LARGEST PROJECTIONS FOR RANDOM WALKS AND SHORTEST CURVES IN RANDOM MAPPING TORI

ALESSANDRO SISTO AND SAMUEL J. TAYLOR

Abstract. We show that the largest subsurface projection distance between a marking and its image under the nth step of a random walk grows logarithmically in n, with probability approaching 1 as $n \to \infty$. Our setup is general and also applies to (relatively) hyperbolic groups and to $\mathrm{Out}(F_n)$.

We then use this result to prove Rivin's conjecture that for a random walk (w_n) on the mapping class group, the shortest geodesic in the hyperbolic mapping torus M_{w_n} has length on the order of $1/\log^2(n)$.

1. Introduction

In [14], Maher proved that a random walk generated by a nonelementary measure on the mapping class group $\mathrm{Mod}(S)$ has *positive drift* with respect to the action $\mathrm{Mod}(S) \curvearrowright \mathcal{C}(S)$ on the curve complex $\mathcal{C}(S)$ of an orientable surface S. Informally, this means that for a fixed curve $\alpha \in \mathcal{C}(S)$, the displacement $d_{\mathcal{C}(S)}(\alpha, w_n\alpha)$ between α and its image under the nth step of the random walk (w_n) grows linearly in n with probability approaching 1 as $n \to \infty$. Generalizations were made to actions of countable groups on hyperbolic spaces by Calegari–Maher [8] and Maher–Tiozzo [17], see also [19].

In this note, we give finer information about the behavior of random walks on mapping class groups, by studying largest subsurface projections. Subsurface projections are very useful for example in view of the distance formula [18] for mapping class groups, as well as their connections with the geometry of hyperbolic 3–manifolds, see e.g. [22] and Section 7 below. We show that the largest distance between subsurface projection of 1 and w_n grows logarithmically in n with probability approaching 1 as $n \to \infty$. Since curve complex distance is "stalled" while progress is being made in a subsurface, this complements Maher's result.

In fact, our methods apply to several other examples, which we list in the theorem below.

Theorem 1.1. *Consider one of the following setups:*

- *Let G be hyperbolic group, Q an infinite index quasiconvex subgroup, and fix a finite generating set S for G which contains a generating set for Q. Let $\{\mathcal{C}(Y)\}_{Y\in\mathcal{S}}$ be the set of induced subgraphs of left cosets of Q in the Cayley graph $\mathrm{Cay}_S(G)$. For $Y \in \mathcal{S}$, let $\pi_Y : G \to \mathcal{C}(Y)$ denote a closest point projection in $\mathrm{Cay}_S(G)$.*
- *Let G be the mapping class group of a closed connected oriented surface of genus at least 2, and let $\{\mathcal{C}(Y)\}_{Y\in\mathcal{S}}$ be the set of curve complexes of either all proper subsurfaces or all proper subsurfaces of a given topological type. For $Y \in \mathcal{S}$, let $\pi_Y : G \to \mathcal{C}(Y)$ denote the subsurface projection, i.e. we let $\pi_Y(g)$ be the subsurface projection to Y of $g\lambda$ for a fixed marking λ [18].*
- *Let G denote $\mathrm{Out}(F_n)$, $n \geq 3$, and let $\{\mathcal{C}(Y)\}_{Y\in\mathcal{S}}$ be the set of free factor complexes of either all proper free factors of F_n or all proper free factors of a given rank ≥ 2.*

For $Y \in \mathcal{S}$, let $\pi_Y : G \to \mathcal{C}(Y)$ denote the subfactor projection, i.e. $\pi_Y(g)$ is the subfactor projection to the factor complex of Y of $g\lambda$ for a fixed marked graph λ [5, 31].

- *Let G be hyperbolic relative to its proper subgroups $H_1, \ldots, H_n$, each containing an undistorted element*[1]*. Fix a finite generating set S for G which contains generating sets for each H_i, and let $\{\mathcal{C}(Y)\}_{Y\in\mathcal{S}}$ be the set of induced subgraphs of left cosets of H_1 in $\mathrm{Cay}_S(G)$. For $Y \in \mathcal{S}$, let $\pi_Y : G \to \mathcal{C}(Y)$ denote a closest point projection in $\mathrm{Cay}_S(G)$.*

Let μ be a symmetric measure whose support is finite and generates G, and let (w_n) denote the random walk driven by μ. Then there exists C so that

$$\mathbb{P}\left(\sup_{Z\in\mathcal{S}} d_{\mathcal{C}(Z)}(\pi_Z(1), \pi_Z(w_n)) \in [C^{-1}\log n, C\log n]\right) \to 1,$$

as n tends to infinity.

We will treat all cases simultaneously by using common features of the mentioned setups. We explain those conditions in Section 2, where we also state the theorem that covers all cases (Theorem 2.3).

In fact, we remark that Theorem 2.3 covers even more cases. For example, it can be applied in the contexts of hyperplanes in cube complexes, of hierarchically hyperbolic spaces [3, 4], and of hyperbolically embedded subgroups [9]. Also, there is no need for the support of the measure to generate G, as long as it generates a "large enough" semigroup of G (in the mapping class group case, it suffices that the semigroup contains a pseudo-Anosov and an infinite order reducible element). In the interest of brevity, we have chosen Theorem 1.1 as stated.

Application to random fibered 3–manifolds. Let S be a closed connected orientable surface of genus at least 2 and $\mathrm{Mod}(S)$ its mapping class group. For each $f \in \mathrm{Mod}(S)$, we denote the corresponding mapping torus $M(f)$. This is the 3–manifold constructed by starting with $S \times [0,1]$ and glueing $S \times \{1\}$ to $S \times \{0\}$ via a homeomorphism in the class of f. Thurston's celebrated hyperbolization theorem for fibered 3–manifolds states that $M(f)$ has a (unique by Mostow Rigidity) hyperbolic structure if and only if f is pseudo-Anosov [33, 25].

The property of being pseudo-Anosov, and hence of determining a hyperbolic mapping torus, is typical. In fact, Rivin [26] and Maher [14] proved that for an appropriate random walk (w_n) on $\mathrm{Mod}(S)$ the probability that w_n is pseudo-Anosov goes to 1 as $n \to \infty$. Hence, it make sense to address questions about the typical geometry of these random hyperbolic 3-manifolds. For example, in [27, Conjecture 5.10], Rivin conjectures:

Conjecture 1.2 (Rivin)**.** *The injectivity radius of a random mapping torus of a surface S decays as $1/\log^2(n)$.*

Rivin also proves his conjecture in the case where S is a punctured torus. In the case of punctured surfaces, injectivity radius should be interpreted as $1/2 \cdot \mathrm{sys}(M)$, where $\mathrm{sys}(M)$ is the *systole* of the hyperbolic manifold M, i.e. the length of the shortest geodesic in M.

In Section 7 we prove Rivin's conjecture by establishing:

[1]We call an g undistorted if $n \mapsto g^n$ is a quasi-isometric embedding. One can weaken the assumption to H_i being infinite, but that requires adjustments to the proof, and we opted to keep the proof as simple and as uniform as possible.

Theorem 1.3. *Let μ be a symmetric measure whose support is finite and generates* $\mathrm{Mod}(S)$*, and let (w_n) denote the random walk driven by μ. Then there exists $C > 0$ so that*

$$\mathbb{P}\left(\frac{C^{-1}}{\log^2(n)} \leq \mathrm{sys}(M(w_n)) \leq \frac{C}{\log^2(n)}\right) \to 1,$$

as $n \to \infty$.

Outline. In Section 2 we introduce the general setup that covers all the cases of Theorem 1.1. In Section 3 we prove a result useful to establish both the lower and the upper bound on projections, namely that it is exponentially unlikely that the projection of the random walk to Z moves a large distance. The proof is based on ideas from [30, 29], but we simplify the arguments from those papers. In Section 4 we prove the lower bound on projections. Besides our application to short curves on random mapping tori, this is the main and most original contribution of this paper. The proof is based on the second moment method, and to the best of our knowledge this is the first time that the method is applied in a similar context. In Section 5 we prove the upper bound on projections, using ideas from [1] and [30] (in which the upper bound is proved in the case of mapping class groups). In Section 6 we verify that the examples in Theorem 1.1 fit into the general framework.

Finally, in Section 7 we review the connection between subsurface projections and lengths of curves in hyperbolic mapping tori. When then use this to prove Theorem 1.3 verifying Rivin's conjecture.

Acknowledgements. We are grateful to Igor Rivin, as well as one of the referees, for pointing out the connection between our work and Conjecture 1.2. We also thank Yair Minsky and David Futer for helpful conversations about hyperbolic geometry. Finally, we thank the anonymous referees for valuable suggestions that improved our exposition.

This material is based upon work supported by the National Science Foundation under grant No. DMS-1440140 while the authors were in residence at the Mathematical Sciences Research Institute in Berkeley, California, during the Fall 2016 semester. The second author is also partially supported by NSF DMS-1400498.

2. Statement of the main theorem

We now describe the common features of the projections in the various setups of Theorem 1.1.

Definition 2.1. A *projection system* $(\mathcal{S}, Y_0, \{\pi_Z\}_{Z\in\mathcal{S}}, \pitchfork)$ on a finitely generated group G endowed with the word metric d_G is a quadruple consisting of the following objects.

