and reducible mapping classes are elliptic and pseudo-Anosov mapping classes are loxodromic [73].

When X is a Gromov hyperbolic space or $\mathcal{T}(\Sigma)$ and G is its isometry group, we say that a subgroup or a subsemigroup H is *non-elementary* if it contains two loxodromics that do not share fixed points at infinity. We say that a probability measure μ is *non-elementary* if the subsemigroup generated by its support is non-elementary. In this case, we also say that the random walk ω generated by μ is also non-elementary. We say that a probability measure μ is *non-arithmetic* if some convolution $(\operatorname{supp} \mu)^N$ of the support of μ contains two loxodromics with different translation lengths.

A standing assumption of this paper is that

X is either a Gromov hyperbolic space or the Teichmüller space $\mathcal{T}(\Sigma)$ of a closed orientable surface Σ of genus at least 2. The isometry group G is a countable group acting on X as isometries, μ is a non-elementary probability measure on G and ω is the random walk generated by μ.

3. Early works and ergodic theorems

We can apply ergodic theory to random walks by using the ergodicity of the step shift map. This strategy is observed in, for example, the pioneering work by Furstenberg and Kesten [36] of the product of random matrices. This is generalized to the so-called Kingman's subadditive ergodic theorem [56–58]. We present one version of this theorem as follows.

Theorem 3.1 ([91, Theorem 8.10]). *Let $(\Omega, \mathbb{P})$ be a probability space and $U \colon \Omega \to \Omega$ be a measure-preserving transformation. If W_n is a non-negative real-valued random variables on Ω satisfying the subadditivity $W_{n+m} \leq W_n + W_m \circ U^n$ for all $m, n \in \mathbb{N}$, and W_1 has finite first moment, then there is a U-invariant random variable W_∞ such that*

$$\lim_{n \to \infty} \frac{1}{n} W_n = W_\infty$$

almost surely and in $L^1(\Omega, \mathbb{P})$. If U is ergodic in addition, then W_∞ is constant a.e.

The subadditive ergodic theorem is particularly useful in non-commutative settings, where one cannot directly bring the results about Euclidean random walks. For example, let us consider a random walk $\omega = (\omega_n)$ on G. After fixing a reference point $o \in X$, the displacement $d(o, \omega_n o)$ becomes subadditive. Then the existence of the escape rate

follows from the subadditive ergodic theorem. Additional properties of G and its action on X, such as the non-amenability of G and the properness of the action, guarantee that the escape rate is strictly positive and the random walk escapes to infinity. Hence, one can discuss the hitting measure induced on a suitable boundary of X, which is useful to investigate the asymptotics of the random walk.

In fact, such a boundary can be constructed without referring to the action of G on X. This program was initiated by Furstenberg in [32]. Furstenberg considered the set $\mathcal{A}$ of all functions on $G \times \Omega$ of the form

$$F(g, \omega) := \lim_{n \to \infty} f(g \cdot \omega_n)$$

for some left uniformly continuous function f such that the above limit exists almost surely. Furstenberg then showed that

(1) $\mathcal{A}$ has a 1-1 correspondence with the set of μ-harmonic functions, and

(2) $\mathcal{A}$ corresponds to the space of all continuous functions on a compact space Π by using the Gelfand representation.

Then every left uniformly continuous, μ-harmonic function on G admits a Poisson representation regarding a continuous function on Π. More generally, the space of bounded μ-harmonic functions on G and the space of bounded measurable functions on Π has 1-1 correspondence via the Poisson representation. This space Π is called a *Poisson space* by Furstenberg.

The construction of Poisson spaces is purely measure-theoretical, but Furstenberg also proved that Poisson spaces for a semi-simple Lie group are always finite covers of a maximal boundary of G endowed with a stationary measure. Here, a boundary of G refers to a compact homogeneous (or minimal, in general) G-space whose points can be approximated by elements of G when probed with measures. More precisely, for every probability measure ν on M, there exists a sequence $\{g_n\}$ in G such that $g_n * \nu \to \delta_h$ for some $h \in M$. Using this observation, Furstenberg deduced the following rigidity result.

Theorem 3.2 ([34, Theorem 1]). *For $d \geq 2$ and $n \geq 3$, no countable group can become a cocompact lattice of* $\mathrm{Isom}(\mathbb{H}^d)$ *and* $SL(n, \mathbb{R})$ *simultaneously.*

We now return to the case that G is a countable group acting on X. Recall the above definition of boundaries. A pair (B, ν) of a G-space B and a μ-stationary measure ν on B is called a μ-*boundary* if for a.e. sample path $\omega = (\omega_n)_n$, $\omega_n * \nu$ converges to an atom on B. With this a.e. convergence, we can attach B to G and say that (B, ν) captures the eventual fate of a.e. path arising from the random walk ω on G. From a measure-theoretic viewpoint, all μ-boundaries arise as quotients of a maximal μ-boundary, which we call the *Poisson boundary* of (G, μ). While Furstenberg constructed it using the Gelfand representation, one can instead define it as the space of ergodic components of the Bernoulli shift (see [35, Theorem 3.1] and [54]). Moreover, this is the Poisson space described before: the Poisson formula gives a 1-1 correspondence between μ-harmonic functions on G and bounded measurable functions on the Poisson boundary of (G, μ).

After this measure-theoretic discussion, one can ask whether a Poisson boundary of a group G can be modeled on the boundary ∂X of the space X that G is acting on. Kaimanovich and Masur brought this perspective to mapping class groups in their work [54]. Kaimanovich and Masur asked whether $\mathcal{PMF}(\Sigma)$ is the correct boundary of $\mathcal{T}(\Sigma)$ in the measure-theoretical viewpoint, that means, it hosts a μ-stationary measure ν such that $(\mathcal{PMF}, \nu)$ becomes the Poisson boundary of (Mod, μ). We present Kaimanovich–Masur's result below.

Theorem 3.3 ([54, Theorems 2.2.4, 2.3.1]). *Let G be the mapping class group $\mathrm{Mod}(\Sigma)$ acting on $X = \mathcal{T}(\Sigma)$ and μ be a non-elementary measure on G. Then there exists a unique μ-stationary ν on $\mathcal{PMF}(\Sigma)$, which is purely non-atomic and concentrated on $\mathcal{UE} \subseteq \mathcal{PMF}$, and $(\mathcal{UE}, \nu)$ is a μ-boundary. In fact, ν is the hitting measure of μ on $\mathcal{PMF}$; for any $x \in \mathcal{T}(\Sigma)$ and $\mathbb{P}$-a.e. sample path $\omega = (\omega_n)$, $\omega_n x$ converges in $\mathcal{PMF}$ to a limit $F = F(\omega) \in \mathcal{UE}$, and the distribution of the limits $F(\omega)$ is given by ν.*

If μ has a finite entropy and finite first logarithmic moment with respect to the Teichmüller metric in addition, then $(\mathcal{PMF}, \nu)$ is the Poisson boundary of (Mod, μ).

The proof relies on the geometry of $\mathcal{T}(\Sigma)$ and the structure of $\mathcal{PMF}$, which we briefly sketch now. As the escape to infinity is not established *a priori*, the μ-stationary measure ν is not constructed as the hitting measure of μ; rather, ν is first defined as a μ-stationary measure on a compact space, and the escape to infinity follows from the property of ν. The existence of ν relies on the fact that $\mathcal{PMF}(\Sigma)$ is a compact sphere. Using the structure of $\mathcal{PMF}$, namely, that $\mathcal{PMF} \setminus \mathcal{MIN}$ admits a countably infinite partition that respects $\mathrm{Mod}(\Sigma)$-action, one can deduce that ν is concentrated on $\mathcal{MIN}$.

Let us now consider the quotient measure $\tilde{\nu}$ of ν on $\widetilde{\mathcal{MIN}}$, the set of equivalence classes of minimal foliations. Then a.e. path (ω_n) has the limit point $F(\omega_n) \in \widetilde{\mathcal{MIN}}$ in the sense that $\omega_n \tilde{\nu} \to \delta_{F(\omega)}$ weakly ([54, Lemma 2.2.3]). This map $F \colon \omega \mapsto \widetilde{\mathcal{MIN}}$ pushes the measure $\mathbb{P} = \mu^{\mathbb{Z}}$ on $G^{\mathbb{Z}}$ forward to $\tilde{\nu}$, which is non-atomic, so the Poisson boundary of $(\mathrm{Mod}(\Sigma), \mu)$ becomes non-trivial. This implies that the random walk escapes to infinity almost surely. The final technical step is to show that actually ν is concentrated on $\mathcal{UE}$ so that $\tilde{\nu}$ coincides with ν. This relies on Masur's divergence result [71, Theorem 1.1] that Teichmüller geodesics heading to $\xi \notin \mathcal{UE}$ escapes every thick part of $\mathcal{T}(\Sigma)$ (and hence the mapping class group orbit). The uniqueness of ν now follows from the integral representation: F actually pushes $\mathbb{P}$ forward to ν. Now the orbit convergence $\omega_n o \to F(\omega)$ for a.e. ω follows from the properties of uniquely ergodic foliations and universally convergent sequences.

In order to show that $(\mathcal{PMF}, \nu)$ is maximal, we invoke the strip approximation criterion introduced in [53]. Namely, given that the entropy is finite, the maximality of $(\mathcal{PMF}, \nu)$ follows once we construct a measurable $\mathrm{Mod}(\Sigma)$-equivariant 'strips'

$$S \colon (F_-, F_+) \in \mathcal{UE} \times \mathcal{UE} \mapsto S(F_-, F_+) \subseteq \mathrm{Mod}(\Sigma)$$

such that for all $g \in G$ and $\nu_- \otimes \nu_+$-a.e. $(F_-, F_+) \in \mathcal{UE} \times \mathcal{UE}$, we have

$$\frac{1}{n} \log \#\big(S(F_-, F_+)g \cap B(\mathrm{id}, |\omega_n|)\big) \to 0 \quad \text{as } n \to \infty$$

in probability, where $|\omega_n|$ denotes the word metric norm of ω_n with respect to a finite generating set of $\mathrm{Mod}(\Sigma)$ and $B(\mathrm{id}, |\omega_n|)$ denotes the ball of radius $|\omega_n|$ about the identity element id. In this argument, we take

$$S(F_-, F_+) = \{h \in \mathrm{Mod}(\Sigma) : d(ho, [F_-, F_+]) \le M\}$$

for some suitable M. As $\mathrm{Mod}(\Sigma)$ acts on $\mathcal{T}(\Sigma)$ properly discontinuously, balls of radius M (with respect to the Teichmüller metric) in $\mathcal{T}(\Sigma)$ contain finitely many translates ho of o and $\#(S(F_-, F_+)g \cap B(e, k))$ grows at most linearly along k. Given this, the bottleneck of the growth of $\#(S(F_-, F_+)g \cap B(e, |\omega_n|))$ becomes the growth rate of $|\omega_n|$ in probability. It turns out that finite logarithmic moment condition controls the overall growth in a subexponential manner.

Recall now Theorem 3.2 of Furstenberg: there, the dichotomy between two types of lattices is as given as follows. If G is a lattice in a semi-simple Lie group of rank at least 2, then there exists a measure μ with $\mathrm{supp}\,\mu = G$ and a number $\varepsilon > 0$ such that the following holds. If μ-harmonic functions f_1, f_2 on G satisfy that

(1) $0 \le f_1, f_2 \le 1$ on G, and

(2) $f_1(\mathrm{id}), f_2(\mathrm{id}) \ge 0.5 - \varepsilon$,

then (3) $\min[f_1(g), f_2(g)]$ does not tend to zero as $g \to \infty$. In contrast, if G is a lattice in $\mathrm{Isom}(\mathbb{H}^d)$, then for any μ and $\varepsilon > 0$ we can construct μ-harmonic functions f_1, f_2 that satisfy (1) and (2), but not (3). Kaimanovich and Masur similarly constructed such μ-harmonic functions on $\mathrm{Mod}(\Sigma)$ and deduced that $\mathrm{Mod}(\Sigma)$ is not a lattice in semi-simple Lie groups.

The storyline so far already shows the interplay between phenomena inside $\mathcal{T}(\Sigma)$ and limiting phenomena on $\mathcal{PMF}(\Sigma)$. More specifically, (a) the escape to infinity and (b) the finite growth rate of strips are intimately related to (c) the characterization of the Poisson boundary. Here, a part of (c) helped establish (a), while (b) contributed to establishing another part of (c). Later we will see opposite situations, where (a) directly leads to a part of (c) and (c) helps establish a variant of (b).

A final remark is on the usage of Teichmüller geometry. Kaimanovich–Masur's work was motivated by the Gromov boundary of Gromov hyperbolic spaces that possess the visual measure and the Patterson–Sullivan measure [78, 86] of negatively curved manifolds. It is expected that the harmonic measure and these measures are mutually singular in compact manifolds with non-constant negative curvature [60, 61]. Since the action of $\mathrm{Mod}(\Sigma)$ on $\mathcal{T}(\Sigma)$ has co-finite volume and $\mathcal{T}(\Sigma)$ has variable curvature in some sense, we can also expect the mutual singularity of those measures. We will revisit this topic later.

The main obstacle to this problem was that both $\mathrm{Mod}(\Sigma)$ and $\mathcal{T}(\Sigma)$ are not Gromov hyperbolic. Nevertheless, Kaimanovich and Masur exploited partial hyperbolicity of the Teichmüller space in order to establish the escape to infinity (see [54, Section 1.5], for example). This alludes to a unified storyline for random walks on mapping class groups and hyperbolic groups, which we will see in the subsequent sections.

4. Curve complex

Another space on which $\mathrm{Mod}(\Sigma)$ acts is the curve complex $\mathcal{C}(\Sigma)$. Introduced by Harvey in [48] as an analogy for the Tits building, this complex became a central object in the study of mapping class groups thanks to Masur–Minsky's theory ([72, 73]). In particular, Masur and Minsky proved that

(1) The shortest curve projection $\pi\colon \mathcal{T}(\Sigma) \to \mathcal{C}(\Sigma)$ sends geodesics to unparametrized quasi-geodesics (of uniform quality),

(2) $\mathcal{C}(\Sigma)$ is δ-hyperbolic for some $\delta = \delta(\Sigma)$, and

(3) $\mathcal{T}(\Sigma)$ and $\mathrm{Mod}(\Sigma)$ are weakly relatively hyperbolic in the sense of Farb [29].

In fact, the constant $\delta(\Sigma)$ can be taken as 17 for any surface Σ ([49]; see also [2, 12, 18]). We warn that mapping class groups are not relatively hyperbolic in general. Note also that $\mathcal{T}(\Sigma)$ and $\mathcal{C}(\Sigma)$ are not quasi-isometric, nor are $\mathrm{Mod}(\Sigma)$ and $\mathcal{C}(\Sigma)$. Furthermore, the orbit map $g \mapsto go$ from $\mathrm{Mod}(\Sigma)$ to $\mathcal{T}(\Sigma)$ is not a quasi-isometric embedding (but see [28]).

As $\mathcal{T}(\Sigma)$ does, $\mathcal{C}(\Sigma)$ also captures the dynamics of each mapping class of $\mathrm{Mod}(\Sigma)$. In particular, pseudo-Anosov mapping classes have positive translation lengths on $\mathcal{C}(\Sigma)$ and $\mathcal{T}(\Sigma)$. Comparing these two translation lengths is an interesting question.

To investigate random walks on $\mathcal{C}(\Sigma)$, let us recall the work of Kaimanovich on the Poisson boundary of hyperbolic groups [53]. Note that hyperbolic groups are finitely generated groups whose Cayley graph is Gromov hyperbolic with respect to a finite generating set. Consequently, the given Cayley graph X is locally finite and the Gromov boundary ∂X is a compact metrizable space.

