COMPOSITIO MATHEMATICA

Random walks and contracting elements I: deviation inequality and limit laws

Inhyeok Choi

Compositio Math. **161** (2025), 1512–1575.

doi: 10.1112/S0010437X25007511

FOUNDATION COMPOSITIO MATHEMATICA

LONDON MATHEMATICAL SOCIETY EST. 1865

Compositio Math. 161 (2025) 1512–1575
doi:10.1112/S0010437X25007511

Random walks and contracting elements I: deviation inequality and limit laws

Inhyeok Choi

Abstract

We study random walks on metric spaces with contracting isometries. In this first article of the series, we establish sharp deviation inequalities by adapting Gouëzel's pivotal time construction. As an application, we establish the exponential bounds for deviation from below, central limit theorem, law of the iterated logarithms, and the geodesic tracking of random walks on mapping class groups and CAT(0) spaces.

Contents

Received 31 October 2023, accepted in final form 26 February 2025.
2020 Mathematics Subject Classification 20F67, 30F60, 57M60, 60G50 (primary).
Keywords: random walk; CAT(0) space; mapping class group; central limit theorem; geodesic tracking.

1. Introduction

This is the first in the series of articles concerning random walks on metric spaces with contracting elements. This series is a reformulation of the preprint [Cho22a] announced by the author, aiming for a more concise and systematic exposition.

Let G be a countable group of isometries of a metric space (X,d) with basepoint $o\in X$. We consider the random walk generated by a probability measure μ on G, which entails the product $Z_n=g_1\cdots g_n$ of independent random isometries g_i chosen with law μ. We are interested in the asymptotic behavior of a random path $(Z_n)_{n>0}$ seen by X, or in other words, the behavior of a random *orbit path* $(Z_n o)_{n>0}$ on X. For instance, we can ask the following questions:

– Does the random variable $(1/n)d(o,Z_n o)$ converge to a constant almost surely?
– Does the random variable $(1/\sqrt{n})d(o,Z_n o)$ converge in law to a Gaussian law?
– How fast does $\mathbb{P}(an\leqslant d(o,Z_n o)\leqslant bn)$ decay for $0\leqslant a\leqslant b$?

These questions are associated with the so-called *moment conditions*. For each $p>0$ we define the pth moment of μ by

$$\mathbb{E}_\mu[d(o,go)^p]=\int_G d(o,go)^p\,d\mu,$$

and the exponential moment (with a parameter $K>0$) of μ by

$$\mathbb{E}_\mu \exp(Kd(o,go))=\int_G e^{Kd(o,go)}\,d\mu.$$

In the classical setting of $X=\mathbb{R}$, the previous three questions are answered when μ has finite first moment, finite second moment and finite exponential moment, respectively.

A particularly interesting examples come from isometric actions on non-positively curved spaces. This setting includes Gromov hyperbolic groups [BQ16, BMSS22, Gou22]; relatively hyperbolic groups [Sis17, QRT20]; groups with nontrivial Floyd boundary [GGPY21]; the mapping class group of a finite-type hyperbolic surface acting on Teichmüller space [KM96, Hor18, DH18, BCK21] or the curve complex [Mah10a, Mah10b, Mah11]; the outer automorphism group of a finite-rank free group acting on the Culler–Vogtmann outer space [Hor18, DH18] and the free factor complex; groups acting on CAT(0) spaces [KM99, KL06, Fer18, LB22b, LB22a].

In this paper, we propose a unified theory for random walks on the aforementioned spaces. We first study the case where X possesses strongly contracting isometries (see Convention 2.11), and μ is non-elementary (see § 2.4). This condition is mild enough to cover all the aforementioned spaces (except for outer space, which will be studied carefully in [Cho22b] due to the asymmetry issue). At the same time, this is just the right amount of restriction that leads to limit laws under optimal moment conditions.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

We also present a parallel theory for metric spaces with weakly contracting isometries (see Convention 7.2). As a result, we obtain limit laws on hierarchically hyperbolic groups (HHGs) with optimal moment conditions. We describe the case of mapping class group for concreteness.

Theorem 1.1. *Let G be the mapping class group of a finite-type surface, let d be a word metric on G, let $(Z_n)_{n\geq 0}$ be the random walk generated by a non-elementary probability measure μ on G, and let*

$$\lambda = \lambda(\mu) := \lim_{n\to\infty} \frac{1}{n}\mathbb{E}[d(\mathrm{id}, Z_n)]$$

be the drift of μ on G. Then, for each $0 < L < \lambda$, the probability $\mathbb{P}(d(\mathrm{id}, Z_n) \leqslant Ln)$ decays exponentially as n tends to infinity.

This is an analog of the result of Gouëzel [Gou22, Theorem 1.3], who established the exponential bound for Gromov hyperbolic spaces. Note that, for every admissible probability measure μ on the mapping class group G, the spectral radius of μ is strictly smaller than 1 due to the non-amenability of G [Kes59]. Combining this with the exponential growth of G, one can obtain $L > 0$ for which $\mathbb{P}(d(\mathrm{id}, Z_n) \leqslant Ln)$ decays exponentially. Hence, the nontrivial part of Theorem 1.1 is that L can be as close to λ as we want.

We also obtain the deviation inequalities with optimal moment conditions (see Proposition 4.13). Combining this with Mathieu–Sisto's theory [MS20], we establish the central limit theorem (CLT) and law of the iterated logarithms (LIL) on mapping class groups.

Theorem 1.2. *Let G be the mapping class group of a finite-type hyperbolic surface, let d be a word metric on G, and let $(Z_n)_{n\geqslant 0}$ be the random walk generated by a non-elementary probability measure μ on G with finite second moment. Then, there exists $\sigma(\mu) \geqslant 0$ such that $(1/\sqrt{n})(d(\mathrm{id}, Z_n) - n\lambda(\mu))$ converges in law to the Gaussian law $\mathcal{N}(0, \sigma(\mu))$ of variance $\sigma(\mu)^2$, and moreover,*

$$\limsup_{n\to\infty} \frac{d(\mathrm{id}, Z_n) - n\lambda(\mu)}{\sqrt{2n\log\log n}} = \sigma(\mu) \quad \textit{almost surely.}$$

In acylindrically hyperbolic groups, Mathieu and Sisto established CLT for random walks with finite exponential moment [MS20, Theorem 13.4]. We strengthen their result by weakening the moment condition.

Lastly, we address the geodesic tracking of random paths.

Theorem 1.3. *Let G be the mapping class group of a finite-type surface, let d be a word metric on G, and let $(Z_n)_{n\geqslant 0}$ be the random walk generated by a non-elementary measure μ on G.*

(i) *Let $p > 0$ and suppose that μ has finite pth moment. Then for almost every sample path $(Z_n)_{n\geqslant 0}$, there exists a geodesic γ on G such that*

$$\lim_{n\to\infty} \frac{1}{n^{1/p}} d(Z_n, \gamma) = 0.$$

(ii) *If μ has finite exponential moment, then there exists $K < \infty$ such that the following holds. For almost every sample path $(Z_n)_{n\geqslant 0}$, there exists a geodesic γ on G such that*

$$\limsup_{n\to\infty} \frac{1}{\log n} d(Z_n, \gamma) < K.$$

For finitely supported random walks, Sisto established the deviation rate $d(Z_n, \gamma) = O(\sqrt{n \log n})$ [Sis17, Theorem 1.2]. Later, Qing, Rafi, and Tiozzo obtained the rate $d(Z_n, \gamma) =$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

$O(\log^{3g-3+b}(t))$, where g and b denote the genus and the number of punctures of the surface [QRT20, Theorem C]. We refine these results by suggesting the deviation rate $O(\log(t))$ for random walks with finite exponential moment.

In full generality, the main results hold in the setting of Conventions 2.11 and 7.2. In particular, Theorems 1.1, 1.2 and 1.3 apply to random walks on rank-one CAT(0) spaces. This extends the author's previous work [Cho23] that deals with Gromov hyperbolic spaces and Teichmüller space, and recovers several results by Le Bars [LB22b, LB22a].

To obtain the main theorems, we blend the pioneering theories due to Gouëzel [Gou22] and due to Mathieu and Sisto [MS20]. Gouëzel's method effectively captures the alignment of the orbit path on X (see § 4.1), while Mathieu–Sisto's technique provides the desired limit theorems when appropriate deviation inequalities are given. Both of these theories rely on the Gromov hyperbolicity of the ambient space. Our contribution is to replace the Gromov hyperbolicity with weaker notion of hyperbolicity. In particular, we obtain large deviation principle, CLT and geodesic tracking on (possibly non-proper) CAT(0) spaces. Moreover, we generalize Mathieu–Sisto's theory by lifting the moment condition, leading to the exponential bounds for the escape to infinity and CLT for random walks without finite exponential moment.

1.1 Context

Random walks on groups have often been studied via their actions on Gromov hyperbolic spaces. For instance, random walks on Teichmüller space and outer space have been understood by coupling them with the curve complex and the free factor complex, respectively [Hor18, DH18]. A similar strategy was recently pursued for proper CAT(0) spaces by Le Bars [LB22b, LB22a], building upon a new hyperbolic model for CAT(0) spaces [PSZ24].

These strategies eventually depend on the following ingredients:

- the non-atomness of the stationary measure on the Gromov boundary [MT18, Proposition 5.1];
- CLT for martingales arising from Busemann cocycles [BQ16, Theorem 4.7];
- linear progress with exponential decay [Mah12, Theorem 1.2]; or
- linear progress using the acylindricity of the action [MS20, Theorem 9.1, Proposition 9.4].

The first two items require a nice (e.g., compact) boundary structure of X. These boundary structures are also available in some class of non-Gromov-hyperbolic spaces (such as Teichmüller space, outer space and finite-dimensional CAT(0) cube complices, see [Fer18, Fer18, FLM24]), but are hard to come by in the general case.

To establish the third item, Maher considered a stopping time that arises when a random path penetrates nested shadows, which relies on moment conditions: see [Mah12] and [Sun20]. For the last item, Mathieu and Sisto assumed finite exponential moment condition to couple the random paths on G with the corresponding paths on G in probability.

It is not straightforward to apply the aforementioned strategies to, say, random walks on (non-proper) CAT(0) spaces. Even in well-known settings such as Gromov hyperbolic groups, moment conditions are often necessary. Our goal is to lift these restrictions: we want a structure for random walks on a wide class of spaces X that:

- does not assume global Gromov hyperbolicity of X;
- does not rely on any boundary structure of X;
- does not assume any moment condition *a priori*; and
- effectively captures the 'alignment' of a sample path on X.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

The first goal was studied by Sisto in [Sis18]. Not assuming global Gromov hyperbolicity of X, Sisto presented a random walk theory using strongly contracting isometries, which are found in both Gromov hyperbolic spaces and CAT(0) spaces. Note that the existence of strongly contracting isometries also has implications on the growth problem and counting problem [ACT15, Yan14, Yan19, Yan20, Leg22, Cou22].

The second goal was pursued by Mathieu and Sisto for acylindrically hyperbolic groups in [MS20], establishing deviation inequalities without referring to the boundary of X.

The first and the second goals were also pursued by Boulanger, Mathieu, Sert, and Sisto in [BMSS22]. They discuss Corollary 6.5 for Gromov hyperbolic spaces and pointed out the versatility of Schottky sets in other spaces. For more detail, see § 6.

All the goals except the first one were achieved in Gouëzel's recent paper [Gou22]. In [Gou22], Gouëzel establishes Theorem 1.1 for Gromov hyperbolic spaces by recording the Schottky directions aligned along a random path. Such a recording, called the set of pivotal times, grows linearly with exponential decay. More importantly, this growth is uniform and *is independent of the intermediate non-Schottky steps*.

Our theory achieves the four goals in the setting of Convention 2.11. For this purpose, we combine Gouëzel's pivotal time construction with Sisto's theory of random walks involving strongly contracting isometries. This was indirectly pursued for Teichmüller space in [Cho23]. Our usage of strongly contracting isometries is also hugely influenced by Yang's series of papers [Yan14, Yan19, Yan20] in the context of counting problems.

Although strongly contracting isometries are found in various groups, it is not known whether the Cayley graph of a mapping class group possesses strongly contracting isometries. A related issue arises when one considers a group G that is quasi-isometric to another group H. Having a strongly contracting isometry is not passed through quasi-isometries: it is even not preserved under the change of finite generating set of a group [ACGH19, Theorem 4.19].

This is why we provide a parallel theory in the language of weakly contracting isometries. We note that having a weakly contracting *infinite quasigeodesic* is stable under quasi-isometry. Strictly speaking, our setting is not stable under quasi-isometry: we consider two coarsely equivariant G-actions, one involving weak contraction and the other one involving strong contraction. Nevertheless, the present theory is an attempt towards quasi-isometry (QI)-invariant random walk theory. We record recent breakthrough in this direction by Goldborough and Sisto [GS21], showing that certain QI-invariant group-theoretic property (that involves an action on a hyperbolic space) guarantees a CLT for simple random walks.

1.2 Strategy

Morally, contracting directions constitute a tree-like structure. As a toy model, consider

$$G = F_2 * \mathbb{Z}^2 = \langle a, b, c, d \mid cd = dc \rangle$$

acting on its Cayley graph X. A geodesic $\gamma = abaaba$ in X is composed of edges $e_1 = [\mathrm{id}, a]$, $e_2 = [a, ab]$, $e_3 = [ab, aba]$, and so on. The geodesicity of γ forces the local alignment among e_i: e_i projects onto e_{i+1} at the beginning point of e_{i+1} and e_{i+1} projects onto e_i at the ending point of e_i. Conversely, this local alignment implies that γ is geodesic. (This is false when e_i are directions in a flat, e.g., $e_1 = [\mathrm{id}, c]$, $e_2 = [c, cd]$, and $e_3 = [cd, cdc^{-1}]$.) The same conclusion holds even if we insert edges in the flats in between e_i. For example, consider

$$e_1 = [c, ca], \quad e_2 = [cacd, cacdb^2], \quad g = cacdb^2cd.$$

Observe that (id, e_1), (e_1, e_2), and (e_2, g) satisfy the local alignment conditions. This forces that e_1 and e_2 are subsegments of any geodesic between id and g even if such a geodesic is not unique due to flat parts. We will formulate this more precisely in the alignment lemma in § 3. We will then construct many independent 'tree-like' directions. In our example, the set

$$S_{M,m} = \{(s_1 s_2 \cdots s_M)^m : s_i \in \{a, b\}\}$$

consists of 2^M directions in the free factor. We have the following properties.

(i) For any $x \in X$, $d(\mathrm{id}, [x, s^{\pm 1}]) < M$ for all but at most 1 element $s \in S_{M,m}$.
(ii) For all $s \in S_{M,m}$, the geodesic $[s^{-1}, s]$ passes through id.

These properties will be captured by the notion of Schottky sets (Definition 3.15). Note that one can increase the cardinality of $S_{M,m}$ by taking larger M.

Let us now consider the random walk Z generated by a probability measure μ with $\mu(a), \mu(b) > 0$. Then, for any $M, m > 0$, each element of the Schottky set $S_{M,m}$ is admitted by μ^{*Mm}. By decomposing μ^{*Mm} into a uniform measure on $S_{M,m}$ and the remainder, a random path $(Z_n)_n$ can be modelled by the concatenation of some non-Schottky isometries w_i and Schottky isometries s_i, where the timings for Schottky progresses are given by a renewal process. That means, for a large K, a random word $Z_n = g_1 \cdots g_n$ is of the form

$$Z_n = w_0 s_1 w_1 \cdots s_{n/K} w_{n/K},$$

where the s_i are drawn from $S_{M,m}$. Now Gouëzel's construction of pivotal times provides a large K' such that the following holds: among $\{1, \ldots, n/K\}$, we can pick indices $i(1) < \cdots < i(n/KK')$ at which the Schottky segment is aligned along the entire progress, i.e., $w_0 s_1 \cdots w_{i(k)}[\mathrm{id}, s_{i(k)}]$ are subsegments of $[\mathrm{id}, Z_n]$ $(*)$. Now pick $x \in X$. We have plenty of Schottky isometries available for the slot $s_{i(k)}$. By choosing the right choice among them (i.e., by pivoting), we can also assure that $(x, w_0 s_1 \cdots w_{i(k)}[\mathrm{id}, s_{i(k)}])$ is aligned. Combined with $(*)$, this means that we have a bound of $d(\mathrm{id}, [x, Z_n])$ in terms of an initial subsegment $w_0 s_1 \cdots w_{i(k)}$ of the random path. All these phenomena are exponentially generic (see Lemma 4.10). We subsequently obtain deviation inequalities (Proposition 4.13), CLT, and geodesic tracking. A more involved combinatorial model for random paths leads to the large deviation principle.

In this example, the contracting property of a tree-like edge e is as strong as possible: any geodesic γ connecting the left and the right of e passes through e. We study two variants of such a contracting property. If we require that γ passes through a bounded neighborhood of e, we say that e is strongly contracting. If we require that γ passes through a $\log(\mathrm{diam}(e))$-neighborhood of e, then we say that e is weakly contracting. The argument so far also works for strongly contracting directions, up to a finite error. A more delicate argument is required for weakly contracting isometries. We will deal with these notions in Part I and Part II, respectively.

2. Preliminaries

Before entering Part I, we review basic notions and lemmata. We fix a metric space (X, d) and a basepoint $o \in X$. For $x, y, z \in X$, we define the *Gromov product* of x and z with respect to y by

$$(x, z)_y := \tfrac{1}{2}(d(x, y) + d(y, z) - d(x, z)).$$

2.1 Paths

Let A and B be subsets of X; *A is K-coarsely contained in B* if A is contained in the K-neighborhood of B; A and B are *K-coarsely equivalent* if A is K-coarsely contained in B and *vice*

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

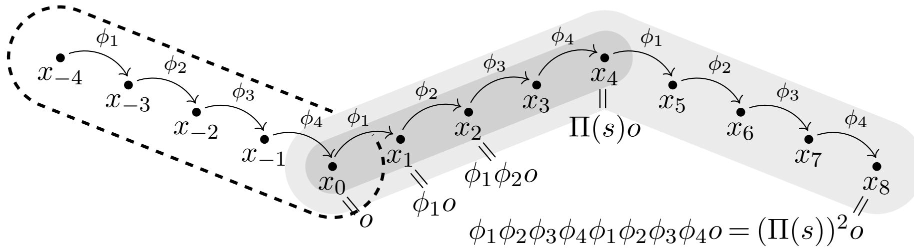

FIGURE 1. Axes associated with a sequence of isometries $s = (\phi_1, \phi_2, \phi_3, \phi_4)$. Points inside the darker shadow constitute $\Gamma^+(s)$, and those inside the lighter shadow constitute $\Gamma^2(s)$. Points in the dashed region constitute $\Gamma^-(s)$.

versa; A is *K-coarsely connected* if for every $x, y \in A$ there exists a chain $x = a_0, a_1, \dots, a_n = y$ of points in A such that $d(a_i, a_{i+1}) \leqslant K$ for each i.

A *path* on X is a map $\gamma : I \to X$ from a 1-coarsely connected subset I of $\mathbb{R}$, called a domain, to X. A *subdomain* J of I is of the form $I \cap [a, b]$ for some $a, b \in \mathbb{R}$. The restriction of γ on J is called a *subpath* of γ. We denote this subpath by $\gamma|_{[a,b]}$.

For paths $\gamma : I \to X$ and $\gamma' : I' \to X$ we say that γ' is a *reparameterization of* γ when there exists a non-decreasing map $\rho : I' \to I$ such that $\gamma' = \gamma \circ \rho$. We say that two paths $\kappa : I \to X$ and $\eta : J \to X$ are *K-fellow traveling* if there exists a reparameterization $\kappa' : J \to X$ of κ such that $d(\kappa'(t), \eta(t)) \leqslant K$ for every $t \in J$. In this case, note that the images of κ and η are within Hausdorff distance K and the endpoints of κ and η are pairwise K-near. By abuse of notation, for a $\kappa : I \to X$ path, κ will often refer to the set-theoretical image $\kappa(I)$ of κ. For instance, when we say that a path $\kappa : I \to X$ is K-close to a point x, it means $d(\kappa(t), x) < K$ for some $t \in I$.

We say that X is *geodesic* if for each pair of points $x, y \in X$ there exists a geodesic connecting x to y. Given two points $x, y \in X$, we denote by $[x, y]$ an arbitrary geodesic connecting x to y.

Let $[x, y]$ be a geodesic on X and $A_1, \dots, A_N$ be subsets of $[x, y]$. We say that $A_1, \dots, A_N$ are *in order from left to right* if $d(x, x_1) \leqslant d(x, x_2) \leqslant \cdots \leqslant d(x, x_N)$ for any choices of $x_i \in A_i$.

We will construct a path for a sequence of isometries as follows. Given a sequence $\alpha = (\phi_1, \cdots, \phi_k)$ of isometries of X, we denote the product of its entries $\phi_1 \cdots \phi_k$ by $\Pi(\alpha)$. Now let

$$x_{mk+i} := \Pi(s)^m \phi_1 \cdots \phi_i o = (\phi_1 \cdots \phi_k)^m \phi_1 \cdots \phi_i o$$

for each $m \in \mathbb{Z}$ and $i = 0, \dots, k-1$; see Figure 1. We let $\Gamma^m(\alpha) := (x_0, x_1, \dots, x_{mk})$ when $m \geqslant 0$ and $\Gamma^m(\alpha) := (x_0, x_{-1}, \dots, x_{mk})$ when $m < 0$. For $m = \pm 1$, we also use a simpler notation

$$\Gamma^+(s) := (x_0, x_1, \dots, x_k),$$
$$\Gamma^-(s) := (x_0, x_{-1}, \dots, x_{-k}).$$

In other words, we write

$$\Gamma^+(\phi_1, \dots, \phi_k) := (o, \phi_1 o, \phi_1\phi_2 o, \dots, \phi_1\phi_2 \cdots \phi_k o),$$
$$\Gamma^-(\phi_1, \dots, \phi_k) := (o, \phi_k^{-1} o, \phi_k^{-1}\phi_{k-1}^{-1} o, \dots, \phi_k^{-1} \cdots \phi_1^{-1} o).$$

Given a path $\gamma = (y_1, \dots, y_N)$, we denote by $\bar\gamma$ its *reversal*, defined by

$$\bar\gamma := (y_N, \dots, y_1).$$

For example, the reversal of $\Gamma^-(\alpha)$ is denoted by $\bar\Gamma^-(\alpha)$, which is

$$\begin{aligned}\bar\Gamma^-(\phi_1, \dots, \phi_k) &:= (x_{-k}, x_{-(k-1)}, \dots, x_0)\\ &= (\phi_k^{-1} \cdots \phi_1^{-1} o, \dots, \phi_k^{-1}\phi_{k-1}^{-1} o, \phi_k^{-1} o, o).\end{aligned}$$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

2.2 Strong contraction

Given a subset A of X, we define the closest point projection $\pi_A : X \to 2^A$ onto A by

$$\pi_A(x) := \{a \in A : d(x,a) = d(x,A)\}.$$

Note that $\pi_A(x)$ is nonempty for each $x \in X$ when A is a closed and locally compact set.

Definition 2.1. Let $K > 0$. A subset A of X is *K-strongly contracting* if the following holds for the closest point projection π_A:

$$\operatorname{diam}_X(\pi_A(x) \cup \pi_A(y)) \leqslant K$$

for all $x, y \in X$ that satisfy $d_X(x,y) \leqslant d_X(x,A)$.

A K-strongly contracting K-quasigeodesic is called a *K-contracting axis*. A lemma follows.

Lemma 2.2. *Let A be a K-strongly contracting subset of X. Then the closest point projection $\pi_A : X \to A$ is $(1, 4K)$-coarsely Lipschitz, i.e., for each $x, y \in X$ we have*

$$\operatorname{diam}(\pi_A(x) \cup \pi_A(y)) < d(x,y) + 4K.$$

This lemma is well known in various forms ([ACT15, Lemma 2.11], [Sis18, Lemma 2.4] and [Yan19, Proposition 2.4(4)]). The explicit constant $4K$ is given as a consequence of Lemma 3.1.

Lemma 2.3 [Yan20, Proposition 2.2 (3)]. *For each $K > 1$ there exists a constant $K' = K'(K)$ such that any subpath of a K-contracting axis is a K'-contracting axis.*

Lemma 2.4 ([ACT15, Lemma 2.15] and [Yan20, Proposition 2.2(2)]). *Let A and A' be coarsely equivalent subsets of X. Then A is strongly contracting if and only if A' is strongly contracting.*

Definition 2.5. An isometry g of X is *strongly contracting* if its orbit $\{g^i o\}_{i \in \mathbb{Z}}$ is a strongly contracting quasigeodesic.

Definition 2.6. We say that isometries g and h of X are *independent* if for any $x \in X$ the map

$$(m,n) \mapsto d(g^m o, h^n o)$$

is proper, i.e., $\{(m,n) : d(g^m o, h^n o) < M\}$ is bounded for each $M > 0$.