(1) A set $\mathcal{S}$ together with an action of G of $\mathcal{S}$ with one orbit, and a specified element $Y_0 \in \mathcal{S}$.
(2) For each $Z \in \mathcal{S}$, a metric space $\mathcal{C}(Z)$ and an L–Lipschitz map $\pi_Z : (G, d_G) \to \mathcal{C}(Z)$. We set $d_Z(x, y) = d_{\mathcal{C}(Z)}(\pi_Z(x), \pi_Z(y))$.
(3) We have $d_{gZ}(gx, gy) = d_Z(x, y)$ for each $g, x, y \in G$ and $Z \in \mathcal{S}$.
(4) An equivariant symmetric relation $\pitchfork$ on $\mathcal{S}$. We require that there exists $B > 0$ so that whenever $Y_0 \pitchfork gY_0$, we have $\min\{d_{Y_0}(g, h), d_{gY_0}(1, h)\} < B$ for every $h \in G$.
(5) There exists an integer s so that for each pairwise distinct $Y_1, \dots, Y_s \in \mathcal{S}$ there exist i, j so that $Y_i \pitchfork Y_j$.

Item 4 above is probably most important, and in the context of mapping class groups, where $\mathcal{C}(Z)$ is the curve complex of a subsurface, it is called the Behrstock inequality [2].

See Figure 1. The reader may consult Section 6.1 to see how this notion of a projection system fits into the well-known context of subsurface projections for mapping class groups.

It might be worth noting that there is no (Gromov-)hyperbolic space involved, even though we will make use of hyperbolic spaces to check a condition from Definition 2.2 below.

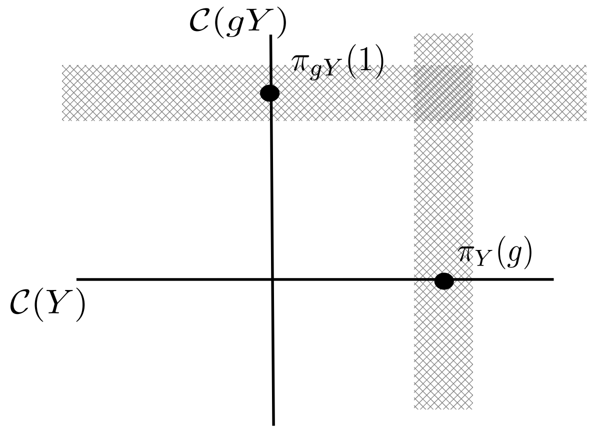

Figure 1. An illustration of the Behrstock inequality (similar to Figure 6 of [2]). Here, $Y_0 \pitchfork gY_0$ and the shaded region indicates the allowable projection of $h \in G$ to $\mathcal{C}(Y_0) \times \mathcal{C}(gY_0)$.

We now give the hypotheses that we will need on the measure. Notice that we did not require the stabilizer of Y_0 to act on $\mathcal{C}(Y_0)$, which is the reason why item 2 below might look strange at first.

Definition 2.2. Fix a projection system $(\mathcal{S}, Y_0, \{\pi_Z\}_{Z\in\mathcal{S}}, \pitchfork)$ on the group G. We call a finitely supported measure μ *semi-admissible* if the random walk (w_n) driven by μ satisfies the following.

1. (1) There are $h_1, h_2 \in G$ in the support of μ such that for every $Z \in \mathcal{S}$ either $h_1 Y_0 \pitchfork Z$ or $h_2 Y_0 \pitchfork Z$.
2. (2) There are $x_1, x_2 \in G$ in the support of μ such that $d_{Y_0}(x_1 h, x_2 h) \geq 2B$ for each $h \in G$ and for B as in Definition 2.1.4.
3. (3) There exists $C_0 \geq 1$ so that $\mathbb{P}\big(w_n \in Stab(Y_0), d_{Y_0}(1, w_n) \geq n/C_0\big)$ is positive for each $n \geq 1$, where $Stab(Y_0)$ denotes the stabilizer of Y_0 in G.
4. (4) There exists $C_0 \geq 1$ so that $\mathbb{P}(w_n Y_0 \not\pitchfork Y_0) \leq C_0 e^{-n/C_0}$ for each $n \geq 1$.

We call a measure μ *admissible* if both μ and the reflected measure $\hat{\mu}(g) = \mu(g^{-1})$ are semi-admissible.

We remark that in our applications, we show that a convolution power of some initial measure is admissible. (See Section 6.) The main theorem of the paper is the following.

Theorem 2.3. *Let G be a group and let $(\mathcal{S}, Y_0, \{\pi_Z\}_{Z\in\mathcal{S}}, \pitchfork)$ be a projection system on G. Let (w_n) be a random walk driven by an admissible measure. Then there exists $C \geq 1$ so that*

$$\mathbb{P}\left(\sup_{Z\in\mathcal{S}} d_Z(1, w_n) \in [C^{-1}\log n, C\log n]\right)$$

goes to 1 as n goes to ∞.

The upper bound and lower bound on random projections are proved separately in Theorem 4.3 and Theorem 5.2 respectively, which together give Theorem 2.3.

3. Exponential decay of projection distance

In this section we fix the projection system $(\mathcal{S}, Y_0, \{\pi_Z\}_{Z\in\mathcal{S}}, \pitchfork)$ on the finitely generated group G endowed with the word metric d_G. Also, we let μ be a semi-admissible measure generating the random walk (w_n).

The following lemma tells us, thinking of g as an intermediate step of the random walk, that there is a definite probability that the projection onto Z of the random walk does not change from some point on.

Lemma 3.1. *There exist $\epsilon, C > 0$ so that for any n the following holds. For any $g \in G$ and $Z \in \mathcal{S}$, we have $\mathbb{P}(d_Z(g, gw_n) \geq C) \leq 1 - \epsilon$.*

The proof exploits a "replacement" trick, where we start with w_n having $d_Z(g, gw_n) \geq C$ and (thinking of w_n as a random path) replace its initial subpath of length 2 to lower the d_Z.

Proof. We first note that it suffice to prove the lemma for $Z = Y_0$. This is because $Z = zY_0$ for some $z \in G$ and $d_Z(g, gw_n) = d_{zY_0}(g, gw_n) = d_{Y_0}(z^{-1}g, z^{-1}gw_n)$.

Fix h_1, h_2 as in Definition 2.2.1, and x_1, x_2 as in Definition 2.2.2. If $n = 1$, the statement of the lemma holds if we choose $C > L\sup_{h\in supp(\mu)} d_G(1, h)$, where L is the Lipschitz constant of π_{Y_0}. For $n \geq 2$ we have

$$\mathbb{P}(d_{Y_0}(g, gw_n) \geq C) = \sum_{h\in G} \mathbb{P}(d_{Y_0}(g, gw_2h) \geq C)\mathbb{P}(w_2^{-1}w_n = h).$$

Fix any $h \in G$. There exists i so that $g^{-1}Y_0 \pitchfork h_iY_0$ and hence $Y_0 \pitchfork gh_iY_0$. Also, since $d_{gh_iY_0}(gh_ix_1h, gh_ix_2h) \geq 2B$, there exists j so that $d_{gh_iY_0}(1, gh_ix_jh) \geq B$. Hence $d_{Y_0}(gh_i, gh_ix_jh) < B$ by Definition 2.1.4, so that $d_{Y_0}(g, gh_ix_jh) < B + Ld_G(1, h_i)$. This proves $\mathbb{P}(d_{Y_0}(g, gw_2h) \geq C) \leq 1 - \epsilon$ if $C > B + L\max\{d_G(1, h_1), d_G(1, h_2)\}$, where we set $\epsilon = \min_{i,j}\{\mu(h_i)\mu(x_j)\} > 0$. To finish up,

$$\mathbb{P}(d_{Y_0}(g, gw_n) \geq C) \leq (1 - \epsilon)\sum_{h\in G} \mathbb{P}(w_2^{-1}w_n = h) = 1 - \epsilon.$$

□

The following key proposition says that it is (exponentially) unlikely that a random walk projects far away on Z. To prove the lower bound on the largest random projection we just need the case $R = 0$, i.e. we do not need the conditional probability.

Proposition 3.2. *There exists $M \geq 1$ so that for each $Z \in \mathcal{S}$, each positive integer n and each $t, R \geq 0$ we have*

$$\mathbb{P}\big(d_Z(1, w_n) \geq t + R \mid d_Z(1, w_n) \geq R\big) \leq Me^{-t/M}.$$

The idea of proof is to show that the probability that the projection to Z is much further away than $s \geq 0$ plus some constant is a bounded multiple of the probability that it is about s. This is because, in view of Lemma 3.1, once an intermediate step of the random walk projects further than s, there is a definite probability that the projection does not change.

Proof. First of all, let us rephrase the statement. Let $f_{Z,n}(s) = \mathbb{P}(d_Z(1, w_n) \geq s)$. We have to find M, not depending on Z, n, so that $f_{Z,n}(t + R) \leq f_{Z,n}(R)Me^{-t/M}$ for each $t, R \geq 0$. For the proof, fix Z, n and set $f(s) = f_{Z,n}(s)$.

We now fix some constants. Let $\epsilon, C > 0$ be as in Lemma 3.1; in particular

$$\mathbb{P}(d_Z(g, gw_m) \geq C) \leq \frac{1-\epsilon}{\epsilon}\,\mathbb{P}(d_Z(g, gw_m) < C)$$

for each $g \in G$ and $m \geq 0$, since $\mathbb{P}(d_Z(g, gw_m) < C) \geq \epsilon$. We increase C to ensure that $Ld_G(1, g) \leq C$ for each $g \in \mathrm{supp}(\mu)$, where L is the Lipschitz constant of π_Z. In particular, any given step of the random walk moves the projection by at most C.