Given a non-elementary measure μ on a hyperbolic group G, one can first obtain a μ-stationary measure ν on ∂X by considering a limit point of the Cesàro mean

$$\mu_n := \frac{1}{n}(\mu + \mu * \mu + \cdots + \mu^{*n}),$$

where such a limit point always exists by the compactness of ∂X (cf. [54, Lemma 2.2.1]). One can show that ν is non-atomic, and for almost every sample path $\omega = (\omega_n)_n$ the sequence of measures $(\omega_n \nu)$ has a weak-$*$ limit ν_ω on ∂X. Meanwhile, since μ is non-elementary, the random walk is transient and almost every orbit path $(\omega_n o)_n$ has limit points on ∂X. Lemma 2.2 of [53] asserts that the Dirac measures on such limit points are limit points of the sequence $(\omega_n \nu)_n$. It follows that the orbit path $(\omega_n o)_n$ also converges

to a unique limit $\xi_\omega \in \partial X$ and ν is indeed the hitting measure, i.e.,

$$\nu(A) = \mathbb{P}\{\omega : \xi_\omega \in A\}.$$

Kaimanovich then suggests two criteria for Poisson boundaries: the strip approximation criterion mentioned before and the ray approximation criterion. As in the previous discussion, balls of radius M in the Cayley graph X of G contain a uniformly bounded number of orbit points. The strip approximation criterion then implies that $(\partial G, \nu)$ is maximal when μ has finite entropy and finite logarithmic moment.

The above argument works for *locally compact* Gromov hyperbolic spaces in general, where $X \cup \partial X$ becomes a compact space. This compactness is critical as it guarantees the μ-stationary measure ν and the a.e. convergence of $\omega_n \nu$ to a Dirac measure. Unfortunately, $\mathcal{C}(\Sigma)$ is not locally compact since every vertex has infinite valency.

Let us now investigate the ray approximation (also known as geodesic tracking) criterion. This asserts that if there exist nice rays $(\pi_1(\xi), \pi_2(\xi), \ldots)$ on G for each $\xi \in \partial G$ of a μ-boundary $(\partial G, \nu)$, then $(\partial G, \nu)$ is the Poisson boundary of (G, μ). Here the niceness condition is that the ray projections are measurable and that for a.e. $\omega = (\omega_n)$, if ω heads to $\xi \in \partial G$, then $\pi_n(\xi)$ and ω_n deviate from each other sublinearly with respect to some gauge function $\mathcal{G}$.

When G is properly discontinuously acting on a proper space X, then $|g|_{\mathcal{G}} := d(o, go)$ serves as a gauge. In this situation, if the progress $d_X(o, \omega_n o)$ of the random walk is sublinear, then the trivial μ-boundary becomes the Poisson boundary and G becomes amenable. Conversely, random walks on non-amenable groups show positive escape rate. Nonetheless, the strategy does not apply to $\mathrm{Mod}(\Sigma)$ acting on $\mathcal{C}(\Sigma)$, as $\mathcal{C}(\Sigma)$ is not proper and each $\alpha \in \mathcal{C}(\Sigma)$ has infinite stabilizer. These limitations necessitated a radically different approach for $\mathcal{C}(\Sigma)$.

Maher and Tiozzo made the breakthrough in [67]. They considered weakly hyperbolic groups, the case where X is a separable geodesic Gromov hyperbolic space and G is a non-elementary subgroup of $\mathrm{Isom}(X)$. In this setting, they constructed the Poisson boundary for (G, μ) on the Gromov boundary of X. Before delving into their work, let us explore preceding observations by Maher.

Maher first deduced in [65] the following: when $\mathrm{supp}\,\mu$ generates a non-elementary subgroup of $\mathrm{Mod}(\Sigma)$, the probability that ω_n is not pseudo-Anosov converges to zero as $n \to \infty$. Maher mixes the probabilistic feature of harmonic measure and the group-theoretical structure of $\mathrm{Mod}(\Sigma)$ to achieve this. The first ingredient is the relative conjugacy bounds of non-pseudo-Anosovs [65, Lemma 5.5]. Given this, it would suffice to show the transience of the set R of elements whose relative conjugacy length is bounded by some constant (this is proved in [65, Section 5.2] and it would lead to an even stronger result in the almost sure sense).

One possible approach is to use the fact that the harmonic measure $\nu(\partial R)$ of the limit set of R dominates the probability of recurrence of R. Unfortunately, the limit set of R is the entire $\mathcal{PMF}$ so this strategy fails. Instead, we consider a subset R_k of R that is

contained in the union of centralizers of elements with word norm $1, \ldots, k$. We now claim that

$$\textit{each of these centralizers has infinite copies in } \langle\!\langle \operatorname{supp} \mu \rangle\!\rangle. \qquad (*)$$

Given this claim, a similar argument to [54, Lemma 2.2.2] shows that the centralizers have harmonic measure 0. It is now argued that $\limsup_n \mathbb{P}(\omega_n \in R \setminus R_k)$ decreases down to zero as $k \to \infty$. The final ingredient is to ensure $(*)$: it can be guaranteed by passing to a finite cover of Σ, if necessary.

Maher then proved in [64] that random walks show linear progress in the curve complex metric. Note that in contrast with $\operatorname{Mod}(\Sigma)$ with the word metric, $\mathcal{C}(\Sigma)$ is not locally compact and thus the standard non-amenability argument does not apply. Maher's idea was to construct a sequence of nested *halfspaces* associated with each trajectory and show that the length of the sequence grows linearly in probability. This notion records a useful partial history of the random walk, rather than merely recording the final product, in the form of stopping times. This strategy also helped improve the result of [65] into the exponential decay of the probability of non-pseudo-Anosovs, given that the transition probability μ is finitely supported [66].

Since these results were established based on the action on $\mathcal{C}(\Sigma)$, the key challenge was to remove the properness of X and instead rely on the Gromov inequality among points only. Another important ingredient was that $\mathcal{T}(\Sigma)$ and $\mathcal{C}(\Sigma)$ partially share the same boundary structure, which coincides when we are concerned with μ-boundaries. Recall Theorem 3.3 again: $(\mathcal{PMF}, \nu)$ is a μ-boundary concentrated on $\mathcal{UE} \subseteq \widetilde{\mathcal{MIN}}$. Here, $\mathcal{UE}$ is not only a subset of $\partial \mathcal{T}(\Sigma)$ but also a subset of $\partial \mathcal{C}(\Sigma) = \widetilde{\mathcal{MIN}}$. Thanks to this coincidence, one can use the μ-boundary obtained from the dynamics on $\mathcal{T}(\Sigma)$ to investigate the asymptotic behavior of the random walk on $\mathcal{C}(\Sigma)$.

This effort culminated in Maher–Tiozzo's extensive work in [67]. There, X is only assumed to be separable, geodesic and Gromov hyperbolic, and G is a non-elementary countable subgroup of $\operatorname{Isom}(X)$. Not assuming that X is proper, this setting includes $\operatorname{Mod}(\Sigma)$ acting on $\mathcal{C}(\Sigma)$ and $\operatorname{Out}(F_n)$ acting on the complex of free factors. As hinted at by the preceding works, we obtain the μ-boundary of G not directly from X but from other spaces. We then extract the asymptotic phenomenon from the coinciding boundary structure.

To this end, Maher and Tiozzo exploit the horofunction boundary $\bar{X}_\infty^h$, whose usage is motivated from Calegari–Maher's work [14]. Using the local minimum map $\phi \colon \bar{X}_\infty^h \to \partial X$ onto the Gromov boundary, one may

(1) push measures on $\bar{X}_\infty^h$ forward to ∂X, and

(2) transfer the (weak-$*$) convergence of measures on $\bar{X}_\infty^h$ to that on ∂X.

As the horofunction compactification $\bar{X}^h$ of a separable metric space X is always compact, the existence of a μ-boundary ν on $\bar{X}^h$ and its concentration on the boundary $\bar{X}_\infty^h$ follow. Also deduced is a.e. convergence of measures $(\omega_n \nu)$ on $\bar{X}_\infty^h$. All of these results can be pushed forward to ∂X; for example, the invariant measure ν on $\bar{X}^h$ gives rise to another measure $\widetilde{\nu}$ on ∂X.

Meanwhile, given the convergence of $(\omega_n \widetilde{\nu})$ in ∂X, its convergence to an atom follows from the Gromov hyperbolicity. This leads to the fact that a.e. $\omega = (\omega_n)$ has a limit point $\lambda = \lambda(\omega) \in \partial X$ such that $(\omega_n \widetilde{\nu}) \to \delta_\lambda$ weakly. The final touch is to deduce the orbit convergence from the weak convergence using shadows and the Gromov hyperbolicity; this is not available on $\bar{X}^h$ but on $X \cup \partial X$. It also follows that $\widetilde{\nu}$ is non-atomic; if not, wandering of the maximum atom yields a contradiction with the assumption that G is non-elementary.

Having established the relationship between the invariant, non-atomic measure $\widetilde{\nu}$ and the escape to infinity of sample paths, plenty of dynamical properties of the random walk can be obtained. These include positive escape rate, geodesic tracking and the linear growth of translation length. We summarize these below.

Theorem 4.1 (cf. [67, Theorems 1.2, 1.3, 1.4]). *Let G be a countable group acting on a separable Gromov hyperbolic space X and μ be a non-elementary measure on G. Then the following hold:*

(1) (Convergence to the boundary) *For a.e. sample path $\omega = (\omega_n)_n$, there exists $\xi \in \partial G$ such that $\lim_n \omega_n o = \xi$.*

(2) (Positive drift) *There exists $L > 0$ such that*

$$\liminf_{n \to \infty} \frac{d_X(o, \omega_n o)}{n} \geq L \quad a.s.$$

(3) (Geodesic tracking) *If μ has finite first moment, for a.e. $\omega = (\omega_n)$, there exists a quasi-geodesic ray γ such that*

$$\lim_{n \to \infty} \frac{d_X(\omega_n o, \gamma)}{n} = 0.$$

(4) (Growth of translation length) *There exists $L > 0$ such that*

$$\mathbb{P}(\tau(\omega_n) \leq Ln) \to 0 \quad as\ n \to \infty\ a.s.$$

The final term is related to our main concern. Since a mapping class is pseudo-Anosov if and only if it acts on the curve complex loxodromically, we deduce that random walks on mapping class groups eventually become loxodromic in probability.

A distinction should be made for random walks with bounded support. In this case, the argument of [66] indicates that the probability of a shadow decreases exponentially as its distance from the origin increases. Using this, the above results are promoted into the following form.

Theorem 4.2 (cf. [67, Theorems 1.2, 1.3, 1.4]). *Let X, G, μ be as in Theorem 4.1, and suppose further that μ has bounded support. Then the following hold:*

(1) *There exists $L, K > 0$ and $0 < c < 1$ such that*

$$\mathbb{P}(d_X(o, \omega_n o) \leq Ln) \leq Kc^n$$

for all n.

(2) *For a.e. $\omega = (\omega_n)$, there exists a quasi-geodesic ray γ such that*

$$\limsup_{n \to \infty} \frac{d_X(\omega_n o, \gamma)}{\log n} < \infty.$$

(3) *There exists $L, K > 0$ and $0 < c < 1$ such that*

$$\mathbb{P}(\tau(\omega_n) \leq Ln) \leq Kc^n$$

for all n.

In particular, the escape to infinity and the linear growth of translation length occur almost surely, rather than in probability; the geodesic tracking occurs in a logarithmic manner, which is stronger than the sublinear one. In fact, combining Maher–Tiozzo's theory with Benoist–Quint's theory (to be explained later) yields the almost sure phenomena under finite second moment condition, as Dahmani and Horbez remark. We note that [14, 81, 85] are also concerned with finitely supported random walks and deduce that the probability of non-pseudo-Anosov elements decays exponentially.

We have yet to discuss the ultimate goal that $(\partial X, \tilde{\nu})$ is indeed the Poisson boundary of (G, μ). This requires a mild geometric condition on the action, namely, the acylindricity. The statement holds given that G acts on X acylindrically and μ has finite entropy and first moment.

Maher–Tiozzo's work shows that if loxodromic isometries are abundant, the coupling of the group structure and the space is not strictly required for investigating the dynamical features of random walks. Note that Theorem 4.1 does not require the action of G on X to be cocompact or properly discontinuous (which would also restrict the geometry of X). Yet, the Gromov hyperbolicity of X plays a significant role throughout the argument. We also require X to be separable in order to control the topology of the horofunction compactification (in fact, by [10, Theorem 4.1] and [46, Remark 4], one can remove the separability assumption). In the next section, we will examine how critical these conditions are.

5. Teichmüller spaces

We now discuss random walks on $\text{Mod}(\Sigma)$ acting on $\mathcal{T}(\Sigma)$. Recall that the translation length $\tau_{\mathcal{T}(\Sigma)}(g)$ of a pseudo-Anosov mapping class g with respect to the Teichmüller metric and the stretch factor $\lambda(g)$ of g have the relationship $\lambda(g) = \log \tau_{\mathcal{T}(\Sigma)}(g)$. Hence, investigating the asymptotics of the translation length on Teichmüller spaces can reveal the topological/dynamical properties of a generic mapping class.

An immediate difficulty is that Teichmüller spaces are not Gromov hyperbolic. To observe this, consider a geodesic triangle with vertices $o, T_A^n o, T_B^{-n} o$ for a point $o \in \mathcal{T}(\Sigma)$ and Dehn twists T_A, T_B along disjoint curves A, B; the Hausdorff distance of $[T_A^n o, T_B^{-n} o]$ from $[o, T_A^n o] \cup [o, T_B^{-n} o]$ increases logarithmically [74]. Another evidence is that the

part of $\mathcal{T}(\Sigma)$ where a collection of disjoint curves $\{\gamma_1, \ldots, \gamma_n\}$ is pinched resembles the product space $\mathcal{T}(\Sigma \setminus \{\gamma_1, \ldots, \gamma_n\}) \times \prod_{i=1}^{n} \mathbb{H}^2$ (see [76]).

Despite this failure, Teichmüller spaces (Mod(Σ), resp.) share many aspects with negatively curved spaces (hyperbolic groups, resp.). For example, Margulis' work on the exponential growth of volumes and deck transformation orbits of a negatively curved manifold has an analogy in the setting of Teichmüller spaces and mapping class group orbits [4]. The uniform exponential growth of hyperbolic groups is also copied onto mapping class groups ([1, 68]). In the same vein, many efforts have been made to copy the 'thin triangle phenomenon' from Gromov hyperbolic geodesic spaces onto Teichmüller spaces.

Let us begin with Duchin's work on the geodesic tracking à la Kaimanovich, who suggested two criteria for determining the Poisson boundary of groups, the sublinear geodesic tracking and the strip approximation. Given that the strip approximation was effective enough to determine the Poisson boundary of mapping class groups, Kaimanovich asked whether the other criterion works, i.e., random walks on Mod(Σ) acting on $\mathcal{T}(\Sigma)$ show sublinear geodesic tracking (cf. [53, Section 0]).

Kaimanovich–Masur's work already suggests the right candidate for the approximating geodesic. Namely, almost every sample path $\omega = (\omega_n)$ possesses the limit point $F(\omega) \in \mathcal{UE}$ such that $\omega_n o$ converges to $F(\omega)$ when viewed in Thurston's compactification. Each geodesic $[o, \omega_n o]$ is recorded at the initial point o with the initial quadratic differential $\varphi_n \in QD_o$. Using Masur's comparison of Thurston and visual boundaries [70], it follows that $\varphi_n \to \varphi$ in QD_o and the geodesic $\gamma = \gamma(\omega)$ with the initial quadratic differential φ converges to $F(\omega)$. Here, we are critically using the fact that $F(\omega)$ belongs to $\mathcal{UE}$ almost surely; compare this with the result in [63].