The following lemma will be proved in § 3.1.

Lemma 2.7. *Two strongly contracting isometries g and h of X are independent if and only if $\pi_{\{g^i o : i \in \mathbb{Z}\}}(\{h^i o : i \in \mathbb{Z}\})$ and $\pi_{\{h^i o : i \in \mathbb{Z}\}}(\{g^i o : i \in \mathbb{Z}\})$ have finite diameters.*

2.3 Weak contraction

This subsection only matters in Part II; readers interested in Part I may skip this subsection.

Definition 2.8. Let $K > 0$ and $A \subset X$. A *K-projection* onto A is a K-coarsely Lipschitz map $\pi : X \to A$ such that $d(a, \pi(a)) \leqslant K$ for each $a \in A$. Note that for each $x \in X$ we have

$$\begin{aligned} d(x, \pi(x)) &\leqslant \inf_{a \in A} [d(x,a) + d(a, \pi(a)) + d(\pi(a), \pi(x))] \\ &\leqslant \inf_{a \in A} [(K+1) d(x,a) + 2K] \leqslant (K+1) d(x,A) + 2K. \end{aligned} \tag{1}$$

A set A is *K-weakly contracting* if there exists a K-projection π_A such that

$$\operatorname{diam}_X(\pi_A(x) \cup \pi_A(y)) \leqslant K \tag{2}$$

holds for all $x, y \in X$ that satisfy $d(x,y) \leqslant d(x,A)/K$.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Lemma 2.9. *For each $K, M > 1$ there exists $K' > K$ such that the following holds.*

Let $x, y \in X$. Let A be a K-weakly contracting set such that $d(x, A) \geqslant K'$ and such that $\operatorname{diam}(\pi_A(x) \cup \pi_A(y)) \geqslant K'$. Then there exists $p \in [x, y]$ such that $\operatorname{diam}(\pi_A(x) \cup \pi_A(p)) \leqslant 2K'$ and such that either

$$d(x, A) \geqslant M d(p, A) \quad \text{or} \quad M d(x, A) \leqslant d(p, A).$$

Proof. We set $K' = K^2(M(M+7)(K+1)+1)$.

Let $\eta : [0, L] \to X$ be a geodesic connecting x to y. Note that for

$$\tau := \inf\{0 \leqslant t \leqslant L : \operatorname{diam}(\pi_A(x) \cup \pi_A(\eta(t))) > K' + K\}.$$

The K-coarse Lipschitzness of π_A and inequality (2) imply

$$\lim_{\epsilon \to 0+} \operatorname{diam}(\pi_A(\eta(\tau)) \cup \pi_A(\eta(\tau + \epsilon))) \leqslant K,$$
$$\operatorname{diam}(\pi_A(x) \cup \pi_A(\eta(\tau))) \geqslant (K' + K) - K = K'.$$

Hence, by replacing y with $\eta(\tau)$, we may assume $\operatorname{diam}(\pi_A(x) \cup \pi_A(\eta(t))) \leqslant K' + K$ for $t \in [0, L]$. If $d(\eta(t), A) < (1/M) d(x, A)$ for some $t \in [0, L]$, then we are done; suppose not. We inductively take

$$t_0 := 0, \quad t_i := \min\left\{t_{i-1} + \frac{1}{MK} d(x, A), L\right\} \quad (i > 0).$$

The process halts at step N when t_N reaches L. We then have

$$d(\eta(t_{i-1}), \eta(t_i)) = t_i - t_{i-1} \leqslant \frac{1}{MK} d(x, A) \leqslant \frac{1}{K} d(\eta(t_{i-1}), A)$$

for each i. Using inequality (2), we deduce

$$\operatorname{diam}(\pi_A(x) \cup \pi_A(y)) \leqslant \sum_{i=1}^{N} \operatorname{diam}(\pi_A(\eta(t_{i-1})) \cup \pi_A(\eta(t_i))) \leqslant NK.$$

Since the left-hand side is at least K', we have $N \geqslant K'/K \geqslant 2KM(M+7)(K+1)+1$.

Meanwhile, $t_i - t_{i-1} = d(x, A)/MK$ holds for $i \leqslant N-1$. This implies

$$d(x, y) \geqslant t_{N-1} - t_0 \geqslant (N-1) \frac{1}{MK} d(x, A)$$

and considering the assumption $d(x, A) \geqslant K' \geqslant K$ we deduce

$$\begin{aligned} d(x, y) - \operatorname{diam}(\pi_A(x) \cup \pi_A(y)) &\geqslant (N-1) \frac{1}{MK} d(x, A) - (K' + K) \\ &\geqslant MK(M+7)(K+1) \cdot \frac{1}{MK} d(x, A) - 2K' \\ &= (M+7)(K+1) d(x, A) - 2K' \\ &\geqslant (M+1)(K+1) d(x, A) + 4K. \end{aligned}$$

Now using inequality (1) twice, we get

$$\begin{aligned} d(y, A) &\geqslant \frac{1}{K+1}[d(y, \pi_A(y)) - 2K] \\ &\geqslant \frac{1}{K+1}[d(y, x) - d(x, \pi_A(x)) - \operatorname{diam}(\pi_A(x) \cup \pi_A(y)) - 2K] \\ &\geqslant \frac{1}{K+1}[d(x, y) - \operatorname{diam}(\pi_A(x) \cup \pi_A(y)) - (K+1) d(x, A) - 4K] \\ &\geqslant M d(x, A). \end{aligned}$$

□

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

LEMMA 2.10. *For each $K>1$ there exists $K'>0$ satisfying the following.*

Let A be a K-weakly contracting set, let $x,y \in X$, let p be a point on $[x,y]$ and define

$$D_1 := \operatorname{diam}(\pi_A(x) \cup \pi_A(p)), \quad D_2 := \operatorname{diam}(\pi_A(y) \cup \pi_A(p)).$$

Then we have

$$d(p,A) \leqslant K' e^{-D_1/K'} d(x,A) + K' e^{-D_2/K'} d(y,A) + K'. \tag{3}$$

Proof. Let $M := 2K+4$, let $K_1 := K'(K,M)$ be as in Lemma 2.9, and let $K' := 9MK_1$.

Suppose on the contrary that inequality (3) does not hold. Our goal is to find a triple x', y', z' on $[x,y]$, in order from left to right, such that

$$\begin{aligned} d(y',A) &> \max(Md(x',A), Md(z',A), K'), \\ 4K_1 &\geqslant \operatorname{diam}(\pi_A\{x',y',z'\}). \end{aligned}$$

If we find such triple, then we have

$$\begin{aligned} d(y',A) &> \frac{K+1}{2K+4} \cdot M \cdot d(x',A) + \frac{K+1}{2K+4} \cdot M \cdot d(z',A) + \frac{2}{2K+4} \cdot K' \\ &\geqslant (K+1)d(x',A) + (K+1)d(z',A) + 18K_1 \\ &> (K+1)d(x',A) + (K+1)d(z',A) + (2K_1 + 4K). \end{aligned}$$

This will then lead to the contradiction

$$\begin{aligned} d(x',z') &\leqslant d(x',\pi_A(x')) + \operatorname{diam}(\pi_A(x') \cup \pi_A(z')) + d(\pi_A(z'),z') \\ &\leqslant ((K+1)d(x',A) + 2K) + 4K_1 + ((K+1)d(z',A) + 2K) \\ &< 2d(y',A) - (K+1)d(x',A) - (K+1)d(z',A) - 4K \\ &\leqslant \left[d(y',A) - d(x',\pi_A(x'))\right] + \left[d(y',A) - d(z',\pi_A(z'))\right] \\ &\leqslant d(x',y') + d(y',z'). \end{aligned}$$

Let $\eta : [0,L] \to X$ be the geodesic connecting p to x and let $t_0 = 0$. Given $t_{i-1} \in [0,L)$, we pick $t_i \in [t_{i-1}, L]$ such that

$$\operatorname{diam}(\pi_A(\eta(t_{i-1})) \cup \pi_A(\eta(t_i))) \leqslant 2K_1, \quad d(\eta(t_i),A) \geqslant Md(\eta(t_{i-1}),A). \tag{4}$$

If such t_N does not exist at step N, we let $t_N = L$ and stop.

Recall that we are assuming

$$d(\eta(t_0),A) \geqslant K' e^{-D_1/K'} d(x,A) + K' e^{-D_2/K'} d(y,A) + K' \geqslant K'.$$

Hence, $d(\eta(t_i),A) \geqslant M^i K' \geqslant K'$ for $i = 0, \ldots, N-1$. $(*)$ Since η is bounded, the process must stop at some N. We always have $t_N = L$ and $\eta(t_N) = x$. We discuss possible scenarios.

(i) $d(\pi_A(\eta(t_{N-1})) \cup \pi_A(\eta(t_N))) > 2K_1$. Recall Lemma 2.9: there exists $\tau \in [t_{N-1}, t_N]$ such that $\operatorname{diam}(\pi_A(\eta(t_{N-1}) \cup \eta(\tau)) \leqslant 2K_1$ and either $d(\eta(\tau),A) \geqslant Md(\eta(t_{N-1}),A)$ or $d(\eta(\tau),A) \leqslant (1/M)d(\eta(t_{N-1}),A)$. Since the first possibility is excluded, we conclude that $d(\eta(\tau),A) \leqslant (1/M)d(\eta(t_{N-1}),A)$. There are two subcases.

  (a) $N \geqslant 2$: in this case, $d(\eta(t_{N-1}),A) \geqslant Md(\eta(t_{N-2}),A)$ holds by our choice in display (4). By $(*)$, we also know that $d(\eta(t_{N-1}),A) \geqslant K'$. Lastly, $\pi_A\{\eta(t_{N-2}), \eta(t_{N-1}), \eta(\tau)\}$ has diameter at most $4K_1$. Hence, we can take $x' = \eta(\tau)$, $y' = \eta(t_{N-1})$ and $z' = \eta(t_{N-2})$.

  (b) $N = 1$: in this case, we have $d(\eta(\tau),A) \leqslant (1/M)d(p,A)$. We first pick $x' = \eta(\tau)$ and will pick y' and z' later.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

(ii) $\operatorname{diam}(\pi_A(\eta(t_{N-1})), \pi_A(\eta(t_N)) \leqslant 2K_1$. Then we have

$$D_1 = \operatorname{diam}(\pi_A(\eta(0)), \pi_A(\eta(t_N)) \leqslant \sum_{i=1}^{N} \operatorname{diam}(\pi_A(\eta(t_{i-1})), \pi_A(\eta(t_i))) \leqslant 2K_1 N.$$

Since $K' \geqslant 4K_1$, $K' \geqslant e^2$ and $e < 4 < 2K+4 = M$, we deduce

$$d(x, A) \leqslant \frac{1}{K'} e^{D_1/K'} d(p, A) \leqslant \frac{1}{K'} e^N d(p, A) \leqslant (2K+4)^{N-2} d(\eta(t_0), A) \leqslant \frac{1}{M} d(\eta(t_{N-1}), A).$$

Given this, when $N \geqslant 2$, we can pick $x' = x = \eta(t_N)$, $y' = \eta(t_{N-1})$ and $z' = \eta(t_{N-2})$ and deduce a similar contradiction. When $N = 1$, we set $x' = x$.

So far, we have obtained either the desired triple (x', y', z'), or a point $x' \in [x, p]$ such that

$$\operatorname{diam}(\pi_A(x') \cup \pi_A(p)) \leqslant 2K', \quad d(x', A) \leqslant (1/M) d(p, A).$$

A similar discussion on $[p, y]$ also gives either the desired triple, or a point $z' \in [p, y]$ such that $d(\pi_A(z'), \pi_A(p)) \leqslant 2K'$ and $d(z', A) \leqslant \frac{1}{M} d(p, A)$. If we fall into the latter cases in both discussions, we let $y' = p$ and deduce the contradiction. □

2.4 Random walks

Let μ be a probability measure on a discrete group G acting on a metric space (X, d). We denote by $\check{\mu}$ the *reflected version of* μ, which by definition satisfies $\check{\mu}(g) := \mu(g^{-1})$. The *random walk* generated by μ is the Markov chain on G with the transition probability $p(g, h) := \mu(g^{-1}h)$.

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*, and there is a corresponding *(bi-infinite) sample path* $(Z_n)_{n \in \mathbb{Z}}$ under the correspondence

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

We also introduce the notation $\check{g}_n = g_{-n+1}^{-1}$ and $\check{Z}_n = Z_{-n}$. Note that we have an isomorphism $(G^{\mathbb{Z}}, \mu^{\mathbb{Z}}) \to (G^{\mathbb{Z}_{>0}}, \mu^{\mathbb{Z}_{>0}}) \times (G^{\mathbb{Z}_{>0}}, \check{\mu}^{\mathbb{Z}_{>0}})$ by $(g_n)_{n \in \mathbb{Z}} \mapsto ((g_n)_{n>0}, (\check{g}_n)_{n>0})$. In view of this, we sometimes write the bi-infinite sample path as $((Z_n)_{n \geq 0}, (\check{Z}_n)_{n \geq 0})$, where the distributions of $(Z_n)_n$ and $(\check{Z}_n)_n$ are independent.

In certain circumstances, it is beneficial to consider a probability space $(\Omega, \mathbb{P})$ where the step distributions for the random walk are defined, together with some other random variables (RVs). For this purpose, we say that $(\Omega, \mathbb{P})$ is a *probability space for* μ if there is a measure-preserving map from $(\Omega, \mathbb{P})$ to $(G^{\mathbb{Z}_{>0}}, \mu^{\mathbb{Z}_{>0}})$, or equivalently, if independent step RVs $\{g_n(\omega)\}_{n>0}$ are defined and distributed according to μ. We similarly define a probability space $(\check{\Omega}, \mathbb{P})$ for $\check{\mu}$, together with RVs $\{g_n(\check{\omega})\}_{n>0}$. Then the product space $(\Omega \times \check{\Omega}, \mathbb{P})$ models the (bi-infinite) random walk generated by μ. We often omit ω while writing e.g., $g_n = g_n(\omega)$ and $Z_n = Z_n(\omega)$. To make a distinction, we mark RVs on $\check{\Omega}$ with the 'check' sign, e.g., $\check{g}_n := g_n(\check{\omega})$, $\check{Z}_n := Z_n(\check{\omega})$.

We define the *support* of μ, denoted by $\operatorname{supp}\mu$, as the set of elements in G that are assigned nonzero values of μ. We denote by μ^N the product measure of N copies of μ, and by μ^{*N} the Nth convolution measure of μ. We say that μ is *non-elementary* if the subsemigroup generated by the support of μ contains two independent strongly contracting isometries g, h of X. By taking suitable powers, we may assume that g and h belong to the same $\operatorname{supp}\mu^{*N}$ for some $N > 0$.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

When a constant M_0 (to be fixed later) is understood, we use the notation

$$\mathbf{Y}_i(\omega) := (Z_{i-M_0}(\omega)o, Z_{i-M_0+1}(\omega)o, \ldots, Z_i(\omega)o).$$

Similarly, we denote $(Z_{i-M_0}(\check{\omega})o, \ldots, Z_i(\check{\omega})o)$ by $\mathbf{Y}_i(\check{\omega})$.

PART I. Random walks with strongly contracting isometries

In Part I, we develop a theory of random walks that involve strongly contracting isometries. The following convention is employed throughout Part I.

CONVENTION 2.11. *We assume that:*

- (X, d) *is a geodesic metric space;*
- G *is a countable group of isometries of* X*; and*
- G *contains two independent strongly contracting isometries.*

We also fix a basepoint $o \in X$.

We emphasize that no further requirements (properness, weak proper discontinuity, etc.) are imposed on X or G. Convention 2.11 includes the following situations:

(i) (X, d) is a geodesic Gromov hyperbolic space and G contains independent loxodromics, e.g.,
  (a) (X, d) is the curve complex of a finite-type hyperbolic surface and G is the mapping class group; or
  (b) (X, d) is the complex of free factors of the free group of rank $N \geqslant 3$ and G is the outer automorphism group $\mathrm{Out}(F_N)$;
(ii) X is a Teichmüller space of finite type, G is the corresponding mapping class group, and d is either the Teichmüller metric $d_{\mathcal{T}}$ [Min96] or the Weil–Petersson metric d_{WP} [BF02];
(iii) (X, d) is the Cayley graph of a braid group modulo its center $B_n/Z(B_n)$ with respect to its Garside generating set, and G is the braid group B_n [CW21];
(iv) (X, d) is the Cayley graph of a group G with nontrivial Floyd boundary [Kar03, GP13];
(v) (X, d) is the Cayley graph of a $Gr'(1/6)$-labeled graphical small cancelation group G [ACGH19];
(vi) (X, d) is a (not necessarily proper nor finite-dimensional) CAT(0) space and G contains independent rank-one isometries; e.g., G is an irreducible right-angled Artin group and (X, d) is the universal cover of its Salvetti complex.

3. Alignment I: strongly contracting axes

In this section, we will formulate and prove the following claim. Let $(\kappa_i)_{i=1}^n$ be a sequence of long enough contracting axes. Suppose that each pair of consecutive axes is aligned: κ_i (κ_{i+1}, respectively) projects onto κ_{i+1} (κ_i, respectively) near the beginning point of κ_{i+1} (the ending point of κ_i, respectively). Then the axes are globally aligned: κ_i projects onto κ_j near the beginning point (ending point, respectively) of κ_j when $i < j$ ($i > j$, respectively).

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

3.1 Contracting geodesics

The goal of this subsection is to establish Corollary 3.5. We begin by recalling a lemma that appeared as [ACT15, Lemma 2.14], [Sis18, Lemma 2.4] and [Yan19, Lemma 2.4(4)]. For a version with an explicit constant, see the work by Chawla et al. [CCT25, Lemma 2.2].

Lemma 3.1. *Let A be a K-strongly contracting set and let $\eta : I \to X$ be a geodesic such that $\operatorname{diam}(\pi_A(\eta)) > K$. Then there exist $t < t'$ in I such that $\pi_A(\eta)$ and $\eta|_{[t,t']}$ are $4K$-coarsely equivalent, and moreover, such that*

$$\operatorname{diam}(\pi_A(\eta|_{(-\infty,t]}) \cup \eta(t)) < 2K \quad \textit{and} \quad \operatorname{diam}(\pi_A(\eta|_{[t',+\infty)}) \cup \eta(t')) < 2K.$$

Lemma 3.2. *For each $K > 1$ there exists $K' = K'(K) > K$ that satisfies the following.*

Let $\eta : J \to X$ be a K-quasigeodesic whose endpoints are x and y, let A be a subset of η such that $d(x, A) < K$, $d(y, A) < K$, and let $\gamma : J' \to X$ be a geodesic that is K-coarsely equivalent to A. Then, η and γ are also K'-coarsely equivalent, and, moreover, there exists a K'-quasi-isometry $\varphi : J \to J'$ such that $d(\eta(t), (\gamma \circ \varphi)(t)) < K'$ for each $t \in J$.

Proof. Without loss of generality, let $J = [a, b]$, $J' = [c, d]$ and $\eta(a) = x$, $\eta(b) = y$. For each $s \in J'$, we can pick $t_s \in J$ such that $d(\gamma(s), \eta(t_s)) < K$ as γ is coarsely contained in A. Note that

$$|t_{s_1} - t_{s_2}| \leqslant K d(\eta(t_{s_1}), \eta(t_{s_2})) + K^2 \leqslant K d(\gamma(s_1), \gamma(s_2)) + 3K^2 = K|s_1 - s_2| + 3K^2.$$

Similarly, $|t_{s_1} - t_{s_2}| \geqslant (1/K)|s_1 - s_2| - 1 - 2K$ holds. Hence, $s \mapsto t_s$ is a $3K^2$-quasi-isometric embedding.

It remains to show that $\{t_s : s \in J'\}$ is coarsely equivalent to J. Note that A is $3K$-coarsely connected, as it is K-coarsely contained in a 1-connected set γ. It follows that $\eta^{-1}(A)$ is $4K^2$-coarsely connected subset of $[a, b]$. Moreover, since x and y are K-close to A, we have $d(a, \eta^{-1}(A)), d(b, \eta^{-1}(A)) < 2K^2$. Combined together, $[a, b]$ is $4K^2$-coarsely contained in $\eta^{-1}(A)$.

Next, for each $p \in A$ there exists $s \in J'$ such that $d(\gamma(s), p) < K$. This implies $d(\eta(t_s), p) < 2K$ and $\operatorname{diam}(t_s, \eta^{-1}(p)) < 3K^2$. Hence, $\eta^{-1}(A)$ is $3K^2$-coarsely contained in $\{t_s : s \in J'\}$. □

The K-quasi-isometries between intervals are K'-coarsely equivalent to a monotone map for some $K' = K'(K)$ (for an explicit K', see the proof of [San06, Theorem 1.2]). Hence, we have:

Corollary 3.3. *For each $K > 1$ there exists $K' = K'(K) > K$ that satisfies the following.*

Let $\eta : J \to X$ be a K-quasigeodesic connecting x to y, let A be a subset of η such that $d(x, A) < K$ and $d(y, A) < K$, and let $\gamma : J' \to X$ be a geodesic that is K-coarsely equivalent to A. Then η and γ are K'-fellow traveling.

Combining Lemmas 3.1 and 3.2, we observe an instance of the Morseness of contracting axes ([ACGH17, Theorem 1.3], [Sis18, Lemma 2.8.(2)] and [Yan14, Lemma 2.2]).

Corollary 3.4. *For each $K > 1$ there exists a constant $K' > K$ that satisfies the following. Let $\eta : J \to X$ be a K-contracting axis and $\gamma : J' \to X$ be a geodesic that share the endpoints. Then η and γ are K'-fellow traveling.*

Corollary 3.5. *For each $K > 1$ there exists a constant $K' = K'(K)$ that satisfies the following.*

Let $\kappa : I \to X$ and $\eta : J \to X$ be K-contracting axes. Suppose that $\operatorname{diam}(\pi_\kappa(\eta)) > K'$. Then, there exist $t < t'$ in I and $s < s'$ in J such that the following sets are all K'-coarsely equivalent:

$$\kappa|_{[t,t']}, \quad \eta|_{[s,s']}, \quad \pi_\kappa(\eta), \quad \pi_\eta(\kappa).$$

Moreover, we have

$$\operatorname{diam}(\pi_\kappa(\eta|_{(-\infty,s]}) \cup \eta(s)) < K', \quad \operatorname{diam}(\pi_\kappa(\eta|_{[s',+\infty)}) \cup \eta(s')) < K'.$$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Proof. For simplicity, we focus on the case where κ, η have endpoints.

Let $\gamma: J' \to X$ be a geodesic that connects the endpoints of η. Then γ and η are coarsely equivalent by Corollary 3.4. Lemma 2.2 tells us that $\pi_\kappa(\gamma)$ is coarsely equivalent to $\pi_\kappa(\eta)$ and hence large. By Lemma 3.1, there exist $u < u'$ in J' such that $\pi_\kappa(\gamma)$ and $\gamma|_{[u,u']}$ are coarsely equivalent and such that $\gamma|_{(-\infty,u]}$ and $\gamma|_{[u',+\infty)}$ project onto κ near $\gamma(u)$ and $\gamma(u')$, respectively.

Note again that η and γ are fellow traveling by Corollary 3.4 and π_κ is coarsely Lipschitz. This enables us to replace γ with η: there exist $s < s'$ in J such that $\pi_\kappa(\eta)$ and $\eta|_{[s,s']}$ are coarsely equivalent and such that $\eta|_{(-\infty,s]}$ and $\eta|_{[s',+\infty)}$ project onto κ near $\eta(s)$ and $\eta(s')$, respectively.

Since $\pi_\kappa(\eta) \subseteq \kappa$ and $\eta|_{[s,s']}$ are nearby, each point $\eta(t)$ in $\eta|_{[s,s']}$ is near a point $\kappa(s_t)$ of κ. This $\kappa(s_t)$ projects onto η near $\eta(t)$. It follows that $\pi_\eta(\kappa)$ coarsely contains $\eta|_{[s,s']}$ and hence $\pi_\kappa(\eta)$.

This implies that $\pi_\eta(\kappa)$ is also large, and we have another round: there exist $t < t'$ in I such that $\pi_\eta(\kappa)$ and $\kappa|_{[t,t']}$ are coarsely equivalent. Moreover, $\pi_\kappa(\eta)$ coarsely contains $\pi_\eta(\kappa)$. Hence, the two projections are coarsely equivalent, and

$$\pi_\eta(\kappa), \quad \kappa|_{[t,t']}, \quad \pi_\kappa(\eta), \quad \eta|_{[s,s']}$$

are all coarsely equivalent. □

We now digress to the proof of Lemma 2.7.