For $g \in G$ and m an integer, denote by $E_{g,m}$ the event where $w_{n-m} = g$ and $n - m = \min\{j : d_Z(1, w_j) \geq s + C\}$. Then

$$\begin{aligned}
& f(s + 3C) \\
\leq & \sum_{\substack{d_Z(1,g)\in[s+C,s+2C]\\ m\leq n}} \mathbb{P}(d_Z(1, gw_m) \geq s + 3C)\mathbb{P}(E_{g,m}) \\
\leq & \sum \mathbb{P}(d_Z(g, gw_m) \geq C)\mathbb{P}(E_{g,m}) \\
\leq & \frac{1-\epsilon}{\epsilon} \sum \mathbb{P}(d_Z(g, gw_m) < C)\mathbb{P}(E_{g,m}) \\
\leq & \frac{1-\epsilon}{\epsilon} \left(f(s) - f(s + 3C)\right),
\end{aligned}$$

where we used $f(s) - f(s + 3C) = \mathbb{P}\big(d_Z(1, w_n) \in [s, s + 3C)\big)$.

Hence, $f(s + 3C) \leq (1 - \epsilon)f(s)$, and in turn we get $f(R + 3Ci) \leq (1 - \epsilon)^i f(R)$ for each integer $i \geq 0$. This implies the required exponential decay of f. □

4. Lower bound on projections via the second moment method

4.1. **Heuristic discussion.** In this section we prove that there exists a projection of at least logarithmic size with high probability. The reason why one expects a logarithmic size projection is that, roughly speaking, a random word of length n contains all subwords of length $\epsilon \log n$, and in particular it will contain a subword that creates a logarithmic size projection. The remaining parts of the word should not affect this projection too much in view of Proposition 3.2. This heuristic alone, however, only gives that the expected number of logarithmic size projections diverges, but it does not say anything about the probability that one exists.

The actual proof relies on the second moment method, i.e. the estimate that for a random variable $X \geq 0$ with finite variance and $\mathbb{E}(X) > 0$, we have

$$\mathbb{P}(X > 0) \geq \frac{\mathbb{E}(X)^2}{\mathbb{E}(X^2)}.$$

The second moment method is especially suited for dealing with random variables X that can be written as $\sum_{i=1}^n Y_i$, where "most pairs" Y_i, Y_j are "mostly uncorrelated", meaning that $\mathbb{E}(Y_iY_j)$ is approximately $\mathbb{E}(Y_i)\mathbb{E}(Y_j)$. In this case the numerator $\sum_{i,j} \mathbb{E}(Y_i)\mathbb{E}(Y_j)$ is approximately equal to the denominator $\sum_{i,j} \mathbb{E}(Y_iY_j)$, implying that $\mathbb{P}(X > 0)$ is close to 1.

4.2. **The proof.** We use the notation of Theorem 2.3 as well as Definitions 2.1 and 2.2. In this section we prove that there exists $\epsilon_0 > 0$ so that

$$\mathbb{P}\left(\sup_{Z\in\mathcal{S}} d_Z(1,w_n) \geq \epsilon_0 \log n\right) \to 1, \quad (*)$$

as n tends to ∞.

Let c be the minimal probability of an element in $\mathrm{supp}(\mu)$. Then for each k and each $g \in G$ we have that $\mathbb{P}(w_k = g)$ is either 0 or at least c^k. We fix a positive $\epsilon_1 < \frac{1}{\log(1/c)}$ from now until the end of the section and set $k(n) = \lfloor \epsilon_1 \log n \rfloor$. To simplify the notation, we fix n and set $k = k(n)$, and stipulate that all constants appearing below do not depend on n.

In view of the discussion above and Definition 2.2.3, there exists $\epsilon_2 \in (0,1)$ so that for every sufficiently large n we can choose $x_n \in G$ with the properties that

(1) $p_n := \mathbb{P}(w_k = x_n) \geq n^{\epsilon_1 \log(c)}$,
(2) $x_n Y_0 = Y_0$, and
(3) $d_Y(1, x_n) \geq \epsilon_2 \log(n)$.

For $i \leq n-k$, let W_i be the indicator function of the event $w_i^{-1} w_{i+k} = x_n$. Also, let L_i be the indicator function for the event that $d_{Y_0}(w_i^{-1}, 1) \leq \epsilon_2 \log(n)/3$ and let R_i be the indicator function for the event that $d_{Y_0}(1, w_{i+k}^{-1} w_n) \leq \epsilon_2 \log(n)/3$.

Set $Y_i = L_i W_i R_i$ and note that if $Y_i = 1$, then

$$\begin{aligned} d_{w_iY_0}(1,w_n) &\geq d_{w_iY_0}(w_i, w_i x_n) - d_{w_iY_0}(1,w_i) - d_{w_i x_n Y_0}(w_i x_n, w_n) \\ &\geq d_{Y_0}(1,x_n) - d_{Y_0}(w_i^{-1},1) - d_{Y_0}(1, w_{i+k}^{-1} w_n) \\ &\geq \epsilon_2 \log(n) - 2\frac{\epsilon_2}{3}\log(n) \\ &= \frac{\epsilon_2}{3}\log(n). \quad (**) \end{aligned}$$

Hence, what we want to show is that with high probability there exists i with $Y_i = 1$.

Lemma 4.1. $\mathbb{E}(Y_i) = p_n(1 - O(n^{-\epsilon_3}))$ *for each* i*, where* $\epsilon_3 > 0$.

Proof. Since $Y_i = L_i W_i R_i$ and L_i, W_i, R_i are independent, it suffices to show that $\mathbb{P}(L_i = 0)$ and $\mathbb{P}(R_i = 0)$ decay polynomially as $n \to \infty$. Indeed, by Proposition 3.2,

$$\begin{aligned} \mathbb{P}\left(d_{Y_0}(w_i^{-1},1) > \frac{\epsilon_2}{3}\log(n)\right) &\leq M e^{-\frac{\epsilon_2}{3M}\log(n)} \\ &\leq M n^{-\frac{\epsilon_2}{3M}}. \end{aligned}$$

as required. The case for $\mathbb{P}(R_i = 0)$ is similar. □

The following proposition is the key one to apply the second moment method.

Proposition 4.2. $\mathbb{E}(Y_i Y_j) = p_n^2(1 - O(n^{-\epsilon_4}))$ *whenever* $|i-j| \geq \log n$*, where* $\epsilon_4 > 0$.

The idea of proof is the following. We have to prove that if in two specified spots along the random path we see the word x_n, then it is very likely that $Y_i = Y_j = 1$, i.e. that certain projections are not too big. In order to show this, we consider the 3 remaining subpaths of the random path. See Figure 2. Such paths and their inverses give small projection to Y by Proposition 3.2. But then it is easy to control all projections we are interested in using the Behrstock inequality.

Proof. Fix $i \le j - \log n$. Let $A_1 = \{W_i = 1\}$, $A_2 = \{W_j = 1\}$, and A_3 be the event that either $Y_0 \not\pitchfork w_{i+k}^{-1} w_j Y_0$ or one of the following distances is larger than $\frac{\epsilon_2}{3}\log(n) - B$:

(1) $d_{Y_0}(1, w_i^{-1})$,
(2) $d_{Y_0}(1, w_{i+k}^{-1} w_j)$,
(3) $d_{Y_0}(1, w_j^{-1} w_{i+k})$,
(4) $d_{Y_0}(1, w_{j+k}^{-1} w_n)$.

Notice that A_1, A_2, A_3 are independent. We claim that $W_i W_j 1_{A_3^c} \le Y_i Y_j \le W_i W_j$. Once we establish the claim, we have that $p_n^2(1 - \mathbb{P}(A_3)) \le \mathbb{E}(Y_i Y_j) \le p_n^2$. Since the probability of each of the 5 events making up A_3 decays polynomial in n, the first by the admissibility condition Definition 2.2.4 and the last four by Proposition 3.2, this will complete the proof of the proposition.

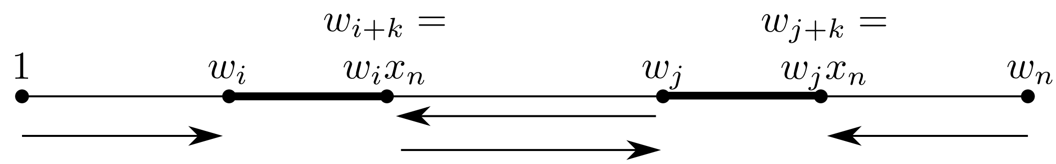

Figure 2. The line represents the random path under consideration, with the thicker segments being the occurrences of x_n. The arrows suggest the paths that have small enough projection onto Y_0 in the event A_3^c (which is often).

To prove the claim, we show that $A_1 \cap A_2 \cap A_3^c \subset \{Y_i = 1\} \cap \{Y_j = 1\}$, i.e. that if

- $w_i^{-1} w_{i+k} = x_n$, $w_j^{-1} w_{j+k} = x_n$,
- $Y_0 \pitchfork w_{i+k}^{-1} w_j Y_0$, and
- all distances in $(1) - (4)$ listed above are at most $\frac{\epsilon_2}{3}\log(n) - B$

then for l equal to i or j, $d_{Y_0}(w_l^{-1}, 1) \le \frac{\epsilon_2}{3}\log(n)$ and $d_{Y_0}(1, w_{l+k}^{-1} w_n) \le \frac{\epsilon_2}{3}\log(n)$. We show this for $l = i$ since the other case is similar.