As hinted before, the nuisance is the thin part of $\mathcal{T}(\Sigma)$. If $\gamma(\omega)$ were always living inside a thick part of $\mathcal{T}(\Sigma)$, then one could apply the theory for Gromov hyperbolic spaces explained in Section 4. Although γ is approximated by geodesics connecting thick points, γ may take a long excursion in the thin part of $\mathcal{T}(\Sigma)$. This led Duchin to focus on the phenomenon inside thick parts [26]. More precisely, Duchin showed that when μ has finite first moment, a.e. $\omega = (\omega_n)$ possesses a geodesic $\gamma : [0, \infty) \to \mathcal{T}(\Sigma)$ beginning from o such that

$$\frac{1}{n} d(\omega_n o, \gamma) 1_K\big(\gamma(d(o, \omega_n o))\big) \to 0 \quad \text{as } n \to 0.$$

Here K denotes a thick part of $\mathcal{T}(\Sigma)$.

Duchin's idea was to bring one particular property of 'thin triangles' in Gromov hyperbolic spaces to some collection of triangles in $\mathcal{T}(\Sigma)$. In order for a random walk to be aligned along a geodesic, it is favored that consecutive orbits $\omega_n o$ form a sort of 'highly obtuse triangles'; in such a case, $d(\omega_{n-k} o, \omega_n o) + d(\omega_n o, \omega_{n+k} o)$ would be comparable to $d(\omega_{n-k} o, \omega_{n+k} o)$. Assuming such distance relations, we now conversely hope that each $\omega_n o$ is not far away from the limiting geodesic γ. Motivated by this, Duchin required the following property. Let us first fix $A > 0$, and consider a geodesic triangle $\triangle xyz$ with the longest side $[y, z]$. Let $w \in [y, z]$ be such that $d(x, y) = d(w, y)$. Then the desired

property is

$$d(w, x) < A[d(y, x) + d(x, z) - d(y, z)].$$

For example, $A = 2$ works for triangles in $\mathbb{R}$-trees. In general, $A = 2$ works for geodesic triangles in a Gromov hyperbolic space with a sidelength threshold. Duchin showed that geodesic triangles such that $w \in K$, together with a side length threshold, satisfy this property for some $A = A(K)$. The condition $w \in K$ led to the subsequence restriction in the theorem.

Before explaining how Rafi strengthened this approach, we digress to the complete sublinear geodesic tracking proved by Tiozzo [87]. Tiozzo's approach applies not only to $\mathrm{Mod}(\Sigma)$ but also to groups acting on proper Gromov hyperbolic spaces, groups with infinitely many ends, and groups acting on CAT(0) spaces. In the case of $\mathrm{Mod}(\Sigma)$ acting on $\mathcal{T}(\Sigma)$, we rely on the following fact: for a.e. $\omega = (\omega_n)$, the forward limit η and the backward limit ξ are distinct points in $\mathcal{UE}$, and they are connected by a unique Teichmüller geodesic. Given this, Tiozzo applies the subadditive ergodic theorem to deduce the conclusion.

Let us now discuss Rafi's analysis on thin triangles of $\mathcal{T}(\Sigma)$ in [80]. Masur–Minsky's theory motivated Rafi to investigate Teichmüller geodesics with subsurface projection. A Teichmüller geodesic γ in $\mathcal{T}(\Sigma)$ for some surface Σ can be cut into distinct subsegments γ_α, each behaving like a Teichmüller geodesic on some subsurface Y_α isolated during that time. Using such a decomposition, Rafi deduced the following two instances of hyperbolicity in $\mathcal{T}(\Sigma)$.

The first item is fellow traveling. Consider two geodesics $\gamma \colon [a, b] \to X$, $\eta \colon [a', b'] \to X$ with $d(\gamma(a), \eta(a')) < C$, $d(\gamma(b), \eta(b')) < C$. If X were δ-hyperbolic, then γ and η $K(C, \delta)$-fellow travel. We also expect $K(C, \varepsilon)$-fellow traveling between such geodesics inside the ε-thick part of $\mathcal{T}(\Sigma)$. However, there exists no *a priori* uniform bound K for every pair of geodesics in $\mathcal{T}(\Sigma)$ having pairwise near endpoints. Rafi's theorem ([80, Theorem 7.1]) asserts that the geodesics $K(C, \varepsilon)$-fellow travel if the pairwise near endpoints are ε-thick, even if the geodesics are not entirely ε-thick and visit the ε-thin part.

The second item is as follows. Consider a geodesic triangle $\triangle xyz$ in X and $p \in [yz]$. If X is δ-hyperbolic, then p is within distance $K(\delta)$ from either $[x, y]$ or $[x, z]$. This is not guaranteed in $\mathcal{T}(\Sigma)$ in general, but there instead exist $K_1(\varepsilon)$, $K_2(\varepsilon)$ satisfying the following. If $p \subseteq I \subseteq [y, z]$ for some ε-thick subsegment I that is longer than $K_1(\varepsilon)$, then the distance between p and $[x, y] \cup [x, z]$ is at most $K_2(\varepsilon)$.

As we will see in the next section, these results are useful to compare the concatenation of geodesic segments $[x_0, x_1], [x_1, x_2], \ldots, [x_{N-1}, x_N]$ with the direct one $[x_0, x_N]$. In detail, suppose for each i that

(1) $[x_i, x_{i+1}]$ begins and ends in thick directions, and

(2) the beginning direction of $[x_i, x_{i+1}]$ and the ending direction of $[x_{i-1}, x_i]$ are almost aligned.

Then we conclude that each $[x_i, x_{i+1}]$ fellow travels with a subsegment of $[x_0, x_N]$. This fact is utilized by Baik–Choi–Kim's pivoting, which we explain later.

On the other hand, Rafi's approach involving subsurface projections was continued by Horbez, Dahmani–Horbez and Mathieu–Sisto.

Horbez's approach in [51] and Dahmani–Horbez's approach in [22] begin with transferring sample paths on $\mathcal{T}(\Sigma)$ to $\mathcal{C}(\Sigma)$ via the shortest curve projection $\pi \colon \mathcal{T}(\Sigma) \to \mathcal{C}(\Sigma)$. Here, the preimage of each point $p \in \mathcal{C}$ by π is of infinite diameter. Nevertheless, for a Teichmüller geodesic γ that is long in terms of both the Teichmüller metric and the curve complex metric, the (rough) preimage of $\pi(\gamma)$ may have stricter restriction. The following observation is motivated by the work of Dowdall–Duchin–Masur improving Rafi's thin triangle result [25, Theorem A].

Proposition 5.1 ([22, Proposition 3.7]). *For all $\kappa > 0$, there exist $B, D > 0$ such that the following holds. Let $[x, y]$ be a Teichmüller geodesic that contains a subsegment γ with sufficient progress on $\mathcal{C}(\Sigma)$, that means, $\mathrm{diam}_{\mathcal{C}(\Sigma)}(\pi(\gamma)) > B$. If $z \in \mathcal{T}(\Sigma)$ satisfies that $\pi([x, z])$ and $\pi(\gamma)$ κ-fellow travel up to a reparametrization, then there exists a subsegment $\eta \subseteq [x, z]$ such that the Hausdorff distance of γ and η in $\mathcal{T}(\Sigma)$ is at most D and $\mathrm{diam}_{\mathcal{C}(\Sigma)}(\pi(\eta)) \geq \mathrm{diam}_{\mathcal{C}(\Sigma)}(\pi(\gamma)) - B$.*

In other words, the fellow-traveling among projections of long enough Teichmüller geodesics onto the curve complex can be lifted to the Teichmüller space. Recall also the result of Masur and Minsky that π is coarsely $\mathrm{Mod}(\Sigma)$-equivariant, is coarsely Lipschitz, and sends Teichmüller geodesics to $K(\Sigma)$-quasi-geodesics. In this framework, we now explain how Dahmani and Horbez lifted the behavior of random walks on $\mathcal{C}(\Sigma)$ to $\mathcal{T}(\Sigma)$.

First, Masur–Minsky's theory provides a constant K_0 such that the shortest curve projections of Teichmüller geodesics are unparametrized K_0-quasi-geodesics on $\mathcal{C}(\Sigma)$ ([72, Theorems 2.3, 2.6]). Since $\mathcal{C}(\Sigma)$ is Gromov hyperbolic, we have the following Morse lemma for some $K_1 > 0$: any pair of K_0-quasi-geodesics that share the beginning and the ending points K_1-fellow travel up to a reparametrization. Moreover, there exist $K_2 > 0$ such that for a pair of K_0-quasi-geodesics K_2-fellow travel up to a reparametrization if their beginning points and ending points are $(K_1 + 2\delta)$-close, respectively.

Let us temporarily fix a reference point o' in $\mathcal{T}(\Sigma)$, and let $o = \pi(o')$. We recall a result of Maher–Tiozzo: given a finitely supported measure μ on $\mathrm{Mod}(\Sigma)$, almost every sample path $\omega = (\omega_n)$ of the random walk satisfies that

$$\lim_{n \to \infty} \frac{\tau_{\mathcal{C}(\Sigma)}(\omega_n)}{n} = \lambda,$$

where λ is the escape rate of the random walk in $\mathcal{C}(\Sigma)$.

See Figure 1. Let us consider geodesics $[o, \omega_n o], [\omega_n o, \omega_n^2 o], \ldots, [\omega_n^{k-1} o, \omega_n^k o]$ in $\mathcal{C}(\Sigma)$. Suppose, say, that $1000(\delta + B(10K_2)) \leq d_{\mathcal{C}(\Sigma)}(o, \omega_n o) \leq 2\lambda n$ and

$$d_{\mathcal{C}(\Sigma)}(o, \omega_n o) - \tau_{\mathcal{C}(\Sigma)}(\omega_n) \leq 0.01 d_{\mathcal{C}(\Sigma)}(o, \omega_n o)$$

(this will happen eventually in a.e. path ω). Then $[\omega_n^l o, \omega_n^{l-1} o]$ and $[\omega_n^l o, \omega_n^{l+1} o]$ should deviate early, at distance within $0.005 d_{\mathcal{C}(\Sigma)}(o, \omega_n o)$. By δ-hyperbolicity of $\mathcal{C}(\Sigma)$, there
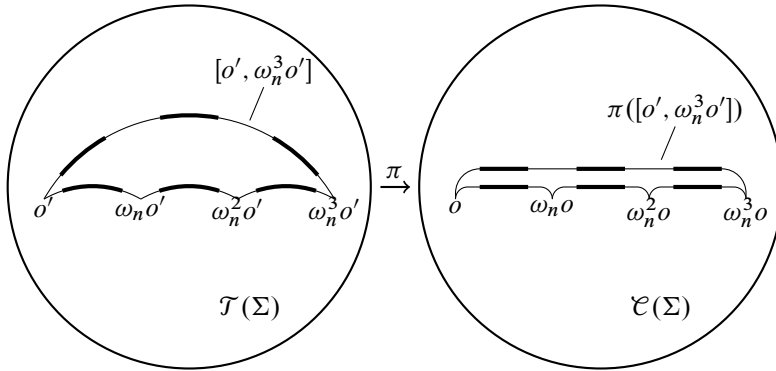
Figure 1. Schematics for the proof of Theorem 5.2. Curves on $\mathcal{T}(\Sigma)$ are Teichmüller geodesics, while their projections on $\mathcal{C}(\Sigma)$ are K_0-quasi-geodesics. The corresponding thick segments on $[o', \omega_n^3 o']$ and $[\omega_n^{l-1} o', \omega_n^l o']$ projects to sufficiently long, fellow-traveling thick segments on $\mathcal{C}(\Sigma)$. This forces the original segments to also fellow-travel.

exist disjoint subsegments $[x_l, y_l]$ of $[o, \omega_n^k o]$ that 2δ-fellow travel with the middle 99% of $[\omega_n^{l-1} o, \omega_n^l o]$. (∗)

We now lift the situation to $\mathcal{T}(\Sigma)$ using the ingredients below:

(1) First, curve complex geodesics $[\omega_n^{l-1} o, \omega_n^l o]$ are K_0-close enough to the projections $\pi([\omega_n^{l-1} o', \omega_n^l o'])$ of the Teichmüller geodesics.

(2) Similarly, $[o, \omega_n^k o]$ and $\pi([o', \omega_n^k o'])$ are K_0-close.

(3) (1), (2) and (∗) imply that $\pi([o', \omega_n^k o'])$ crosses the middle 98% of each projection $\pi([\omega_n^{l-1} o', \omega_n^l o'])$ up to distance K_2.

We now apply Proposition 5.1 twice to obtain subsegments $(\eta_l)_{l=1}^{k-1}$ of $[o', \omega_n^k o']$ that satisfy the following. Let γ_0 be a subsegment of $[o', \omega_n o']$ that projects onto the middle 96% of $\pi([o', \omega_n o'])$. Then η_l and $\omega_n^l \gamma_0$ are within Hausdorff distance $D(K)$ on $\mathcal{T}(\Sigma)$. Therefore, we have $d_{\mathcal{T}(\Sigma)}(\omega_n^l o', [o', \omega_n^k o']) \leq d(o', \gamma_0)$ for each l and

$$d_{\mathcal{T}(\Sigma)}(o, \omega_n o) - \tau_{\mathcal{T}(\Sigma)}(\omega_n) = d_{\mathcal{T}(\Sigma)}(o, \omega_n o) - \lim_k \frac{1}{k} d(o, \omega_n^k o)$$

$$\leq 2 d_{\mathcal{T}(\Sigma)}(o', \gamma_0) + D(K).$$

It now suffices to control the final term, the *Teichmüller* length of the left 2% portion of $[o', \omega_n o']$ with respect to the *curve complex* distance. Although two distances are not comparable in general, the linear escape and sublinear tracking of a.e. sample path on both $\mathcal{C}(\Sigma)$ enable this. Using this type of argument, Dahmani and Horbez obtain the following theorem.

Theorem 5.2 ([22, Theorem 0.2]). *Suppose that μ is finitely supported. Then for a.e. sample path (ω_n), we have*

$$\lim_{n \to \infty} \sqrt[n]{\lambda(\omega_n)} = \lambda,$$

where $\lambda(\omega_n)$ is the stretch factor of ω_n and $\log \lambda$ is the escape rate of the random walk.

The finite support assumption in Theorem 5.2 originates from the spectral theorem for $\mathcal{C}(\Sigma)$. As Dahmani and Horbez explain, the arguments of Benoist–Quint and Maher–Tiozzo give rise to a spectral theorem for $\mathcal{C}(\Sigma)$ with finite second moment assumption. Given this, the rest of the Dahmani–Horbez argument relies on the sublinear tracking and the subadditive ergodic theorem that only requires finite first moment.

Another way to relate the actions of $\mathrm{Mod}(\Sigma)$ on $\mathcal{C}(\Sigma)$ and $\mathcal{T}(\Sigma)$ was suggested by Mathieu and Sisto [75]. They first argue that non-elementary random walks on acylindrically hyperbolic groups are almost additive. This almost additivity reduces the limit laws on such random walks to those of commutative random walks on $\mathbb{R}$. In the course of the argument, they establish deviation inequalities, logarithmic geodesic tracking (see also [84]), and many more ingredients that we will observe soon.

Meanwhile, although it is true that $\mathrm{Mod}(\Sigma)$ is acting on $\mathcal{C}(\Sigma)$ and $\mathcal{T}(\Sigma)$ acylindrically, $\mathcal{T}(\Sigma)$ is not Gromov hyperbolic. Hence, one needs to bring the results on $\mathcal{C}(\Sigma)$ to $\mathcal{T}(\Sigma)$, which motivated Mathieu and Sisto to show the existence of $o \in \mathcal{C}(\Sigma)$ and $L \geq 0$ that satisfy the following. For $l_1, l_2, t \geq 0$ and $g, h \in \mathrm{Mod}(\Sigma)$ such that $d_{\mathcal{C}(\Sigma)}(go, ho) \geq L + l_1 + l_2$, we have

$$\mathrm{diam}^{\mathcal{T}(\Sigma)}\big[\pi^{-1}(B^{\mathcal{C}(\Sigma)}(go, l_1)) \cap N_t^{\mathcal{T}(\Sigma)}\big(\pi^{-1}(B^{\mathcal{C}(\Sigma)}(ho, l_2))\big)\big] \leq Lt,$$

where π denotes the shortest curve projection and diam^X, N_t^X, B^X refer to the diameter, neighborhood and the ball with respect to d_X, respectively. This property follows from the coarse distance formula of the Teichmüller metric in terms of (truncated) curve complex distances on subsurfaces [79], and the bounded geodesic image theorem on curve complexes of subsurfaces with uniform constant (see [72, 89]). Note that this property promotes the bounded distance of $\pi(p)$ from a long enough quasi-geodesic $\pi(\gamma)$ to the bounded distance of p from γ.