Proof of Lemma 2.7. Let η and κ denote the axes of g and h, i.e., $\eta: i \mapsto g^i o$ and $\kappa: j \mapsto h^j o$. Let η and κ be K-contracting axes for some $K > 0$.

Suppose that $\pi_\kappa(\eta)$ has finite diameter, i.e., there exists M such that

$$\pi_\kappa(\eta) \subseteq \{\kappa(-M), \kappa(-M+1), \ldots, \kappa(M-1), \kappa(M)\}.$$

Then, for each $i \in \mathbb{Z}$ and $|j| > M + 2K^2$, the diameter of $\pi_\kappa(\eta(i)) \cup \kappa(j)$ is greater than K and $[\eta(i), \kappa(j)]$ is $2K$-close to $\pi_\kappa(\eta(i))$. This forces that

$$d(\eta(i), \kappa(j)) \geqslant \inf_{|t| \leqslant M} d(\kappa(j), \kappa(t)) - 2K \geqslant \frac{1}{K}|j| - \frac{M}{K} - 3K.$$

Similarly, if $\pi_\eta(\kappa)$ has finite diameter, then there exists M' such that $d(\eta(i), \kappa(j)) \geqslant (1/K)|i| - M'$ holds for all j and $|i| > M'$. Hence, $d(g^i o, h^j o)$ is a proper function, and g and h are independent.

Now suppose that $\pi_\kappa(\eta)$ has infinite diameter. By Corollary 3.5, η and κ have subpaths η' and κ', respectively, that are coarsely equivalent to $\pi_\kappa(\eta)$, of infinite diameter. This means that η and κ are not independent. □

3.2 Alignment

Let us now define the notion of alignment. See Figure 2.

Definition 3.6. For $i = 1, \ldots, n$, let κ_i be a path on X whose beginning and ending points are x_i and y_i, respectively. We say that $(\kappa_1, \ldots, \kappa_n)$ is C-aligned if

$$\operatorname{diam}_X(y_i \cup \pi_{\kappa_i}(\kappa_{i+1})) < C, \quad \operatorname{diam}_X(x_{i+1} \cup \pi_{\kappa_{i+1}}(\kappa_i)) < C$$

hold for $i = 1, \ldots, n-1$.

Note that, if $(\kappa_i, \ldots, \kappa_j)$ and $(\kappa_j, \ldots, \kappa_k)$ are C-aligned, then $(\kappa_i, \ldots, \kappa_j, \ldots, \kappa_k)$ is also C-aligned. We allow degenerate paths, e.g., the case where κ_1 or κ_n is a point.

Combining Lemma 3.1 and Corollary 3.3, we obtain the following consequence of alignment.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

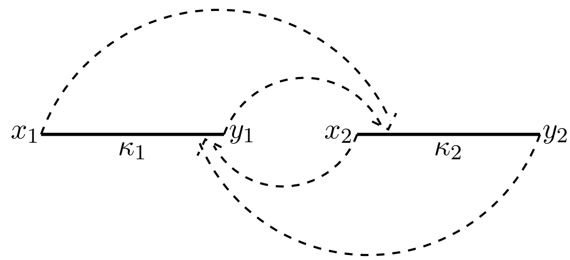

Figure 2. Schematics for an aligned sequence of paths.

Corollary 3.7. *For each $C, K > 1$, there exists $K' = K'(K, C) > \max(K, C)$ such that the following holds.*

Let $x, y \in X$ and let κ be a K-contracting axis such that $\operatorname{diam}(\kappa) > K + 2C$ and such that (x, κ, y) is C-aligned. Then $[x, y]$ contains a subsegment η that is $4K$-coarsely contained in κ and is K'-fellow traveling with κ.

Our first lemma states that the alignment of two strongly contracting axes is governed by the projections of their endpoints to the other axis.

Lemma 3.8. *For each $C, K > 1$, there exists $D = D(K, C) > \max(K, C)$ such that the following holds.*

Let κ, η be K-contracting axes. If $(\kappa, (\text{beginning point of } \eta))$ and $((\text{beginning point of } \kappa), \eta)$ are each C-aligned, then (κ, η) is D-aligned.

Proof. For simplicity, let us assume that the domains of κ and η are closed intervals, say, $I = [t_0, t_1]$ and $J = [s_0, s_1]$, respectively.

It suffices to show that $\pi_\kappa(\eta)$ and $\pi_\eta(\kappa)$ are both small. Suppose not. Then Corollary 3.5 provides $t < t'$ in I and $s < s'$ in J such that

$$\pi_\kappa(\eta), \quad \pi_\eta(\kappa), \quad \kappa|_{[t,t']}, \quad \eta|_{[s.s']}$$

are all coarsely equivalent and large. Moreover, $\pi_\eta(\kappa(t_0))$ is near $\kappa(t)$ and $\pi_\eta(\kappa(t_1))$ is near $\kappa(t')$. Similarly, $\pi_\kappa(\eta(s_0))$ is near $\eta(s)$ and $\pi_\kappa(\eta(s_1))$ is near $\eta(s')$.

Since $\pi_\kappa(\eta)$ and $\pi_\eta(\kappa)$ are large, both $t' - t$ and $s' - s$ are large. Since $[\kappa(t), \kappa(t')]$ and $[\eta(s), \eta(s')]$ are coarsely equivalent, one of the following is true:

– $\kappa(t)$ is near $\eta(s)$ and $\kappa(t')$ is near $\eta(s')$; or
– $\kappa(t)$ is near $\eta(s')$ and $\kappa(t')$ is near $\eta(s)$.

This leads to the following contradictions.

– If $\kappa(t)$ is near $\eta(s)$, then $\eta(s_0)$ projects onto κ near $\kappa(t)$. Since $t_1 - t \geqslant t' - t$ is large, this projection cannot be near $\kappa(t_1) = y$.
– If $\kappa(t)$ is near $\eta(s')$, then $\kappa(t_0)$ projects onto η near $\eta(s')$. Since $s' - s_0 \geqslant s' - s$ is large, this projection cannot be near $\eta(s_0) = x'$.

Hence, $\pi_\kappa(\eta)$ and $\pi_\eta(\kappa)$ cannot be large and the conclusion follows. ☐

The following lemma was inspired by Behrstock's inequality for subsurface projections and curve complexes [Beh06, Theorem 4.3].

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

LEMMA 3.9 [Sis18, Lemma 2.5]. *For each $D, K > 1$, there exists $E = E(K, D) > \max(K, D)$ that satisfies the following.*

Let κ, η be K-contracting axes in X. Suppose that (κ, η) is D-aligned. Then for any $p \in X$, either (p, η) is E-aligned or (κ, p) is E-aligned.

We are now ready to prove the main result of this section.

PROPOSITION 3.10. *For each $D, K > 1$, there exist $E = E(K, D) > \max(K, D)$ and $L = L(K, D) > \max(K, D)$ that satisfy the following.*

Let $x, y \in X$ and let $\kappa_1, \ldots, \kappa_n$ be K-contracting axes whose domains are longer than L. Suppose that $(x, \kappa_1, \ldots, \kappa_n, y)$ is D-aligned. Then (x, κ_i, y) is E-aligned for each i.

Proof. Let $E = E(K, D)$ be as in Lemma 3.9 and let $L = 3KE + K^2$. Our claim is that (x, κ_i) and (κ_i, y) are E-aligned for each i. By symmetry, it suffices to prove the alignment of (x, κ_i).

Let κ be a K-contracting axis whose domain is longer than L. Then the endpoints of κ are at least $3E$-apart. Consequently, no point p in X satisfies the following at the same time:

$$(p, \kappa) \text{ is } E\text{-aligned}, \quad (\kappa, p) \text{ is } E\text{-aligned}.$$

From this observation, we inductively deduce

$$(x, \kappa_i) \text{ is } E\text{-aligned} \Rightarrow (\kappa_i, x) \text{ is not } E\text{-aligned} \Rightarrow (x, \kappa_{i+1}) \text{ is } E\text{-aligned},$$

for $i = 1, \ldots, n$, where the latter implication follows from Lemma 3.9. □

The above proposition can be strengthened as follows. First, we record an immediate consequence of the definition of fellow traveling.

LEMMA 3.11. *Let $E > 0$ and $x, y \in X$. Let κ be a path that E-fellow travels with a subsegment of $[x, y]$. Then (x, κ, y) is $4E$-aligned.*

PROPOSITION 3.12. *For each $D, K > 1$, there exist $E = E(K, D) > \max(K, D)$ and $L = L(K, D) > \max(K, D)$ that satisfy the following.*

Let $x, y \in X$ and let $\kappa_1, \ldots, \kappa_n$ be K-contracting axes whose domains are longer than L and such that $(x, \kappa_1, \ldots, \kappa_n, y)$ is D-aligned. Then the geodesic $[x, y]$ has subsegments $\eta_1, \ldots, \eta_n$, in order from left to right, that are longer than $100E$ and such that η_i and κ_i are $0.1E$-fellow traveling for each i. In particular, (x, κ_i, y) are E-aligned for each i.

Proof. Let $E_1 = E(K, D)$ and $L_1 = L(K, D)$ be as in Proposition 3.10. Let $K_1 = E_1 + 8K$ and let $E = 10K'(K, K_1)$, where $K'(K, K_1)$ is as in Corollary 3.7. Let also $L = L_1 + 101K(K + E) + 2K$.

We will inductively prove a variant of the given statement, namely: First, we know that (κ_1, y) is E_1-aligned by Proposition 3.10. Since (x, κ_1, y) is K_1-aligned and κ_1 is long enough, Corollary 3.7 provides a subsegment $\eta_1 = [x_1', y_1']$ of $[x, y]$ that is $4K$-coarsely contained in κ_1 and is $0.1E$-fellow traveling with κ_1. We then have

$$d(x_1', y_1') \geqslant \operatorname{diam}(\kappa_1) - 0.2E \geqslant \frac{L}{K} - K - 0.2E \geqslant 100E.$$

If $n = 1$, this finishes the proof. If not, note that y_1' is $4K$-close to κ_1. Lemma 2.2 implies that (y_1', κ_2) is $(D + 8K)$-aligned, and hence K_1-aligned. Now the induction hypothesis implies that $[y_1', y]$ has subsegments $\eta_2, \ldots, \eta_n$, in order from left to right, that are longer than $100E$ and such that η_i and κ_i $0.1E$-fellow travel for $i \geqslant 2$. Then $\eta_1, \ldots, \eta_n$ become the desired subsegments. □

Using Proposition 3.12, we can recover the following results by Yang.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Lemma 3.13 [Yan14, Lemma 4.4, Yan19, Proposition 2.9]. *For each $D, M > 0$ and $K > 1$, there exist $E = E(K, D, M) > D$ and $L = L(K, D) > D$ that satisfy the following.*

Let $\kappa_1, \ldots, \kappa_n$ be K-contracting axes whose domains are longer than L. Suppose that $(\kappa_1, \ldots, \kappa_n)$ is D-aligned and $d(\kappa_i, \kappa_{i+1}) < M$ for each i. Then the concatenation $\kappa_1 \cup \cdots \cup \kappa_n$ of $\kappa_1, \ldots, \kappa_n$ is an E-contracting axis.

Lemma 3.14 [Yan14, Corollary 3.2]. *For each $D > 0$ and $K > 1$, there exist $E = E(K, D) > D$ and $L = L(K, D) > D$ that satisfies the following.*

For each $i \in \mathbb{Z}$, let κ_i be a K-contracting axis whose beginning and ending points are x_i and y_i, respectively, and whose domain is longer than L. Suppose that $(\ldots, \kappa_i, \kappa_{i+1}, \ldots)$ is D-aligned. Then the concatenation of $(\ldots, [x_{i-1}, y_{i-1}], [y_{i-1}, x_i], [x_i, y_i], [y_i, x_{i+1}], \ldots)$ is an E-quasigeodesic.

3.3 Schottky sets

Using the previous concatenation lemmata, we will construct arbitrarily many independent contracting isometries. Recall again the notation introduced in § 2.1.

Definition 3.15 (Cf. [Gou22, Definition 3.11]). Let $K > 0$ and let $S \subseteq G^n$ be a set of sequences of isometries. We say that S is *K-Schottky* if:

(i) $\Gamma^+(s)$ and $\Gamma^-(s)$ are K-contracting axes for all $s \in S$;
(ii) for each $x \in X$ we have
$$\#\{s \in S : (x, \Gamma^+(s)) \text{ and } (x, \Gamma^-(s)) \text{ are } K\text{-aligned}\} \geqslant \#S - 1;$$
(iii) for each $s \in S$, $(\bar{\Gamma}^-(s), \Gamma^+(s))$ is K-aligned.

Once a Schottky set S is understood, its element s is called a *Schottky sequence* and the translates of $\Gamma^{\pm}(s)$ are called *Schottky axes*. We say that S is *large enough* if its cardinality is at least 400.

Let μ be a probability measure on G. If each element s of S is attained by the product measure of μ, i.e., $S \subseteq (\mathrm{supp}\mu)^n$, then we say that S *is a Schottky set for* μ.

An intuitive example was given in the introduction. Consider $S_M := \{s_1 s_2 \cdots s_M : s_i \in \{a, b\}\}$ in $F_2 = \langle a, b \rangle$. For any infinite ray on F_2, at most 1 element $s \in S_m$ heads into the direction
$$\#\{s \in S_M : (\xi, s)_{\mathrm{id}} \geqslant M \text{ or } (\xi, s^{-1})_{\mathrm{id}} \geqslant M\} \leqslant 1,$$
for each infinite ray ξ. Moreover, s and s^{-1} diverge early for any $s \in S_M$
$$(s^{-1}, s)_{\mathrm{id}} < 1 \quad \text{for each } s \in S.$$
These properties are also satisfied by the set of Nth powers of elements of S_M
$$S_{N,M} := \{s^N : s \in S_M\}.$$

Definition 3.16. Given a constant $K_0 > 0$, we define:

– $D_0 = D(K_0, K_0)$ as in Lemma 3.8;
– $E_0 = E(K_0, D_0)$, $L_0 = L(K_0, D_0)$ as in Proposition 3.12.

A K_0-Schottky set S whose elements have domains longer than L_0 is called a *long enough K_0-Schottky set*. In other words, when $S \subseteq G^n$ is K_0-Schottky and $n > L_0$, S is called a long enough K_0-Schottky set. In this case, note that the endpoints of $\Gamma^+(s)$ are $100E_0$-apart for each $s \in S$.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

This definition is motivated by the alignment lemmata. Note that the D_0-alignment of a sequence of Schottky axes $(\gamma_1, \ldots, \gamma_N)$ is a *local* condition, between consecutive pairs of axes. Proposition 3.12 then promotes this into the global alignment, i.e., the E_0-alignment of (γ_i, γ_j) for any $i < j$, given that the involved Schottky set is long enough. The following definition is designed to capture this local-to-global phenomenon.

DEFINITION 3.17. Let S be a Schottky set, let $x, y \in X$ and let $\kappa_1, \ldots, \kappa_N$ be Schottky axes. We say that $(x, \kappa_1, \ldots, \kappa_N, y)$ is *C-semi-aligned* if it is a subsequence of a C-aligned sequence of x, y, and Schottky axes, i.e., if there exist Schottky axes $\eta_1, \ldots, \eta_{N'}$ and $1 \leqslant i(1) < \cdots < i(N) \leqslant N'$ such that:

(i) $(x, \eta_1, \ldots, \eta_{N'}, y)$ is C-aligned;
(ii) $\kappa_k = \eta_{i(k)}$ for $k = 1, \ldots, N$.

Here, we also say that $(x, \kappa_1, \ldots, \kappa_N)$ and $(\kappa_1, \ldots, \kappa_N, y)$ are C-semi-aligned.

LEMMA 3.18. *Let S be a long-enough K_0-Schottky set. Let $x, y \in X$, and for each $i = 1, \ldots, N$, let κ_i be a Schottky axis whose beginning and ending points are x_i and y_i, respectively.*

(i) *If (x_1, κ_2) and (κ_1, x_2) are K_0-aligned, then (κ_1, κ_2) is D_0-aligned.*
(ii) *If $(x, \kappa_1, \ldots, \kappa_N, y)$ is D_0-semi-aligned, then (κ_i, κ_j) is E_0-aligned for each $i < j$. Moreover, κ_i is $0.1E_0$-coarsely contained in $[x, y]$ and (x, κ_i, y) is E_0-aligned for each i. We also have*

$$d(x, x_1) + \sum_{i=1}^{N} d(x_i, y_i) + \sum_{i=1}^{N-1} d(y_i, x_{i+1}) + d(y_N, y) \leqslant d(x, y) + E_0 N,$$
$$d(x, x_1) + \sum_{i=1}^{N-1} d(y_i, x_{i+1}) + d(y_N, y) \leqslant d(x, y) - 50E_0 N.$$

Proof. (1) By Lemma 3.8.

(2) Proposition 3.12 explains the first two claims in item (ii). More explicitly, $[x, y]$ contains subsegments $[x'_1, y'_1], \ldots, [x'_N, y'_N]$, in order from left to right, such that:

(a) $[x'_i, y'_i]$ and κ_i are $0.1E_0$-coarsely equivalent;
(b) $d(x'_i, x_i) < 0.1E_0$ and $d(y'_i, y_i) < 0.1E_0$;
(c) $d(x'_i, y'_i) > 100E_0$;

for each i. This implies that

$$\begin{aligned}
d(x, y) &= d(x, x'_1) + \sum_{i=1}^{N} d(x'_i, y'_i) + \sum_{i=1}^{N-1} d(y'_i, x'_{i+1}) + d(y'_N, y) \\
&\geqslant d(x, x_1) + \sum_{i=1}^{N} d(x_i, y_i) + \sum_{i=1}^{N-1} d(y_i, x_{i+1}) + d(y_N, y) - 2\sum_{i=1}^{N} (d(x_i, x'_i) + d(y_i, y'_i)) \\
&\geqslant d(x, x_1) + \sum_{i=1}^{N} d(x_i, y_i) + \sum_{i=1}^{N-1} d(y_i, x_{i+1}) + d(y_N, y) - E_0 N \\
&\geqslant d(x, x_1) + \sum_{i=1}^{N-1} d(y_i, x_{i+1}) + d(y_N, y) + 50E_0 N. \qquad \square
\end{aligned}$$

We now associate long enough and large Schottky sets with non-elementary measures.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Proposition 3.19 (Cf. [Gou22, Proposition 3.12]). *Let μ be a non-elementary probability measure on G. Then for each $N>0$, there exists $K=K(N)>0$ such that for each $L>0$ there exists a K-Schottky set of cardinality N in $(\mathrm{supp}\mu)^n$ for some $n>L$.*

Proof. Since μ is non-elementary, the semigroup generated by $\mathrm{supp}\mu$ contains independent strongly contracting isometries a and b. By taking suitable powers, we may assume that $a=\Pi(\alpha)$ and $b=\Pi(\beta)$ for some $\alpha,\beta\in(\mathrm{supp}\mu)^{L_0}$ for some $L_0>0$. There exists $K_0>0$ such that:

(i) $\Gamma^+(\alpha)$, $\Gamma^-(\beta)$ are K_0-contracting axes; and
(ii) $\mathrm{diam}(o\cup\pi_\gamma(\eta))<K_0$ for distinct axes γ, η among

$$\Gamma^+(\alpha),\quad \Gamma^-(\alpha),\quad \Gamma^+(\beta),\quad \Gamma^-(\beta).$$

The above statements still hold with the same K_0 when α and β are replaced with their self-concatenations, thanks to Lemmas 2.3 and 2.7. Let:

– $K_1=E(K_0,K_0)>K_0$ be as in Lemma 3.9;
– $K_2=E(K_0,K_1)>K_1$, $L_2=L(K_0,K_1)$ be as in Proposition 3.12;
– $K_3=E(K_0,K_2)>K_2$, $L_3=L(K_0,K_2)$ be as in Proposition 3.12;
– $L_3'=3K_0K_3$;
– $K_4=E(K_0,K_0,0)$, $L_4=L(K_0,K_0)$ be as in Lemma 3.13.

By self-concatenating α and β if necessary, we may assume that

$$L_0>L_2+L_3+L_3'+L_4.$$

Since $\Gamma^+(\alpha)$ is a K_0-quasigeodesic whose domain L_0-long, the endpoints of $\Gamma^+(\alpha)$ are at least (L_0/K_1-K_1)-apart. Since L_0 is greater than $3K_0K_3\geqslant 2K_0K_1+K_0^2$, the endpoints of $\Gamma^+(\alpha)$ are $2K_1$-far. In particular, no set $A\subseteq X$ can be simultaneously contained in the K_1-neighborhoods of the two endpoints of $\Gamma^+(\alpha)$. Hence, the statements

$$(x,\Gamma^{\pm}(\alpha))\text{ is }K_1\text{-aligned},\quad (\Gamma^{\pm}(\alpha),x)\text{ is }K_1\text{-aligned}$$

are mutually exclusive for any $x\in X$. Similarly, the statements

$$(x,\Gamma^{\pm}(\beta))\text{ is }K_1\text{-aligned},\quad (\Gamma^{\pm}(\beta),x)\text{ is }K_1\text{-aligned}$$

are mutually exclusive.

Let S_0 be the set of sequences of NL_0 isometries that are concatenations of α and β, i.e.,

$$S_0:=\left\{(\phi_1,\ldots,\phi_{NL_0})\in G^{NL_0}:(\phi_{L_0(i-1)+1},\ldots,\phi_{L_0 i})\in\{\alpha,\beta\}\text{ for }i=1,\ldots,N\right\}.$$

Note that $\#S_0=2^N$ is greater than N. We claim that for each $m>0$, the set

$$S_0^{(m)}:=\Big\{m\text{-self-concatenations of }s\in S_0\Big\}=\Big\{(\underbrace{s,\ldots,s}_{m\text{ times}}):s\in S_0\Big\}$$

is a $((K_0L_0+K_0)N+K_2+K_4)$-Schottky set.

Step 1: *Investigating* $\Gamma^m(s)$. Pick $s=(\phi_1,\ldots,\phi_{NL_0})$ and $s'=(\phi_1',\ldots,\phi_{NL_0}')$ in S_0. Recall the notation

$$x_{nNL_0+i}(s):=(\phi_1\cdots\phi_{NL_0})^n\phi_1\cdots\phi_i o,$$

for $n\in\mathbb{Z}$ and $i=0,\ldots,NL_0-1$. We now define 'sub-axes'

$$\Gamma_i(s):=(x_{L_0(i-1)}(s),\ldots,x_{L_0 i}(s)),$$
$$\Gamma_{-i}(s):=(x_{-L_0(i-1)}(s),\ldots,x_{-L_0 i}(s)),$$

for each $i > 0$. These are translates of $\Gamma^{\pm}(\alpha)$ and $\Gamma^{\pm}(\beta)$. Our initial choices of K_0 and L_0 guarantee that:

- $\Gamma_i(s)$ is a K_0-contracting axis whose domain is longer than L_2, L_3, L_3' and L_4 for each $i \in \mathbb{Z}$;
- $(\Gamma_i(s), \Gamma_{i+1}(s))$ and $(\Gamma_{-i}(s'), \Gamma_{-(i+1)}(s'))$ are K_0-aligned for each $i > 0$. Moreover, $(\bar{\Gamma}_{-1}(s'), \Gamma_1(s))$ is K_0-aligned.

Lemma 3.13 tells us that $\cup_{i>0}\Gamma_i(s)$ is a K_4-contracting axis. In particular, $\Gamma^m(s)$ is a K_4-contracting axis for each $m > 0$. Similarly, $\Gamma^{-m}(s')$ is a K_4-contracting axis for each $m > 0$.

Now note that the following sequence of sub-axes is K_0-aligned:

$$(\ldots, \bar{\Gamma}_{-2}(s'), \bar{\Gamma}_{-1}(s'), \Gamma_1(s), \Gamma_2(s), \ldots).$$

Let $i > 0$ and let $p \in \Gamma_{-i}(s')$. Then Proposition 3.12 tells us that $d(p, \Gamma_1(s)) < d(p, \Gamma_j(s))$ for each $j > 1$ and that $(p, \Gamma_1(s))$ is K_2-aligned. It follows that $(p, \cup_{i>0}\Gamma_i(s))$ is K_2-aligned. For this reason (and its symmetric counterpart), $(\bar{\Gamma}^{-m}(s'), \Gamma^m(s))$ is K_2-aligned for each $m > 0$.