First, note that by assumption $w_{i+k} = w_i x_n$ and $w_{j+k} = w_j x_n$. Hence,

$$\begin{aligned} d_{Y_0}(1, w_{i+k}^{-1} w_n) &= d_{w_i Y_0}(w_i x_n, w_n) \\ &\le d_{w_i Y_0}(w_i x_n, w_j) + d_{w_i Y_0}(w_j, w_n) \\ &\le \frac{\epsilon_2}{3}\log(n) - B + d_{w_i Y_0}(w_j, w_n), \end{aligned}$$

and so it suffices to show that $d_{w_i Y_0}(w_j, w_n) \le B$. If not, then since $w_i x_n Y_0 = w_i Y_0 \pitchfork w_j Y_0$, we must have that $d_{w_j Y_0}(w_i x_n, w_n) \le B$. On the other hand, by the triangle inequality,

$$\begin{aligned} d_{w_j Y_0}(w_i x_n, w_n) &\ge d_{w_j Y_0}(w_j, w_j x_n) - d_{w_j Y_0}(w_i x_n, w_j) - d_{w_j x_n Y_0}(w_j x_n, w_n) \\ &= d_{Y_0}(1, x_n) - d_{Y_0}(1, w_{i+k}^{-1} w_j) - d_{Y_0}(1, w_{j+k}^{-1} w_n) \\ &\ge \epsilon_2 \log n - 2\frac{\epsilon_2}{3}\log n + 2B, \end{aligned}$$

a contradiction. □

We are ready to prove the lower bound $(*)$.

Theorem 4.3. *Let $(\mathcal{S}, Y, \{\pi_Z\}_{Z\in\mathcal{S}}, \pitchfork)$ be a projection system on the finitely generated group G and let μ be a an admissible measure generating the random walk (w_n). Then there exists $\epsilon_0 > 0$ so that*

$$\mathbb{P}\left(\sup_{Z\in\mathcal{S}} d_Z(1, w_n) \geq \epsilon_0 \log(n)\right) \to 1,$$

as $n \to \infty$.

Proof. Let Y_i be as above and set $X = X_n = \sum_{i=1}^{n-k} Y_i$. We show that $\mathbb{P}(X > 0)$ approaches 1 as $n \to \infty$, which suffices in view of the estimate $(**)$.

Since $\mathbb{E}(Y_iY_j) \leq \mathbb{E}(Y_i) \leq p_n(1 - O(n^{-\epsilon_3}))$ by Lemma 4.1, Proposition 4.2 implies that

$$\begin{aligned}\sum_{i,j} \mathbb{E}(Y_iY_j) &= \sum_{|i-j|<\log n} \mathbb{E}(Y_iY_j) + \sum_{|i-j|\geq\log n} \mathbb{E}(Y_iY_j) \\ &\leq 3n\log(n)p_n(1 - O(n^{-\epsilon_3})) + n^2p_n^2(1 - O(n^{-\epsilon_4})) \\ &\leq n^2p_n^2(1 - o(1)),\end{aligned}$$

where we used that $n\log(n)p_n/n^2p_n^2 = \log(n)/np_n$ tends to 0. This holds since $np_n \geq n^{1+\epsilon_1\log(c)}$, which grows polynomially since $\epsilon_1 < \frac{1}{\log(1/c)}$.

By the second moment method (and Lemma 4.1), we have

$$\begin{aligned}\mathbb{P}(X > 0) &\geq \frac{\mathbb{E}(X)^2}{\mathbb{E}(X^2)} = \frac{\sum_{i,j} \mathbb{E}(Y_i)\mathbb{E}(Y_j)}{\sum_{i,j} \mathbb{E}(Y_iY_j)} \\ &\geq \frac{(n-k)^2p_n^2(1 - o(1))}{n^2p_n^2(1 - o(1))} = 1 - o(1),\end{aligned}$$

as required. □

5. Upper bound on projections via distance formula lower bound

We start with a proposition that provides a distance-formula-type lower bound on the distance in G between two elements.

This will be useful for us because in order to show that with high probability there is a logarithmic upper bound on $\sup_Z d_Z(1, w_n)$ the idea is the following. For each given Z we have the required upper bound in view of Proposition 3.2, and in view of the following proposition we only need to check a controlled number of Z.

Proposition 5.1. *Let $(\mathcal{S}, Y_0, \{\pi_Z\}_{Z\in\mathcal{S}}, \pitchfork)$ be a projection system on the finitely generated group G endowed with the word metric d_G. Then there are $K, C \geq 0$ so that for all $h \in G$ we have*

$$\sum_{Z\in\mathcal{S}} \{\{d_Z(1, h)\}\}_K \leq C \cdot d_G(1, h),$$

where $\{\{x\}\}_K = x$ if $x \geq K$ and 0 otherwise.

In the setting of the mapping class groups, the metric spaces $\mathcal{C}(Z)$ are curve complexes of subsurfaces and Proposition 5.1 follows from the Masur–Minsky distance formulas [18]. However, the proof we give does not rely on the Masur–Minsky hierarchy machinery and applied to our general notion of a projection system. Our proof is a generalization of the argument appearing in [1].

Proof. We use the notation and constants in Definitions 2.1.

Let $K = 5B + 3L$ and fix a geodesic $1 = h_0, h_1, \ldots, h_N = h$ in G. Let Ω be the set of $Z \in \mathcal{S}$ with $d_Z(1, h) \geq K$. For each $Z \in \Omega$ choose $i_Z, t_Z \in \{0, \ldots, N\}$ as follows: i_Z is the largest index k with $d_Z(h_0, h_k) \leq 2B + L$ and t_Z is the smallest index k greater than i_Z with $d_Z(h_k, h_N) \leq 2B + L$. Write $I_Z = [i_Z, t_Z] \subset \{0, 1, \ldots, N\}$ and note that this subinterval is well-defined. Further since the projection of adjacent vertices in the geodesic have d_Z less than or equal to L, $d_Z(h_0, h_k), d_Z(h_k, h_N) \geq 2B$ for all $k \in I_Z$ and $d_Z(h_{i_Z}, h_{t_Z}) \geq B + L$.

Claim. If $aY_0, bY_0 \in \Omega$ and $aY_0 \pitchfork bY_0$ then $I_{aY_0} \cap I_{bY_0} = \emptyset$.

Proof. Toward a contradiction, take $k \in I_{aY_0} \cap I_{bY_0}$. Since $aY_0 \pitchfork bY_0$ we have either $d_{aY_0}(b, h_0) < B$ or $d_{bY_0}(a, h_0) < B$. Assume the former; the latter case is proven by exchanging the occurrences of a and b in the proof. By the triangle inequality,

$$\begin{aligned} d_{aY_0}(b, h_k) &\geq d_{aY_0}(h_0, h_k) - d_{aY_0}(b, h_0) \\ &\geq 2B - B = B. \end{aligned}$$

So, since $aY_0 \pitchfork bY_0$, we have $d_{bY_0}(a, h_k) < B$ and

$$\begin{aligned} d_{bY_0}(a, h_N) &\geq d_{bY_0}(h_k, h_N) - d_{bY_0}(a, h_k) \\ &\geq 2B - B = B. \end{aligned}$$

and we conclude, again since $aY_0 \pitchfork bY_0$, that $d_{aY_0}(b, h_N) \leq B$. This, together with our initial assumption, implies

$$d_{aY_0}(h_0, h_N) \leq d_{aY_0}(h_0, b) + d_{aY_0}(b, h_N) \leq 2B < K$$

contradicting that $aY_0 \in \Omega$. □

Returning to the proof of the proposition, we have the intervals $\{I_Z : Z \in \Omega\}$ of $\{0, 1, \ldots, N\}$. By the claim above and our assumption on the number of pair-wise $\pitchfork$-incomparable elements of $\mathcal{S}$, each $k \in \{0, 1, \ldots, N\}$ is contained in at most s intervals. Hence,

$$\sum_{Z \in \Omega} |t_Z - i_Z| \leq s \cdot d_G(1, h).$$

Finally, using the Lipschitz condition on the projections,

$$\begin{aligned} d_Z(1, h) &\leq d_Z(h_{i_Z}, h_{t_Z}) + 4B + 2L \\ &\leq L|t_Z - i_Z| + 4B + 2L. \end{aligned}$$

Since, we have $d_Z(1, h) \geq 5B + 3L$ for each $Z \in \Omega$, we get $\frac{1}{5L} \cdot d_Z(1, h) \leq |t_Z - i_Z|$ and so putting this with the inequality above

$$\sum_{Z \in \Omega} d_Z(1, h) \leq 5sL \cdot d_G(1, h)$$

as required. □

Theorem 5.2. *Let $(\mathcal{S}, Y_0, \{\pi_Z\}_{Z \in \mathcal{S}}, \pitchfork)$ be a projection system on the finitely generated group G and let μ be a semi-admissible measure generating the random walk (w_n). Then there exists $A \geq 1$ so that*

$$\mathbb{P}(\exists Z \in \mathcal{S} \ : \ d_Z(1, w_n) \geq A \log n)$$

tends to 0 as n tends to ∞.