Both Dahmani–Horbez's and Mathieu–Sisto's approaches are concerned with geometric properties of π that are not expected for arbitrary pairs of points or geodesics on $\mathcal{T}(\Sigma)$ but expected in almost every sample path. One partial reason, although not complete, is that $\mathrm{Mod}(\Sigma)$ acts on $\mathcal{T}(\Sigma)$ as isometries that translates each ε-thick reference point to another ε-thick point, rather than arbitrary points. This implies that the randomness from random walks and other types of randomness in $\mathcal{T}(\Sigma)$ may show different behavior.

In this spirit, Gadre, Maher and Tiozzo captured the contrast between the harmonic measure on $\mathcal{PMF}(\Sigma)$ arising from random walks and the Lebesgue measure [37]. A similar contrast holds between the Lebesgue measure on $\partial\mathbb{H}^2$ and the harmonic measure from a random walk on a cusped Fuchsian group. Gadre–Maher–Tiozzo considered the following quantity: for a boundary point $p \in \mathcal{PMF}(\Sigma)$, we first take a geodesic γ tending to p,

and approximate thick points $\gamma(t)$ with mapping class group orbits $h_t o$ of the reference point o. Then we compare the word norm of h_t on $\mathrm{Mod}(\Sigma)$ and the displacement of h_t with respect to the curve complex metric. In terms of the Lebesgue measure, the ratio between $d_{\mathrm{Mod}(\Sigma)}(1, h_t)$ and $d_{\mathcal{C}(\Sigma)}(o, h_t o)$ tends to infinity in almost every choice of p; on the other hand, in terms of the harmonic measure for μ with finite first moment in the word metric on $\mathrm{Mod}(\Sigma)$, $d_{\mathrm{Mod}(\Sigma)}(1, h_t)$ and $d_{\mathcal{C}(\Sigma)}(o, h_t o)$ are comparable and their ratio converges to a uniform constant in almost every choice of p.

Let us finish this section by explaining a consequence of Rafi's theorems that will be used later on. Consider a geodesic triangle $\triangle xyz$ in $\mathcal{T}(\Sigma)$. A priori, $\triangle xyz$ is not δ-thin and $[yz]$ need not be contained in a bounded neighborhood of $[xy] \cup [xz]$. However, suppose that $[x, y]$ initially fellow travels with a thick segment $[x, y']$. This forces that $[x, y]$ is initially thick also, and Rafi's theorem asserts that this beginning portion should be contained in a bounded neighborhood of $[x, y] \cup [y, z]$.

Let us similarly suppose that $[x, z]$ initially fellow travels with a thick segment $[x, z']$, and $[x, y']$ and $[x, z']$ are heading in different directions, i.e., $(y', z')_x$ is bounded. Then the initial segment of $[x, y]$ cannot be contained in the neighborhood of $[x, z]$ and vice versa. Finally, if we further suppose that points y, z are also thick, then Rafi's fellow traveling theorem implies that $\triangle xyz$ is an obtuse thin triangle: $[y, z]$ and $[x, y] \cup [x, z]$ are within bounded Hausdorff distance.

6. Limit theorems I: Displacement

In the Euclidean setting, stronger moment assumptions lead to finer description of random walks. For example, strong laws of large numbers (SLLN) are linked with the finitude of first moment; central limit theorems (CLT) and laws of the iterated logarithm (LIL) are relevant to the finitude of second moment; when the random walk has finite exponential moment, large deviation principles (LDP) is also available. Many recent work in this topic tried to bring these results to hyperbolic settings under suitable moment conditions. Among them we explain the results of Benoist–Quint, Horbez, Mathieu–Sisto, Boulanger–Mathieu–Sert–Sisto, Gouëzel, Baik–Choi–Kim and Choi.

In the setting of general metric spaces, two meaningful quantities arise from random walks $\omega = (\omega_n)$ on the isometry group: the displacement $d(o, \omega_n o)$ of a reference point $o \in X$ and the translation length $\tau(\omega_n)$. The first one is subadditive while the latter one is not; this complicates the investigation of translation length. We will first discuss the theorems for displacement and then move on to the case of translation length.

The theorem in hyperbolic settings that corresponds to laws of large number is the subadditive ergodic theorem. For completeness, we spell out the following statement.

Theorem 6.1 (cf. [53, Theorem 5.5], [67, Theorem 1.3]). *Let $X = \mathcal{T}(\Sigma)$ or $\mathcal{C}(\Sigma)$ and suppose that μ is a non-elementary probability measure on G has finite first moment. Then*

there exists $\lambda > 0$, called the escape rate of μ, such that the random variables $\frac{1}{n} d_X(o, \omega_n o)$ converge to λ in L^1 and almost surely.

Here, the strict positivity of the escape rate is due to the non-amenability of $\mathrm{Mod}(\Sigma)$ in the case of $X = \mathcal{T}(\Sigma)$, whereas it follows from the existence of 'persistent joints' from Maher–Tiozzo's argument in the case of $X = \mathcal{C}(\Sigma)$ (which ultimately relies on the fact that the harmonic measure for μ on $\partial\mathcal{C}(\Sigma)$ is atom-free). We will revisit the notion of persistent joints in Section 7.

We remark that this is not a consequence of the Borel–Cantelli argument. Indeed, for example, the exponential decay of $\mathbb{P}(\frac{1}{n} d(o, \omega_n o) \geq \lambda + \varepsilon)$ for $\varepsilon > 0$ implies that μ has finite exponential moment. In contrast, $\mathbb{P}(\frac{1}{n} d(o, \omega_n o) \leq \lambda - \varepsilon)$ does decay exponentially even without any moment condition due to the recent work of [44]. We note that Gouëzel's technique is powerful enough to deduce other results including the continuity of the escape rate.

The next natural goal is CLTs of the following form.

Theorem 6.2 ([7, Theorem 1.1], [51, Theorem 0.1]). *Let $X = \mathcal{T}(\Sigma)$ or $\mathcal{C}(\Sigma)$. Suppose that μ is a non-elementary, non-arithmetic probability measure on G with finite second moment. Then there exists $\sigma > 0$ such that $\frac{1}{\sqrt{n}}[d_X(o, \omega_n o) - \lambda n]$ converges to the Gaussian law $\mathcal{N}(0, \sigma)$ in law, where $\lambda > 0$ is the escape rate of the random walk. Explicitly, for any $a < b$, we have*

$$\lim_{n \to \infty} \mathbb{P}\left[a\sqrt{n} \leq d_X(o, \omega_n o) - \lambda n \leq b\sqrt{n} \right] = \int_a^b \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} \, dx.$$

This direction dates back to Sawyer–Steger's investigation [83] on the random walks on free groups, which was also discussed by Ledrappier [62]. Its generalization to Gromov hyperbolic groups under the finite exponential moment assumption is due to Björklund (see [8]). The current CLT under the finite second moment assumption was proven by Benoist and Quint in [7], using the machinery of their previous work on linear groups. Finally, by the lifting principle that we explained before, Horbez generalized this CLT to Teichmüller spaces [51].

Benoist–Quint's theory applies to non-elementary groups G acting on a proper, quasi-convex, Gromov hyperbolic space X and non-elementary, non-arithmetic Borel measures on G. Note that G and μ need not be discrete here. The properness of X is assumed for constructing the Gromov compactification $X \cup \partial X$ and the Busemann compactification $X \cup \partial_B X$. The main strategy is to find an alternative for the random variable $d(o, \omega_n o)$, which can be expressed as martingales with step differences controlled in L^2 and in probability. The first trick is to use Busemann functions

$$\sigma(g, x) := \lim_{n \to \infty} [d(g^{-1}o, x_n) - d(o, x_n)]$$

for $x \in \partial_B X$ and $x_n \to x$ instead of the displacement. In contrast with displacement, Busemann functions satisfy the cocycle condition

$$\sigma(gg', x) = \sigma(g, g'x) + \sigma(g', x). \tag{1}$$

At the cost of this advantage, however, one should pay attention to several details. First, the choice of the boundary point x causes some asymmetry. Moreover, the decomposition of $\sigma(\omega_n, x)$ into $\sigma(g_{k+1}, \omega_k x)$ as in equation (1) involves different boundary points $\omega_k x$; hence, the argument requires analysis on the boundary action of G and the stationary measure on $\partial_B X$ or ∂X. Second, $\sigma(\omega_n, x)$ is nonetheless different from $d(o, \omega_n o)$ and the discrepancy

$$d(o, \omega_n o) - \sigma(\omega_n, x)$$

should be controlled for large enough n in probability. Finally, the quantities $\sigma(g_{k+1}, \omega_k x)$ still may not be adequate for martingale CLTs, as they are not 'centered' at the right value, namely, the escape rate λ. One should therefore solve the cohomological equation and center $\sigma(g_{k+1}, \omega_k x)$ by subtracting bounded random variables. After all these preliminary steps, we can control the step differences in L^2 and in probability using the finitude of second moment of μ and conclude. Note again that the spirit of this proof is 'from the infinity', rather than 'working inside the space'.

As we have seen before, the lifting argument often promotes phenomena in $\mathcal{C}(\Sigma)$ to the corresponding ones in $\mathcal{T}(\Sigma)$. Horbez's strategy in [51] was to lift the ingredients for Benoist–Quint's CLT, including the centerability of Busemann functions and the summable decay of shadows in particular directions, from $\mathcal{C}(\Sigma)$ to $\mathcal{T}(\Sigma)$. Although $\mathcal{C}(\Sigma)$ is not proper and thus Benoist–Quint's proof does not apply as is, these ingredients are available on non-proper spaces by Maher–Tiozzo's work. Once the ingredients are lifted using Proposition 5.1, Benoist–Quint's argument applies to Busemann functions on $\mathcal{T}(\Sigma)$ and the desired CLT follows.

Another approach to the CLT for displacement was proposed by Mathieu and Sisto in [75]. In fact, they provide a much more general framework, requiring the control on the defects of the form

$$Q_{n+m}(\omega) - Q_n(\omega) - Q_m(T^n \omega)$$

(the Gromov products $(\omega_{n+m} o, o)_{\omega_n o}$ in our setting, for example) and yielding quantitative estimates on the lack of additivity of such sequences.

To see the principle behind this, let G be acting on any metric space X and let us assume that $\mathbb{E}[(\omega_n o, \check{\omega}_n o)_o^2]$ is bounded by some constant B for all n. We now estimate the distances among $o, \omega_n o, \omega_{2n} o, \ldots, \omega_{2^k n} o$. If the points were always perfectly aligned, then $\mathbb{E}[d(o, \omega_{2^k n} o)]$ and $\mathrm{Var}[d(o, \omega_{2^k n} o)]$ would grow linearly with respect to 2^k (note that the family $\{\omega_{in}^{-1} \omega_{(i+1)n}\}_i$ consists of independent RVs). We would also have

$$d(o, \omega_{2^k n} o) = \sum_{i=1}^{2^k} d(\omega_{i-1} o, \omega_i o)$$

and the classical CLT would imply that

$$\frac{1}{\sqrt{2^k n}}\big[d(o, \omega_{2^k n}o) - \mathbb{E}[d(o, \omega_{2^k n}o)]\big]$$

converges in law to a Gaussian law $\mathcal{N}(0, \sigma_n)$, where $\sigma_n = \sqrt{\mathrm{Var}[d(o, \omega_n o)]/n}$.

However, the addition is not exact and the deficit is recorded in the form

$$2(\omega_{in}o, \check{\omega}_{kn}o)_{\omega_{jn}o}.$$

Note that in order to make $d(o, \omega_{2^k n}o)$ out of $d(o, \omega_n o), \dots, d(\omega_{(2^k-1)n}o, \omega_{2^k n}o)$, we need to add up $2^k - 1$ deficits

$$2(\omega_{2^t(2j-2)}o, \omega_{2^t(2j-1)}o)_{\omega_{2^{t+1}j}o} \quad (t = 0, \dots, k-1, \ j = 1, \dots, 2^{k-t-1} - 1).$$

For a fixed t, these form a family of $(2^{k-t-1} - 1)$ i.i.d. with variance less than B. Summing them up and dividing by $1/\sqrt{2^k n}$, the error from these terms is bounded by $7/\sqrt[3]{n}$ outside an event of probability at most $8B/\sqrt[3]{n}$. By taking dyadic $n = 2^m$, we deduce that

$$\frac{1}{\sqrt{2^m}}\big[d(o, \omega_{2^m}o) - \mathbb{E}[d(o, \omega_{2^m}o)]\big]$$

is Cauchy and $\sigma_m \to \sigma > 0$ (here is required at least linear growth of $Var[d(o, \omega_{2^m}o)]$, which is deduced from the non-arithmeticity of μ). A similar argument can handle non-dyadic steps also, if $\mathbb{E}[(\omega_n o, \check{\omega}_m o)_0^2]$ is uniformly controlled for arbitrary n, m.

It remains to control $\mathbb{E}[(\omega_n o, \check{\omega}_n o)_o^2]$ as promised, for which Mathieu–Sisto's argument requires two assumptions: (1) that μ has finite exponential moment, and (2) that the action is acylindrical. Since the action of $\mathrm{Mod}(\Sigma)$ on $\mathcal{C}(\Sigma)$ is acylindrical, Mathieu–Sisto's theory applies to $\mathcal{C}(\Sigma)$ and deduces the CLT. Moreover, $\mathcal{T}(\Sigma)$ also fits into this scheme since it is acylindrically intermediate for $(\mathrm{Mod}(\Sigma), \mathcal{C}(\Sigma))$.

Let us explain how Mathieu–Sisto's viewpoint of 'almost exact addition' was expanded in the works of Boulanger–Mathieu–Sert–Sisto, Gouëzel and Choi. All these works use the following same modification of random walks. Given a measure μ on G and $S \subseteq G$ such that $\alpha := \min\{\mu(g) : g \in S\} > 0$, there exists a measure η such that $\mu = \alpha\mu_S + (1-\alpha)\eta$, where μ_S is the uniform measure on S. We then consider:

- Bernoulli RVs ρ_i (with $\mathbb{P}(\rho_i = 1) = \alpha$ and $\mathbb{P}(\rho_i = 0) = 1 - \alpha$),

- η_i with the law η, and

- ν_i with the law μ_S,

all independent, and define

$$\gamma_i = \begin{cases} \eta_i & \text{when } \rho_i = 1, \\ \nu_i & \text{when } \rho_i = 0. \end{cases}$$

Then γ_i are i.i.d. of the law μ that models the random walk on G generated by μ. Let us also define $\mathcal{N}(k) := \sum_{i=1}^{k} \rho_i$ and $\vartheta(i) := \min\{j \geq 0 : B(j) = i\}$ for convenience. In

this model, a random trajectory consists of relatively usual steps $(\gamma_{\vartheta(i)+1}, \ldots, \gamma_{\vartheta(i+1)-1})$ and special steps $\gamma_{\vartheta(i)}$ in an alternating way, the first one being chosen with law ν and the second one being chosen with law μ_S.

Morally, ν is designed to behave almost like μ: they share the same moment condition and similar moment values. The displacements made by these usual steps are then linked with the special steps from S. The desired property of special steps is that they can align consecutive displacements with high probability. This is encoded in the notion of a *Schottky set*, which stems from the classical Schottky decomposition.