Next, fix $x \in X$ and consider the condition

$$(x, \Gamma_N(s)) \quad \text{is } K_2\text{-aligned}. \tag{5}$$

If Condition (5) holds, then for each $i > N$

$$(x, \Gamma_N(s), \Gamma_{N+1}(s) \ldots, \Gamma_i(s))$$

is K_2-aligned and $d(x, \Gamma_N(s)) < d(x, \Gamma_i(s))$ holds by Proposition 3.12. Hence, $\pi_{\cup_{i>0}\Gamma_i(s)}(x)$ is contained in $\Gamma_1(s) \cup \cdots \cup \Gamma_N(s)$. Meanwhile, recall that, for each i, $\Gamma_i(s)$ is a K_0-quasigeodesic whose domain is L_0-long. Hence, we have

$$\mathrm{diam}(\Gamma_i(s)) \leqslant K_0 \cdot (\text{length of the domain of } s) + K_0 = K_0 L_0 + K_0.$$

Combining these ingredients, we observe that

$$\mathrm{diam}(\pi_{\Gamma^m(s)}(x) \cup o) \leqslant \mathrm{diam}(\Gamma_1(s)) + \cdots + \mathrm{diam}(\Gamma_N(s)) \leqslant (K_0 L_0 + K_0)N$$

holds for every $m > 0$. For a similar reason, the condition

$$(x, \Gamma_{-N}(s)) \quad \text{is } K_2\text{-aligned} \tag{6}$$

implies $\mathrm{diam}(\pi_{\Gamma^m(s)}(x) \cup o) \leqslant (K_0 L_0 + K_0)N$ for all $m < 0$. In summary,

Observation 3.20. If $s \in S_0^{(m)}$ satisfies Condition (5) and (6), then $(x, \Gamma^m(s))$ is $(K_0 L_0 + K_0)N$-aligned for all $m \in \mathbb{Z}$.

Step 2: *Comparing two distinct axes.* We now pick $m > 0$ and consider an element of $S_0^{(m)}$ which violates these conditions.

Observation 3.21. If $s = (\phi_1, \ldots, \phi_{mNL_0}) \in S_0^{(m)}$ violates Condition (5), then all the other elements $s' = (\phi_1', \ldots, \phi_{mNL_0}') \in S_0^{(m)}$ satisfy Condition (5) and (6).

To show this, let $k \in \{1, \ldots, N\}$ be the first index such that $(\phi_{L_0(k-1)+1}, \ldots, \phi_{L_0 k})$ and $(\phi'_{L_0(k-1)+1}, \ldots, \phi'_{L_0 k})$ differ. Let us denote $x_i(s)$ by x_i and $x_i(s')$ by x_i'. Note that the path

$$(x_{NL_0}, x_{NL_0-1}, \ldots, x_{(k-1)L_0} = x'_{(k-1)L_0}, x'_{(k-1)L_0+1}, \ldots, x'_{NL_0})$$

is the concatenation of K_0-aligned K_0-contracting axes

$$(\eta_i)_{i=1}^{2(N-k+1)} := (\bar{\Gamma}_N(s), \bar{\Gamma}_{N-1}(s), \ldots, \bar{\Gamma}_k(s), \Gamma_k(s'), \ldots, \Gamma_N(s')).$$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Recall that s violates Condition (5): $(\bar{\Gamma}_N(s), x) = (\eta_1, x)$ is not K_2-aligned. Since (η_1, η_2) is K_0-aligned, Lemma 3.9 tells us that (x, η_2) is K_1-aligned. Then $(x, \eta_2, \ldots, \eta_{2(N-k+1)})$ is K_1-aligned and Proposition 3.12 tells us that $(x, \eta_{2(N-k+1)}) = (x, \Gamma_N(s'))$ is K_2-aligned. Hence, s' satisfies Condition (5).

Similarly, by considering the K_0-aligned sequence

$$(\bar{\Gamma}_N(s), \bar{\Gamma}_{N-1}(s), \ldots, \bar{\Gamma}_1(s), \Gamma_{-1}(s'), \Gamma_{-2}(s'), \ldots, \Gamma_{-N}(s')),$$

we can deduce that $(x, \Gamma_{-N}(s'))$ is K_2-aligned as desired.

A similar argument leads to the following.

Observation 3.22. If $s \in S_0^{(m)}$ violates Condition (6), then all the other elements in $S_0^{(m)}$ satisfy Conditions (5) and (6).

Step 3: *Summary*. We claim that $S_0^{(m)}$ is $((K_0L_0 + K_0)N + K_2 + K_4)$-Schottky. The first and the third requirements for Schottky sets were already observed before, so it remains to discuss the second requirement. Considering Observation 3.20, it suffices to show that Conditions (5) and (6) are satisfied by all but at most 1 element of $S_0^{(m)}$. Observations 3.21 and 3.22 imply that this is the case.

Given these observations, we can finish the proof by taking $K = (K_0L + K_0)N + K_2 + K_4$, $m = L$ and by taking any subset $S \subseteq S_0^{(m)}$ such that $\#S = N$. □

4. Pivoting and limit laws

In this section, we establish the notion of pivotal times and pivoting. We will then deduce CLT, LIL and geodesic tracking of random walks using probabilistic estimates about pivotal times. The proof of a key probabilistic estimate will be postponed to § 5.

4.1 Pivotal times: statement

Let μ be a non-elementary probability measure on G and let S be a long enough and large Schottky set for μ. Then for sufficiently small $\epsilon > 0$, an n-step random path $(g_1, \ldots, g_n)$ in the μ-random walk contains at least ϵn subsegments

$$(g_{j(i)-M_0+1}, \ldots, g_{j(i)}) \in S \quad (i = 1, \ldots, \epsilon n).$$

The appearance of Schottky sequences in a random path does not necessarily imply something about $Z_n = g_1 \cdots g_n$. For example, every Schottky sequence might be canceled out with the next step, resulting in $Z_n = \mathrm{id}$. We nonetheless claim that, for a high probability, a certain number of Schottky axes survive. More explicitly, we seek indices $j(1) < \cdots < j(M)$, called the *pivotal times*, such that the Schottky axes arising at these indices are aligned along $[o, Z_n o]$:

$$(o, \mathbf{Y}_{j(1)}, \ldots, \mathbf{Y}_{j(M)}, Z_n o) \text{ is aligned}, \quad \text{where } \mathbf{Y}_{j(k)} = (Z_{j(k)-M_0} o, \ldots, Z_{j(k)} o).$$

We will observe that, for a high probability, a random path has sufficiently many pivotal times. Then, we will freeze the steps except at the pivotal slots and choose the Schottky sequences at the pivotal times from S. More explicitly, we will realize a structure where $\mathbf{Y}_{j(k)}$ are identical and independent distributions (i.i.d.s) on the uniform measure on $\{\Gamma(s) : s \in S\}$: once this is guaranteed, we can control the direction $[o, Z_n o]$ and establish the deviation inequality.

We now formulate the discussion above.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

DEFINITION 4.1. Let μ be a non-elementary probability measure on G, let $(\Omega, \mathbb{P})$ be a probability space for μ, let $K_0, M_0 > 0$ and let S be a long enough K_0-Schottky set contained in $(\mathrm{supp}\mu)^{M_0}$, i.e., M_0 is as large as described in Definition 3.16.

A subset $\mathcal{E}$ of Ω, accompanied by the choice of a subset $\mathcal{P}(\mathcal{E}) = \{j(1) < j(2) < \cdots\} \subseteq M_0\mathbb{Z}_{>0}$, is called a *pivotal equivalence class* if:

(i) for each $i \notin \{j(k) - l : k \geqslant 1, l = 0, \ldots, M_0 - 1\}$, $g_i(\omega)$ is fixed on $\mathcal{E}$;
(ii) for each $\omega \in \mathcal{E}$ and $k \geqslant 1$, the following is a Schottky sequence:

$$s_k(\omega) := (g_{j(k)-M_0+1}(\omega), g_{j(k)-M_0+2}(\omega), \ldots, g_{j(k)}(\omega)) \in S;$$

(iii) for each $\omega \in \mathcal{E}$, $(o, \mathbf{Y}_{j(1)}(\omega), \mathbf{Y}_{j(2)}(\omega), \ldots)$ is D_0-semi-aligned; and
(iv) on $\mathcal{E}$, $\{s_1(\omega), s_2(\omega), \ldots\}$ are i.i.d.s distributed according to the uniform measure on S.

We say that $\mathcal{P}(\mathcal{E})$ is the *set of pivotal times* for $\mathcal{E}$.

When a pivotal equivalence class $\mathcal{E} \subseteq \Omega$ is understood, with the set of pivotal times $\mathcal{P}(\mathcal{E})$, for each element ω of $\mathcal{E}$ we call $\mathcal{P}(\mathcal{E})$ the *set of pivotal times for* ω and write it as $\mathcal{P}(\omega)$.

When the probability space $(\Omega, \mathbb{P})$ for μ is partitioned into pivotal equivalence classes $\{\mathcal{E}_\alpha\}_\alpha$, then belonging to the same $\mathcal{E}_\alpha$ becomes an equivalence relation. Choosing a different element from the same pivotal equivalence class is called *pivoting*. But note that the choice of pivotal equivalence classes is not canonical: given an $\omega \in \Omega$, there are several ways to define the pivotal equivalence class for ω. Proposition 4.2 below describes a particular choice of pivotal equivalence classes that will be useful.

Let k be a positive integer. We say that a pivotal equivalence class $\mathcal{E}$ *avoids* k if k is not in $\{j - l : j \in \mathcal{P}(\mathcal{E}),\ l = 0, \ldots, M_0 - 1\}$; in this case, g_k is fixed on $\mathcal{E}$.

PROPOSITION 4.2. *Let μ be a non-elementary probability measure on G and let S be a long enough and large Schottky set for μ. Then there exist a probability space $(\Omega, \mathbb{P})$ for μ and a constant $K > 0$ such that, for each $n \geqslant 0$, we have a measurable partition $\mathscr{P}_n = \{\mathcal{E}_\alpha\}_\alpha$ of Ω into pivotal equivalence classes avoiding $1, \ldots, \lfloor n/2 \rfloor + 1$ and $n + 1$ that satisfies*

$$\mathbb{P}(\omega : \#(\mathcal{P}(\omega) \cap \{1, \ldots, k\}) \leqslant k/K \mid g_1, \ldots, g_{\lfloor n/2 \rfloor + 1}, g_{n+1}) \leqslant K e^{-k/K}, \tag{7}$$

for each choice of $g_1, \ldots, g_{\lfloor n/2 \rfloor + 1}, g_{n+1} \in G$ and $k \geqslant n$.

We postpone the proof of Proposition 4.2 to the next section and first see its consequence.

4.2 Pivoting

Let $K_0, N_0 > 0$ and let S be a long enough K_0-Schottky set with cardinality N_0. Given isometries u_i, let us draw a choice $s = (s_1, s_2, \ldots, s_n)$ from S^n with the uniform measure and define

$$U_n = u_0 \Pi(s_1) u_1 \Pi(s_2) u_2 \cdots \Pi(s_n) u_n.$$

Let $\kappa_i := U_{i-1}\Gamma^+(s_i) = u_0 \Pi(s_1) \cdots u_{i-1} \Gamma^+(s_i)$. We claim the following.

LEMMA 4.3. *We have*

$$\mathbb{P}((x, \kappa_i) \text{ is } K_0\text{-aligned for some } i \leqslant k) \geqslant 1 - (1/N_0)^k,$$
$$\mathbb{P}((\kappa_{n-i+1}, U_n x) \text{ is } K_0\text{-aligned for some } i \leqslant k) \geqslant 1 - (1/N_0)^k,$$

for each $1 \leqslant k \leqslant n$ and $x \in X$.

Proof. We prove the first estimate only; the second one follows similarly. Consider the statement

$$(u_0^{-1} x, \Gamma^+(s_1)) \quad \text{is } K_0\text{-aligned}.$$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Thanks to the Schottky property, at most one choice of s_1 from S violates this statement. Fixing that bad choice, consider the statement

$$((u_0\Pi(s_1)u_1)^{-1}x, \Gamma^+(s_2)) \quad \text{is } K_0\text{-aligned.}$$

Again, at most one choice of s_2 from S violates this. Keeping this manner, we conclude the following: except at most one bad choice among S^k

$$((u_0\Pi(s_1)\cdots u_{i-1})^{-1}x, \Gamma^+(s_i)) \text{ is } K_0\text{-aligned}, \quad \text{i.e., } (x,\kappa_i) \text{ is } K_0\text{-aligned},$$

holds for at least one $i \leqslant k$. This happens for probability at least $1-(1/N_0)^k$. □

Now fix another set of isometries $\check{u}_i$ and another K_0-Schottky set $\check{S}$ with cardinality N_0. We draw $\check{s} = (\check{s}_1, \check{s}_2, \ldots, \check{s}_n)$ from $\check{S}^n$ with the uniform measure, independently from s, and define

$$\check{U}_n = \check{u}_0\Pi(\check{s}_1)\check{u}_1\cdots\Pi(\check{s}_n)\check{u}_n.$$

Let $\eta_i := \check{U}_{i-1}\Gamma^+(\check{s}_i)$. Recall that $\bar{\eta}_i$ denotes the reversal of η_i.

Lemma 4.4. *We have*

$$\mathbb{P}((\bar{\eta}_i,\kappa_i) \textit{ is } D_0\textit{-aligned for some } i \leqslant k) \geqslant 1-(2/N_0)^k,$$

for each $1 \leqslant k \leqslant n$.

Proof. Consider the statements

$$(u_0^{-1}\check{u}_0\cdot o,\ \Gamma^+(s_1)) \quad \text{is } K_0\text{-aligned},$$
$$(\check{u}_0^{-1}u_0\Pi(s_1)\cdot o,\ \Gamma^+(\check{s}_1)) \quad \text{is } K_0\text{-aligned.}$$

Thanks to the Schottky property, at most one choice of s_1 from S violates the first statement. Similarly, given s_1, at most one choice of $\check{s}_1$ from $\check{S}$ violates the second statement. In short, the two statements hold for all but at most $2N_0$ choices of $(s_1,\check{s}_1) \in S\times\check{S}$.

Fixing a bad choice $(s_1,\check{s}_1)$, consider the statements

$$((u_0\Pi(s_1)u_1)^{-1}\check{u}_0\Pi(\check{s}_1)\check{u}_1\cdot o,\ \Gamma^+(s_2)) \quad \text{is } K_0\text{-aligned},$$
$$((\check{u}_0\Pi(\check{s}_1)\check{u}_1)^{-1}u_0\Pi(s_1)u_1\Pi(s_2)\cdot o,\ \Gamma^+(\check{s}_2)) \quad \text{is } K_0\text{-aligned.}$$

Again, at most $2N_0$ choices of $(s_2,\check{s}_2)\in S\times\check{S}$ violate the statements. Keeping this manner, we conclude the following: for probability at least $1-(2/N_0)^k$, there exists $i \leqslant k$ such that

$$((u_0\Pi(s_1)\cdots u_{i-1})^{-1}\check{u}_0\Pi(\check{s}_1)\cdots\check{u}_{i-1}\cdot o,\ \Gamma^+(s_i)) \quad \text{is } K_0\text{-aligned},$$
$$((\check{u}_0\Pi(\check{s}_1)\cdots\check{u}_{i-1})^{-1}u_0\Pi(s_1)\cdots u_{i-1}\Pi(s_i)\cdot o,\ \Gamma^+(\check{s}_i)) \quad \text{is } K_0\text{-aligned.}$$

In other words, (ending point of $\bar{\eta}_i,\kappa_i$) and ($\bar{\eta}_i$, ending point of κ_i) are K_0-aligned. Lemma 3.8 then tells us that $(\bar{\eta}_i,\kappa_i)$ is D_0-aligned. □

Applying Lemmas 4.3 and 4.4 to pivotal equivalence classes, we obtain the following corollaries.

Corollary 4.5. *Let* μ *be a non-elementary probability measure on* G*, let* $K_0, N_0 > 0$*, and let* S *be a long enough* K_0*-Schottky set for* μ *with cardinality* N_0*. Let* $\mathcal{E}$ *be a pivotal equivalence class for* μ *with* $\mathcal{P}(\mathcal{E}) = \{j(1) < j(2) < \cdots\}$ *and let* $x \in X$*. Then for each* $k \geqslant 1$ *we have*

$$\mathbb{P}((x, \mathbf{Y}_{j(k)}(\omega), \mathbf{Y}_{j(k+1)}(\omega), \ldots) \textit{ is } D_0\textit{-semi-aligned} \,|\, \mathcal{E}) \geqslant 1-(1/N_0)^k.$$

Moreover, for any $m \geqslant 1$, $n \geqslant j(m)$ *and* $k = 1, \ldots, m$*, we have*

$$\mathbb{P}((\mathbf{Y}_{j(1)}(\omega), \ldots, \mathbf{Y}_{j(m-k+1)}(\omega), Z_n(\omega)o) \textit{ is } D_0\textit{-semi-aligned} \,|\, \mathcal{E}) \geqslant 1-(1/N_0)^k.$$

COROLLARY 4.6. *Let μ be a non-elementary probability measure on G and let $\check{\mu}$ be its reflected version, let $K_0, N_0 > 0$ and let S and $\check{S}$ be long enough K_0-Schottky sets for μ and $\check{\mu}$, respectively, with cardinality N_0. Let $\mathcal{E}$ be a pivotal equivalence class for μ with $\mathcal{P}(\mathcal{E}) = \{j(1) < j(2) < \cdots\}$, and let $\check{\mathcal{E}}$ be a pivotal equivalence class for $\check{\mu}$ with $\mathcal{P}(\check{\mathcal{E}}) = \{\check{j}(1) < \check{j}(2) < \cdots\}$. Then we have*

$$\mathbb{P}((\bar{\mathbf{Y}}_{j(k)}(\omega), \mathbf{Y}_{\check{j}(k)}(\check{\omega})) \text{ is } D_0\text{-semi-aligned} \,|\, \mathcal{E}) \geqslant 1 - (2/N_0)^k \quad (\forall k > 0).$$

We now record a small consequence of pivoting.

COROLLARY 4.7. *Let $(Z_n)_{n>0}$ be the random walk generated by a non-elementary probability measure μ on G with finite first moment. Then there exists a strictly positive quantity $\lambda(\mu) \in (0, +\infty]$, called the drift of μ, such that*

$$\lambda(\mu) := \lim_{n\to\infty} \tfrac{1}{n} d(o, Z_n o) \quad \text{almost surely.}$$

Remark 4.8. The statement in Corollary 4.7 holds true even without the moment condition. This will be the consequence of Theorem 6.4 in § 6.

Proof. By Kingman's subadditive ergodic theorem, $\lambda(\mu) = \lim_n (1/n) d(o, Z_n o)$ exists and is constant almost surely. It remains to show that $\lambda(\mu) > 0$.

Since μ is non-elementary, Proposition 3.19 provides a long enough and large Schottky set S for μ. Given this, Proposition 4.2 provides a constant $K > 0$ and a measurable partition $\mathscr{P} = \{\mathcal{E}_\alpha\}_\alpha$ into pivotal equivalence classes such that

$$\mathbb{P}(\omega : \#(\mathcal{P}(\omega) \cap \{1, \ldots, k\}) \leqslant k/K) \leqslant K e^{-k/K},$$

for each k. Now let $n > 0$ and let $\mathcal{E}$ be a pivotal equivalence class with $\mathcal{P}(\mathcal{E}) = \{j(1) < j(2) < \cdots\}$ such that $\#(\mathcal{P}(\mathcal{E}) \cap \{1, \ldots, n\}) \geqslant n/K$, i.e., $j(\lfloor n/K \rfloor) \leqslant n$. Corollary 4.5 tells us that

$$\mathbb{P}((o, \mathbf{Y}_{j(1)}(\omega), \ldots, \mathbf{Y}_{j(\lfloor n/2K \rfloor)}(\omega), Z_n o) \text{ is } D_0\text{-semi-aligned} \,|\, \mathcal{E}) \geqslant 1 - (1/\#S_0)^{-n/2K+1}.$$

By Lemma 3.18, we then have

$$\mathbb{P}(d(o, Z_n o) < 50 E_0 n/2K \,|\, \mathcal{E}) \leqslant (1/\#S_0)^{-n/2K+1}.$$

We sum up these conditional probabilities on $\{\omega : \#(\mathcal{P}(\mathcal{E} \cap \{1, \ldots, n\}) \geqslant n/K\}$ to conclude

$$\mathbb{P}(d(o, Z_n o) < 50 E_0 n/2K) \leqslant (1/\#S_0)^{-n/2K+1} + K e^{-n/K}.$$

The Borel–Cantelli lemma then implies $d(o, Z_n o) \geqslant 50 E_0 n/2K$ eventually almost surely. $\square$

4.3 Deviation inequality

Let μ be a non-elementary probability measure on G and let S be a long enough and large K_0-Schottky set contained in $(\mathrm{supp}\mu)^{M_0}$ for some $K_0, M_0 > 0$. Consider a bi-infinite path $((Z_n(\omega))_{n>0}, (Z_n(\check{\omega}))_{n>0})$ arising from the random walk generated by μ. Recall:

$$\mathbf{Y}_i(\omega) := (Z_{i-M_0} o, Z_{i-M_0+1} o, \ldots, Z_i o),$$
$$\mathbf{Y}_i(\check{\omega}) := (\check{Z}_{i-M_0} o, \check{Z}_{i-M_0+1} o, \ldots, \check{Z}_i o).$$

For each $k \geqslant M_0$, we investigate whether there exists $M_0 \leqslant i \leqslant k$ such that:

(i) $(g_{i-M_0+1}, \ldots, g_i)$ is a Schottky sequence;
(ii) $(\check{Z}_m o, \mathbf{Y}_i(\omega), Z_n o)$ is D_0-semi-aligned for all $n \geqslant k$ and $m \geqslant 0$.

We define $\upsilon = \upsilon(\check{\omega}, \omega)$ as the minimal index k with the auxiliary index $i \leqslant k$ as described above. See Figure 3.

A motivating observation for the definition of $\upsilon(\check{\omega}, \omega)$ is as follows.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

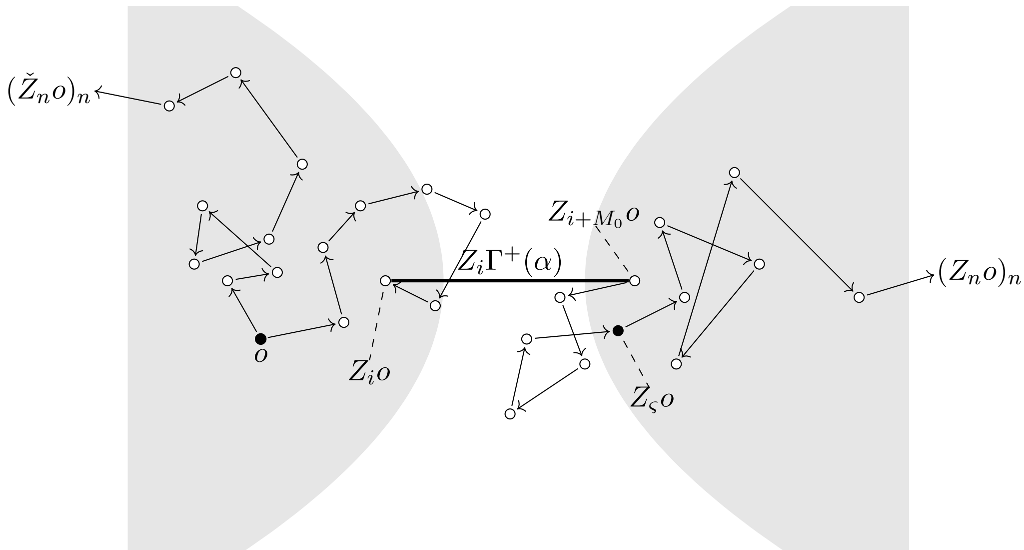

Figure 3. Persistent progress and υ. Here, all of the backward loci $(\check{Z}_n o)_{n \geqslant 0}$ are on the left of the persistent progress $Z_i \Gamma^+(\alpha)$, while the forward loci after $Z_\varsigma o$ are all on the right.

Lemma 4.9. *Let $\Omega = G^{\mathbb{Z}_{>0}} \times G^{\mathbb{Z}_{>0}}$ be the space of (bi-directional) step paths in G, let $K_0 > 0$ and let S be a long enough K_0-Schottky set. Then for each $(\check{\omega}, \omega) \in \Omega$, we have*

$$(\check{Z}_m o, Z_n o)_o \leqslant d(o, Z_k o),$$

for all $m \geqslant 0$ and $n, k \geqslant \upsilon(\check{\omega}, \omega)$.