Proof. Let C, K be as in Proposition 5.1, let M be as in Proposition 3.2, and set $l = \sup_{h \in supp(\mu)} d_G(1,h)$. Then we notice that $\sum_{Z \in \mathcal{S}} \mathbb{P}(d_Z(1,w_n) \geq K) \leq Cl \cdot n$ for each $n \geq 0$. In fact $\mathbb{P}(d_Z(1,w_n) \geq K)$ is the expected value of the indicator function of the event $A_Z = \{d_Z(1,w_n) \geq K\}$, so that the aforementioned sum equals the expected value of the random variable $|\{Z \in \mathcal{S} : d_Z(1,w_n) \geq K\}|$. By Proposition 5.1 we have $|\{Z \in \mathcal{S} : d_Z(1,g) \geq K\}| \leq C d_G(1,g)$ for each $g \in G$, hence the estimate follows.

Choose any $A > M$. For n large enough that $A \log n \geq K$ we have

$$\begin{aligned}
&\mathbb{P}\big(\exists Z \in \mathcal{S} \ : d_Z(1,w_n) \geq A \log n\big) \\
&\qquad \leq \sum_{Z \in \mathcal{S}} \mathbb{P}\big(d_Z(1,w_n) \geq A \log n \mid d_Z(1,w_n) \geq K\big) \mathbb{P}(d_Z(1,w_n) \geq K) \\
&\qquad \leq M e^{-(A \log n - K)/M} \sum \mathbb{P}\big(d_Z(1,w_n) \geq K\big) \\
&\qquad \leq (MCle^{K/M}) n^{1-A/M},
\end{aligned}$$

where the second inequality follows from Proposition 3.2. Since this quantity tends to 0 by our choice of A, the proof is complete. □

6. Proof of Theorem 1.1

Here we show how Theorem 1.1 follows from Theorem 2.3. We give full details in the case of mapping class groups; the other cases follow a very similar outline and we provide slightly fewer details and the needed references.

6.1. **Mapping class groups and subsurface projections.** In this subsection, we assume that the reader has some familiarity with subsurface projections, as defined by Masur–Minsky in [18]. Set $G = \mathrm{Mod}(S)$. Since there are only finitely many G–orbits of (isotopy classes of) essential subsurfaces, we can deal with each orbit separately. For any proper essential subsurface Y_0 of S, let $\mathcal{S} = \{gY_0 : g \in G\}$, and for $Z \in \mathcal{S}$, let π_Z denote the subsurface projection to the curve complex $\mathcal{C}(Z)$ of Z. If we complete ∂Y_0 to a marking λ (a collection of curves cutting S into disks and once-punctured disks), then we define $\pi_Z \colon G \to \mathcal{C}(Z)$ by $\pi_Z(g) = \pi_Z(g\lambda)$. With this definition, π_Z is L–Lipschitz for some $L > 1$ ([18, Lemma 2.5]). This verifies the first 3 conditions in the definition of a projection system (Definition 2.1).

Subsurfaces Y and Z overlap if, up to isotopy, they are neither disjoint nor nested. In this case, we write $Y \pitchfork Z$. Note that by construction, if Y and Z overlap, then $d_Z(\lambda, Y) \leq L$. Here, $\pi_Z(Y)$ is by definition the projection of the boundary of Y to $\mathcal{C}(Z)$. The usual Behrstock inequality [2] states that there exists $D \geq 0$ such that for any marking η and overlapping subsurfaces Y and Z, $\min\{d_Y(\partial Z, \eta), d_Z(\partial Y, \eta)\} \leq D$. Hence, setting $B = D + L$ verifies the fourth condition of Definition 2.1. To show that subsurface projections give a projection system for the mapping class group, it remains to show the there is a bound on the size of a collection of pairwise nonoverlapping subsurfaces of S. This fact is easily verified since such subsurfaces can be realized simultaneously as either disjoint or nested on S (see the proof of Theorem 6.10 in [18]).

Let μ be a symmetric probability measure on $\mathrm{Mod}(S)$ whose support is finite and generates $\mathrm{Mod}(S)$. We show that there is a k such that $\overline{\mu} = \mu^{*k}$ is admissible. Here, μ^{*k} is the k-fold convolution power of μ, which by definition is

$$\mu^{*k}(g) = \sum_{x_1 \dots x_k = g} \mu(x_1) \cdots \mu(x_k).$$

From this, Theorem 1.1 for $G = \mathrm{Mod}(S)$ follows immediately from Theorem 2.3 (and the fact that the projections are L–Lipschitz). Let Y_0 and λ be as above and fix $f \in \mathrm{Mod}(S)$ which acts as a pseudo-Anosov on Y_0 such that f^2 has translation length at least $2B$ in $\mathcal{C}(Y_0)$. Further, fix a pseudo-Anosov $g \in \mathrm{Mod}(S)$ whose translation length on $\mathcal{C}(S)$ is at least $3L+1$. Choose k so that, up to replacing f and g with appropriate powers, f and g are in the support of $\overline{\mu} = \mu^{*k}$. Note we also have that $f^{-1}, g^{-1} \in \mathrm{supp}(\overline{\mu})$.

Now the first condition of Definition 2.2 hold for $h_1 = g^{-1}, h_2 = g$ since if both gY_0 and $g^{-1}Y_0$ fail to overlap some surface Z and γ is a boundary component of Y_0, then

$$d_S(\gamma, g^2\gamma) = d_S(g^{-1}\gamma, g\gamma) \leq d_S(g^{-1}\gamma, \partial Z) + d_S(\partial Z, g\gamma) \leq 3L,$$

contradicting our choice of g. The second condition of Definition 2.2 is satisfied for $x_1 = f, x_2 = f^{-1}$, while the third condition holds since the support of $\overline{\mu}^{*n}$ contains f^n. The final condition of Definition 2.2 is easily deduced from the fact that, by [15, Theorem 1.2], the random walk makes linear progress with exponential decay in $\mathcal{C}(S)$.

6.2. **Quasiconvex subgroups.** Consider the setup of a quasiconvex subgroup H of a hyperbolic group G, with $\mathcal{S}$ the family of its cosets. The first 3 properties in Definition 2.1 are easy. The relation $\pitchfork$ is having bounded projection onto each other, so that the fourth item follows from basic hyperbolic geometry. The final item follows from finiteness of width [12].

As in the case of mapping class groups, the proof of admissibility uses a hyperbolic space, X, that we now define. We let S be any finite generating set of G, and consider the Cayley graph X of G with respect to the infinite generating set $S \cup H$.

We could not find a reference for the following statement, despite it being implicit in several papers. Recall that an action on a hyperbolic space is non-elementary if there exist two loxodromic elements with no common limit point at infinity.

Lemma 6.1. *X is hyperbolic and the action of G on X is non-elementary.*

Proof. By [11, Theorem 6.4], X is hyperbolic. In fact, by [13, Proposition 2.6] quasigeodesics in G map to unparameterized quasigeodesics in X. In particular, if g is any element of G, then the action of g on X is either elliptic or loxodromic. The combination of [20, Theorem 1-(b)] (which provides an element $x \in G$ not conjugate into H) and [20, Theorem 2] (for $K = \langle x \rangle$) prove that there exists an element of G, that we denote g, that cannot act elliptically, and hence it acts loxodromically. Similarly, we can apply the same reasoning to find an element g' that acts loxodromically on the Cayley graph of G with respect to $S \cup H \cup \langle g \rangle$, and in particular it will also act loxodromically on X. It is easy to see that g and g' cannot have a common limit point at infinity. □

A convolution power of μ will have support containing an infinite order element of the quasiconvex subgroup, as well as an element with large translation distance on X, easily implying the first three items of Definition 2.2.

We can now apply the linear progress result in [17] to get that the random walk we are considering makes linear progress with exponential decay in X, easily implying the final condition of Definition 2.2 (since if two cosets are far away in X then they are, in particular, far away in G and hence they have bounded projection onto each other).

6.3. **Peripheral subgroups.** Consider the setup of a peripheral subgroup with an undistorted element of a relatively hyperbolic group, with $\mathcal{S}$ the family of its cosets. The relation $\pitchfork$ in this case is just being distinct, and the needed properties of projections follow from, e.g., [28].

The last item once again uses a hyperbolic space. In this case, the hyperbolic space X is a coned-off graph: If S is a finite generating set for G, then X is the Cayley graph with respect to the generating set $S \cup H_1 \cup \cdots \cup H_n$. The hyperbolicity of X is part of the definition of relative hyperbolicity from [10], and the action is non-elementary due to results in [23]. More precisely, [23, Lemma 4.5] gives a loxodromic element g for the action. Moreover, g is contained in an elementary subgroup $E(g)$ that can be added to the list of peripheral subgroups [23, Corollary 1.7], so that one can find a loxodromic element g' with respect to the new list of peripherals. Similarly to the hyperbolic group case, it is easy to see that g, g' are the required pair of loxodromic elements.

Admissibility now follows similarly to the other cases, using a sufficiently large power of the undistorted element of H (any undistorted element of a group H has powers with arbitrarily large translation distance on the Cayley graph of H).

6.4. $\mathrm{Out}(F_n)$ **and subfactor projections.** Subfactor projections were introduced by Bestvina–Feighn in [5] and refined in [31]. We refer to these references for definitions and details.