Definition 6.3. Let $K, K', \varepsilon > 0$. A finite set S of isometries of X is said to be (K, K')-*Schottky* if the following hold:

(1) for all $x, y \in X$, $|\{s \in S : (x, s^i y)_o \geq K \text{ for some } i > 0\}| \leq 2$;

(2) for all $x, y \in X$, $|\{s \in S : (x, s^i y)_o \geq K \text{ for some } i < 0\}| \leq 2$;

(3) for all $s \in S$ and $i \neq 0$, $d(o, s^i o) \geq K'$.

When X is Teichmüller space, S is said to be (K, K', ε)-*Schottky* if the following condition holds in addition to the above three:

(4) for all $s \in S$ and $i \in \mathbb{Z}$, the geodesic $[o, s^i o]$ is ε-thick.

By employing Schottky sets for 'linking steps', one can add up step distances almost exactly. In particular, Boulanger–Mathieu–Sert–Sisto recovered the following deviation inequality of Mathieu–Sisto without the acylindricality assumption: if μ has finite exponential moment, there exists $K, \kappa > 0$ such that

$$\mathbb{P}[(o, \omega_n o)_{\omega_i o} \leq R] \leq K e^{-\kappa R}$$

holds for any $0 \leq i \leq n$. This result is subsequently used to establish the following large deviation principle for the random walk.

Theorem 6.4 ([11, Theorem 1.1]). *If μ has finite exponential moment, then there exists a proper convex function $I : \mathbb{R} \to [0, \infty]$ (called the* rate function*) that satisfies*

$$- \inf_{\alpha \in \text{int}(R)} I(\alpha) \leq \liminf_n \frac{1}{n} \ln \mathbb{P}\left[\frac{1}{n} d(o, \omega_n o) \in R\right]$$

$$\leq \limsup_n \frac{1}{n} \ln \mathbb{P}\left[\frac{1}{n} d(o, \omega_n o) \in R\right] \leq - \inf_{\alpha \in \bar{R}} I(\alpha)$$

for any measurable subset R of $\mathbb{R}$. Moreover, I vanishes only at the escape rate λ of the random walk.

Roughly speaking, the probability that $\frac{1}{n} d(o, \omega_n o)$ deviates from λ decays exponentially, the speed of which is precisely encoded by I. We note that Boulanger, Mathieu, Sert and Sisto establish the rate function from above (for values greater than λ) on arbitrary metric spaces. Indeed, the existence of the rate function from above is essentially equivalent to the finitude of exponential moment, rather than the geometric properties of

the underlying space, as mentioned before. Meanwhile, establishing the rate function from below requires the existence of Schottky sets and the Gromov inequalities among points.

It was unexpected, however, that the exponential decay of the deviation from below does not require any moment condition.

Theorem 6.5 ([44, Theorems 1.1, 1.2]). *Let* $X = \mathcal{T}(\Sigma)$ *or* $\mathcal{C}(\Sigma)$ *and* μ *be a non-elementary probability measure on* G.

(1) *If* μ *has finite first moment and* λ *is its escape rate, then*

$$\mathbb{P}[d(o, \omega_n o) \leq (\lambda - \varepsilon)n]$$

decays exponentially for any $\varepsilon > 0$.

(2) *If* μ *has infinite first moment, then there is no finite 'escape rate':*

$$\mathbb{P}[d(o, \omega_n o) \leq rn]$$

decays exponentially for any $r > 0$.

To establish this result, Gouëzel first takes suitable integer N and a Schottky set S so that we have a decomposition

$$\mu^{*N} = \alpha \mu_S^{*2} + (1 - \alpha)\nu. \tag{2}$$

Here the N-th convolution of μ is designed to guarantee sufficiently large size of S; the purpose of the self-convolution of the Schottky measure will become apparent soon. Then the composition γ_i of the Bernoulli variable ρ_i and η_i, ν_i models the convolution of steps $g_{N(i-1)+1}, \ldots, g_{Ni}$. We name some of the special steps $\vartheta(i)$ as *pivotal times*, which are meant to be the crucial moments throughout the history of the random path.

At each pivotal step, we hope that two Schottky segments are directed away from each other and the former (latter, resp.) Schottky segment does not cancel out the previous (upcoming, resp.) progress. To be concrete, suppose that we have chosen $\{m_1 < \cdots < m_k\}$ from the special steps $\{\vartheta(i)\}_{i=1}^{M-1}$ as pivotal times for the path $(g_1, \ldots, g_{\vartheta(M)})$. Let

$$w_i := \omega_{N(m_i+1)}^{-1}\omega_{Nm_{i+1}}, \quad s_i := \omega_{Nm_i}^{-1}\omega_{N(m_i+0.5)}, \quad s_i' := \omega_{N(m_i+0.5)}^{-1}\omega_{N(m_i+1)},$$

and $w_0 := \omega_{Nm_1}$, $s_0 = \text{id}$. See Figure 2. The desired situation that $m_1, \ldots, m_k$ is assumed to satisfy is the following for suitable K:

- $(s_i^{-1}o, s_i'o)_o \leq K$ for $i = 1, \ldots, k$,
- $(s_i'^{-1}o, w_io)_o \leq K$ for $i = 1, \ldots, k$,
- $(w_{i-1}^{-1}s_{i-1}'^{-1}o, s_io)_o \leq K$ for $i = 1, \ldots, k$.

From these conditions, we can deduce small Gromov products among consecutive pivotal loci: $(w_{i-1}o, w_{i+1}o)_{w_io}$ for each i. Note that this alone does not guarantee small Gromov products $(w_io, w_ko)_{w_jo}$ for arbitrary $i < j < k$ that we need to sum up intermediate progresses: recall the theory of Mathieu–Sisto. This is remedied by the fact that small
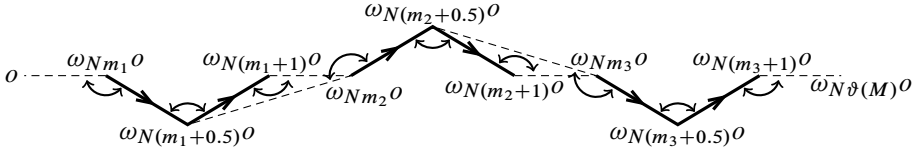
Figure 2. A preliminary definition of pivotal loci.

Gromov products are actually guided by long enough Schottky segments. Hence, each $d(w_{i-1}o, w_{i+1}o)$ is large enough and the Gromov inequality can deduce small arbitrary Gromov products. In $\mathcal{T}(\Sigma)$, we rely on the consequence of Rafi's theorems (cf. Section 5) and similarly deduce small arbitrary Gromov products.

The ideal situation is that all special steps can work as pivotal steps. If it is the case, the intermediate progresses are added up and the overall progress becomes large enough. Unfortunately, there is always a small chance of the undesirable event: the probability that all special steps are pivotal times decays exponentially. Nonetheless, we want to tolerate some error and select a sufficiently large number of special steps that still work as pivotal steps.

To further illustrate this idea, given pivotal times $\{m_1, \ldots, m_k\}$ for $(g_1, \ldots, g_{\vartheta(M)})$, let us determine the pivotal times for the path $(g_1, \ldots, g_{\vartheta(M+1)})$. One possible strategy is just adding $m_{k+1} = \vartheta(M)$ to the original set of pivotal times. Recall again the conditions for $m_1, \ldots, m_k$:

- $(s_i^{-1}o, s_i'o)_o \leq K$ for $i = 1, \ldots, k$,
- $(s_i'^{-1}o, w_io)_o \leq K$ for $i = 1, \ldots, k$,
- $(w_{i-1}^{-1}s_{i-1}'^{-1}o, s_io)_o \leq K$ for $i = 1, \ldots, k$.

Fixing w_i and s_i', there are many other choices for each s_i that satisfies the condition. In particular, the property of Schottky sets implies that at least $(\#S - 2)$ choices out of all choices are valid at each step. This process, fixing w_i and s_i' and modifying the choice of s_i into another valid choice, is called *pivoting*.

If, for example, the additional s_{k+1}, s_{k+1}' and w_{k+1} satisfy the above condition, then we can add it to the set of pivotal times. This already takes up large enough probability, at least $((\#S - 2)/\#S)^2$. In case of failure, however, we do not wish to give up the entire selection $\{m_1, \ldots, m_k\}$ but rather retain a portion that works for $(g_1, \ldots, g_{\vartheta(M+1)})$. For example, can we hope that the set $\{m_1, \ldots, m_{k-1}\}$ itself works for intermediate words $w_0, w_1, \ldots, w_{k-1}s_ks_k'w_ks_{k+1}s_{k+1}'w_{k+1}$? A priori, the final word depends on $s_{k+1}s_{k+1}'$ so this cannot be answered without altering s_{k-1}': this is not what we want. We can however require the following conditions:

- $(s_{k-1}'^{-1}o, w_ko)_o \leq K$,
- $(w_{k-1}^{-1}s_{k-1}'^{-1}o, s_ko)_o \leq K$,
- $(s_k^{-1}o, s_k'w_ks_{k+1}s_{k+1}'w_{k+1}o)_o \leq K$.

The first condition is already achieved by the fixed s'_{k-1} from the assumption, and the latter two conditions can be achieved for any fixed $s'_{k-1}, w_{k-1}, s_{k+1}, s'_{k+1}, w_{k+1}$ by picking valid s_k only. This has high chance, so we have

$$\mathbb{P}\big[\{m_1,\ldots,m_{k-1}\} \text{ works for } (g_1,\ldots,g_{\vartheta(M+1)}) \mid \{w_i,w'_i,s_i\}, \{s_i\}_{i\neq k}, s_k : \text{valid}\big]$$
$$\geq \frac{\#S-3}{\#S}.$$

Note that the estimation is conditioned on each equivalence class of choices that are pivoted at the k-th slot from each other. Summing them over bad choices of s_{k+1}, s'_{k+1}, we have

$$\mathbb{P}\big[\{m_1,\ldots,m_{k-1}\}, \{m_1,\ldots,m_k\}, \{m_1,\ldots,m_{k+1}\} \text{ does not work}\big]$$
$$\leq \left[1 - \left(\frac{\#S-2}{\#S}\right)^2\right] \cdot \frac{3}{\#S}.$$

Inductively, we deduce that the first $k-i$ slots (and possibly some more) among $\{m_1,\ldots,m_{k+1}\}$ can be employed as pivotal times for $(g_1,\ldots,g_{\vartheta(M+1)})$ except an exponentially decaying probability, whose decay rate depends on the size of S. In summary, we can guarantee almost definite increase of the proportion of pivots, as close to 1 as needed, by taking a large enough Schottky set. For the precise definition that includes the modified conditions, see [44] or [16].

Using small Gromov products among pivotal loci, one can show that $d(o,\omega_n o)$ is bounded from below by a multiple of $\#\{\text{pivots for } (g_1,\ldots,g_n)\}$. Hence, we have established the definite progress of random walks outside an event of exponentially decaying probability. In order to push this progress as close to the escape rate as we want, one should modify the decomposition (2) and sandwich an auxiliary variable between Schottky steps. We refer the readers to [44] for details.

7. Limit theorems II: Translation length

We now discuss the theory of translation length. In contrast with the case of displacement, where SLLN with the optimal moment condition was obtained at once, the first result for translation length was the following weak law of large numbers (WLLN).

Theorem 7.1. *Let $X = \mathcal{T}(\Sigma)$ or $\mathcal{C}(\Sigma)$. Then there exists $L > 0$ such that*

$$\lim_n \mathbb{P}\left[\frac{1}{n}\tau_X(\omega_n) \leq L\right] = 0.$$

If μ has finite first moment in addition, then for any $\varepsilon > 0$ we have

$$\lim_n \mathbb{P}\left[\left|\frac{1}{n}\tau_X(\omega_n) - \lambda\right| > \varepsilon\right] = 0,$$

where λ is the escape rate of the random walk.

The WLLN on $\mathscr{C}(\Sigma)$ follows from Maher–Tiozzo's theory, again due to the fact that the harmonic measure is atom-free. The corresponding result on $\mathscr{T}(\Sigma)$ is due to Dahmani–Horbez's lifting argument. Meanwhile, this convergence in probability alone is not enough to deduce the following SLLN.

Theorem 7.2. *Let $X = \mathscr{T}(\Sigma)$ or $\mathscr{C}(\Sigma)$ and suppose that μ satisfies some moment condition. Then almost every random path $\omega = (\omega_n)$ satisfies*

$$\lim_n \frac{1}{n} \tau_X(\omega_n) = \lambda.$$

In [67], Maher and Tiozzo discuss summable estimates of shadows along a direction for random walks on $\mathscr{C}(\Sigma)$ with bounded support. This is generalized to random walks with finite exponential moment, for which Boulanger, Mathieu, Sert and Sisto establish the same exponential decay of harmonic measure along the distance from the reference point (see [11]). Moreover, as Dahmani and Horbez point out, Benoist–Quint's analogue of Hsu–Robbins–Baum–Katz's theorem gives summable estimates for random walks with finite second moment. This leads to Theorem 7.2 under finite second moment, and Dahmani–Horbez's argument brings Theorem 7.2 to the setting of $\mathscr{T}(\Sigma)$. Here the lifting is possible whenever the random walk has finite first moment with respect to $d_{\mathscr{T}(\Sigma)}$, but it is the SLLN on $\mathscr{C}(\Sigma)$ that restricts the moment condition.

Before introducing Baik–Choi–Kim's theory in [5] for the SLLN under finite first moment condition, let us recall Maher–Tiozzo's strategy. In a δ-hyperbolic space such as $\mathscr{C}(\Sigma)$, the discrepancy $d(o, \omega_n o) - \tau(\omega_n)$ is comparable to the quantity $(\omega_n^{-1} o, \omega_n o)_o$. In particular, if $(\omega_n^{-1} o, \omega_n o)_o$ is smaller than half of $d(o, \omega_n o)$ minus a constant, then the discrepancy is bounded by $(\omega_n^{-1} o, \omega_n o)_o$ plus a constant.

In order to control $(\omega_n^{-1} o, \omega_n o)_o$, we now claim that the direction of $[o, \omega_n o]$ ($[o, \omega_n^{-1} o]$, resp.) is almost guided by $[o, g_1 \cdots g_{\lfloor n/2 \rfloor} o]$ ($[o, g_n^{-1} \cdots g_{\lfloor n/2 \rfloor + 1}^{-1} o]$, resp.). Given this claim, $(\omega_n^{-1} o, \omega_n o)_o$ is now correlated with the deviation $(g_n^{-1} \cdots g_{\lfloor n/2 \rfloor + 1}^{-1} o, g_1 \cdots g_{\lfloor n/2 \rfloor} o)$ between two *independent* random paths. Actually, the claim itself also involves quantities of the same nature: $(\omega_n o, \omega_{\lfloor n/2 \rfloor} o)_o$ grows linearly almost surely if $d(o, \omega_{\lfloor n/2 \rfloor} o)$ does grow linearly (which is true by the ergodic theorem) and

$$(o, \omega_n o)_{\omega_{\lfloor n/2 \rfloor} o} = (g_{\lfloor n/2 \rfloor}^{-1} \cdots g_1^{-1} o, g_{\lfloor n/2 \rfloor + 1} \cdots g_n o)_o,$$

being the deviation between another pair of independent random paths, grows sublinearly.

Hence, it suffices to show that $\mathbb{P}[(\check{\omega}_n o, \omega_n o)_o \geq Kn]$ is summable for any $K > 0$. For this Maher and Tiozzo condition on each choice of $\check{\omega}_n$ and regard $(\check{\omega}_n o, \omega_n o)_o$ as the deviation of a random segment $[o, \omega_n o]$ from a fixed direction $[o, \check{\omega}_n o]$. By [7, Lemma 4.5] (used together with the observation that a can be chosen as any positive number for [7, item (4.8)]), the probability is summable when μ has finite second moment and the conclusion follows. However, it is difficult to obtain summable estimates from weaker moment conditions with this strategy, considering Baum–Katz's theorem.

The situation is more complicated in $\mathcal{T}(\Sigma)$ since it lacks δ-hyperbolicity. Here, summable estimates of $\mathbb{P}[(\breve{\omega}_n o, \omega_n o)_o \geq Kn]$ are not enough and the deviation between paths should occur at special loci that enable arguments for δ-hyperbolic spaces.