Proof. Let $i \leqslant \upsilon(\check{\omega}, \omega)$ be the index such that $(\check{Z}_{m'} o, \mathbf{Y}_i(\omega), Z_{n'} o)$ is D_0-semi-aligned for all $n' \geqslant \upsilon(\check{\omega}, \omega)$ and $m' \geqslant 0$. Lemma 3.18 tells us that

$$\begin{aligned} d(\check{Z}_{m'} o, Z_{n'} o) &\geqslant d(\check{Z}_{m'} o, Z_{i-M_0} o) + d(Z_{i-M_0} o, Z_i o) + d(Z_i o, Z_{n'} o) - E_0 \\ d(\check{Z}_{m'} o, Z_{n'} o) &\geqslant d(\check{Z}_{m'} o, Z_{i-M_0} o) + d(Z_i o, Z_{n'} o) + 50 E_0. \end{aligned} \qquad (n' \geqslant \upsilon(\check{\omega}, \omega), m' \geqslant 0)$$

Let us now pick $n, k \geqslant \upsilon(\check{\omega}, \omega)$ and $m \geqslant 0$. Then we have

$$\begin{aligned} d(\check{Z}_m o, Z_n o) &\geqslant d(\check{Z}_m o, Z_{i-M_0} o) + d(Z_{i-M_0} o, Z_i o) + d(Z_i o, Z_n o) - E_0 \\ &\geqslant d(\check{Z}_m o, Z_{i-M_0} o) + d(Z_{i-M_0} o, Z_n o) - E_0, \\ d(o, Z_k o) &\geqslant d(o, Z_{i-M_0} o) + d(Z_i o, Z_k o) + 50 E_0 \geqslant d(o, Z_{i-M_0} o) + 50 E_0. \end{aligned} \tag{8}$$

Hence

$$\begin{aligned} 2(\check{Z}_m o, Z_n o)_o &= d(\check{Z}_m o, o) + d(o, Z_n o) - d(\check{Z}_m o, Z_n o) \\ &\leqslant (d(\check{Z}_m o, Z_{i-M_0} o) + d(Z_{i-M_0} o, o)) + (d(o, Z_{i-M_0} o) + d(Z_{i-M_0} o, Z_n o)) \\ &\quad - (d(\check{Z}_m o, Z_{i-M_0} o) + d(Z_{i-M_0} o, Z_n o) - E_0) \\ &\leqslant 2d(o, Z_{i-M_0} o) + E_0 \leqslant 2d(o, Z_k o). \end{aligned}$$

$\square$

We now provide a probabilistic estimate for $\upsilon(\check{\omega}, \omega)$.

Lemma 4.10. *Let μ be a non-elementary probability measure on G, let $K_0 > 0$ and let S be a long enough and large K_0-Schottky set for μ. Then there exists $K' > 0$ such that*

$$\mathbb{P}(\upsilon(\check{\omega}, \omega) \geqslant k \mid g_{k+1}, \check{g}_1, \ldots, \check{g}_{k+1}) \leqslant K' e^{-k/K'} \tag{9}$$

holds for all $k \geqslant 0$ and all choices of $g_{k+1}, \check{g}_1, \ldots, \check{g}_{k+1} \in G$.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Proof. Let S be a long enough and large K_0-Schottky set in $(\mathrm{supp}\mu)^{M_0}$ for some $M_0 > 0$. Let $\check{S}$ be the reflected version of S, that means

$$\check{S} := \{(s_{M_0}^{-1}, \ldots, s_1^{-1}) : (s_1, \ldots, s_{M_0}) \in S\}.$$

Then $\check{S}$ is a long enough and large K_0-Schottky set for $\check{\mu}$. Let $K > 0$ be the constant determined for S and $\check{S}$ in Proposition 4.2. We now fix k and $g_{k+1}, \check{g}_1, \ldots, \check{g}_{k+1} \in G$.

Let $\mathscr{P}_k = \{\mathcal{E}_\alpha\}_\alpha$ be the partition of Ω into pivotal equivalence classes avoiding $1, \ldots, \lfloor k/2 \rfloor + 1$ and $k+1$, given by Proposition 4.2. Let also $\check{\mathscr{P}}_{2k} = \{\check{\mathcal{E}}_\alpha\}_\alpha$ be the partition of $\check{\Omega}$ into pivotal equivalence classes avoiding $1, \ldots, k+1$ and $2k+1$, given by Proposition 4.2. We have

$$\mathbb{P}(A := \{\omega : \#(\mathcal{P}(\omega) \cap \{1, \ldots, n\}) \geqslant n/K \quad \text{for all } n \geqslant k\}) \geqslant 1 - \frac{K}{1 - e^{-1/K}} e^{-k/K},$$
$$\mathbb{P}(\check{A} := \{\check{\omega} : \#(\mathcal{P}(\check{\omega}) \cap \{1, \ldots, n\}) \geqslant n/K \quad \text{for all } n \geqslant 2k\}) \geqslant 1 - \frac{K}{1 - e^{-1/K}} e^{-2k/K}.$$

Let us enumerate $\mathcal{P}(\omega)$ by $\{j(1) < j(2) < \cdots\}$, and $\mathcal{P}(\check{\omega})$ by $\{\check{j}(1) < \check{j}(2) < \cdots\}$. Let $\mathcal{E} \in \mathscr{P}_k$ and $\check{\mathcal{E}} \in \check{\mathscr{P}}_{2k}$ be pivotal equivalence classes in A and $\check{A}$, respectively. In $\check{\mathcal{E}} \times \mathcal{E}$, let B be the set of $(\check{\omega}, \omega)$ that satisfies the following:

(i) for $x \in \{o, \check{Z}_1 o, \ldots, \check{Z}_{2k} o\}$, the following sequence is D_0-semi-aligned:
$$(x, \mathbf{Y}_{j(\lceil k/3K \rceil)}(\omega), \mathbf{Y}_{j(\lceil k/3K \rceil + 1)}(\omega), \ldots);$$
(ii) for each $n \geqslant k$ and $m \geqslant 2k$, the following are D_0-semi-aligned:
$$(o, \mathbf{Y}_{j(1)}(\omega), \mathbf{Y}_{j(2)}(\omega), \ldots, \mathbf{Y}_{j(\lceil 2n/3K \rceil)}(\omega), Z_n o),$$
$$(o, \mathbf{Y}_{\check{j}(1)}(\check{\omega}), \mathbf{Y}_{\check{j}(2)}(\check{\omega}), \ldots, \mathbf{Y}_{\check{j}(\lceil 2m/3K \rceil)}(\check{\omega}), \check{Z}_m o);$$
(iii) $(\bar{\mathbf{Y}}_{\check{j}(i)}(\check{\omega}),\ \mathbf{Y}_{j(i)}(\omega))$ is D_0-aligned for some $i \leqslant k/3K$.

The first item is handled by Lemma 4.3: it holds for probability at least $1 - 2k \cdot (1/400)^{k/3K}$.

Next, recall that, for each $n \geqslant k$, there are at least n/K pivotal times for $\mathcal{E}$ before n. Also, for each $m \geqslant 2k$, there are at least m/K pivotal times for $\check{\mathcal{E}}$ before m. Hence, we can apply Lemma 4.3 and deduce that the following are D_0-semi-aligned:

$$(\mathbf{Y}_{j(1)}(\omega), \ldots, \mathbf{Y}_{j(\lceil 2n/3K \rceil)}(\omega), Z_n o), \quad (\mathbf{Y}_{\check{j}(1)}(\check{\omega}), \ldots, \mathbf{Y}_{\check{j}(\lceil 2m/3K \rceil)}(\check{\omega}), \check{Z}_m o),$$

for probabilities at least $1 - (1/400)^{n/3K-1}$ and $1 - (1/400)^{m/3K-1}$, respectively. Taking intersection for $n \geqslant k$ and $m \geqslant 2k$, we observe that item (ii) holds for probability at least $1 - 3 \cdot (1/400)^{k/3K-1}$.

Finally, item (iii) is handled by Lemma 4.4: it holds for probability at least $1 - (1/200)^{k/3K-1}$. Combining these, we deduce

$$\mathbb{P}(B \mid \check{\mathcal{E}} \times \mathcal{E}) \geqslant 1 - (2k+4) \cdot (1/200)^{k/3K-1} \geqslant 1 - 200 \cdot (2k+4) \cdot 0.01^{k/3K}.$$

It remains to prove that $v(\check{\omega}, \omega) \leqslant k$ for $(\check{\omega}, \omega) \in B$. First, by definition of A, $j(\lceil k/3K \rceil)$ is smaller than k and

$$s_{\lceil k/3K \rceil} = (g_{j(\lceil k/3K \rceil) - M_0 + 1}, \ldots, g_{j(\lceil k/3K \rceil)})$$

is Schottky. Next, for each $n \geqslant k$, $(o, \mathbf{Y}_{j(1)}(\omega), \mathbf{Y}_{j(2)}(\omega), \ldots, \mathbf{Y}_{j(\lceil 2n/3K \rceil)}(\omega), Z_n o)$ is D_0-semi-aligned. Hence, $(o, \mathbf{Y}_{j(\lceil k/3K \rceil)}(\omega), Z_n o)$ is also D_0-semi-aligned.

We now investigate the alignment of $(\check{Z}_m o, \mathbf{Y}_{j(\lceil k/3K \rceil)}(\omega))$. For $m \leqslant 2k$, this is guaranteed by item (1). When $m \geqslant 2k$, we appeal to items (2) and (3). Namely, the sequence

$$(\check{Z}_m o, \bar{\mathbf{Y}}_{\check{j}(\lceil 2m3/ \rceil)}(\check{\omega}), \ldots, \bar{\mathbf{Y}}_{\check{j}(i+1)}(\check{\omega}), \bar{\mathbf{Y}}_{\check{j}(i)}(\check{\omega}), \mathbf{Y}_{j(i)}(\omega), \ldots, \mathbf{Y}_{j(\lceil k/3K \rceil)}(\omega), \mathbf{Y}_{j(\lceil k/3K \rceil + 1)}, \ldots)$$

is D_0-semi-aligned. In particular, $(\check{Z}_m o, \bar{\mathbf{Y}}_{j(\lceil k/3K \rceil)}(\omega))$ is D_0-semi-aligned. $\square$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Here is a corollary of Lemma 4.10 that we will use in § 6.

Corollary 4.11 [Gou22, Lemma 4.14]. *Let μ be a non-elementary probability measure on G and let $(Z_n)_n$ be the random walk generated by μ. Then for each $\epsilon > 0$, there exists $C > 0$ such that*

$$\mathbb{P}(d(o, gZ_n o) \geqslant d(o, go) - C \text{ for all } n \geqslant 0) \geqslant 1 - \epsilon/2 \quad (\forall\, g \in G).$$

Proof. Let us pick $K_0 > 0$ and a long enough and large K_0-Schottky set S for μ. Let K' be the constant as in Lemma 4.10. Given $\epsilon > 0$, we take $N > 1$ large enough so that $K' e^{-N/K'} \leqslant \epsilon/4$. Then, the definition of the RV $\upsilon(\check{\omega}, \omega)$ and Lemma 4.10 tell us that

$$\mathbb{P}\left(\begin{matrix} \text{there exists } i < N \text{ such that } \mathbf{Y}_i \text{ is a Schottky axis and} \\ (g^{-1}o, \mathbf{Y}_i, Z_n o) \text{ is } D_0\text{-semi-aligned for each } n \geqslant N \end{matrix} \,\middle|\, \check{g}_1 = g^{-1} \right) \geqslant 1 - \epsilon/4.$$

When $(g^{-1}o, \mathbf{Y}_i, Z_n o)$ is D_0-semi-aligned, the second inequality in Lemma 3.18(ii) implies

$$d(g^{-1}o, Z_n o) \geqslant d(g^{-1}o, Z_{i-M_0} o) + 50E_0 N \geqslant d(g^{-1}o, o) - d(Z_{i-M_0}o, o) \geqslant d(o, go) - \sum_{j=1}^{N} d(o, g_j o).$$

This bound also holds for $n \leqslant N$

$$d(g^{-1}o, Z_n o) \geqslant d(g^{-1}o, o) - d(Z_n o, o) \geqslant d(o, go) - \sum_{j=1}^{N} d(o, g_j o).$$

Given these, the proof ends by taking large enough $C > 0$ such that

$$\mathbb{P}\left(\sum_{j=1}^{N} d(o, g_j o) \geqslant C \right) \leqslant \epsilon/4. \qquad \square$$

Corollary 4.12. *Let μ be a non-elementary probability measure on G whose expectation is infinite. Then μ^{*m} has infinite expectation for each $m > 0$. In particular, the drift $\lambda(\mu) := \lim_{m\to\infty}(1/m)\mathbb{E}_{\mu^{*m}}[d(o, go)]$ is infinity.*

Proof. Let $\epsilon = 0.2$ and let $C = C(\mu, \epsilon)$ be as in Corollary 4.11. Let $(g_1, \ldots, g_m)$ be distributed according to μ^m. Then by Corollary 4.11, we have

$$\mathbb{E}[d(o, g_1 g_2 \cdots g_m o) | g_1 = g] \geqslant \mathbb{E}[(d(o, go) - C) \cdot 1_{\{d(o, gg_2\cdots g_m o) \geqslant d(o,go)\}} | g_1 = g] \geqslant 0.9 \cdot (d(o, go) - C).$$

Now integrating over $g_1 \in \mathrm{supp}\mu$ with law μ, we get

$$\mathbb{E}[d(o, g_1 g_2 \cdots g_m o)] \geqslant 0.9\mathbb{E}_\mu[d(o, go) - C] = +\infty. \qquad \square$$

Similarly, fixing the Schottky set S for μ, we similarly define $\check{\upsilon} = \check{\upsilon}(\check{\omega}, \omega)$ as the minimal index k that is associated with another index $i \leqslant k$ such that:

(i) $(\check{g}_i^{-1}, \ldots, \check{g}_{i-M_0+1}^{-1})$ is a Schottky sequence;
(ii) $(\check{Z}_m o, \bar{\mathbf{Y}}_i(\check{\omega}), o)$ is D_0-semi-aligned for all $m \geqslant k$; and
(iii) $(\bar{\mathbf{Y}}_i(\check{\omega}), Z_n o)$ is D_0-semi-aligned for all $n \geqslant 0$.

Then we similarly have

$$\mathbb{P}(\check{\upsilon}(\check{\omega}, \omega) \geqslant k \,|\, \check{g}_{k+1}, g_1, \ldots, g_{k+1}) \leqslant K' e^{-k/K'}. \tag{10}$$

Thanks to these exponential bounds, we can establish the deviation inequality.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

PROPOSITION 4.13. *Let $p > 0$ and let $((\check{Z}_n)_n, (Z_n)_n)$ be the (bi-directional) random walk generated by a non-elementary probability measure μ on G with finite pth moment. Then the random variable $\sup_{n,m\geqslant 0}(\check{Z}_m o, Z_n o)_o$ has finite 2pth moment.*

Note the difference between this proposition and [Cho23, Propositions 5.6, 5.8]; we are taking the global suprema, not the limit suprema.

Proof. Let K' be the constant for μ as in Lemma 4.10 and let

$$D_k := \sum_{i=1}^{k} d(o, g_i o), \quad \check{D}_k := \sum_{i=1}^{k} d(o, \check{g}_i o).$$

By triangle inequality, $d(o, Z_k o) < D_l$ and $d(o, \check{Z}_k o) \leqslant \check{D}_l$ for all $k \leqslant l$. We begin by claiming

$$\sup_{n,m\geqslant 0}(\check{Z}_m o, Z_n o)_o^{2p} \leqslant \sum_{i=0}^{\infty} |\check{D}_{i+1}^p D_{i+1}^p - \check{D}_i^p D_i^p| (1_{\check{D}_i \geqslant D_i} 1_{i<\upsilon} + 1_{\check{D}_i \leqslant D_i} 1_{i<\check{\upsilon}}) \quad \text{almost surely.} \tag{11}$$

Since $\mathbb{P}(\max\{\upsilon, \check{\upsilon}\} \geqslant k)$ is summable by inequality (9) and (10), the Borel–Cantelli lemma implies that

$$l := \min\{i : 1_{\check{D}_i \geqslant D_i} 1_{i<\upsilon} + 1_{\check{D}_i \leqslant D_i} 1_{i<\check{\upsilon}} = 0\} < +\infty \quad \text{almost surely.}$$

Note that the right-hand side of inequality (11) is at least $\check{D}_l^p D_l^p$.

Now at $i = l$, we have either $\check{D}_l \geqslant D_l$ or $\check{D}_l \leqslant D_l$. In the first case $l \geqslant \upsilon$ must hold. Then for $m \geqslant 0$ and $n \geqslant l$, we have

$$(\check{Z}_m o, Z_n o)_o^{2p} \leqslant d(o, Z_l o)^{2p} \leqslant D_l^{2p} \leq \check{D}_l^p D_l^p,$$

by Lemma 4.9. Moreover, for $m \geqslant 0$ and $n \leqslant l$, we have

$$(\check{Z}_m o, Z_n o)_o^{2p} \leqslant d(o, Z_n o)^{2p} \leqslant D_n^{2p} \leqslant D_l^{2p} \leqslant \check{D}_l^p D_l^p.$$

In the second case $l \geqslant \check{\upsilon}$ must hold, and for a similar reason $(\check{Z}_m o, Z_n o)_o^{2p}$ is dominated by $\check{D}_l^p D_l^p$. Inequality (11) now follows.

We now need a small observation.

Fact 4.14. For $s_1, s_2, t_1, t_2 \geqslant 0$, the following holds:

$$\begin{aligned} |t_1^p t_2^p - s_1^p s_2^p| &= |t_1^p(t_2^p - s_2^p) + (t_1^p - s_1^p)s_2^p| \\ &\leqslant 2^{2p}(|t_1 - s_1|^p + s_1^{p-n_p}|t_1 - s_1|^{n_p} + s_1^p) \cdot (|t_2 - s_2|^p + s_2^{p-n_p}|t_2 - s_2|^{n_p}) \\ &\quad + 2^p(|t_1 - s_1|^p + s_1^{p-n_p}|t_1 - s_1|^{n_p})s_2^p \quad (n_p = p \text{ if } 0 \leqslant p \leqslant 1, \ n_p = 1 \text{ otherwise}). \end{aligned}$$

Proof of Fact 4.14. The fact follows from the following inequality in [BQ16, Section 5.4]:

$$|t^p - s^p| \leqslant 2^p(|t - s|^p + s^{p-n_p}|t - s|^{n_p}) \quad (n_p = p \text{ if } 0 \leqslant p \leqslant 1, \ n_p = 1 \text{ otherwise}).$$

We give its proof for completeness. Assume $t \geqslant s$ without loss of generality. When $p \leqslant 1$, the concavity of $f(x) = x^p$ implies the inequality. When $p > 1$, we divide the cases. If $s < t/2$, then

$$t^p - s^p < t^p < (2(t - s))^p \leqslant 2^p|t - s|^p.$$

If $s \geqslant t/2$, then we have

$$\begin{aligned} t^p - s^p &= \int_s^t p x^{p-1}\, dx \leqslant \int_s^t p\Big(\frac{s}{t-s}(x - s) + s\Big)^{p-1} dx && \Big(\because \frac{s}{t-s} \geqslant 1\Big) \\ &= (t - s) \cdot p s^{p-1} \int_1^2 u^{p-1}\, du && \Big(u = \frac{1}{t-s}(x - s) + 1\Big) \\ &= (t - s)s^{p-1}(2^p - 1) \leqslant 2^p s^{p-1}(t - s). \end{aligned}$$

□

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

By Fact 4.14, the expectations of $|\check{D}_{i+1}^p D_{i+1}^p - \check{D}_i^p D_i^p|(1_{\check{D}_i \geqslant D_i} 1_{i<\upsilon} + 1_{\check{D}_i \leqslant D_i} 1_{i<\check{\upsilon}})$ for $i \geqslant 0$ are summable as soon as there exists $K'' > 0$ such that

$$\mathbb{E}[d(o,\check{g}_{i+1})^{n_1} d(o,g_{i+1})^{n_2} \check{D}_i^{p-n_1} D_i^{p-n_2} (1_{\check{D}_i \geqslant D_i} 1_{i<\upsilon} + 1_{\check{D}_i \leqslant D_i} 1_{i<\check{\upsilon}})] < K'' i^{2p+2} e^{-i/K''}, \tag{12}$$

for each $0 \leqslant n_1, n_2 \leqslant p$ with $n_1 + n_2 \geqslant \min(p,1)$. We discuss the case $n_2 > 0$; the other case $n_1 > 0$ can be handled in the same way.

We will take advantage of the fact that $\mathbb{E}[\check{D}_i^p D_i^p]$ is bounded. Namely, the expectation of $\check{D}_i^{p-n_1} D_i^{p-n_2}$ on the set $\{D_i > c\}$ is small for large c. Next, on the set $\{D_i \leqslant c\}$, we will bound the expectation of $\check{D}_i^{p-n_1} D_i^{p-n_2} 1_{D_i<c} 1_{i<\upsilon}$ by using the exponential bound on $\mathbb{P}(i<\upsilon)$ (that suppresses $D_i^{p-n_2} < c^{p-n_2}$) independent of the distribution of $\check{D}_i$.

We first discuss the term $\mathbb{E}[d(o,\check{g}_{i+1})^{n_1} d(o,g_{i+1})^{n_2} \check{D}_i^{p-n_1} D_i^{p-n_2} \cdot 1_{\check{D}_i \geqslant D_i} 1_{i<\upsilon}]$. Let us fix $\check{g}_{i+1}$ and g_{i+1} for the moment, and let $c := e^{i/2pK'}$. We then have a decomposition

$$\begin{aligned}
&\mathbb{E}[\check{D}_i^{p-n_1} D_i^{p-n_2} 1_{\check{D}_i \geqslant D_i} 1_{i<\upsilon} \mid \check{g}_{i+1}, g_{i+1}] \\
&\quad = \mathbb{E}[\check{D}_i^{p-n_1} D_i^{p-n_2} 1_{D_i>c} 1_{\check{D}_i \geqslant D_i} 1_{i<\upsilon} \mid \check{g}_{i+1}, g_{i+1}] + \mathbb{E}[\check{D}_i^{p-n_1} D_i^{p-n_2} 1_{D_i \leqslant c} 1_{\check{D}_i \geqslant D_i} 1_{i<\upsilon} \mid \check{g}_{i+1}, g_{i+1}].
\end{aligned} \tag{13}$$

The first term is controlled as follows:

$$\begin{aligned}
&\mathbb{E}[\check{D}_i^{p-n_1} D_i^{p-n_2} 1_{D_i>c} 1_{\check{D}_i \geqslant D_i} 1_{i<\upsilon} \mid \check{g}_{i+1}, g_{i+1}] \\
&\quad \leqslant \mathbb{E}[\check{D}_i^{p-n_1} D_i^{p-n_2} 1_{D_i>c} \mid \check{g}_{i+1}, g_{i+1}] \\
&\quad \leqslant \mathbb{E}[\check{D}_i^{p-n_1} D_i^{p} \cdot c^{-n_2}] \leqslant \mathbb{E}[\check{D}_i^{p-n_1}] \cdot \mathbb{E}[D_i^p] \cdot c^{-n_2} \qquad (\because D_i^{-n_2} \leqslant c^{-n_2}) \\
&\quad \leqslant i^{p-n_1+1} \mathbb{E}_\mu[d(o,go)^{p-n_1}] \cdot i^{p+1} \mathbb{E}_\mu[d(o,go)^p] \cdot c^{-n_2}.
\end{aligned}$$

In the final step, we used the following fact for each $r > 0$ and $i > 0$:

$$\mathbb{E}\left[\left(\sum_{j=1}^{i} d(o,g_j o)\right)^r\right] \leqslant \mathbb{E}\left[\left(i \cdot \max_{1 \leqslant j \leqslant i} d(o,g_j o)\right)^r\right] \leqslant \mathbb{E}\left[i^r \cdot \sum_{j=1}^{i} d(o,g_j o)^r\right] \leqslant i^{r+1} \mathbb{E}_\mu[d(o,go)^r]. \tag{14}$$

Next, we apply Lemma 4.10 to the second term of the right-hand side of Equation (13) and observe

$$\begin{aligned}
\mathbb{E}[\check{D}_i^{p-n_1} D_i^{p-n_2} 1_{D_i \leqslant c} 1_{\check{D}_i \geqslant D_i} 1_{i<\upsilon} \mid \check{g}_{i+1}, g_{i+1}] &\leqslant \mathbb{E}[\check{D}_i^{p-n_1} \cdot \mathbb{E}[c^{p-n_2} 1_{i<\upsilon} \mid \check{g}_1, \ldots, \check{g}_{i+1}, g_{i+1}]] \\
&\leqslant \mathbb{E}[\check{D}_i^{p-n_1} \cdot c^{p-n_2} \mathbb{P}[\upsilon > i \mid \check{g}_1, \ldots, \check{g}_{i+1}, g_{i+1}]] \\
&\leqslant i^{p-n_1+1} \mathbb{E}_\mu[d(o,go)^{p-n_1}] \cdot c^{p-n_2} \cdot K' e^{-i/K'}.
\end{aligned}$$

Here, c^{p-n_2} is dominated by $c^p = e^{i/2K'}$. Overall, we have

$$\begin{aligned}
&\mathbb{E}[\check{D}_i^{p-n_1} D_i^{p-n_2} 1_{\check{D}_i \geqslant D_i} 1_{i<\upsilon} | \check{g}_{i+1}, g_{i+1}] \\
&\quad \leqslant K' \mathbb{E}_\mu[d(o,go)^{p-n_1}](1 + \mathbb{E}_\mu[d(o,go)^p]) \cdot i^{2p} \max(e^{-i/2K'}, e^{-n_2 i/2pK'}).
\end{aligned}$$