For a rank ≥ 2 free factor Y_0 of F_n, let π_{Y_0} denote the subfactor projection to $\mathcal{C}(Y_0)$, the free factor graph of Y_0. Set $G = \mathrm{Out}(F_n)$ and let $\mathcal{S} = \{gY_0 : g \in G\}$. Finally, let λ be a F_n–marked graph, i.e. a graph with a fixed isomorphism $F_n \to \pi_1(\lambda)$, containing a subgraph λ_{Y_0} with $\pi_1(\lambda_{Y_0}) = Y_0$.

For any $Z \in \mathcal{S}$, define $\pi_Z \colon G \to \mathcal{C}(Z)$ by $\pi_Z(g) = \pi_Z(g\lambda)$. Here, $g\lambda$ denotes the image of λ under $g \in G$ with respect to the natural left action of G on the set of marked graphs. Free factors X and Z of F_n are said to be disjoint if, up to conjugation, $F_n = W * X * Z$ for some (possibly trivial) free factor W. The factors X and Z overlap, written $X \pitchfork Z$, if they are neither disjoint nor nested. By [31, Theorem 1.1], when free factors X and Z overlap, there is a well-defined coarse projection $\pi_Z(X) \subset \mathcal{C}(Z)$ and that the natural version of the Behrstock inequality holds (see also [5, Proposition 4.18]). This, together with the fact that $d_Z(Y_0, \lambda)$ is bounded for all $Z \in \mathcal{S}$ with $Y_0 \pitchfork Z$, implies the fourth condition of Definition 2.1. Finally, condition (5) follows, for example, from [5, Lemma 4.14]. Hence, subfactor projections form a projection system.

From this, it follows just as in the situation of $\mathrm{Mod}(S)$ that if μ is a symmetric probability measure on $\mathrm{Out}(F_n)$ whose support is finite and generates $\mathrm{Out}(F_n)$, then $\overline{\mu} = \mu^{*k}$ is admissible for some $k \geq 1$. The only needed modifications are that one chooses g to be fully irreducible with large translation length on the free factor complex of F_n, chooses f to fix Y_0 and have large translation length on $\mathcal{C}(Y_0)$, and applies the general linear progress result of [17, Theorem 1.2].

7. Shortest curves in random mapping tori

Let S be a closed connected orientable surface of genus at least 2. Hereafter, we assume that the reader is familiar with the subsurface projection machinery of [18] and refer them to the terminology established in Section 6.1. Throughout this section, we fix a symmetric measure μ whose support is finite and generates $\mathrm{Mod}(S)$, and let (w_n) denote the random walk driven by μ. In this setting, we have that for almost every sample path (w_n), w_n is pseudo-Anosov for sufficiently large $n \geq 0$ [15]. We will use this fact freely without further comment.

For a pseudo-Anosov $f \in \mathrm{Mod}(S)$, let $\lambda^+(f)$ and $\lambda^-(f)$ denote its stable and unstable laminations. For a subsurface Y of S, define

$$d_Y(f) = d_Y(\lambda^+(f), \lambda^-(f)),$$

to be the distance in the curve graph of Y between the projections of the stable and unstable laminations of f. Note that by invariance of the laminations, $d_{f^iY}(f) = d_Y(f)$ for all $i \in \mathbb{Z}$. See [18, 21] for details. When Y is an annulus about the curve α, we use the notation d_α rather than d_Y. For a curve α in S, let $\mathcal{Y}_\alpha$ be the (nonannular) subsurfaces with α as a boundary component. Finally, in the following proposition we will need to make use of the *bounded geodesic image theorem* of [18]. This states that there is a constant $M > 0$ such that if γ is a geodesic in the curve graph $\mathcal{C}(S)$ that does not meet the 1–neighborhood of ∂X for some subsurface X of S, then the diameter of the projection of γ to $\mathcal{C}(X)$ is at most M.

When writing expressions such as $\sum_{Z \subset X}$, for X a subsurface, we mean that we are summing over all isotopy classes of (essential) subsurfaces of X. Recall that the notation $\{\{x\}\}_L$ denotes x if $x \geq L$ and 0 otherwise. Using Theorem 1.1 for mapping class groups, we can prove the following:

Proposition 7.1. *For the random walk* (w_n) *on* $\mathrm{Mod}(S)$ *as above, there are* $K_2, C_2 \geq 1$ *such that*

(1) $\mathbb{P}\left(\sup_{X \subsetneq S} \sum_{Z \subset X} \{\{d_Z(w_n)\}\}_{K_2} \leq C_2 \cdot \log(n)\right) \to 1$, *as* $n \to \infty$.

(2) $\mathbb{P}\left(\exists \text{ a curve } \alpha_n \text{ in } S : d_{\alpha_n}(w_n) \geq C_2^{-1} \cdot \log(n) \text{ and } \sup_{X \in \mathcal{Y}_{\alpha_n}} d_X(w_n) \leq K_2\right) \to 1$, *as* $n \to \infty$.

Proof. The main idea of the argument is that the quantities $d_Y(w_n)$ are typically closely related to the distances $d_Y(1, w_n)$, as studied in the first part of the paper. Hence, we first prove the corresponding statements for $d_Y(1, w_n)$, and then we will translate them into statements about $d_Y(w_n)$. Just as in Section 6.1, it suffices to assume that the measure μ is admissible.

Item 1 for the $d_Y(1, w_n)$ is [30, Lemma 5.12].

Item 2 for the $d_Y(1, w_n)$ is a slight variation of Theorem 4.3, and can be shown by modifying the proof as follows. We fix Y to be an annular subsurface with core curve α. For any subsurface $Z \subset S$ (including the case $Z = S$), $\pi_Z(g) \subset \mathcal{C}(Z)$ is the subsurface projection of $g\lambda$, where λ is a fixed marking chosen to contain α. As in Section 6.1, we have $\mathrm{diam}(\pi_Z(g)) \leq L$ for all $Z \subseteq S$. Recall that the first step we took in Section 4 is to choose some $x_n \in G = \mathrm{Mod}(S)$ whose most important property is that $d_Y(1, x_n) \geq \epsilon_2 \log(n)$. In this case, we want 3 large projections instead of 1, and 2 of them will serve as "buffer" for the "middle" one. We now fix once and for all $g \in \mathrm{Mod}(S)$ so that $d_{\mathcal{C}(S)}(x, gx) \geq 10$ for all $x \in \mathcal{C}(S)$ (in particular $g^{\pm 1}Y' \pitchfork Y$ for $Y' = Y$ or Y' disjoint from Y), and choose $x_n = s_n g t_n g u_n$ so that

(1) $p_n := \mathbb{P}(w_k = x_n) \geq n^{\epsilon_1 \log c}$,
(2) s_n, t_n, u_n are powers of Dehn twists around α,
(3) $d_Y(1, s_n), d_Y(1, t_n), d_Y(1, u_n) \geq \epsilon_2 \log n$, for some fixed $\epsilon_2 > 0$.

We then define W_i, L_i, R_i, Y_i as in Subsection 4.2. Similarly to the claim below Figure 2, one can show that if $Y_i = 1$ and n is sufficiently large, then $d_Z(1, w_n)$ is log–large whenever Z is one of the annuli $w_iY, w_i'Y = w_i s_n g Y$, or $w_i s_n g t_n g Y$, and that, moreover, all projection distances to subsurfaces X contained in the complement of $w_i'Y$ are uniformly bounded.

In fact, we are in the situation where w_n can be written as a product of group elements $g_1 g_2 \dots g_7$ (where $g_1 = w_i$, $g_2 = s_n$, $g_3 = g$, $g_4 = t_n$, $g_5 = g$, $g_6 = u_n$) so that for k odd we have that $g_k Y \pitchfork Y$ and g_k has controlled projection on Y, while for k even we have that g_k is a large power of a Dehn twist supported of Y. Just as before, in this situation the projections created by the even g_k "persist." Similarly, if X is contained in the complement of $w_i' Y = g_1 g_2 g_3 Y$, then, first of all, $X \pitchfork g_1 Y$ and $X \pitchfork g_1 \dots g_5 Y$ because of our hypothesis on g. Then, by the triangle inequality,

$$d_X(1, w_n) \leq d_X(1, g_1) + d_X(g_1, g_1 \dots g_5) + d_X(g_1 \dots g_5, w_n) + 2L,$$

where the middle term is at most $d_{X'}(g^{-1}, g) + 2L$, for X' a subsurface contained in the complement of Y. Indeed, since X is disjoint from $g_1 g_2 g_3 Y$, $X' = (g_1 g_2 g_3)^{-1} X$ is disjoint from Y. Hence, $d_X(g_1, g_1 \dots g_5) = d_{X'}(g_3^{-1} g_2^{-1}, g_4 g_5) = d_{X'}(g^{-1}, g)$, where we have used that g_2 and g_4 are twists about α and that $\alpha \subset \lambda$, so that $\mathrm{diam}(\pi_{X'}(\lambda) \cup \pi_{X'}(g_2^{-1}\lambda)) \leq 2L$ and $\pi_{X'}(g_4 g_5) = g_4 \pi_{X'}(g_5) = \pi_{X'}(g_5)$. Since this middle term is uniformly bounded, we just need to bound the other two terms. We will focus on the first term, as the last term is handled similarly. Now if $d_X(1, g_1)$ is larger than $B + L$ (where B is from the Behrstock inequality stated in Section 6.1) then $d_X(1, g_1 \partial Y)$ is larger than B and so $d_{g_1 Y}(1, \partial X) \leq B$. But ∂X and $w_i' \partial Y$ are disjoint and so the Lipschitz property of subsurface projections implies that $d_{g_1 Y}(1, w_i') \leq B + L$, contradicting that the log–large projection to $g_1 Y$ persists.