Interestingly, Baik–Choi–Kim's strategy in [5] does not rely on the estimates for $\mathbb{P}[(\breve{\omega}_n o, \omega_n o)_o \geq Kn]$. Their idea is to utilize the notion of persistent joints of Maher and Tiozzo. Fixing suitable L, we say that a persistent joint arises at step $3kL$ if:

- the steps $(g_{3(k-1)L+1}, \ldots, g_{3kL})$ constitute one of two Schottky-type paths,
- the forward subpath $(\omega_{3kL} o, \omega_{3kL+1} o, \ldots)$ is contained in a shadow centered at $\omega_{3kL} o$ viewed from $\omega_{(3k-1)L} o$, and
- the backward subpath $(\ldots, \omega_{3(k-1)L-1} o, \omega_{3(k-1)L} o)$ is contained in a shadow centered at $\omega_{(3k-2)L} o$.

Both this construction and Gouëzel's construction pinpoints the steps for pivoting and derive almost sure phenomena. Nevertheless, the desired phenomena are different: Baik, Choi and Kim intend to guarantee large translation length by pivoting on the event of small translation length, while Gouëzel intends to guarantee definite progress from the prevalence of pivotal loci and one performs pivoting to establish this prevalence.

Moreover, Baik, Choi and Kim's pivots entail technicalities that are not shared with Gouëzel's pivots. First, the prevalence of persistence joints is guaranteed outside events of summable probabilities. Another complication is that persistent joints are not independent variables. Nonetheless, persistent joints at different steps are linked by the ergodic shift and the subadditive ergodic theorem does guarantee the eventual prevalence of persistence joints for almost every path. We remark that Gouëzel's construction equally works with stronger implications.

We explain how Baik–Choi–Kim achieved the almost sure linear growth of $\tau_X(\omega_n)$ without moment condition. We hope to declare an equivalence relation among random paths by pivoting: two paths are equivalent if they are identical except at the middle Schottky segment of the first N persistent joints. Here comes one technical issue that the original and the pivoted path may not have exactly the same persistent joint steps. This is because persistent joints are random variables depending on entire ω (not on finitely many steps near that joint) and single pivoting may alter entire distribution of persistent joints. To avoid this issue, Baik–Choi–Kim redefines pivots so that pivoting does not alter the pivot distribution. Moreover, according to their definition, persistent joints are incorporated in these pivots so the number of pivots also linearly grows almost surely.

Given this, the more pivots a path has, the smaller conditional probability that the path possesses inside its equivalence class. We now observe that if a path ω has small $\tau_X(\omega_n)$ and has enough number of early pivots within distance $\frac{1}{2}[d(o, \omega_n o) - \tau_X(\omega_n)]$ from o, then the early pivoting $\omega \mapsto \overline{\omega}$ results in large $\tau_X(\overline{\omega}_n)$. Similar discussion holds for the late pivoting; in this case, enough number of late pivots within distance $\frac{1}{2}[d(o, \omega_n o) - \tau_X(\omega_n)]$ from $\omega_n o$ are needed. It remains to show that random paths either have enough number of early or late pivots near o or near $\omega_n o$, respectively. This follows from distant allocation

of pivotal loci: pivotal loci cannot be concentrated within distance Ln for some suitable $L > 0$, so those paths ω with $\tau_X(\omega_n) \leq Ln$ necessarily fall into the above two categories.

The above argument can be improved if μ has finite first moment. For example, a linearly growing number of pivots for $(g_1, \ldots, g_n)$ should arise before $0.01n$ almost surely due to the subadditive ergodic theorem. Then the SLLN for displacement, another consequence of the ergodic theorem, asserts that a linearly growing number of pivots appear within distance $0.02\lambda n$ from o, where λ is the escape rate. Thus, one can rely only on the early pivoting and bound the probability of $\{\tau_X(\omega_n) \leq d(o, \omega_n o) - 0.04\lambda n\}$. Since $d(o, \omega_n o)/n$ also converges to λ, we obtain that $\limsup_n \tau_X(\omega_n)/n \geq 0.96\lambda$ almost surely when $X = \mathcal{C}(\Sigma)$.

In the case of $\mathcal{T}(\Sigma)$, one should keep in mind that small $(\omega_n^{-1}o, \omega_n o)_o$ will not automatically imply small $(\omega_n^{-m}o, \omega_n^k o)_o$ for all $m, k > 0$. Nonetheless, Baik–Choi–Kim exploit Rafi's results on thin triangles and fellow-traveling with thick ingredients and deduce the following fact. If the pivots are constructed with a (K, K', ε)-Schottky set for a sufficiently large constant K', then each middle Schottky segment at the pivotal steps fellow travels with some subsegment of $[o, \omega_n o]$. Moreover, when a random path ω satisfies

$$\frac{1}{2}[d(o, \omega_n o) - \tau(\omega_n)] \geq d(o, z) + C$$

for some pivotal locus z, then the directions of $[o, \check{\omega}_m o]$ and $[o, \omega_n o]$ near z are guided by the same Schottky segment. If one pivots the path at z by choosing a different Schottky direction, then $[o, \check{\bar{\omega}}_m o]$ and $[o, \bar{\omega}_n o]$ deviate at z and $[\check{\bar{\omega}}_m o, \bar{\omega}_n o]$ passes nearby z. This in turn implies that

$$\tau(\omega_n) \geq d(o, \omega_n o) - 2d(o, z).$$

Therefore, the SLLN in δ-hyperbolic spaces is copied to $\mathcal{T}(\Sigma)$.

After Gouëzel and Baik–Choi–Kim's work, Choi tried to incorporate two notions of pivots in [16]. As a result, Choi explained how accurately displacement and translation length match from the prevalence of pivots except for an event of exponentially decaying probability. This implies the following deviation inequality.

Proposition 7.3. *Let $X = \mathcal{T}(\Sigma)$ or $\mathcal{C}(\Sigma)$ and μ be a non-elementary probability measure on G with finite p-moment for some $p > 0$. Let also $q \leq p$ be a non-negative integer. Then there exists $K > 0$ such that*

$$\mathbb{E}\big[(\check{\omega}_m o, \omega_{m'} o)_o^{p+q}\big] < K + Ke^{-m/K}(m' - m)^q,$$
$$\mathbb{E}\big[d(o, [\check{\omega}_m o, \omega_{m'} o])^{p+q}\big] < K + Ke^{-m/K}(m' - m)^q$$

for all $0 \leq m \leq m'$, respectively.

In the special case $m = m'$, we obtain uniform control on $\mathbb{E}[(\check{\omega}_m o, \omega_{m'} o)_o^{2p}]$ from the finite p-moment of μ. While Maher–Tiozzo's argument first fixes one of random paths and considers the deviation of the other path from that fixed direction, Choi performs pivoting on both random paths to make the estimate more effective and obtains exponent doubling.

In particular, when μ has finite p-moment for some $p > 1/2$, the above estimate implies that

$$\mathbb{P}\big[(\check{\omega}_m o, \omega_m o)_o \geq Cm\big] \leq \frac{\mathbb{E}[(\check{\omega}_m o, \omega_m o)_o^{2p}]}{(Cm)^{2p}} \leq \frac{K}{(Cm)^{2p}}$$

is summable for any $C > 0$. Hence, Choi's result (together with Gouëzel's weak LDP from below) implies the SLLN for translation length in Gromov hyperbolic spaces including $\mathcal{C}(\Sigma)$ when μ has finite p-moment for some $p > 1/2$. Nevertheless, this still requires a moderate moment condition; the SLLN for translation length without moment condition relies on the pivoting itself, as in Baik–Choi–Kim's argument. Choi deduces another consequence of the pivoting, the control of the discrepancy between displacement and translation length with greater precision, as follows.

Theorem 7.4. *Suppose that μ has finite first moment. Then there exists a constant $K < \infty$ such that*

$$\limsup_n \frac{1}{\log n} |\tau(\omega_n) - d(o, \omega_n o)| < K$$

for almost every ω.

Meanwhile, the deviation inequality of Choi also turns out to be useful. One application is the improvement of the moment condition for geodesic tracking. Using the eventual version of Proposition 7.3, one can prove that sublinear geodesic tracking occurs in random walks with finite $(1/2)$-th moment. Moreover, a similar result for random walks with finite exponential moment implies logarithmic geodesic tracking. We note that logarithmic tracking was previously discussed on free groups, Gromov hyperbolic spaces and relatively hyperbolic spaces, the last two dealing with the case of bounded support (see [9, 62, 67, 84]).

Recall also that one can complete Mathieu–Sisto's approach to the CLT for displacement and translation length with this deviation inequality for $p = 2$, hence achieving the optimal moment condition. Moreover, Choi established via explicit pivoting the converse of CLTs: the convergence of $\frac{1}{\sqrt{n}}[d(o, \omega_n o) - c_n]$ or $\frac{1}{\sqrt{n}}[\tau(\omega_n) - c_n]$ in law for some constant c_n implies that the random walk has finite second moment. By adapting de Acosta's proof of the LIL for real-valued variables, Choi also establishes the LIL for displacement and translation length.

Let us now explain why the pivoting method is so effective. First, phenomena *in probability* are correlated to certain probabilities that decay to zero, which ultimately relies on the non-atomness of the harmonic measure. This is due to the fact that G is non-elementary: if a boundary point has the maximal atom, then all of its translations by G should also have maximal atom and the boundary point should have finite orbit by G. This technique has been employed by many authors, including Woess [90], Kaimanovich–Masur [54], Maher [65] and Maher–Tiozzo [67].

Nonetheless, this is not enough to deduce *almost sure* phenomena and more accurate information is needed. Specifically, we need to elaborate the decay rate of the harmonic

measure corresponding to shadows, in terms of the distance of the shadows from the reference point. The first success was achieved by Maher, who deduced in [66] exponential decay rate for random walks on $X = \mathcal{C}(\Sigma)$ with bounded support. We here explain a slight variation of Maher's argument.

Let us define the shadow $S_x(y, r)$ by the set $\{z : (y, z)_x \geq r\}$. Observe the following.

Lemma 7.5. *For $x, y, z \in G \cdot o$, $\xi \in X \cup \partial X$ and sufficiently large $R, R' > 0$, we have the following:*

(1) *if $y \notin S_x(\xi, R')$ and $z \in S_x(\xi, R' + R)$, then $y \in S_z(o, R')$ and $z \in S_y(\xi, R')$;*

(2) *$\nu(S_x(\xi, R)) := \mathbb{P}[(\omega_n x, \xi)_x \geq R \text{ eventually}] \leq 0.1$;*

(3) *$H(S_x(\xi, R)) := \mathbb{P}[(\omega_n x, \xi)_x \geq R \text{ at least once}] \leq 0.12$.*

The first item in fact holds for arbitrary R and R'; it follows from the inequalities

$$(z, \xi)_y, (x, y)_z \geq (z, \xi)_x - (y, \xi)_x,$$

which are equivalent to the triangle inequality. The second item is due to the fact that ν is atom-free. For the last item, we should correlate the once-hitting event and the eventual event. Let N_0 be the first hitting time for $S_x(\xi, R)$, i.e., the earliest step at which $(\omega_n x, \xi)_o \geq R$; this is a stopping time and the Markov property can be applied. Now with respect to any point $p \in S_x(\xi, R)$, we have

$$S_x(\xi, 0.5R)^c \subseteq S_p(x, 0.5R)$$

by the second item. Since $\nu(S_p(o, 0.5R)) \leq 0.1$ for sufficiently large R, we have

$$\nu[S_x(\xi, 0.5R) \mid \omega_{N_0} x = p] \geq 0.9$$

for each $p \in S_x(\xi, R)$. Consequently, we obtain

$$\nu[S_x(\xi, 0.5R)] \geq 0.9 H(S_x(\xi, R)), H(S_x(\xi, R)) \leq 0.12.$$

Let us now fix $\xi \in X \cup \partial X$ and a sufficiently large number $R > 0$, and estimate the hitting measure of $S_o(\xi, kR)$. We establish k 'intermediate rivers'

$$R_i = S_o\left(\xi, \frac{3i - 2}{3} R\right) \setminus S_o\left(\xi, \frac{3i - 1}{3} R\right)$$

that satisfy the following properties:

(1) each R_i separates $X \setminus R_i$ into two part, X_i^+ and X_i^-, such that $d(X_i^+, X_i^-) > M$;

(2) for each i, $R_1, \ldots, R_{i-1}$ are contained in X_i^- and $R_{i+1}, \ldots, R_k$ are contained in X_i^+;

(3) R_{i+1} is contained in a shadow of distance $R/4$ with respect to any point in R_i.
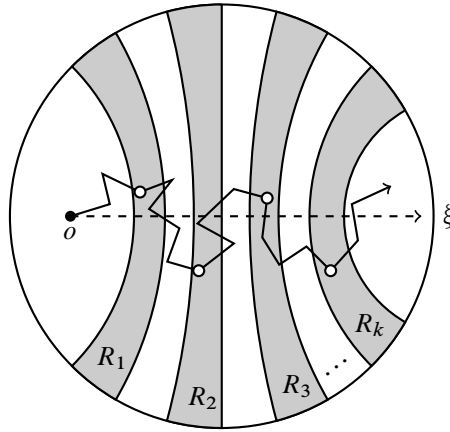
Figure 3. Schematics of rivers R_i in Maher's argument. Each white dots represents the sample locus at $N_i(\omega)$.

Due to properties (1), (2), and the fact that each random path consists of bounded steps, it is necessary to enter each R_i at least once to reach beyond R_k. This motivates us to consider the first hitting time $N_i(\omega)$ at which step $\omega_n o$ first enters R_i and define $E_i := \{\omega : N_1(\omega), \ldots, N_i(\omega) < \infty\}$. In order to calculate $\mathbb{P}[E_{i+1} \mid E_i]$, we condition on each choice of ω until N_i and fix $p = \omega_{N_i} o \in R_i$. As $N_i(\omega)$ is a stopping time, one can then apply the Markov property for estimation. Namely,

$$R_{i+1} \subseteq S_o\left(\xi, \frac{3i+1}{3}R\right) \subseteq S_p(\xi, 2R/3)$$

has hitting measure at most 0.12, and we have

$$\mathbb{P}[E_{i+1}] = \sum_{a_1,\ldots,a_n \in G} H(R_{i+1}) \cdot \mathbb{P}\big[\omega : N_i(\omega) = n, g_i = a_i \text{ for each } i = 1, \ldots, n\big]$$

$$\leq \sum_{a_1,\ldots,a_n \in G} H(S_{a_1 \cdots a_n o}(\xi, 2R/3)) \cdot \mathbb{P}\big[\omega : N_i(\omega) = n, g_i = a_i \text{ for each } i = 1, \ldots, n\big]$$

$$\leq 0.12 \sum_{a_1,\ldots,a_n \in G} \mathbb{P}\big[\omega : N_i(\omega) = n, g_i = a_i \text{ for each } i = 1, \ldots, n\big] = 0.12\mathbb{P}[E_i].$$

This implies $H(S_o(\xi, kR)) \leq \mathbb{P}[E_k] \leq 0.12^k$, as desired.

The role of Lemma 7.5 is to correlate the probability of progress in a specific direction with the probability of the rest. Note that this process does not require moment condition. However, in order to correlate those probabilities with the distance, one needs to quotient out the path space into measurable equivalence classes at regular distances. This is realized using hitting times, which crucially depend on the boundedness of each step so that no path jumps over and skips any river. We remark that random walks with finite exponential

moment exhibit similar behavior. Although the hitting time for each river is not exactly realized, the probability of the error case that one jumps over n river decays exponentially and we have similar exponential decay of the harmonic measure (cf. [11, Corollary 2.13]).