We now multiply $d(o,\check{g}_{i+1})^{n_1} d(o,g_{i+1})^{n_2}$ and integrate. As a result, we observe

$$\begin{aligned}
&\mathbb{E}[d(o,\check{g}_{i+1})^{n_1} d(o,g_{i+1})^{n_2} \check{D}_i^{p-n_1} D_i^{p-n_2} \cdot 1_{\check{D}_i \geqslant D_i} 1_{i<\upsilon}] \\
&\quad = \mathbb{E}[d(o,g_{i+1}o)^{n_1} d(o,\check{g}_{i+1}o)^{n_2} \cdot \mathbb{E}[\check{D}_i^{p-n_1} D_i^{p-n_2} 1_{\check{D}_i \geqslant D_i} 1_{i<\upsilon} \mid \check{g}_{i+1}, g_{i+1}]] \\
&\quad \leqslant \mathbb{E}[d(o,g_{i+1}o)^{n_1} d(o,\check{g}_{i+1}o)^{n_2} \cdot \mathbb{E}_\mu[d(o,go)^{p-n_1}](1 + \mathbb{E}_\mu[d(o,go)^p]) \cdot i^{2p} K' e^{-(n_2/2(p+1)K')i}] \\
&\quad \leqslant C(\mu) \cdot i^{2p} K' e^{-(n_2/2(p+1)K')i},
\end{aligned}$$

for some constant $C(\mu) < +\infty$ determined by the distribution of μ, independent of i. Note that μ has finite qth moment for every $0 \leqslant q \leqslant p$ thanks to Jensen's inequality.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

We similarly deal with the term $\mathbb{E}[d(o,\check{g}_{i+1})^{n_1}d(o,g_{i+1})^{n_2}\check{D}_i^{p-n_1}D_i^{p-n_2}\cdot 1_{\check{D}_i\leqslant D_i}1_{i<\check{\upsilon}}]$. Fixing g_{i+1} and $\check{g}_{i+1}$ first, we split the expectation based on the dichotomy for $\check{D}_i$:

$$\begin{aligned}&\mathbb{E}[\check{D}_i^{p-n_1}D_i^{p-n_2}1_{\check{D}_i\leqslant D_i}1_{i<\check{\upsilon}}\mid \check{g}_{i+1},g_{i+1}]\\&\quad=\mathbb{E}[\check{D}_i^{p-n_1}D_i^{p-n_2}1_{\check{D}_i>c}1_{\check{D}_i\leqslant D_i}1_{i<\check{\upsilon}}\mid \check{g}_{i+1},g_{i+1}]\\&\qquad+\mathbb{E}[\check{D}_i^{p-n_1}D_i^{p-n_2}1_{\check{D}_i\leqslant c}1_{\check{D}_i\leqslant D_i}1_{i<\check{\upsilon}}\mid \check{g}_{i+1},g_{i+1}].\end{aligned}$$

Here, a crucial observation is that $\check{D}_i^{p-n_1}D_i^{p-n_2}1_{\check{D}_i>c}1_{\check{D}_i\leqslant D_i}1_{i<\check{\upsilon}}$ is dominated by $\check{D}_i^{p-n_1}D_i^{p-n_2}1_{D_i>c}$. The remaining step is analogous to the previous computations

$$\begin{aligned}&\mathbb{E}[\check{D}_i^{p-n_1}D_i^{p-n_2}1_{\check{D}_i\leqslant D_i}1_{i<\check{\upsilon}}\mid \check{g}_{i+1},g_{i+1}]\\&\quad\leqslant\mathbb{E}[\check{D}_i^{p-n_1}D_i^{p-n_2}1_{D_i>c}\mid \check{g}_{i+1},g_{i+1}]+\mathbb{E}[\check{D}_i^{p-n_1}D_i^{p-n_2}1_{\check{D}_i\leqslant c}1_{i<\check{\upsilon}}\mid \check{g}_{i+1},g_{i+1}]\\&\quad\leqslant\mathbb{E}[\check{D}_i^{p-n_1}D_i^{p}\cdot c^{-n_2}\mid \check{g}_{i+1},g_{i+1}]+\mathbb{E}[D_i^{p-n_2}\cdot\mathbb{E}[c^{p-n_1}1_{i<\check{\upsilon}}\mid \check{g}_{i+1},g_1,\ldots,g_{i+1}]]\\&\quad\leqslant\mathbb{E}[\check{D}_i^{p-n_1}]\cdot\mathbb{E}[D_i^{p}]\cdot c^{-n_2}+\mathbb{E}[D_i^{p-n_2}]\cdot c^{p-n_1}\mathbb{P}[\check{\upsilon}>i\mid \check{g}_{i+1},g_1\ldots,g_{i+1}]\\&\quad\leqslant i^{p-n_1+1}\mathbb{E}_\mu[d(o,go)^{p-n_1}]\cdot i^{p+1}\mathbb{E}_\mu[d(o,go)^{p}]\cdot c^{-n_2}+i^{p-n_2+1}\mathbb{E}_\mu[d(o,go)^{p-n_2}]\cdot c^{p-n_1}\cdot K'e^{-i/K'}.\end{aligned}$$

We then multiply $d(o,\check{g}_{i+1})^{n_1}d(o,g_{i+1})^{n_2}$ and integrate over $\check{g}_{i+1}$ and g_{i+1} to obtain a summable bound. This concludes the inequality (12). $\square$

The previous proof also yields the following corollary.

COROLLARY 4.15. *Let $p>0$ and let $((\check{Z}_n)_{n>0},(Z_n)_{n>0})$ be the (bi-directional) random walk generated by a non-elementary probability measure μ on G with finite pth moment. Then there exists $K>0$ such that*

$$\mathbb{E}[\min\{d(o,Z_\upsilon o),d(o,\check{Z}_{\check{\upsilon}}o)\}^{2p}]<K.$$

Proof. In view of the previous proof, it suffices to check

$$\min\{d(o,Z_\upsilon o),d(o,\check{Z}_{\check{\upsilon}}o)\}^{2p}\leqslant\sum_{i=0}^{\infty}|\check{D}_{i+1}^pD_{i+1}^p-\check{D}_i^pD_i^p|(1_{\check{D}_i\geqslant D_i}1_{i<\upsilon}+1_{\check{D}_i\leqslant D_i}1_{i<\check{\upsilon}}).$$

The right-hand side is at least $\check{D}_l^pD_l^p$ for $l=\min\{i:1_{\check{D}_i\geqslant D_i}1_{i<\upsilon}+1_{\check{D}_i\leqslant D_i}1_{i<\check{\upsilon}}=0\}$. Note that either $\check{D}_l\geqslant D_l$ or $\check{D}_l\leqslant D_l$ holds. In the first case, we are forced to have $l\geqslant\upsilon$; then

$$\min\{d(o,Z_\upsilon o),d(o,\check{Z}_{\check{\upsilon}}o)\}^{2p}\leqslant d(o,Z_\upsilon o)^{2p}\leqslant D_\upsilon^{2p}\leqslant D_l^{2p}\leqslant\check{D}_l^pD_l^p.$$

In the second case, we are forced to have $l\geqslant\check{\upsilon}$; then

$$\min\{d(o,Z_\upsilon o),d(o,\check{Z}_{\check{\upsilon}}o)\}^{2p}\leqslant d(o,\check{Z}_{\check{\upsilon}}o)^{2p}\leqslant\check{D}_{\check{\upsilon}}^{2p}\leqslant\check{D}_l^{2p}\leqslant\check{D}_l^pD_l^p.\qquad\square$$

We now discuss random walks with finite exponential moment.

COROLLARY 4.16. *Let $((\check{Z}_n)_{n>0},(Z_n)_{n>0})$ be the (bi-directional) random walk generated by a non-elementary probability measure μ on G with finite exponential moment. Then there exists $K>0$ such that*

$$\mathbb{E}[\exp(d(o,Z_\upsilon o)/K)]<K.$$

Proof. Let K' be as in Lemma 4.10 and $D_i=\sum_{k=1}^{i}d(o,g_ko)$. Then $e^{d(o,Z_\upsilon o)/K}$ is dominated by $\sum_{i\leqslant\upsilon}e^{D_i/K}$. Hence, we need to show that $\mathbb{E}[e^{D_i/K}1_{i<\upsilon}]$ is summable. Let $K,c>0$ and observe

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

$$\begin{aligned}\mathbb{E}[e^{D_i/K}1_{i<v}] &\leqslant \mathbb{E}[e^{D_i/K}1_{D_i<c}1_{i<v}] + \mathbb{E}[e^{D_i/K}1_{D_i\geqslant c}1_{i<v}] \\ &\leqslant \mathbb{E}[e^{c/K}1_{i<v}] + \mathbb{E}[e^{2D_i/K}e^{-c/K}] \\ &\leqslant e^{c/K}K'e^{-i/K'} + e^{-c/K}\cdot(\mathbb{E}_\mu[\exp(2d(o,go)/K)])^i.\end{aligned}$$

By taking K large enough, we can make $\mathbb{E}_\mu[\exp(2d(o,go)/K)] \leqslant e^{1/4K'}$. Then we take $c = iK/2K'$ and conclude $\mathbb{E}[e^{D_i/K}1_{i<v}] < (K'+1)e^{-i/4K'}$. □

4.4 Limit theorems

The second-moment deviation inequality implies the following CLT.

Theorem 4.17. *Let (X, G, o) be as in Convention 2.11 and let $(Z_n)_{n>0}$ be the random walk generated by a non-elementary probability measure μ on G with finite second moment. Then the following limit (called the asymptotic variance of μ) exists:*

$$\sigma^2(\mu) := \lim_{n\to\infty} \tfrac{1}{n} \mathrm{Var}[d(o, Z_n o)],$$

and the random variable $(1/\sqrt{n})[d(o, Z_n o) - \lambda(\mu)n]$ converges in law to the Gaussian law $\mathcal{N}(0, \sigma(\mu))$ with zero mean and variance $\sigma^2(\mu)$.

Proof. Since μ has finite second moment, Proposition 4.13 implies that $\sup_{n,m\geqslant 0}(\check{Z}_m o, Z_n o)_o$ has finite fourth moment, and hence finite second moment. Now Theorems 4.1 and 4.2 of [MS20] lead to the conclusion. □

Remark 4.18. In fact, the following non-degeneracy statement holds.

Fact 4.19. Let (X, G, o) be as in Convention 2.11 and let $(Z_n)_n$ be the random walk generated by a non-elementary probability measure μ on G. Then the asymptotic variance $\sigma^2(\mu) := \lim_n \frac{1}{n}\mathrm{Var}[d(o, Z_n o)]$ is nonzero if and only if μ is *non-arithmetic*, i.e., there exists $N > 0$ and two elements $g, h \in (\mathrm{supp}\mu^{*N})$ of $\mathrm{supp}\mu^{*N}$ with distinct translation lengths.

The strict positivity of $\sigma^2(\mu)$ for non-arithmetic random walks on Gromov hyperbolic spaces and Teichmüller space was discussed in [Cho23]; see Theorem B and Claim 6.2 of [Cho23]. Since the argument in [Cho23] also applies to the general case, we omit the proof here.

We next discuss the LIL.

Theorem 4.20. *Let (X, G, o) be as in Convention 2.11 and let $(Z_n)_{n>0}$ be the random walk generated by a non-elementary probability measure μ on G with finite second moment. Then for almost every sample path $(Z_n)_n$ we have*

$$\limsup_{n\to\infty} \frac{d(o, Z_n o) - \lambda(\mu)n}{\sqrt{2n\log\log n}} = \sigma(\mu),$$

where $\lambda(\mu)$ is the drift of μ and $\sigma^2(\mu)$ is the asymptotic variance of μ.

We proved the LIL based on the uniform fourth-order deviation inequality in [Cho23]. We give another argument because we will only have second-order deviation inequality in Part II.

Proof. In the proof of the LIL in [Cho23] (see [Cho23, Claim 7.1]), the author proved the following lemma.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

LEMMA 4.21. *Let $K > 0$ and let $\{U_{k,i}\}_{i,k\in\mathbb{Z}_{>0}}$ be RVs such that, for each k, $\{U_{k,i}\}_i$ are i.i.d.s with zero mean and variance at most K. Then for each $\epsilon > 0$, there exists $M > 0$ such that*

$$\mathbb{P}\bigg(\limsup_n \frac{1}{\sqrt{2n\log\log n}}\bigg|\sum_{k=M}^{\lfloor \log_2 n\rfloor}\sum_{i=1}^{\lfloor n/2^{k+1}\rfloor} U_{k,i}\bigg| > \epsilon\bigg) \leqslant \epsilon.$$

We now set

$$Y_{k,i} := d(Z_{2^k(i-1)}o, Z_{2^k i}o), \quad b_{k,i} := (Z_{2^k(2i-2)}o, Z_{2^k\cdot 2i}o)_{Z_{2^k(2i-1)}o}.$$

Equivalently, we have $Y_{k+1,i} = Y_{k,2i-1} + Y_{k,2i} - 2b_{k,i}$. Note that $\{b_{k,i}\}_{k,i}$ have uniformly bounded variance by Lemma 4.10 and $\{b_{k,i} - \mathbb{E}[b_{k,i}]\}_i$ are i.i.d.s with zero mean for each k. We also set

$$b_{k;n} = \begin{cases} (Z_{2^{k+1}\lfloor n/2^{k+1}\rfloor}o, Z_n o)_{Z_{2^k(2\lfloor n/2^{k+1}\rfloor+1)}o}, & \text{if } 2^k(2\lfloor n/2^{k+1}\rfloor + 1) < n, \\ 0, & \text{otherwise.}\end{cases}$$

We then observe the decomposition

$$d(o, Z_n o) = \sum_{i=1}^{\lfloor n/2^M\rfloor} Y_{M,i} + d(Z_{2^M\lfloor n/2^M\rfloor}o, Z_n o) - 2\sum_{k=M}^{\lfloor \log_2 n\rfloor}\bigg(b_{k;n} + \sum_{i=1}^{\lfloor n/2^{k+1}\rfloor} b_{k,i}\bigg) \quad (\forall n, M > 0). \tag{15}$$

Indeed, the right-hand side is unchanged when M increases by one and is equal to $d(o, Z_n o)$ at $M > \lfloor \log_2 n\rfloor$.

Now, fixing an $\epsilon > 0$, we take $M > 0$ for $\{b_{k,i} - \mathbb{E}[b_{k,i}]\}_{i,k}$ using Lemma 4.21. We balance each term in display (15) by subtracting its expectation, normalize with the denominator $\sqrt{2n\log\log n}$ and then examine the almost sure limit supremum. The classical LIL tells us that

$$\limsup_{n\to\infty}\frac{1}{\sqrt{2n\log\log n}}\sum_{i=1}^{\lfloor n/2^M\rfloor}(Y_{M,i} - \mathbb{E}[Y_{M,i}]) = \frac{1}{\sqrt{2^M}}\sqrt{\mathrm{Var}(Y_{M,1})}.$$

Regarding the second term, note that $d(Z_{2^M\lfloor n/2^M\rfloor}, Z_n o)$ is dominated by the sum of at most 2^M independent steps distributed according to μ. This implies that

$$\mathbb{P}(d(Z_{2^M\lfloor n/2^M\rfloor}, Z_n o) > \epsilon\sqrt{n}) \leqslant \mathbb{P}\bigg(\sum_{i=1}^{2^M} d(o, g_i o) > \epsilon\sqrt{n}\bigg),$$

and the right-hand side is summable in n because μ has finite second moment. By the Borel–Cantelli lemma,

$$\frac{1}{\sqrt{2n\log\log n}}|d(Z_{2^M\lfloor n/2^M\rfloor}o, Z_n o)| = 0 \quad \text{almost surely.}$$

Next, Lemma 4.21 implies that the term

$$\frac{1}{\sqrt{2n\log\log n}}\sum_{k=M}^{\lfloor \log_2 n\rfloor}\sum_{i=1}^{\lfloor n/2^{k+1}\rfloor}(b_{k,i} - \mathbb{E}[b_{k,i}])$$

eventually falls into the interval $[-\epsilon, +\epsilon]$ outside a set of probability ϵ.

It remains to deal with $(1/\sqrt{2n\log\log n})\sum_{k=M}^{\lfloor \log_2 n\rfloor}(b_{k;n} - \mathbb{E}[b_{k;n}])$. Let

$$b_j := \sup_{i,i'\geqslant 0}(Z_{j-i}o, Z_{j+i'}o)_o.$$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Then, for each k and n, $0 \leqslant b_{k;n} \leqslant b_{2^k(2\lfloor n/2^{k+1}\rfloor+1)}$ holds. Moreover, b_j values are identically distributed with finite variance (and hence finite expectation). This implies that

$$0 \leqslant \frac{1}{\sqrt{2n \log\log n}} \sum_{k=1}^{\lfloor \log_2 n\rfloor} \mathbb{E}[b_{k;n}] \leqslant \frac{\log_2 n}{\sqrt{2n\log\log n}} \mathbb{E}[b_0]$$

tends to 0 as n goes to infinity.

We now estimate the summation

$$\sum_{k=0}^{\infty}\sum_{i=1}^{\infty} \mathbb{P}(b_{2^k(2i-1)}^2 \geqslant \epsilon^2 \sqrt{2}^k (2i-2)).$$

To estimate this, for each $y>0$ let us count the number of pairs $(i,k) \in \mathbb{Z}_{\geqslant 0}^2$ such that $\epsilon^2\sqrt{2}^k(2i-2) < y$. For each $k \in \mathbb{Z}_{\geqslant 0}$, there exist at most $y/(\sqrt{2}^k\epsilon^2)$ candidates for i. Summing them up, there are at most $C_\epsilon y$ such pairs (i,k), where $C_\epsilon > 0$ is a constant. This implies that

$$\begin{aligned}
\sum_{k=0}^{\infty}\sum_{i=1}^{\infty} \mathbb{P}(b_{2^k(2i-2)}^2 \geqslant \epsilon^2\sqrt{2}^k(2i-2)) &\leqslant \sum_{k,i}\sum_{y=0}^{\infty} \mathbb{P}(y-1 \leqslant b_{2^k(2i-2)}^2 < y) 1_{y \geqslant \epsilon^2\sqrt{2}^k(2i-2)} \\
&= \sum_{k,i}\sum_{y=0}^{\infty} \mathbb{P}(y-1 \leqslant b_1^2 < y) 1_{y\geqslant \epsilon^2\sqrt{2}^k(2i-2)} \\
&\leqslant \sum_{y=0}^{\infty} \mathbb{P}(y-1\leqslant b_1^2 < y)\cdot \#\{(i,k) : y \geqslant \epsilon^2\sqrt{2}^k(2i-2)\} \\
&\leqslant \sum_{y=0}^{\infty} \mathbb{P}(y-1 \leqslant b_1^2 < y)\cdot C_\epsilon y \leqslant \mathbb{E}[C_\epsilon b_1^2] < +\infty.
\end{aligned}$$

By the Borell–Cantelli lemma, for almost every sample path $b_{2^k(2i-1)} < \epsilon \cdot 2^{k/4}\sqrt{2i-2}$ holds for all but finitely many (i,k). In particular, for sufficiently large n, we have

$$b_{k;n} \leqslant b_{2^k(2\lfloor n/2^{k+1}\rfloor+1)} \leqslant \epsilon\cdot 2^{k/4}\sqrt{2n/2^{k+1}} = \epsilon\sqrt{n}/2^{k/4},$$

for each $k = 1, \ldots, \lfloor \log_2 n\rfloor$. Hence, we have

$$\frac{1}{\sqrt{2n\log\log n}} \sum_{k=M}^{\lfloor \log_2 n\rfloor} b_{k;n} \leqslant \epsilon \sum_{k=1}^{\infty} 1/2^{k/4} \leqslant 10\epsilon.$$

Combining these estimates with Equation (15), we observe that for probability at least $1-\epsilon$,

$$\limsup_{n\to\infty} \frac{d(o,Z_n o) - \mathbb{E}[d(o,Z_n o)]}{\sqrt{2n\log\log n}} \in \left[\sqrt{\frac{\mathrm{Var}[d(o,Z_{2^M}o)]}{2^M}} - 20\epsilon, \sqrt{\frac{\mathrm{Var}[d(o,Z_{2^M}o)]}{2^M}} + 20\epsilon\right].$$

By decreasing ϵ while increasing M, we arrive at the desired conclusion. □

We finally prove the geodesic tracking by random walks.

Theorem 4.22. *Let (X,G,o) be as in Convention 2.11, let $p>0$, and let $(Z_n)_n$ be the random walk generated by a non-elementary probability measure μ on G with finite pth moment. Then there exists $K>0$ such that, for almost every sample path $(Z_n)_{n\geqslant 0}$, there exists a K-quasigeodesic γ on X satisfying*

$$\lim_{n\to\infty} \frac{1}{n^{1/2p}} d(Z_n o, \gamma) = 0.$$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Proof. Recall Definition 3.16. Given $K_0 > 0$, we have defined:

- $D_0 = D(K_0, K_0) > K_0$ as in Lemma 3.8;
- $E_0 = E(K_0, D_0) > D_0$, $L_0 = L(K_0, D_0)$ as in Proposition 3.12.

In addition to these, we define:

- $E_1 = E(K_0, D_0)$, $L_1 = L(K_0, D_0)$ as in Lemma 3.14;
- $E_2 = E(K_0, E_0 + 5K_0)$, $L_2 = L(K_0, E_0 + 5K_0)$ as in Proposition 3.12.

Since μ is non-elementary, Proposition 3.19 guarantees that there exist $K_0 > 0$, $M_0 > L_0 + L_1 + L_2$ and a large enough K_0-Schottky set $S \subseteq (\mathrm{supp}\mu)^{M_0}$. We fix this S from now on.

By Proposition 4.2, there exists a probability space $(\Omega, \mathbb{P})$ with RV $\mathcal{P}(\omega) = \{j(1) < j(2) < \cdots\} \subseteq M_0\mathbb{Z}_{>0}$, the set of pivotal times, such that $(o, \mathbf{Y}_{j(1)}(\omega), \mathbf{Y}_{j(2)}(\omega), \ldots)$ is D_0-semi-aligned.

We now define $\Gamma(\omega)$ as the concatenation of $[o, Z_{j(1)-M_0}o]$, $[Z_{j(1)-M_0}o, Z_{j(1)}o]$, $[Z_{j(1)}o, Z_{j(2)-M_0}o]$, $[Z_{j(2)-M_0}o, Z_{j(2)}o]$, ... By Lemma 3.14, $\Gamma(\omega)$ is an E_1-quasigeodesic for almost every $\omega \in \Omega$. It remains to prove $\lim_n d(Z_n(\omega)o, \Gamma(\omega))/n^{1/2p} = 0$ almost everywhere.

By Corollary 4.15, $\min[d(o, Z_\upsilon o), d(o, \check{Z}_{\check{\upsilon}}o)]^{2p}$ is dominated by an integrable RV. This implies

$$\sum_k \mathbb{P}(\min(d(o, Z_\upsilon o), d(o, \check{Z}_{\check{\upsilon}}o)) > g(k)) < +\infty, \tag{16}$$

for some function g such that $\lim_k g(k)/k^{1/2p} = 0$. Also, Lemma 4.10 tells us that

$$\sum_k \mathbb{P}(\max(\upsilon, \check{\upsilon}) \geqslant k - M_0) < +\infty. \tag{17}$$

Now, for each $k \in \mathbb{Z}_{>0}$, we consider the following sets:

$$A_k := \left\{(\check{\omega}, \omega) : \begin{array}{c} \text{there exists } M_0 \leqslant i \leqslant k - M_0 \text{ such that } d(o, Z_i o) \leqslant g(k) \text{ and} \\ (\check{Z}_k o, (Z_{i-M_0}o, \ldots, Z_i o), Z_n o) \text{ is } D_0\text{-semi-aligned for all } n \geqslant k \end{array}\right\},$$

$$B_k := \left\{(\check{\omega}, \omega) : \begin{array}{c} \text{there exists } M_0 \leqslant i \leqslant k - M_0 \text{ such that } d(o, \check{Z}_i o) \leqslant g(k) \text{ and} \\ (\check{Z}_k o, (\check{Z}_i o, \ldots, \check{Z}_{i-M_0}o), Z_n o) \text{ is } D_0\text{-semi-aligned for all } n \geqslant k \end{array}\right\}.$$

Then the definition of the RV $\upsilon(\check{\omega}, \omega)$ and $\check{\upsilon}(\check{\omega}, w)$, together with inequality (8), tells us that

$$A_k^c \cap B_k^c \subseteq \{(\check{\omega}, \omega) : \min(d(o, Z_\upsilon o), d(o, \check{Z}_{\check{\upsilon}}o)) > g(k) \text{ or } \max(\upsilon, \check{\upsilon}) \geqslant k - M_0\}.$$

Thanks to display (16) and (17), we observe that $\mathbb{P}(A_k^c \cap B_k^c)$ is also summable.