With our current setup, Lemma 4.1 holds as stated, for the same reason that W_i, L_i, R_i are independent, and Proposition 4.2 also holds with a similar proof based on alternating small and large projections as above. These are all the needed ingredients to conclude the proof that, given a large n, with high probability there exists i with $Y_i = 1$, as we did for Theorem 4.3.

We now translate the statement for the $d_Y(1, w_n)$ into a statement for the $d_Y(w_n)$. We being by showing that, with probability going to 1, for any proper subsurface X of S, we have $d_X(w_n) \leq 3 \sup_i d_{w_n^i X}(1, w_n) + 2M + 2L$, were M is the constant from the bounded geodesic image theorem. This in particular will prove item 1.

We regard w_n as a product of two shorter random walks u_n, v_n, of length approximately $n/2$ each. As explained in, e.g., [16, Lemma 23], the following is a consequence of results in [14] and [19]. With probability going to 1 as n goes to infinity,

- w_n acts hyperbolically on $\mathcal{C}(S)$ with (quasi-)axis A_n,
- writing $\gamma = [\pi_S(1), \pi_S(w_n)]$, $\gamma' = [\pi_S(1), \pi_S(u_n)]$, and $\gamma'' = u_n \cdot [\pi_S(1), \pi_S(v_n)]$, the axis A_n has Hausdorff distance $O(\log n)$ from both

$$\bigcup w_n^k \cdot \gamma \quad \text{and} \quad \bigcup w_n^k \cdot (\gamma' \cup \gamma''),$$

  where with a slight abuse of notation we regarded the various $\pi_S(\cdot)$ as vertices of $\mathcal{C}(S)$,
- and each of γ, γ', and γ'' have length at least $\epsilon' n$ for some uniform $\epsilon' > 0$.

Going forward, we will assume that w_n has this form.

Now if $d_X(w_n) \geq 2M$, then the bounded geodesic image theorem implies that ∂X lies within a uniformly bounded distance from the axis A_n. Hence, there is an $i \in \mathbb{Z}$ such that the boundary of $X_i = w_n^i X$ has bounded distance from γ, and we see that

$$\begin{aligned} d_X(w_n) &= d_{X_i}(w_n) \\ &\le d_{X_i}(\lambda^-(w_n), w_n^{-1}) + d_{X_i}(w_n^{-1}, w_n^2) + d_{X_i}(w_n^2, \lambda^+(w_n)) + 2L \\ &\le M + 3\sup_i d_{X_i}(1, w_n) + M + 2L, \end{aligned}$$

where the last inequality follows from the triangle inequality and another applications of the bounded geodesic image theorem. In a bit more detail, with our setup, the geodesics from w_n^{-1} to $\lambda^-(w_n)$ and from w_n^2 to $\lambda^+(w_n)$ have $\mathcal{C}(S)$–distance from ∂X_i growing linearly in n since the length of γ is greater than $\epsilon' n$ and $d_{\mathcal{C}(S)}(w_n^i, A_n) = O(\log(n))$ for $i \in \mathbb{Z}$. Hence, the bounded geodesic image theorem implies that, with probability going to 1, their images in $\mathcal{C}(X_i)$ have diameter at most M.

Let us now prove item 2. Using what we have already shown applied to the random walk u_n, with probability going to 1 as n goes to infinity, there is a curve α on S such that for some $1 \le i < n$ and $\epsilon, K > 0$:

- the axis of w_n has the form described above,
- for $h \in \{u_i, u_i s_n g, u_i s_n g t_n g\}$, each projection $d_{h\alpha}(1, u_n)$ is greater than $\epsilon \log(n)$ and $d_X(1, u_n) \le K$ for all X disjoint from $u_i s_n g \cdot \alpha$, and
- for $h \in \{u_i, u_i s_n g, u_i s_n g t_n g\}$, each projection $d_{h\alpha}(v_n^{-1}, 1)$ and $d_{h\alpha}(u_n, w_n)$ is bounded by $\epsilon/10 \log(n)$.

The third item holds because of an application of Proposition 3.2 and a simple conditioning argument. We explain the bound on the term $d_{h\alpha}(u_n, w_n)$ and leave the other to the reader. Let $\mathcal{Y}_n = \{f \in \mathrm{Mod}(S) : \exists Y = Y(f) \text{ annulus with } d_Y(1, f) \ge \epsilon \log n\}$. The probability that the third item holds is at least

$$\begin{aligned} &\sum_{f \in \mathcal{Y}_n} \mathbb{P}(d_{Y(f)}(u_n, w_n) \le \epsilon/10 \log(n)) \mathbb{P}(u_n = f) \\ &= \sum_{f \in \mathcal{Y}_n} \mathbb{P}(d_{f^{-1}Y(f)}(1, v_n) \le \epsilon/10 \log(n)) \mathbb{P}(u_n = f) \end{aligned}$$

The terms $\mathbb{P}(d_{f^{-1}Y(f)}(1, v_n) \le \epsilon/10 \log n)$ are arbitrarily close to 1 as n goes to ∞ by Proposition 3.2, and $\sum_{f \in \mathcal{Y}_n} \mathbb{P}(u_n = f) = \mathbb{P}(u_n \in \mathcal{Y}_n)$ goes to 1 by the version of item 2 that we proved above (if ϵ is sufficiently small).

We are now ready to conclude: by the second and third bullet points and the triangle inequality, for any n large enough we have

$$d_{h\alpha}(v_n^{-1}, w_n) \ge 8\epsilon/10 \log(n) - 2L,$$

for each $h \in \{u_i, u_i s_n g, u_i s_n g t_n g\}$ with probability going to 1, as $n \to \infty$. Just as in our previous application of the bounded geodesic image theorem, since γ'' has length at least $\epsilon' n$, each of $h\alpha$ (for h as above) is at bounded distance from γ', and the distances from both v_n^{-1} and w_n to the axis A_n is $O(\log(n))$, we get that $d_{h\alpha}(w_n) \ge 8\epsilon/10 \log(n) - 2L - 2M$. Now exactly as in the first part of the proof, the 3 large projections to $h\alpha$ for $h \in \{u_i, u_i s_n g, u_i s_n g t_n g\}$ guarantee that $d_X(w_n)$ is uniformly bounded whenever X is a subsurface of S disjoint from $u_i s_n g \cdot \alpha$. This completes the proof. □

Informally, Proposition 7.1 controls the size of subsurface distances along the axis of the random pseudo-Anosov w_n. These distances in turn control the lengths of curves in the hyperbolic manifold $M(w_n)$.

Length estimates. Following Minsky [21], we set

$$\omega_\alpha(f) = d_\alpha(f) + i \cdot \big(1 + \sum_{Y \in \mathcal{Y}_\alpha} \{\{d_Y(f)\}\}_{K_3}\big),$$

for some constant $K_3 \geq 0$ as determined in [21]. By [21, Section 9.5], we may suppose that $K_3 \geq K_2$ for K_2 as in Proposition 7.1.

For a pseudo-Anosov $f \in \mathrm{Mod}(S)$, the complex length of α in the hyperbolic mapping torus $M(f)$ is

$$\lambda_\alpha(f) = \ell_\alpha(f) + i \cdot \theta_\alpha(f),$$

where $\ell_\alpha(f)$ is the usual hyperbolic length of α in $M(f)$ and $\theta_\alpha(f) \in (-\pi, \pi]$ is the rotational part of α. Recall that we are interested in smallest $l_\alpha(w_n)$ along the random walk (w_n).

We regard $\omega_\alpha(f)$ and $2\pi i/\lambda_\alpha(f)$ as points in the upper-half plane model of the hyperbolic plane $\mathbb{H}$. The following theorem is part of the Brock–Canary–Minsky proof of Thurston's Ending Lamination Conjecture [7]:

Theorem 7.2 (Length Bound Theorem [7])**.** *There are $D, \epsilon \geq 0$, depending only of S, such that for any pseudo-Anosov $f \in \mathrm{Mod}(S)$ and any curve α in S with $\ell_\alpha(f) \leq \epsilon$:*

$$d_{\mathbb{H}}\left(\omega_\alpha(f), \frac{2\pi i}{\lambda_\alpha(f)}\right) \leq D.$$

Moreover, if $|\omega_\alpha| \geq M$, for $M \geq 0$ depending only on S, then $\ell_\alpha(f) \leq \epsilon$.

Corollary 7.3. *With notation as above and the hypotheses of Theorem 7.2, there is a $D_1 \geq 0$ such that*

$$D_1^{-1} \cdot \frac{2\pi\, S_\alpha(f)}{d_\alpha^2(f) + S_\alpha^2(f)}. \leq \ell_\alpha(f) \leq D_1 \cdot \frac{2\pi\, S_\alpha(f)}{d_\alpha^2(f) + S_\alpha^2(f)}, \tag{1}$$

where $S_\alpha(f) = 1 + \sum_{Y \in \mathcal{Y}_\alpha} \{\{d_Y(f)\}\}_{K_3}$.