Meanwhile, Benoist–Quint's martingale version of Hsu–Robbins–Baum–Katz's theorem can deal with measures with finite p-moment, beyond those with bounded support. Let $x \in \partial X$ be approximated by a sequence $\{x_n\}_n \subseteq X$. We then define the cocycle arising from a horofunction

$$\sigma(g, x) := h_x(g^{-1}o) = \lim_{n \to \infty} \big[d(g^{-1}o, x_n) - d(o, x_n) \big].$$

We then estimate the concentration of $\mathbb{E}_{\mu^{*n}}[\sigma(g, x)]$ near the average λn. In this perspective, one adds up n martingale differences $\sigma_0(g_n^{-1}, g_{n-1}^{-1} \cdots g_1^{-1}o)$, which is a balanced version of $d(x, \omega_n o) - d(x, \omega_{n-1}o)$ and is bounded by $2d(o, g_n o)$. Since this step is L^p-bounded, one obtains an L^{p-2}-convergence rate: for each $\varepsilon > 0$, there exist constants D_n such that $\sum_n n^{p-2} D_n < \infty$ and

$$\mathbb{P}\big[\omega : \lambda - \varepsilon \leq \sigma(\omega_n^{-1}, x) \leq \lambda + \varepsilon \big] \leq D_n.$$

The advantage of this method is that it applies to L^p-integrable cocycles on an arbitrary compact metric space, where the p-th moments of the steps are uniformly bounded but not summable: it becomes summable only after modulating by order 2. In our setting of Gromov hyperbolic spaces or Teichmüller spaces, however, one can expect further efficiency from below since the p-th moment of the steps are not only bounded but exponentially decaying. More precisely, one does not observe $d(x, \omega_n o) - d(x, o)$ but observes its counterpart $(\omega_n o, x)_o$: this conversion requires Gromov hyperbolicity or its analogy on Teichmüller spaces.

This alternative strategy is also pursued in [7], beginning from the spectral gap of amenable groups acting on compact spaces and the exponential growth. Recall that another approach to this part of the argument, suggested in [16], removes the cocompactness assumption. Given this preliminary estimates, the final step is to bound $\lim_n (\omega_n o, x)_o^p$ with the sum of the terms

$$d(o, g_{k+1}o)^p 1_{\lim(x, \omega_n o) \geq d(o, \omega_k o)} \quad \text{(when } 0 < p < 1)$$

or

$$2^p \big[d(o, g_{k+1}o)^p + d(o, \omega_k o)^{p-1} d(o, g_{k+1}o) \big] 1_{\lim(x, \omega_n o) \geq d(o, \omega_k o)} \quad \text{(when } p \geq 1)$$

and control each expectation. In plain language, this counts the contribution of each step of the form $d(o, g_{k+1}o)^p$ or $d(o, \omega_k o)^{p-1} d(o, g_{k+1}o)$ only when the progress of the random path is toward x, whose probability decays exponentially and results in summable contribution to the p-th moment.

The description so far of Maher's and Benoist–Quint's theories are by no means complete; for a fuller analysis, including the martingale version of Hsu–Robbins–Baum–Katz's theorem, see [6, 7, 66].

To sum up, the above strategies estimate the decay rate of the harmonic measure by packing random paths into effective and ineffective cases, in terms of the intermediate steps, and by counting the effective cases only. This philosophy is maximized in the notion of pivots in Gouëzel's, Baik–Choi–Kim's and Choi's work. Each equivalent class of the same pivots consists mostly of the desirable paths and a small portion of undesirable paths; their probability can be compared by pivoting and can be summed using the Markov property. The notion of pivots also fit into the realm of random walks with infinite support; moment conditions are not necessary for the punctual appearance of pivots, and are used only to synchronize the time and distance progress of pivots.

8. Distance and counting

So far, we have discussed various methods to study random mapping classes that arise from random walks, especially with respect to their action on $\mathcal{C}(\Sigma)$ or $\mathcal{T}(\Sigma)$. This philosophy essentially differs from studying the random mapping classes on $\mathrm{Mod}(\Sigma)$ itself, since $\mathrm{Mod}(\Sigma)$ equipped with the word metric is not quasi-isometric to $\mathcal{C}(\Sigma)$ nor $\mathcal{T}(\Sigma)$. Further, counting problems are sensitive to the choice of the finite generating set of $\mathrm{Mod}(\Sigma)$. This is because a property P that is generic with respect to a metric d may not be generic with respect to another metric d' that is quasi-isometric to d. Hence, counting elements in $\mathrm{Mod}(\Sigma)$ with respect to the word metric is an essentially different problem than analogous problems on $\mathcal{C}(\Sigma)$ or $\mathcal{T}(\Sigma)$. A stereotypical problem in this direction is whether pseudo-Anosovs are generic in $\mathrm{Mod}(\Sigma)$ when counted with respect to a word metric.

One possible solution is to use properties of the action of $\mathrm{Mod}(\Sigma)$ on $\mathcal{T}(\Sigma)$ or $\mathcal{C}(\Sigma)$ beyond non-elementariness. In this direction, we mentioned that Mathieu–Sisto exploited the acylindrical hyperbolicity of $\mathrm{Mod}(\Sigma)$ to bring the aforementioned results (including definite progress, CLT, etc.) on $\mathcal{C}(\Sigma)$ to $\mathrm{Mod}(\Sigma)$ or $\mathcal{T}(\Sigma)$ (see [75]).

Another solution is to realize a Markov process on the group itself. Here we use the automatic structure of groups, first hinted at by Cannon [15] and later formulated by W. Thurston. For general reference, see [27]. An automatic structure of a group models (quasi-)geodesics on the group with paths on a directed graph. By considering a Markov process on this graph, we can utilize the techniques for random walks to describe the asymptotic behavior in the counting setting. In particular, if the graph possesses suitable hyperbolicity (e.g., exponential growth, independent directions, etc.), then the counting problem (guided by the Patterson–Sullivan measure) mingles with the random walk theory (guided by the harmonic measure). One can also interpret pivots as a partial realization of an automatic structure, by recording pivotal times (as if they represent specific vertices on the graph structure) and pivoting the choices at pivots (as if we distinguish cone types).

Notable examples of (geodesic) automatic groups include hyperbolic groups, relatively hyperbolic groups, right-angled Artin/Coxeter groups and many more. In particular, hyperbolic groups have a geodesic automatic structure with respect to any finite generating set, allowing the WLLN [40] and CLT [41] for displacement and translation length. Still,

the theory is not applicable for the entire mapping class group at the moment: although mapping class groups have quasi-geodesic automatic structures, it is not known whether they have a geodesic automatic structure. Nonetheless, if the generating set is nicely populated with some Schottky set, one can partially realize this strategy on weakly hyperbolic groups and Mod(Σ) that lack a geodesic automatic structure. This will be explained further in the forthcoming preprint [17] of the second author.

9. Mapping class groups and Out(F_n)

Lastly, let us mention another family of groups related to mapping class groups. By Dehn–Nielsen–Baer's theorem, one can interpret the mapping class group Mod(Σ) of a closed orientable surface Σ of genus at least 2 as an index 2 subgroup of the outer automorphism group of $\pi_1(\Sigma)$. This relationship does not hold true for punctured surface, and the outer automorphism group Out(F_n) of the free group of rank n becomes a separate object to study.

Nonetheless, mapping class groups and Out(F_n) have strong analogy. For example, Out(F_n) also canonically acts on a moduli space of a geometric object, namely, the space of simplicial graphs whose fundamental group is F_n and whose all vertices have valency at least 3 (see [21]). We call it the Culler–Vogtmann outer space of rank n (denoted by CV_n). Another impetus to the study of Outer space was the introduction of an asymmetric metric structure called the Lipschitz metric. This metric correctly captures the dynamics of outer automorphisms, as the Teichmüller metric does for mapping classes. For a systematic exposition on this metric, see [31].

Using the action of Out(F_n) on CV_n and its compactification $\overline{CV}_n$, Horbez proved the following theorem.

Theorem 9.1 ([50, Theorem 0.1]). *Let $G = $ Out(F_n), $X = CV_n$, and μ be a non-elementary probability measure on G. Then for an a.e. sample path $\omega := (\omega_n)_n$ of the random walk generated by μ, there exists a simplex $\xi(\omega)$ of free arational trees such that for all $x \in CV_N$, $(\omega_n x)_n$ converges to $\xi(\omega)$. This gives rise to a non-atomic hitting measure ν on the space $\mathcal{F}l$ of simplices of free arational trees.*

Moreover, when μ has finite first logarithmic moment with respect to the word metric on G and finite entropy, then $(\mathcal{F}l, \nu)$ is the Poisson boundary of (G, μ).

The proof follows the same storyline as in [54]. At the moment, ∂CV_n plays the role of $\mathcal{PMF}$ and rational/arational trees play the role of non-minimal/minimal foliations. First, a μ-stationary measure ν is constructed on ∂CV_n via the Cesàro mean, where ν is supported on $\mathcal{F}l$ due to the maximal atom trick as in [54, Lemma 2.2.2]: each rational tree fixes the conjugacy class of an element of F_n and the set of rational trees in ∂CV_n can be divided into countably infinitely many stabilizers. The non-atomness of ν follows similarly. A lemma of Kaimanovich and Masur ([54, Lemma 2.2.3]) then asserts that $(\mathcal{F}l, \nu)$ is a μ-boundary. The orbit convergence is also argued by bringing the concept of universally

convergent sequences. Finally, Kaimanovich's strip approximation criterion is satisfied by Hamenstädt's lines of minima [47] and leads to the identification of the Poisson boundary.

Here, we did not make an analogy of a part of Kaimanovich–Masur's argument. Among free arational trees, some admit a non-trivial simplex of length measures while the others admit only one length measure. Trees of the latter category are said to be *uniquely ergometric*, which partially correspond to uniquely ergodic measured foliations in $\mathcal{PMF}$. Kaimanovich and Masur argued that the hitting measure $\widetilde{\nu}$ on $\widetilde{\mathcal{MIN}}$ is actually supported on $\mathcal{UE}$. Lacking a powerful divergence result as in [71], we are unable to conclude that a.e. sample path heads to a uniquely ergometric tree. This was supplemented by other authors; namely, Namazi, Pettet and Reynolds proved that the hitting measure ν is supported on the set of uniquely ergometric trees [77], when μ is further assumed to have finite first moment on CV_n. In the same paper, they also showed the sublinear geodesic tracking on CV_n.

Yet another important analogy between the study of $\mathcal{T}(\Sigma)$ and CV_n is the existence of auxiliary Gromov hyperbolic spaces, namely, the curve complex $\mathcal{C}(\Sigma)$ and the free factor complex $\mathcal{FF}$. Just as we pick the shortest curve on a point $x \in \mathcal{T}(\Sigma)$ to define the shortest curve projection $\pi_{\mathcal{C}}(x) \in \mathcal{C}(\Sigma)$, we pick properly embedded, non-contractible, connected subgraphs of $x \in CV_N$ to define the projection $\pi_{\mathcal{FF}}(x) \in \mathcal{FF}$. Just as $\pi_{\mathcal{C}}$, $\pi_{\mathcal{FF}}$ that sends geodesics to quasi-geodesics. Although there are some technicalities, largely due to the asymmetry of the Lipschitz metric on CV_n, the strategies of Horbez [51] and Dahmani–Horbez [22] apply here. Namely, we can bring the limit laws on a Gromov hyperbolic space $\mathcal{FF}$ (due to [67] and [7]) to CV_n, as we did for $\mathcal{C}(\Sigma)$ and $\mathcal{T}(\Sigma)$. As a result, we obtain the SLLN and CLT on CV_n.

10. Further directions

We have discussed random walks on $\mathrm{Mod}(\Sigma)$ from different perspectives. Several questions arise from the difference among groups and spaces. First, it is known that random walks on hyperbolic groups also satisfy the local limit theorem [43]. The ingredient of Gouëzel's argument that depends on the Gromov hyperbolicity is to establish Ancona's inequality (see also [39] for a related work of Gekhtman–Gerasimov–Potyagailo–Yang). Considering the parallel theory of pivoting on Gromov hyperbolic spaces and $\mathcal{T}(\Sigma)$, one might hope for a similar result on $\mathcal{T}(\Sigma)$.

Despite partial achievements, the complete Patterson–Sullivan theory on $\mathrm{Mod}(\Sigma)$ is not attained yet. Once achieved, this will serve as another perspective for the counting problem in $\mathrm{Mod}(\Sigma)$. For instance, see Gekhtman's analysis on the stable type of mapping class groups [38].

References

[1] J. W. Anderson, J. Aramayona, and K. J. Shackleton, Uniformly exponential growth and mapping class groups of surfaces. In *In the tradition of Ahlfors-Bers. IV*, pp. 1–6, Contemp. Math. 432, American Mathematical Society, Providence, RI, 2007 Zbl 1128.20027 MR 2342801

[2] T. Aougab, Uniform hyperbolicity of the graphs of curves. *Geom. Topol.* **17** (2013), no. 5, 2855–2875 Zbl 1273.05050 MR 3190300

[3] G. N. Arzhantseva, C. H. Cashen, and J. Tao, Growth tight actions. *Pacific J. Math.* **278** (2015), no. 1, 1–49 Zbl 1382.20046 MR 3404665

[4] J. Athreya, A. Bufetov, A. Eskin, and M. Mirzakhani, Lattice point asymptotics and volume growth on Teichmüller space. *Duke Math. J.* **161** (2012), no. 6, 1055–1111 Zbl 1246.37009 MR 2913101

[5] H. Baik, I. Choi, and D. M Kim, Linear growth of translation lengths of random isometries on Gromov hyperbolic spaces and Teichmüller spaces. 2021, arXiv:2103.13616

[6] Y. Benoist and J.-F. Quint, Central limit theorem for linear groups. *Ann. Probab.* **44** (2016), no. 2, 1308–1340 Zbl 1341.22006 MR 3474473

[7] I. Benua and Z. F. Kén, Central limit theorem on hyperbolic groups. *Izv. Ross. Akad. Nauk Ser. Mat.* **80** (2016), no. 1, 5–26 Zbl 1341.22006 MR 3462675

[8] M. Björklund, Central limit theorems for Gromov hyperbolic groups. *J. Theoret. Probab.* **23** (2010), no. 3, 871–887 Zbl 1217.60019 MR 2679960

[9] S. Blachère, P. Haïssinsky, and P. Mathieu, Harmonic measures versus quasiconformal measures for hyperbolic groups. *Ann. Sci. Éc. Norm. Supér. (4)* **44** (2011), no. 4, 683–721 Zbl 1243.60005 MR 2919980

[10] M. Bonk and O. Schramm, Embeddings of Gromov hyperbolic spaces. *Geom. Funct. Anal.* **10** (2000), no. 2, 266–306 Zbl 0972.53021 MR 1771428

[11] A. Boulanger, P. Mathieu, Ç. Sert, and A Sisto, Large deviations for random walks on hyperbolic spaces. *Ann. Sci. Éc. Norm. Supér. (4)*, to appear.