Finally, consider

$$C_k := \left\{(\check{\omega}, \omega) : \begin{array}{c} \text{there exists } M_0 \leqslant i \leqslant 2k - M_0 \text{ such that } d(Z_k, Z_i o) \leqslant g(k) \text{ and} \\ (o, (Z_{i-M_0}o, \ldots, Z_i o), Z_n o) \text{ is } D_0\text{-semi-aligned for all } n \geqslant 2k \end{array}\right\}.$$

Then C_k contains $T^k(A_k \cup B_k)$, where T is the Bernoulli shift operator on the bi-infinite sample paths, which is measure preserving. Hence, $\mathbb{P}(C_k^c) \leqslant \mathbb{P}(A_k^c \cap B_k^c)$ is summable. The Borel–lemma implies that, for almost every sample path, for each sufficiently large k there exists $M_0 \leqslant j'(k) \leqslant 2k$ such that $\mathrm{diam}(Z_k o \cup \mathbf{Y}_{j'(k)}o) \leqslant d(Z_k o, Z_{j'(k)}o) + \mathrm{diam}(\mathbf{Y}_{j'(k)}o) \leqslant g(k) + K_0 M_0 + K_0$ and such that $(o, \mathbf{Y}_{j'(k)}, Z_n o)$ is D_0-semi-aligned for $n \geqslant 2k$ $(*)$,

Let us now pick a sample path satisfying $(*)$, pick a sufficiently large k, and let N be an index such that $j(N) \geqslant 2k$. Recall that $(o, \mathbf{Y}_{j(1)}(\omega), \mathbf{Y}_{j(2)}(\omega), \ldots)$ is D_0-semi-aligned. By Proposition 3.12, $[o, Z_{j(N)}o]$ have subsegments $[x_1, y_1], \ldots, [x_N, y_N]$, in order from left to right, such that $[x_i, y_i]$ and $\mathbf{Y}_{j(i)}$ are $0.1E_0$-fellow traveling for $i = 1, \ldots, N$. Moreover, by Corollary 3.4, $\mathbf{Y}_{j(i)}$ and $[Z_{j(i)-M_0}o, Z_{j(i)}o]$ are $0.1E_0$-fellow traveling for $i = 1, \ldots, N$. Finally, since $(o, \mathbf{Y}_{j'(k)}, Z_{j(N)}o)$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

is D_0-semi-aligned, Proposition 3.12 tells us that $[o, Z_{j(N)}o]$ also contains a subsegment $[q_1, q_2]$ that $0.1E_0$-fellow travels with $\mathbf{Y}_{j'(k)}$. For convenience we let $y_0 = o$ and $j(0) = 0$.

If $[q_1, q_2]$ overlaps with some $[x_i, y_i]$, this implies $d(\mathbf{Y}_{j'(k)}, [Z_{j(i)-M_0}o, Z_{j(i)}o]) \leqslant E_0$ and hence $d(Z_k o, \Gamma(\omega)) \leqslant g(k) + E_0 + K_0 M_0 + K_0$. If not, then $[q_1, q_2]$ is a subsegment of $[y_{i-1}, x_i]$ for some i. Lemma 3.11 then tells us that $(y_{i-1}, \mathbf{Y}_{j'(k)}, x_i)$ is $0.4E_0$-aligned. Since $d(y_{i-1}, Z_{j(i-1)}o) \leqslant 0.1E_0$ and $d(x_i, Z_{j(i)-M_0}o) \leqslant 0.1E_0$, Lemma 2.2 implies that $(Z_{j(i-1)}o, \mathbf{Y}_{j'(k)}, Z_{j(i)-M_0}o)$ is $(0.5E_0 + 4K_0)$-aligned. By Proposition 3.12, $[Z_{j(i-1)}o, Z_{j(i)-M_0}o]$ passes through the E_2-neighborhood of $\mathbf{Y}_{j'(k)}$, and $d(Z_k o, \Gamma(\omega)) \leqslant g(k) + E_2$.

In summary, almost every sample path $(\check{\omega}, \omega)$ satisfies $(*)$, which leads to $d(Z_k o, \Gamma(\omega)) \leqslant g(k) + E_0 + E_2 + K_0 M_0 + K_0 = o(k^{1/2p})$ eventually. This ends the proof. $\square$

Recall Corollary 4.16. if μ has finite exponential moment, then $\mathbb{E}[\exp(d(o, Z_\upsilon o)/K)]$ is finite, i.e., $\mathbb{P}(d(o, Z_\upsilon o) > K \log k)$ is summable for some $K > 0$. By replacing $g(k)$ in the previous proof with $K \log k$, we obtain the following theorem.

Theorem 4.23. *Let (X, G, o) be as in Convention 2.11 and let $(Z_n)_n$ be the random walk generated by a non-elementary probability measure μ on G with finite exponential moment. Then there exists $K > 0$ such that, for almost every sample path $(Z_n)_{n \geqslant 0}$, there exists a K-quasigeodesic γ satisfying*

$$\limsup_{n \to \infty} \frac{1}{\log n} d(Z_n o, \gamma) \leqslant K.$$

5. Pivotal time construction

In this section we prove Proposition 4.2 by generalizing Gouëzel's theory in [Gou22, Section 4A] to the setting of Convention 2.11. We first construct and study pivotal times in a discrete model and then realize them on random walks. This strategy is also employed for LDP in § 6.

5.1 Pivotal times: discrete model

Throughout the subsection, we fix a long enough K_0-Schottky set S with cardinality N_0. Given sequences of isometries $\mathbf{w} = (w_i)_{i=0}^{\infty}$ and $\mathbf{v} = (v_i)_{i=1}^{\infty}$ in G, we draw a sequence of Schottky sequences

$$\mathbf{s} = (\alpha_1, \beta_1, \gamma_1, \delta_1, \ldots, \alpha_n, \beta_n, \gamma_n, \delta_n) \in S^{4n},$$

with respect to the uniform measure on S^{4n}. We define isometries

$$a_i := \Pi(\alpha_i), \quad b_i := \Pi(\beta_i), \quad c_i := \Pi(\gamma_i), \quad d_i := \Pi(\delta_i), \tag{18}$$

and study the word

$$w_0 a_1 b_1 v_1 c_1 d_1 w_1 \cdots a_k b_k v_k c_k d_k w_k \cdots.$$

With the base case $w_{0,2}^+ := \mathrm{id}$, we define its subwords for $i > 0$,

$$\begin{aligned} w_{i,2}^- &:= w_{i-1,2}^+ w_{i-1}, \quad & w_{i,1}^- &:= w_{i,2}^- a_i, \quad & w_{i,0}^- &:= w_{i,2}^- a_i b_i, \\ w_{i,0}^+ &:= w_{i,2}^- a_i b_i v_i, \quad & w_{i,1}^+ &:= w_{i,2}^- a_i b_i v_i c_i, \quad & w_{i,2}^+ &:= w_{i,2}^- a_i b_i v_i c_i d_i. \end{aligned}$$

Let us also employ the notations

$$\begin{aligned} \Upsilon(\alpha_i) &:= w_{i,2}^- \Gamma^+(\alpha_i), \quad & \Upsilon(\beta_i) &:= w_{i,1}^- \Gamma^+(\beta_i), \\ \Upsilon(\gamma_i) &:= w_{i,0}^+ \Gamma^+(\gamma_i), \quad & \Upsilon(\delta_i) &:= w_{i,1}^+ \Gamma^+(\delta_i). \end{aligned}$$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

We define the *set of pivotal times* $P_n = P_n(\mathbf{s}; \mathbf{w}, \mathbf{v})$ and an auxiliary moving point $z_n = z_n(\mathbf{s}; \mathbf{w}, \mathbf{v})$ inductively. Let $P_0 = \emptyset$ and $z_0 = o$ as the base case. Given $P_{n-1} \subseteq \{1, \ldots, n-1\}$ and $z_{n-1} \in X$, the data P_n and z_n at step n are determined by the following criteria.

(A) When $(z_{n-1}, \Upsilon(\alpha_n))$, $(\Upsilon(\beta_n), w_{n,1}^+ o)$, $(w_{n,0}^- o, \Upsilon(\gamma_n))$ and $(\Upsilon(\delta_n), w_{n+1,2}^- o)$ are K_0-aligned, we set $P_n = P_{n-1} \cup \{n\}$ and $z_n = w_{n,1}^+ o$ (see Figure 4).

(B) Otherwise, we seek $i \in P_{n-1}$ and an integer $j \in \{i+1, \ldots, n-1\}$ such that $(\Upsilon(\delta_i), \Upsilon(\beta_j))$ is D_0-semi-aligned and such that $(\Upsilon(\beta_j), w_{n+1,2}^- o)$ is K_0-aligned.
If such a pair (i, j) exists, we pick the lexicographically maximal one and let $P_n := P_{n-1} \cap \{1, \ldots, i\}$ and $z_n = w_{j,1}^- o$. If such a pair does not exist, then we let $P_n := \emptyset$ and $z_n := o$.

We note that the set the set P_n depends solely on $(w_i)_{i=0}^n$, $(v_i)_{i=1}^n$ and $(\alpha_i, \beta_i, \gamma_i, \delta_i)_{i=1}^n$; it is independent of $\{w_i, v_i, \alpha_i, \beta_i, \gamma_i, \delta_i : i > n\}$.

The Schottky axes aligned along $[o, w_{n+1,2}^- o]$ are recorded by P_n. More precisely, we have the following.

PROPOSITION 5.1. *Let $P_n = \{i(1) < \cdots < i(m)\}$. Then*

$$(o, \Upsilon(\alpha_{i(1)}), \Upsilon(\beta_{i(1)}), \Upsilon(\gamma_{i(1)}), \Upsilon(\delta_{i(1)}), \ldots, \Upsilon(\alpha_{i(m)}), \Upsilon(\beta_{i(m)}), \Upsilon(\gamma_{i(m)}), \Upsilon(\delta_{i(m)}), w_{n+1,2}^- o),$$

is D_0-semi-aligned.

On Gromov hyperbolic spaces, this corresponds to [Gou22, Lemma 5.3]. Before proving the entire statement, let us prove two small parts of it.

LEMMA 5.2. *For any $\mathbf{s} \in S^{4n}$ and $1 \leqslant i \leqslant n$, $(\Upsilon(\alpha_i), \Upsilon(\beta_i))$ and $(\Upsilon(\gamma_i), \Upsilon(\delta_i))$ are D_0-aligned.*

Proof. Let us prove that $(\Upsilon(\alpha_i), \Upsilon(\beta_i)) = (w_{i,1}^- \bar{\Gamma}^-(\alpha_i), w_{i,1}^- \Gamma^+(\beta_i))$ is D_0-aligned, or equivalently, that $(\bar{\Gamma}^-(\alpha_i), \Gamma^+(\beta_i))$ is D_0-aligned. When $\alpha_i = \beta_i$, this is guaranteed by the definition of K_0-Schottky sets.

Now suppose $\alpha_i \neq \beta_i$. First, $(\bar{\Gamma}^-(\alpha_i), o)$ is 0-aligned. Second, $(a_i^{-1} o, \Gamma^-(\alpha_i))$ is not K_0-aligned as $d(o, a_i^{-1} o) \geqslant 100 E_0 \geqslant K_0$. Then, by the Schottky property of S, $(a_i^{-1} o, \Gamma^+(\beta_i))$ is K_0-aligned. Now Lemma 3.8 tells us that $(\bar{\Gamma}^-(\alpha_i), \Gamma^+(\beta_i))$ is D_0-aligned.

The alignment of $(\Upsilon(\gamma_i), \Upsilon(\delta_i))$ holds for the same reason. □

LEMMA 5.3 [Cho24, Lemma 3.2]. *Let $k \in \mathbb{Z}_{>0}$. Let $l < m$ be consecutive elements in P_k, i.e., $m \in P_k$ and $l = \max(P_k \cap \{1, \ldots, m-1\})$. Then $(\Upsilon(\delta_l), \Upsilon(\alpha_m))$ is D_0-semi-aligned.*

Proof. Here $l, m \in P_k$ implies that $l \in P_l$ and $l, m \in P_m$. In particular, l and m are newly chosen at step l and m, respectively, by fulfilling criterion (A). Hence, $(\Upsilon(\delta_l), w_{l+1,2}^- o)$ and $(z_{m-1}, \Upsilon(\alpha_m))$ are K_0-aligned $(*)$, and $z_l = w_{l,1}^+ o$. Moreover, we have $P_m = P_{m-1} \cup \{m\}$ and $l = \max P_{m-1}$.

If $l = m-1$ and m was newly chosen at step $m = l+1$, then $z_{m-1} = z_l = w_{l,1}^+ o$ holds. Lemma 3.8 and $(*)$ imply that $(\Upsilon(\delta_l), \Upsilon(\alpha_m))$ is D_0-aligned.

If $l < m-1$, then $l = \max P_{m-1}$ has survived at step $m-1$ by fulfilling criterion (B); there exist $j > l$ such that $(\Upsilon(\delta_l), \Upsilon(\beta_j))$ is D_0-semi-aligned and $(\Upsilon(\beta_j), w_{n+1,2}^- o)$ is K_0-aligned. Furthermore, z_{m-1} equals $w_{j,1}^- o$, the beginning point of $\Upsilon(\beta_j)$.

Note that $(z_{m-1}, \Upsilon(\alpha_m))$ is K_0-aligned by $(*)$. Lemma 3.8 then asserts that $(\Upsilon(\beta_j), \Upsilon(\alpha_m))$ is D_0-aligned. Concatenating the two D_0-semi-aligned sequences, we conclude that $(\Upsilon(\delta_l), \Upsilon(\alpha_m))$ is D_0-semi-aligned. □

Proof of Proposition 5.1. Having established Lemma 5.3, it remains to prove that:

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

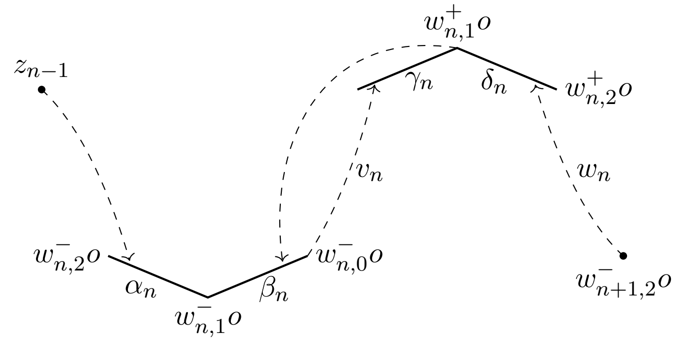

FIGURE 4. Schematics for criteria (19), (20), (21) and (22).

- $(o, \Upsilon(\alpha_{i(1)}))$ is K_0-aligned;
- for $1 \leqslant t \leqslant m$, $(\Upsilon(\alpha_{i(t)}), \Upsilon(\beta_{i(t)}), \Upsilon(\gamma_{i(t)}), \Upsilon(\delta_{i(t)}))$ is D_0-aligned;
- $(\Upsilon(\delta_{i(m)}), w_{n+1,2}^{-}o)$ is D_0-semi-aligned.

Note that for each $t = 1, \ldots, m$, $i(t)$ is newly chosen as a pivotal time at step $i(t)$ by fulfilling criterion (A). In particular, we have that:

- $(\Upsilon(\alpha_{i(t)}), \Upsilon(\beta_{i(t)}))$ is D_0-aligned (Lemma 5.2);
- $(\Upsilon(\beta_{i(t)}), \Upsilon(\gamma_{i(t)}))$ is D_0-aligned since $(\Upsilon(\beta_{i(t)}), w_{n,1}^{+}o)$ and $(w_{i(t),0}^{-}o, \Upsilon(\gamma_{i(t)}))$ are K_0-aligned (Lemma 3.8); and
- $(\Upsilon(\gamma_{i(t)}), \Upsilon(\delta_{i(t)}))$ is D_0-aligned (Lemma 5.2).

This guarantees the second item.

We also note that $P_{i(1)-1} = \emptyset$. Indeed, any j in $P_{i(1)-1}$ is smaller than $i(1)$ and would have survived in $P_{i(1)}$ (since what happened at step $i(1)$ was adding an element, not deleting some). Since $i(1)$ was not deleted at any later step, such j would also not be deleted until the end and should have appeared in P_n. Since $i(1)$ is the earliest pivotal time in P_n, no such j exists. Hence, $z_{i(1)-1} = o$ and criterion (A) for $i(1)$ leads to the first item.

We now observe how $i(m)$ survived in P_n. If $i(m) = n$, then it was newly chosen at step n by fulfilling criterion (A). In particular, $(\Upsilon(\delta_n), w_{n+1,2}^{-}o)$ is K_0-aligned as desired.

If $i(m) \neq n$, then it has survived at step n as the last pivotal time by fulfilling criterion (B). In particular, there exist $j > i(m)$ such that $(\Upsilon(\delta_{i(m)}), \Upsilon(\beta_j))$ is D_0-semi-aligned and such that $(\Upsilon(\beta_j), w_{n+1,2}^{-}o)$ is K_0-aligned. In particular, $(\Upsilon(\delta_{i(m)}), \Upsilon(\beta_j), w_{n+1,2}^{-}o)$ is D_0-semi-aligned. □

Next, we study when $P_n = P_{n-1} \cup \{n\}$ happens, i.e., a new pivotal time is added to the set of pivotal times. This will guide us how to pivot the direction at a pivotal time without affecting the set of pivotal times. Recall that we draw $\alpha_i, \beta_i, \gamma_i, \delta_i$ from S with the uniform measure.

LEMMA 5.4. *Let us fix* $\mathbf{w} = (w_i)_i$, $\mathbf{v} = (v_i)_i$ *and* $\mathbf{s} \in S^{4(n-1)}$. *Then*

$$\mathbb{P}(\# P_n(\mathbf{s}, \alpha_n, \beta_n, \gamma_n, \delta_n) = \# P_{n-1}(\mathbf{s}) + 1) \geqslant 1 - 4/N_0.$$

Proof. Recall criterion (A) for $\# P_n = \# P_{n-1} + 1$. We will investigate the four required conditions one by one.

First, the condition

$$\mathrm{diam}(\pi_{\Upsilon(\gamma_n)}(w_{n,0}^{-}o) \cup w_{n,0}^{+}o) = \mathrm{diam}(\pi_{\Gamma^{+}(\gamma_n)}(v_n^{-1}o) \cup o) < K_0 \tag{19}$$

depends only on γ_n. This holds for at least $(\# S - 1)$ choices in S by the K_0-Schottky-ness of S.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Similarly, the condition

$$\operatorname{diam}(\pi_{\Upsilon(\delta_n)}(w_{n+1,2}^{-}o) \cup w_{n,2}^{+}o) = \operatorname{diam}(\pi_{\Gamma^{-}(\delta_n)}(w_n o) \cup o) < K_0 \tag{20}$$

depends only on δ_n, and holds for at least $(\#S-1)$ choices in S.

Fixing the choice of γ_n, the condition

$$\operatorname{diam}(\pi_{\Upsilon(\beta_n)}(w_{n,1}^{+}o) \cup w_{n,0}^{-}o) = \operatorname{diam}(\pi_{\Gamma^{-}(\beta_n)}(v_n c_n o) \cup o) < K_0 \tag{21}$$

depends only on β_n. This holds for at least $(\#S-1)$ choices in S.

We now additionally fix the choice of $s = (\alpha_1, \beta_1, \gamma_1, \delta_1, \ldots, \alpha_{n-1}, \beta_{n-1}, \gamma_{n-1}, \delta_{n-1})$; in particular, $w_{n,2}^{-}$ and z_{n-1} are now determined. Then the condition

$$\operatorname{diam}(\pi_{\Upsilon(\alpha_n)}(z_{n-1}) \cup w_{n,2}^{-}o) = \operatorname{diam}(\pi_{\Gamma^{+}(\alpha_n)}((w_{n,2}^{-})^{-1} z_{n-1}) \cup o) < K_0 \tag{22}$$

depends on α_n. This holds for at least $(\#S-1)$ choices of α_n.

In summary, the probability that criterion (A) holds is at least

$$\frac{\#S-1}{\#S} \cdot \frac{\#S-1}{\#S} \cdot \frac{\#S-1}{\#S} \cdot \frac{\#S-1}{\#S} \geqslant 1 - \frac{4}{N_0}. \qquad \square$$

We now define the set $\tilde{\mathrm{S}}$ of triples $(\beta, \gamma, v) \in S^2 \times G$ that satisfy Conditions (19) and (21):

$$\tilde{\mathrm{S}} := \{(\beta, \gamma, v) \in S^2 \times G : (\bar{\Gamma}^{-}(\beta), v\Pi(\gamma)o), (v^{-1}o, \Gamma^{+}(\gamma)) \quad \text{are } K_0\text{-aligned}\}.$$

We also define its section for each $v \in G$

$$\tilde{\mathrm{S}}(v) := \{(\beta, \gamma) \in S^2 : (\bar{\Gamma}^{-}(\beta), v\Pi(\gamma)o), (v^{-1}o, \Gamma^{+}(\gamma)) \quad \text{are } K_0\text{-aligned}\}.$$

While checking displays (19) and (21), we observed that $\#\tilde{\mathrm{S}}(v) \geqslant \#S^2 - 2\#S$ for each $v \in G$. We now define pivoting.

Lemma 5.5. *Let $\mathbf{s} = (\alpha_1, \beta_1, \gamma_1, \delta_1, \ldots, \alpha_n, \beta_n, \gamma_n, \delta_n)$ be a choice drawn from S^{4n} and let $\mathbf{w}, \mathbf{v}$ be auxiliary sequences in G.*

Let $k \in P_n(\mathbf{s}; \mathbf{w}, \mathbf{v})$ and let $(\bar{\mathbf{s}}; \mathbf{w}, \bar{\mathbf{v}})$ be obtained from $(\mathbf{s}; \mathbf{w}, \mathbf{v})$ by replacing (β_k, γ_k, v_k) with some $(\bar{\beta}_k, \bar{\gamma}_k, \bar{v}_k)$ chosen from $\tilde{S}$.

Then, $P_l(\mathbf{s}; \mathbf{w}, \mathbf{v}) = P_l(\bar{\mathbf{s}}; \mathbf{w}, \bar{\mathbf{v}})$ for any $1 \leqslant l \leqslant n$.

On Gromov hyperbolic spaces, this corresponds to [Gou22, Lemma 5.7].

Proof. Since $\alpha_1, \beta_1, \gamma_1, \delta_1, \ldots, \alpha_{k-1}, \beta_{k-1}, \gamma_{k-1}, \delta_{k-1}$ are intact, $P_l(\mathbf{s}) = P_l(\bar{\mathbf{s}})$ and $\tilde{\mathrm{S}}_l(\mathbf{s}) = \tilde{\mathrm{S}}_l(\bar{\mathbf{s}})$ hold for $l = 0, \ldots, k-1$. At step k, α_k and δ_k satisfy Condition (22) and Condition (20) since $k \in P_n(\mathbf{s})$. Furthermore, $\bar{\beta}_k$ and $\bar{\gamma}_k$ satisfy Conditions (19) and (21) for the new choice $\bar{v}_k$

$$\operatorname{diam}(\pi_{\Gamma^{+}(\bar{\gamma}_k)}(\bar{v}_k^{-1}o) \cup o) < K_0 \quad \text{and} \quad \operatorname{diam}(\pi_{\Gamma^{-}(\bar{\beta}_k)}(\bar{v}_k \bar{c}_k o) \cup o) < K_0,$$

since $(\bar{\beta}_k, \bar{\gamma}_k, \bar{v}_k) \in \tilde{\mathrm{S}}$. Hence, k is newly added in $P_k(\bar{\mathbf{s}})$ and

$$P_k(\bar{\mathbf{s}}) = P_{k-1}(\bar{\mathbf{s}}) \cup \{k\} = P_{k-1}(\mathbf{s}) \cup \{k\} = P_k(\mathbf{s}).$$

Meanwhile, z_k is modified into $\bar{z}_k = \bar{w}_{k,1}^{+}o = g w_{k,1}^{+}o = g z_k$, where $g := w_{k,2}^{-} a_k \bar{b}_k \bar{v}_k \bar{c}_k \times (w_{k,2}^{-} a_k b_k v_k c_k)^{-1}$. More generally, we have

$$\begin{aligned} \bar{w}_{l,t}^{-} &= g w_{l,t}^{-} \quad (t \in \{0,1,2\},\ l > k), \\ \bar{w}_{l,0}^{+} &= g w_{l,0}^{+} \qquad\qquad\qquad (l > k), \\ \bar{w}_{l,t}^{+} &= g w_{l,t}^{+} \qquad (t \in \{1,2\}, l \geqslant k). \end{aligned} \tag{23}$$

We now claim the following for $k < l \leqslant n$.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

(i) If s fulfills criterion (A) at step l, then so does $\bar{\mathbf{s}}$.
(ii) If not, and if (i,j) is the maximal pair of indices for s in criterion (B) at step l, then it is also the maximal one for $\bar{\mathbf{s}}$ at step l.
(iii) In both cases, we have $P_l(\mathbf{s}) = P_l(\bar{\mathbf{s}})$ and $\bar{z}_l = g z_l$.