Proof. Using Theorem 7.2 together with the $\mathbb{H}$ isometry $z \to -1/z$, we have

$$d_{\mathbb{H}}\left(\frac{-1}{\omega_\alpha(f)}, \frac{i \cdot \lambda_\alpha(f)}{2\pi}\right) \leq D.$$

However, if $d_{\mathbb{H}}(z_1, z_2) \leq D$, then $|\log(\frac{\Im z_1}{\Im z_2})| \leq D$. Since $\Im(-1/\omega_\alpha(f)) = S_\alpha/(d_\alpha^2(f) + S_\alpha^2(f))$ and $\Im(i \cdot \lambda_\alpha(f)/2\pi) = \ell_\alpha(f)/2\pi$, setting $D_1 = e^D$ completes the proof. □

Proof of Theorem 1.3. We break the proof into 3 steps:

Step 1: There is a constant $C_3 > 0$, depending only on S, such that with probability going to 1, there is a curve α_n in S with $\ell_{\alpha_n}(w_n) \leq \frac{C_3}{\log^2(n)}$.

To see this, note that by item (2) of Proposition 7.1, with probability approaching 1, there is a curve α_n such that $d_{\alpha_n}(w_n) \geq C_2^{-1} \cdot \log(n)$ and $\sup_{Y \in \mathcal{Y}_{\alpha_n}} d_Y(w_n) \leq K_2$. By Corollary 7.3,

$$\ell_{\alpha_n}(w_n) \leq D_1 \cdot \frac{2\pi}{d_\alpha^2 + 1} \leq D_1 \cdot \frac{2\pi}{C_2^{-1} \cdot \log^2(n) + 1}, \tag{2}$$

where we have used that $K_3 \geq K_2$. This completes Step 1.

Step 2: There is a constant $C_4 > 0$, depending only on S, such that with probability going to 1, $\ell_\alpha(w_n) \geq \frac{C_4}{\log^2(n)}$ for any curve α in S.

By item (1) of Proposition 7.1, for all subsurface Y of S, $\sup_{Y\subsetneq S}\sum_{Z\subset Y}\{\{d_Z(w_n)\}\}_{K_2}\leq C_2\cdot\log(n)$ with probability going to 1. Hence, with probability going to 1,

$$\begin{aligned}\ell_\alpha(w_n) &\geq D_1^{-1}\cdot\frac{2\pi\, S_\alpha(w_n)}{d_\alpha^2(w_n)+S_\alpha^2(w_n)}\\ &\geq C_2^{-2}D_1^{-1}\cdot\frac{2\pi}{\log^2(n)}.\end{aligned}$$

This finishes Step 2.

Step 3: With probability going to 1, $\mathrm{sys}(M(w_n))$ is realized by a geodesic g_n that is isotopic to a curve β_n in the fiber S.

By Step 1, we know that, with probability approaching 1, $\mathrm{sys}(M(w_n))\to 0$. Hence, it suffices to show that any sufficiently short geodesic in a hyperbolic fibered 3-manifold M with fiber S is isotopic to a curve in S. This is an easy consequence of [6, Lemma 2.1], but an older argument that uses well-known results in the theory of Kleinian groups was already known to experts. We now sketch the argument.

First, any primitive geodesic g which is sufficiently short (depending only on S) must be homotopic into S. Otherwise, g meets the image of any map $f\colon S\to M$ homotopic to the fiber. (In fact, g has nonzero algebraic intersection number with S.) In particular, g meets the image of a pleated surface $f\colon S\to M$ (see [33, 32]). As in [21, Section 3.22], for any $\epsilon_1\geq 0$ there is an $\epsilon_2<\epsilon_1$ such that

$$f(S_{[\epsilon_1,\infty)})\subset M_{[\epsilon_2,\infty)},$$

where $N_{[\epsilon,\infty)}$ is the ϵ-thick part of N, i.e. where the injectivity radius is greater than 2ϵ. Fix ϵ_1 less than the 3-dimensional Margulis constant. (See [32]). If g has length less than ϵ_2, then by the thick-thin decomposition of S, there is a closed curve γ on S with length less than ϵ_1 whose image in M meets g. The Margulis Lemma [32, Chapter 5.10.1] implies that g and γ (at their intersection point) generate a virtually cyclic subgroup $\pi_1(M)$. This, however, contradicts that γ and g cannot be homotopic up to powers.

Now since g is homotopic into S, it lifts to the cover $\widetilde{M}$ corresponding to S. Since $\widetilde{M}\cong S\times\mathbb{R}$, a theorem of Otal [24] implies that if g is sufficiently short, then it is isotopic to a curve β in S. This completes the proof of Step 3 and of the theorem. □

References

[1] T. Aougab, S. J. Taylor, and R. C. Webb. Effectivizing the Masur–Minsky distance formulas and applications to hyperbolic 3-manifolds. In preparation, April 2017.

[2] J. Behrstock. Asymptotic geometry of the mapping class group and Teichmüller space. *Geom. Topol.*, 10:1523–1578, 2006.

[3] J. Behrstock, M. F. Hagen, and A. Sisto. Hierarchically hyperbolic spaces I: curve complexes for cubical groups. *To appear in Geom. Topol.*, 2014.

[4] J. Behrstock, M. F. Hagen, and A. Sisto. Hierarchically hyperbolic spaces II: combination theorems and the distance formula. *arXiv:1509.00632*, 2015.

[5] M. Bestvina and M. Feighn. Subfactor projections. *J. Topol.*, 2014.

[6] I. Biringer and J. Souto. Ranks of mapping tori via the curve complex. *To appear in J. Reine Angew. Math.*, 2017.

[7] J. Brock, R. Canary, and Y. Minsky. The classification of Kleinian surface groups, II: The ending lamination conjecture. *Ann. of Math.*, 176(1):1–149, 2012.

[8] D. Calegari and J. Maher. Statistics and compression of scl. *Ergodic Theory and Dynamical Systems*, 35(01):64–110, 2015.

[9] F. Dahmani, V. Guirardel, and D. Osin. Hyperbolically embedded subgroups and rotating families in groups acting on hyperbolic spaces. *Mem. Amer. Math. Soc.*, 245(1156):v+152, 2017.
[10] B. Farb. Relatively hyperbolic groups. *Geom. Funct. Anal.*, 8(5):810–840, 1998.
[11] S. M. Gersten. Subgroups of word hyperbolic groups in dimension 2. *J. London Math. Soc. (2)*, 54(2):261–283, 1996.
[12] R. Gitik, M. Mitra, E. Rips, and M. Sageev. Widths of subgroups. *Trans. Amer. Math. Soc.*, 350(1):321–329, 1998.
[13] I. Kapovich and K. Rafi. On hyperbolicity of free splitting and free factor complexes. *Groups Geom. Dyn.*, 8(2):391–414, 2014.
[14] J. Maher. Random walks on the mapping class group. *Duke Math. J.*, 156(3):429–468, 2011.
[15] J. Maher. Exponential decay in the mapping class group. *J. Lond. Math. Soc. (2)*, 86(2):366–386, 2012.
[16] J. Maher and A. Sisto. Random subgroups of acylindrically hyperbolic groups and hyperbolic embeddings. arXiv:1701.00253.
[17] J. Maher and G. Tiozzo. Random walks on weakly hyperbolic groups. *To appear in J. Reine Angew. Math.*, 2017.
[18] H. A. Masur and Y. N. Minsky. Geometry of the complex of curves. II. Hierarchical structure. *Geom. Funct. Anal.*, 10(4):902–974, 2000.
[19] P. Mathieu and A. Sisto. Deviation inequalities for random walks. arXiv:1411.7865.
[20] A. Minasyan. Some properties of subsets of hyperbolic groups. *Comm. Algebra*, 33(3):909–935, 2005.
[21] Y. Minsky. The classification of Kleinian surface groups, I: Models and bounds. *Ann. of Math.*, pages 1–107, 2010.
[22] Y. N. Minsky. Bounded geometry for Kleinian groups. *Invent. Math.*, 146(1):143–192, 2001.
[23] D. V. Osin. Elementary subgroups of relatively hyperbolic groups and bounded generation. *Internat. J. Algebra Comput.*, 16(1):99–118, 2006.
[24] J.-P. Otal. Sur le nouage des géodésiques dans les variétés hyperboliques. *Comptes rendus de l'Académie des sciences. Série 1, Mathématique*, 320(7):847–852, 1995.
[25] J.-P. Otal. *The hyperbolization theorem for fibered 3-manifolds*, volume 7. American Mathematical Soc., 2001.
[26] I. Rivin. Walks on groups, counting reducible matrices, polynomials, and surface and free group automorphisms. *Duke Math. J.*, 142(2):353–379, 2008.
[27] I. Rivin. Statistics of random 3-manifolds occasionally fibering over the circle. *arXiv:1401.5736*, 2014.
[28] A. Sisto. Projections and relative hyperbolicity. *Enseign. Math. (2)*, 59(1-2):165–181, 2013.
[29] A. Sisto. Contracting elements and random walks. *J. Reine Angew. Math., to appear*, 2017.
[30] A. Sisto. Tracking rates of random walks. *Israel J. Math., to appear*, 2017.
[31] S. J. Taylor. A note on subfactor projections. *Algebr. Geom. Topol.*, 14(2):805–821, 2014.
[32] W. P. Thurston. The Geometry and Topology of 3-Manifolds. *Princeton University Notes*, 1980.
[33] W. P. Thurston. Hyperbolic structures on 3-manifolds, ii: Surface groups and 3-manifolds which fiber over the circle. *arXiv preprint math/9801045*, 1998.

Department of Mathematics, ETH Zurich, 8092 Zurich, Switzerland
E-mail address: sisto@math.ethz.ch

Department of Mathematics, Temple University, Philadelphia, PA, USA
E-mail address: samuel.taylor@temple.edu