[12] B. H. Bowditch, Uniform hyperbolicity of the curve graphs. *Pacific J. Math.* **269** (2014), no. 2, 269–280 Zbl 1330.57023 MR 3238474

[13] M. R. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*. Grundlehren Math. Wiss. 319, Springer, Berlin, 1999 Zbl 0988.53001 MR 1744486

[14] D. Calegari and J. Maher, Statistics and compression of scl. *Ergodic Theory Dynam. Systems* **35** (2015), no. 1, 64–110 Zbl 1351.37214 MR 3294292

[15] J. W. Cannon, The combinatorial structure of cocompact discrete hyperbolic groups. *Geom. Dedicata* **16** (1984), no. 2, 123–148 Zbl 0606.57003 MR 758901

[16] I. Choi, Central limit theorem and geodesic tracking on hyperbolic spaces and Teichmüller spaces. 2021, arXiv:2106.13017

[17] I. Choi, Pseudo-Anosovs are exponentially generic in mapping class groups. 2021, arXiv:2110.06678

[18] M. Clay, K. Rafi, and S. Schleimer, Uniform hyperbolicity of the curve graph via surgery sequences. *Algebr. Geom. Topol.* **14** (2014), no. 6, 3325–3344 Zbl 1309.57015 MR 3302964

[19] M. Coornaert, Mesures de Patterson–Sullivan sur le bord d'un espace hyperbolique au sens de Gromov. *Pacific J. Math.* **159** (1993), no. 2, 241–270 Zbl 0797.20029 MR 1214072

[20] M. Coornaert, T. Delzant, and A. Papadopoulos, *Géométrie et théorie des groupes*. Lecture Notes in Math. 1441, Springer, Berlin, 1990 Zbl 0727.20018 MR 1075994

[21] M. Culler and K. Vogtmann, Moduli of graphs and automorphisms of free groups. *Invent. Math.* **84** (1986), no. 1, 91–119 Zbl 0589.20022 MR 830040

[22] F. Dahmani and C. Horbez, Spectral theorems for random walks on mapping class groups and out (F_N). *Int. Math. Res. Not. IMRN* (2018), no. 9, 2693–2744 Zbl 1407.37046 MR 3801494

[23] T. Das, D. Simmons, and M. Urbański, *Geometry and dynamics in Gromov hyperbolic metric spaces*. Math. Surveys Monogr. 218, American Mathematical Society, Providence, RI, 2017 Zbl 06729361 MR 3558533

[24] J. L. Doob, Discrete potential theory and boundaries. *J. Math. Mech.* **8** (1959), 433–458; erratum 993 Zbl 0101.11503 MR 0107098

[25] S. Dowdall, M. Duchin, and H. Masur, Statistical hyperbolicity in Teichmüller space. *Geom. Funct. Anal.* **24** (2014), no. 3, 748–795 Zbl 1302.30055 MR 3213829

[26] M. Duchin, Thin triangles and a multiplicative ergodic theorem for Teichmüller geometry. 2005, arXiv:math.GT/0508046

[27] D. B. A. Epstein, J. W. Cannon, D. F. Holt, S. V. F. Levy, M. S. Paterson, and W. P. Thurston, *Word processing in groups*. Jones and Bartlett, Boston, MA, 1992 Zbl 0764.20017 MR 1161694

[28] A. Eskin, H. Masur, and K. Rafi, Large-scale rank of Teichmüller space. *Duke Math. J.* **166** (2017), no. 8, 1517–1572 Zbl 1373.32012 MR 3659941

[29] B. Farb, Relatively hyperbolic groups. *Geom. Funct. Anal.* **8** (1998), no. 5, 810–840 Zbl 0985.20027 MR 1650094

[30] A. Fathi, F. Laudenbach, and V. Poénaru, *Travaux de Thurston sur les surfaces*. Astérisque 66, Société Mathématique de France, Paris, 1979 Zbl 1244.57005 MR 568308

[31] S. Francaviglia and A. Martino, Metric properties of outer space. *Publ. Mat.* **55** (2011), no. 2, 433–473 Zbl 1268.20042 MR 2839451

[32] H. Furstenberg, A Poisson formula for semi-simple Lie groups. *Ann. of Math. (2)* **77** (1963), 335–386 Zbl 0192.12704 MR 146298

[33] H. Furstenberg, Noncommuting random products. *Trans. Amer. Math. Soc.* **108** (1963), 377–428 Zbl 0203.19102 MR 163345

[34] H. Furstenberg, Poisson boundaries and envelopes of discrete groups. *Bull. Amer. Math. Soc.* **73** (1967), 350–356 Zbl 0184.33105 MR 210812

[35] H. Furstenberg, Random walks and discrete subgroups of Lie groups. In *Advances in probability and related topics, Vol. 1*, pp. 1–63, Dekker, New York, 1971 Zbl 0221.22008 MR 0284569

[36] H. Furstenberg and H. Kesten, Products of random matrices. *Ann. Math. Statist.* **31** (1960), 457–469 Zbl 0137.35501 MR 121828

[37] V. Gadre, J. Maher, and G. Tiozzo, Word length statistics for Teichmüller geodesics and singularity of harmonic measure. *Comment. Math. Helv.* **92** (2017), no. 1, 1–36 Zbl 1362.32009 MR 3615033

[38] I. Gekhtman, Stable type of the mapping class group. 2013, arXiv:1310.5364

[39] I. Gekhtman, V. Gerasimov, L. Potyagailo, and W. Yang, Martin boundary covers Floyd boundary. *Invent. Math.* **223** (2021), no. 2, 759–809 Zbl 1498.20102 MR 4209863

[40] I. Gekhtman, S. J. Taylor, and G. Tiozzo, Counting loxodromics for hyperbolic actions. *J. Topol.* **11** (2018), no. 2, 379–419 Zbl 1436.20083 MR 3789828

[41] I. Gekhtman, S. J. Taylor, and G. Tiozzo, Central limit theorems for counting measures in coarse negative curvature. *Compos. Math.* **158** (2022), no. 10, 1980–2013 Zbl 1502.60030 MR 4504671

[42] É. Ghys and P. de la Harpe (eds.), *Sur les groupes hyperboliques d'après Mikhael Gromov (Bern, 1988)*, Progr. Math. 83, Birkhäuser Boston, Boston, MA, 1990 Zbl 0731.20025

[43] S. Gouëzel, Local limit theorem for symmetric random walks in Gromov-hyperbolic groups. *J. Amer. Math. Soc.* **27** (2014), no. 3, 893–928 Zbl 1320.60017 MR 3194496

[44] S. Gouëzel, Exponential bounds for random walks on hyperbolic spaces without moment conditions. *Tunis. J. Math.* **4** (2022), no. 4, 635–671 Zbl 07649064 MR 4533553

[45] M. Gromov, Hyperbolic groups. In *Essays in group theory*, pp. 75–263, Math. Sci. Res. Inst. Publ. 8, Springer, New York, 1987 Zbl 0634.20015 MR 919829

[46] D. Gruber, A. Sisto, and R. Tessera, Random Gromov's monsters do not act non-elementarily on hyperbolic spaces. *Proc. Amer. Math. Soc.* **148** (2020), no. 7, 2773–2782 Zbl 1481.20157 MR 4099767

[47] U. Hamenstädt, Lines of minima in outer space. *Duke Math. J.* **163** (2014), no. 4, 733–776 Zbl 1408.37011 MR 3178431

[48] W. J. Harvey, Boundary structure of the modular group. In *Riemann surfaces and related topics (State Univ. New York, Stony Brook, N.Y., 1978)*, pp. 245–251, Ann. of Math. Stud. 97, Princeton Univ. Press, Princeton, NJ, 1981 Zbl 0461.30036 MR 624817

[49] S. Hensel, P. Przytycki, and R. C. H. Webb, 1-slim triangles and uniform hyperbolicity for arc graphs and curve graphs. *J. Eur. Math. Soc. (JEMS)* **17** (2015), no. 4, 755–762 Zbl 1369.57018 MR 3336835

[50] C. Horbez, The Poisson boundary of Out(F_N). *Duke Math. J.* **165** (2016), no. 2, 341–369 Zbl 1401.20043 MR 3457676

[51] C. Horbez, Central limit theorems for mapping class groups and Out(F_N). *Geom. Topol.* **22** (2018), no. 1, 105–156 Zbl 1439.60008 MR 3720342

[52] N. V. Ivanov, A short proof of non-Gromov hyperbolicity of Teichmüller spaces. *Ann. Acad. Sci. Fenn. Math.* **27** (2002), no. 1, 3–5 Zbl 1017.30058 MR 1884345

[53] V. A. Kaimanovich, The Poisson formula for groups with hyperbolic properties. *Ann. of Math. (2)* **152** (2000), no. 3, 659–692 Zbl 0984.60088 MR 1815698

[54] V. A. Kaimanovich and H. Masur, The Poisson boundary of the mapping class group. *Invent. Math.* **125** (1996), no. 2, 221–264 Zbl 0864.57014 MR 1395719

[55] H. Kesten, Symmetric random walks on groups. *Trans. Amer. Math. Soc.* **92** (1959), 336–354 Zbl 0092.33503 MR 109367

[56] J. F. C. Kingman, The ergodic theory of subadditive stochastic processes. *J. Roy. Statist. Soc. Ser. B* **30** (1968), 499–510 Zbl 0182.22802 MR 254907

[57] J. F. C. Kingman, Subadditive ergodic theory. *Ann. Probability* **1** (1973), 883–909 Zbl 0311.60018 MR 356192

[58] J. F. C. Kingman, Subadditive processes. In *École d'été de Probabilités de Saint-Flour, V-1975*, pp. 167–223, Lecture Notes in Math. 539, Springer, Berlin, 1976 Zbl 0367.60030 MR 0438477

[59] E. Klarreich, The boundary at infinity of the curve complex and the relative Teichmüller space. *Groups Geom. Dyn.* **16** (2022), no. 2, 705–723 Zbl 1507.57021 MR 4502619

[60] F. Ledrappier, Harmonic measures and Bowen–Margulis measures. *Israel J. Math.* **71** (1990), no. 3, 275–287 Zbl 0728.53029 MR 1088820

[61] F. Ledrappier, Applications of dynamics to compact manifolds of negative curvature. In *Proceedings of the International Congress of Mathematicians, Vol. 1, 2 (Zürich, 1994)*, pp. 1195–1202, Birkhäuser, Basel, 1995 Zbl 0841.53037 MR 1404020

[62] F. Ledrappier, Some asymptotic properties of random walks on free groups. In *Topics in probability and Lie groups: boundary theory*, pp. 117–152, CRM Proc. Lecture Notes 28, American Mathematical Society, Providence, RI, 2001 Zbl 0994.60073 MR 1832436

[63] A. Lenzhen, Teichmüller geodesics that do not have a limit in $\mathcal{PMF}$. *Geom. Topol.* **12** (2008), no. 1, 177–197 Zbl 1189.30086 MR 2377248

[64] J. Maher, Linear progress in the complex of curves. *Trans. Amer. Math. Soc.* **362** (2010), no. 6, 2963–2991 Zbl 1232.37023 MR 2592943

[65] J. Maher, Random walks on the mapping class group. *Duke Math. J.* **156** (2011), no. 3, 429–468 Zbl 1213.37072 MR 2772067

[66] J. Maher, Exponential decay in the mapping class group. *J. Lond. Math. Soc. (2)* **86** (2012), no. 2, 366–386 Zbl 1350.37010 MR 2980916

[67] J. Maher and G. Tiozzo, Random walks on weakly hyperbolic groups. *J. Reine Angew. Math.* **742** (2018), 187–239 Zbl 1434.60015 MR 3849626

[68] J. Mangahas, Uniform uniform exponential growth of subgroups of the mapping class group. *Geom. Funct. Anal.* **19** (2010), no. 5, 1468–1480 Zbl 1207.57005 MR 2585580

[69] R. S. Martin, Minimal positive harmonic functions. *Trans. Amer. Math. Soc.* **49** (1941), 137–172 Zbl 0025.33302 MR 3919

[70] H. Masur, Two boundaries of Teichmüller space. *Duke Math. J.* **49** (1982), no. 1, 183–190 Zbl 0508.30039 MR 650376

[71] H. Masur, Hausdorff dimension of the set of nonergodic foliations of a quadratic differential. *Duke Math. J.* **66** (1992), no. 3, 387–442 Zbl 0780.30032 MR 1167101

[72] H. A. Masur and Y. N. Minsky, Geometry of the complex of curves. I. Hyperbolicity. *Invent. Math.* **138** (1999), no. 1, 103–149 Zbl 0941.32012 MR 1714338

[73] H. A. Masur and Y. N. Minsky, Geometry of the complex of curves. II. Hierarchical structure. *Geom. Funct. Anal.* **10** (2000), no. 4, 902–974 Zbl 0972.32011 MR 1791145

[74] H. A. Masur and M. Wolf, Teichmüller space is not Gromov hyperbolic. *Ann. Acad. Sci. Fenn. Ser. A I Math.* **20** (1995), no. 2, 259–267 Zbl 0878.32015 MR 1346811

[75] P. Mathieu and A. Sisto, Deviation inequalities for random walks. *Duke Math. J.* **169** (2020), no. 5, 961–1036 Zbl 1465.60038 MR 4079419

[76] Y. N. Minsky, Extremal length estimates and product regions in Teichmüller space. *Duke Math. J.* **83** (1996), no. 2, 249–286 Zbl 0861.32015 MR 1390649

[77] H. Namazi, A. Pettet, and P. Reynolds, Ergodic decompositions for folding and unfolding paths in outer space. 2020, arXiv:1410.8870

[78] S. J. Patterson, The limit set of a Fuchsian group. *Acta Math.* **136** (1976), no. 3–4, 241–273 Zbl 0336.30005 MR 450547

[79] K. Rafi, A combinatorial model for the Teichmüller metric. *Geom. Funct. Anal.* **17** (2007), no. 3, 936–959 Zbl 1129.30031 MR 2346280

[80] K. Rafi, Hyperbolicity in Teichmüller space. *Geom. Topol.* **18** (2014), no. 5, 3025–3053 Zbl 1314.30082 MR 3285228

[81] I. Rivin, Walks on groups, counting reducible matrices, polynomials, and surface and free group automorphisms. *Duke Math. J.* **142** (2008), no. 2, 353–379 Zbl 1207.20068 MR 2401624

[82] H. L. Royden, Automorphisms and isometries of Teichmüller space. In *Advances in the Theory of Riemann Surfaces (Stony Brook, NY, 1969)*, pp. 369–383, Ann. of Math. Studies 66, Princeton University Press, Princeton, NJ, 1971 Zbl 0218.32011 MR 0288254

[83] S. Sawyer and T. Steger, The rate of escape for anisotropic random walks in a tree. *Probab. Theory Related Fields* **76** (1987), no. 2, 207–230 Zbl 0608.60064 MR 906775

[84] A. Sisto, Tracking rates of random walks. *Israel J. Math.* **220** (2017), no. 1, 1–28 Zbl 1407.60064 MR 3666817

[85] A. Sisto, Contracting elements and random walks. *J. Reine Angew. Math.* **742** (2018), 79–114 Zbl 1499.20108 MR 3849623

[86] D. Sullivan, The density at infinity of a discrete group of hyperbolic motions. *Inst. Hautes Études Sci. Publ. Math.* (1979), no. 50, 171–202 Zbl 0439.30034 MR 556586

[87] G. Tiozzo, Sublinear deviation between geodesics and sample paths. *Duke Math. J.* **164** (2015), no. 3, 511–539 Zbl 1314.30085 MR 3314479

[88] J. Väisälä, Gromov hyperbolic spaces. *Expo. Math.* **23** (2005), no. 3, 187–231 Zbl 1087.53039 MR 2164775

[89] R. C. H. Webb, A short proof of the bounded geodesic image theorem. 2013, arXiv:1301.6187

[90] W. Woess, Boundaries of random walks on graphs and groups with infinitely many ends. *Israel J. Math.* **68** (1989), no. 3, 271–301 Zbl 0723.60009 MR 1039474

[91] W. Woess, *Random walks on infinite graphs and groups*. Cambridge Tracts in Math. 138, Cambridge University Press, Cambridge, 2000 Zbl 0951.60002 MR 1743100

[92] W.-y. Yang, Statistically convex-cocompact actions of groups with contracting elements. *Int. Math. Res. Not. IMRN* (2019), no. 23, 7259–7323 Zbl 07154609 MR 4039013

[93] W.-y. Yang, Genericity of contracting elements in groups. *Math. Ann.* **376** (2020), no. 3–4, 823–861 Zbl 1481.20164 MR 4081104

Hyungryul Baik
Department of Mathematical Sciences, KAIST College of Natural Sciences, 291 Daehak-ro, Yuseong-gu, Daejeon 34141, South Korea; hrbaik@kaist.ac.kr

Inhyeok Choi
Department of Mathematical Sciences, KAIST College of Natural Sciences, 291 Daehak-ro, Yuseong-gu, Daejeon 34141, South Korea; inhyeokchoi48@gmail.com