Assuming the third item for $l-1$: $P_{l-1}(\mathbf{s}) = P_{l-1}(\bar{\mathbf{s}})$ and $\bar{z}_{l-1} = g z_{l-1}$, Equality (23) implies the first item. In this case we deduce $P_l(\mathbf{s}) = P_{l-1}(\mathbf{s}) \cup \{l\} = P_{l-1}(\bar{\mathbf{s}}) \cup \{l\} = P_l(\bar{\mathbf{s}})$ and $\bar{z}_l = \bar{w}^+_{l,1} o = g w^+_{l,1} o = g z_l$, the third item for l.

Furthermore, Equality (23) implies that i in $P_{l-1}(\mathbf{s}) \cap \{k, \ldots, l-1\} = P_{l-1}(\bar{\mathbf{s}}) \cap \{k, \ldots, l-1\}$ and $j > i$ work for $\mathbf{s}$ in criterion (B) if and only if they work for $\bar{\mathbf{s}}$. Such i can be found in $\{k, \ldots, l-1\}$, because k survived in $P_n(\mathbf{s})$ and should not have been erased at step l. Hence, the maximal pair (i,j) for $\mathbf{s}$ is also maximal for $\bar{\mathbf{s}}$. We then deduce $P_l(\mathbf{s}) = P_{l-1}(\mathbf{s}) \cap \{1, \ldots, i\} = P_{l-1}(\bar{\mathbf{s}}) \cap \{1, \ldots, i\} = P_l(\bar{\mathbf{s}})$ and $\bar{z}_l = \bar{w}^-_{j,1} o = g w^-_{j,1} o = g z_l$ (using $j > i$), the third item for l. □

For $\mathbf{s}, \mathbf{s}' \in S^{4n}$ and sequences $\mathbf{w}, \mathbf{v}, \bar{\mathbf{v}}$ in G, we say that $(\bar{\mathbf{s}}; \mathbf{w}, \bar{\mathbf{v}})$ is *pivoted from* $(\mathbf{s}; \mathbf{w}, \mathbf{v})$ if:

– $\alpha_i = \bar{\alpha}_i$, $\delta_i = \bar{\delta}_i$ for all $i \in \{1, \ldots, n\}$;
– $(\bar{\beta}_i, \bar{\gamma}_i, \bar{v}_i) \in \tilde{\mathrm{S}}$ for each $i \in P_n(\mathbf{s}; \mathbf{w}, \mathbf{v})$; and
– $(\beta_i, \gamma_i, v_i) = (\bar{\beta}_i, \bar{\gamma}_i, \bar{v}_i)$ for each $i \in \{1, \ldots, n\} \setminus P_n(\mathbf{s}; \mathbf{w}, \mathbf{v})$.

By Lemma 5.5, being pivoted from each other is an equivalence relation.

Fixing $\mathbf{w}$ and $\mathbf{v}$, for each $\mathbf{s} \in S^{4n}$ let $\mathcal{E}_n(\mathbf{s})$ be the equivalence class of s

$$\mathcal{E}_n(\mathbf{s}) := \{\bar{\mathbf{s}} \in S^{4n} : (\bar{\mathbf{s}}; \mathbf{w}, \mathbf{v}) \text{ is pivoted from } (\mathbf{s}; \mathbf{w}, \mathbf{v})\}.$$

We endow $\mathcal{E}_n(\mathbf{s})$ with the conditional probability of the uniform measure on S^{4n}. We now claim that $\#P_{n+1} - \#P_n$ conditioned on an equivalence class $\mathcal{E}_n(\mathbf{s})$ until step n and the choice at step $n+1$ has uniform exponential tail.

Proposition 5.6. *Fix $\mathbf{w} = (w_i)_{i=0}^{\infty}$ and $\mathbf{v} = (v_i)_{i=1}^{\infty}$. For each $j \geq 0$ and $\mathbf{s} \in S^{4n}$,*

$$\mathbb{P}(\#P_{n+1}(\tilde{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) < \#P_n(\mathbf{s}) - j \mid \tilde{\mathbf{s}} \in \mathcal{E}_n(\mathbf{s}),\ (\alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \in S^4)$$

is less than $(4/N_0)^{j+1}$.

On Gromov hyperbolic spaces, this corresponds to [Gou22, Lemma 5.8].

Proof. An element $\tilde{\mathbf{s}} \in \mathcal{E}_n(\mathbf{s})$ is determined by its coordinates $(\tilde{\beta}_k, \tilde{\gamma}_k)_{k \in P_n(\mathbf{s})}$ subject to the condition $(\tilde{\beta}_k, \tilde{\gamma}_k, v_k) \in \tilde{\mathrm{S}}$. We consider a finer equivalence class by additionally fixing the coordinates γ_k: for $\tilde{\mathbf{s}} \in \mathcal{E}_n(\mathbf{s})$, let $\mathcal{E}'_n(\tilde{\mathbf{s}})$ be the set of $\bar{\mathbf{s}} \in \mathcal{E}_n(\mathbf{s})$ such that $\bar{\gamma}_k = \tilde{\gamma}_k$ for all k. Then $\mathcal{E}_n(\mathbf{s})$ is partitioned into $\{\mathcal{E}'_n(\tilde{\mathbf{s}}) : \tilde{\mathbf{s}} \in \mathcal{E}_n(\mathbf{s})\}$, and it suffices to establish the estimates on each $\mathcal{E}'_n(\tilde{\mathbf{s}})$. Henceforth, we will prove that

$$\mathbb{P}(\#P_{n+1}(\tilde{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) < \#P_n(\mathbf{s}) - j \mid \tilde{\mathbf{s}} \in \mathcal{E}'_n(\mathbf{s}),\ (\alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \in S^4)$$

is less than $(4/N_0)^{j+1}$ for each $\mathbf{s} = (\alpha_1, \beta_1, \gamma_1, \delta_1, \ldots, \alpha_n, \beta_n, \gamma_n, \delta_n) \in S^{4n}$ and $j \geqslant 0$.

Recall that we are fixing the sequences $\mathbf{w}$ and $\mathbf{v}$ throughout the proof. Let us define

$$\tilde{\mathrm{S}}_k := \{\beta \in S : (\bar{\Gamma}^-(\beta), v_k \Pi(\gamma_k) o) \quad \text{is } K_0\text{-aligned}\}.$$

Then $\mathcal{E}'_n(\mathbf{s})$ is parameterized by $\prod_{i \in P_n(\mathbf{s})} \tilde{\mathrm{S}}_k$ with the uniform measure. Let

$$\mathcal{A} := \{(\alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \in S^4 : \#P_{n+1}(\mathbf{s}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) = \#P_n(\mathbf{s}) + 1\}.$$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Lemma 5.4 implies that $\mathbb{P}(\mathcal{A}) \geqslant 1 - 4/N_0$ with respect to the uniform measure on S^4. Note that for each element $(\alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1})$ of $\mathcal{A}$, we have

$$P_n(\mathbf{s}) \subseteq P_n(\mathbf{s}) \cup \{n+1\} = P_{n+1}(\mathbf{s}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}).$$

Hence, for each $\tilde{\mathbf{s}} \in \mathcal{E}'_n(\mathbf{s})$, $(\tilde{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1})$ (as a choice in $S^{4(n+1)}$) is pivoted from $(\mathbf{s}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1})$ and $P_{n+1}(\tilde{\mathbf{s}}) = P_{n+1}(\mathbf{s}) = P_n(\mathbf{s}) \cup \{n+1\} = P_n(\tilde{\mathbf{s}}) \cup \{n+1\}$. Thanks to this, we have

$$\begin{aligned}&\mathbb{P}(\#P_{n+1}(\tilde{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) < \#P_n(\tilde{\mathbf{s}}) \mid \tilde{\mathbf{s}} \in \mathcal{E}'_n(\mathbf{s}), (\alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \in S^4)\\ &\quad \leqslant 1 - \mathbb{P}(\mathcal{A}) \leqslant 4/N_0.\end{aligned}$$

This settles the case $j = 0$.

Now let $j = 1$. The event under discussion becomes void when $\#P_n(\mathbf{s}) < 2$. Excluding such cases, let $l < m$ be the last two elements of $P_n(\mathbf{s})$. We now freeze the coordinates β_k except for $k = m$. Namely, for $\tilde{\mathbf{s}} \in \mathcal{E}'_n(\mathbf{s})$, let $E^{(m)}(\tilde{\mathbf{s}})$ be the set of $\bar{\mathbf{s}} \in \mathcal{E}'_n(\mathbf{s})$ such that $\bar{\beta}_k = \tilde{\beta}_k$ for $k \in P_n(\mathbf{s}) \setminus \{m\}$. Then $\{E^{(m)}(\tilde{\mathbf{s}}) : \tilde{\mathbf{s}} \in \mathcal{E}'_n(\mathbf{s})\}$ becomes a partition of $\mathcal{E}'_n(\mathbf{s})$, and $E^{(m)}(\tilde{\mathbf{s}})$ is parameterized by $\bar{\beta}_m \in \tilde{\mathrm{S}}_m$ with the uniform measure. Note that $\tilde{\mathrm{S}}_m$ has at least $\#S - 1$ elements.

Fixing $(\alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \in S^4$, and let $F^{(m)}(\tilde{\mathbf{s}})$ be the set of $\bar{\mathbf{s}} \in E^{(m)}(\tilde{\mathbf{s}})$ such that $(\Upsilon(\bar{\beta}_m), \bar{w}^-_{n+2,2}o)$ K_0-aligned, or more precisely,

$$\begin{aligned}&\mathrm{diam}(\pi_{\Gamma^{-1}(\bar{\beta}_m)}((\tilde{w}^-_{m,0})^{-1}\tilde{w}^-_{n,2}a_{n+1}b_{n+1}v_{n+1}c_{n+1}d_{n+1}w_{n+1}o) \cup o)\\ &\quad = \mathrm{diam}(o \cup \pi_{\Gamma^{-1}(\bar{\beta}_m)}(v_m\tilde{c}_m\tilde{d}_m w_m \cdots \tilde{c}_n\tilde{d}_n w_n \cdot a_{n+1}b_{n+1}v_{n+1}c_{n+1}d_{n+1}w_{n+1}o)) < K_0.\end{aligned} \tag{24}$$

This amounts to requiring a new Schottky condition to $\bar{\beta}_m$, in addition to the alignment of $(\bar{\Gamma}^{-1}(\bar{\beta}_m), v_m\Pi(\gamma_m)o)$; there are at least $\#S - 2$ choices that additionally satisfy this.

We now claim $\#P_{n+1}(\bar{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \geqslant \#P_n(\mathbf{s}) - 1$ for $\bar{\mathbf{s}} \in F^{(m)}(\tilde{\mathbf{s}})$. First, since $l < m$ are consecutive elements in $P_n(\mathbf{s}) = P_n(\bar{\mathbf{s}})$, Lemma 5.3 asserts that $(\Upsilon(\bar{\delta}_l), \Upsilon(\bar{\alpha}_m))$ is D_0-semi-aligned. Moreover, Lemma 5.2 and Condition (24) imply that

$$(\Upsilon(\bar{\alpha}_m), \Upsilon(\bar{\beta}_m)), (\Upsilon(\bar{\beta}_m), \bar{w}^-_{n+2,2}o),$$

are D_0-aligned and K_0-aligned, respectively. These together imply that

$$(\Upsilon(\bar{\delta}_l), \Upsilon(\bar{\beta}_m)), (\Upsilon(\bar{\beta}_m), \bar{w}^-_{n+2,2}o),$$

are D_0-semi-aligned and K_0-aligned, respectively: the pair (l, m) qualifies criterion (B) at step $n+1$. Hence, $P_{n+1}(\bar{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \supseteq P_n(\bar{\mathbf{s}}) \cap \{1, \ldots, l\}$.

As a result, for each $\tilde{\mathbf{s}} \in \mathcal{E}_n(\mathbf{s})$ and $(\alpha_{n+1}, \ldots, \delta_{n+1}) \in S^4$ we have

$$\begin{aligned}&\mathbb{P}(\#P_{n+1}(\bar{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) < \#P_n(\mathbf{s}) - 1 \mid \bar{\mathbf{s}} \in E^{(m)}(\tilde{\mathbf{s}}))\\ &\quad \leqslant \frac{\#[E^{(m)}(\tilde{\mathbf{s}}) \setminus F^{(m)}(\tilde{\mathbf{s}})]}{\#E^{(m)}(\tilde{\mathbf{s}})} \leqslant \frac{2}{\#S - 1} \leqslant \frac{3}{N_0}.\end{aligned}$$

Since $\{E^{(m)}(\tilde{\mathbf{s}}) : \tilde{\mathbf{s}} \in \mathcal{E}_n(\mathbf{s})\}$ partitions $\mathcal{E}_n(\mathbf{s})$, we deduce

$$\mathbb{P}\left(\#P_{n+1}(\tilde{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) < \#P_n(\mathbf{s}) - 1 \,\middle|\, \tilde{\mathbf{s}} \in \mathcal{E}_n(\mathbf{s})\right) \leqslant \frac{3}{N_0},$$

for each $(\alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \in S^4$. Moreover, this probability vanishes when $(\alpha_{n+1}, \ldots, \delta_{n+1}) \in \mathcal{A}$. Since $\mathbb{P}((\alpha_{n+1}, \ldots, \delta_{n+1}) \in \mathcal{A} \mid (\alpha_{n+1}, \ldots, \delta_{n+1}) \in S^4) \geqslant 1 - 4/N_0$, we deduce that

$$\begin{aligned}&\mathbb{P}(\#P_{n+1}(\tilde{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) < \#P_n(\mathbf{s}) - 1 \mid \tilde{\mathbf{s}} \in \mathcal{E}_n(\mathbf{s}), (\alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \in S^4)\\ &\quad \leqslant \frac{4}{N_0} \cdot \frac{4}{N_0} \leqslant \left(\frac{4}{N_0}\right)^2.\end{aligned} \tag{25}$$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Now let $j=2$. Excluding the void case, we assume that $\#P_n(\mathbf{s}) \geqslant 3$; let $l' < l < m$ be the last three elements. To ease the notation, for $\beta \in S$ we define

$$\mathbf{s}'(\beta) := (\alpha_1, \beta_1, \gamma_1, \delta_1, \ldots, \alpha_m, \beta, \gamma_m, \delta_m, \ldots, \alpha_n, \beta_n, \gamma_n, \delta_n).$$

In other words, $\mathbf{s}'(\beta)$ is obtained from $\mathbf{s}$ by replacing β_m with β. Now let

$$\mathcal{A}_1 := \{(\beta, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \in \tilde{\mathrm{S}}_m \times S^4 : \#P_{n+1}(\mathbf{s}'(\beta), \alpha_{n+1}, \ldots, \delta_{n+1}) \geqslant \#P_n(\mathbf{s}) - 1\}.$$

Equivalently, we are requiring

$$P_n(\mathbf{s}) \cap \{1, \ldots, l\} \subseteq P_{n+1}(\mathbf{s}', \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}).$$

(This equivalence relies on the fact $P_n(\mathbf{s}') = P_n(\mathbf{s})$ due to Lemma 5.5.)

Observation 5.7. For each

$$\tilde{\mathbf{s}} = (\tilde{\alpha}_k, \tilde{\beta}_k, \tilde{\gamma}_k, \tilde{\delta}_k)_{i=1}^n \in \mathcal{E}_n'(\mathbf{s}), \quad (\alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \in S^4,$$

$(\tilde{\beta}_m, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \in \mathcal{A}_1$ if and only if $\#P_{n+1}(\tilde{\mathbf{s}}, \alpha_{n+1}, \ldots, \delta_{n+1}) \geqslant \#P_n(\mathbf{s}) - 1$.

To see this, suppose first that $(\tilde{\beta}_m, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \in \mathcal{A}_1$. Then $(\tilde{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1})$ is pivoted from $(\mathbf{s}'(\tilde{\beta}_m), \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1})$, as they differ only at entries β_k for $k \in P_n(\mathbf{s}) \cap \{1, \ldots, l\} \subseteq P_{n+1}(\mathbf{s}', \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1})$. Lemma 5.5 then implies that

$$P_n(\mathbf{s}) \cap \{1, \ldots, l\} \subseteq P_{n+1}(\mathbf{s}', \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) = P_{n+1}(\tilde{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1})$$

and $\#P_{n+1}(\tilde{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \geqslant \#P_n(\mathbf{s}) - 1$.

Conversely, suppose $\#P_{n+1}(\tilde{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \geqslant \#P_n(\mathbf{s}) - 1$. $(*)$ Recall that $P_n(\tilde{\mathbf{s}}) = P_n(\mathbf{s})$, and recall that $P_{n+1}(\tilde{\mathbf{s}}, \alpha_{n+1}, \ldots, \delta_{n+1})$ is either an initial section of $P_n(\tilde{s})$ or contains $P_n(\tilde{s})$. Considering these, the assumption $(*)$ implies

$$P_n(\mathbf{s}) \cap \{1, \ldots, l\} \subseteq P_{n+1}(\tilde{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}).$$

Then $(\mathbf{s}'(\tilde{\beta}_m), \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1})$ is pivoted from $(\tilde{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1})$, as the former choice differs from the latter choice only at entries $(\tilde{\alpha}_k, \tilde{\beta}_k, \tilde{\gamma}_k)$'s for $k \in P_n(\mathbf{s}) \cap \{1, \ldots, l\} \subseteq P_{n+1}(\tilde{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1})$. Lemma 5.5 then implies that

$$P_n(\mathbf{s}) \cap \{1, \ldots, l\} \subseteq P_{n+1}(\tilde{\mathbf{s}}, \alpha_{n+1}, \ldots, \delta_{n+1}) = P_{n+1}(\mathbf{s}', \alpha_{n+1}, \ldots, \delta_{n+1})$$

and $(\tilde{\beta}_m, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \in \mathcal{A}_1$.

Combining Observation 5.7 and inequality (25), we deduce

$$\begin{aligned}
&\mathbb{P}(\mathcal{A}_1 \mid \tilde{\mathrm{S}}_m \times S^4) \\
&\quad = \mathbb{P}((\tilde{\beta}_m, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \in \mathcal{A}_1 \mid \tilde{\mathbf{s}} \in \mathcal{E}_n'(\mathbf{s}), (\alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \in S^4) \\
&\quad = \mathbb{P}(\#P_{n+1}(\tilde{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \geqslant \#P_n(\mathbf{s}) - 1 \mid \tilde{\mathbf{s}} \in \mathcal{E}_n'(\mathbf{s}), (\alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \in S^4) \\
&\quad \geqslant 1 - \left(\frac{4}{N_0}\right)^2.
\end{aligned}$$

This time, we freeze the coordinates β_k except for $k = l$: for $\tilde{\mathbf{s}} \in \mathcal{E}_n'(\mathbf{s})$, let $E^{(l)}(\tilde{\mathbf{s}})$ be the set of $\bar{\mathbf{s}} \in \mathcal{E}_n'(\mathbf{s})$ such that $\bar{\beta}_k = \tilde{\beta}_k$ for $k \in P_n(\mathbf{s}) \setminus \{l\}$. Then $\{E^{(l)}(\tilde{\mathbf{s}}) : \tilde{\mathbf{s}} \in \mathcal{E}_n'(\mathbf{s})\}$ partitions $\mathcal{E}_n'(\mathbf{s})$ and $E^{(l)}(\tilde{\mathbf{s}})$ is parameterized by $\bar{\beta}_l \in \tilde{\mathrm{S}}_l$ with the uniform measure; note that $\#\tilde{\mathrm{S}}_l \geqslant \#S - 1$.

Fixing $\tilde{\mathbf{s}}$, now pick $(\alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \in S^4$, and let $F^{(l)}(\tilde{\mathbf{s}})$ be the set of $\bar{\mathbf{s}} \in E^{(m)}(\tilde{\mathbf{s}})$ such that $(\Upsilon(\bar{\beta}_l), \bar{w}_{n+2,2}^- o)$ is K_0-aligned, i.e.,

$$\operatorname{diam}(\pi_{\Gamma^{-1}(\bar{\beta}_l)}((\tilde{w}_{l,0}^-)^{-1} \tilde{w}_{n,2}^- a_{n+1} b_{n+1} v_{n+1} c_{n+1} d_{n+1} w_{n+1} o) \cup o) < K_0. \tag{26}$$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

This amounts to requiring another Schottky condition to $\bar{\beta}_l$; there are at least $\#S-2$ choices that additionally satisfy this.

We now claim that $\#P_{n+1}(\bar{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \geqslant \#P_n(\mathbf{s}) - 2$ for $\bar{\mathbf{s}} \in F^{(l)}(\tilde{\mathbf{s}})$. First, since $l' < l$ are consecutive elements in $P_n(\bar{\mathbf{s}})$, Lemma 5.3 asserts that

$$(\Upsilon(\bar{\delta}_{l'}), \Upsilon(\bar{\alpha}_l)),$$

is D_0-semi-aligned. Moreover, Lemma 5.2 and Condition (26) imply that

$$(\Upsilon(\bar{\alpha}_l), \Upsilon(\bar{\beta}_l)), (\Upsilon(\bar{\beta}_l), \bar{w}_{n+2,2}^{-}o),$$

are D_0-aligned and K_0-aligned, respectively. Combining these, we observe that the pair (l', l) qualifies criterion (B) at step $n+1$. This implies $P_{n+1}(\bar{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \supseteq P_n(\bar{\mathbf{s}}) \cap \{1, \ldots, l'\}$, hence the claim.

As a result, for each $\tilde{\mathbf{s}} \in \mathcal{E}_n(\mathbf{s})$ and $(\alpha_{n+1}, \ldots, \delta_{n+1}) \in S^4$ we have

$$\begin{aligned} &\mathbb{P}(\#P_{n+1}(\bar{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) < \#P_n(\mathbf{s}) - 2 \mid \bar{\mathbf{s}} \in E^{(l)}(\tilde{\mathbf{s}})) \\ &\leqslant \frac{\#[E^{(l)}(\tilde{\mathbf{s}}) \setminus F^{(l)}(\tilde{\mathbf{s}})]}{\#E^{(l)}(\tilde{\mathbf{s}})} \leqslant \frac{2}{\#S-1} \leqslant \frac{3}{N_0}. \end{aligned}$$

Moreover, Observation 5.7 asserts that the above probability vanishes for those equivalence classes $E^{(l)}(\tilde{\mathbf{s}})$ such that $(\tilde{\beta}_m, \alpha_{n+1}, \ldots, \delta_{n+1}) \in \mathcal{A}_1$. Since $\mathbb{P}[\mathcal{A}_1 | \tilde{\mathrm{S}}_m \times S^4] \leqslant (4/N_0)^2$, we conclude

$$\begin{aligned} &\mathbb{P}(\#P_{n+1}(\tilde{\mathbf{s}}, \alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) < \#P_n(\mathbf{s}) - 2 \mid \tilde{\mathbf{s}} \in \mathcal{E}_n'(\mathbf{s}), (\alpha_{n+1}, \beta_{n+1}, \gamma_{n+1}, \delta_{n+1}) \in S^4) \\ &\leqslant \left(\frac{4}{N_0}\right)^2 \times \frac{4}{N_0} \leqslant \left(\frac{4}{N_0}\right)^3. \end{aligned} \tag{27}$$

We repeat this procedure for $j < \#P_n(\mathbf{s})$. The case $j \geqslant \#P_n(\mathbf{s})$ is void. □

COROLLARY 5.8. *Let us fix* $\mathbf{w}$ *and* $\mathbf{v}$. *When* $\mathbf{s} = (\alpha_i, \beta_i, \gamma_i, \delta_i)_{i=1}^n$ *is chosen from* S^{4n} *with the uniform measure,* $\#P_n(\mathbf{s})$ *is greater in distribution than the sum of* n *i.i.d.* X_i, *whose distribution is given by*

$$\mathbb{P}(X_i = j) = \begin{cases} (N_0 - 4)/N_0, & \text{if } j = 1, \\ (N_0 - 4)4^{-j}/N_0^{-j+1}, & \text{if } j < 0, \\ 0, & \text{otherwise.} \end{cases} \tag{28}$$

Moreover, we have $\mathbb{P}(\#P_n(\mathbf{s}) \leqslant (1 - 10/N_0)n) \leqslant e^{-n/K}$ *for some* $K > 0$.

Proof. Let $\{X_i\}_i$ be the family of i.i.d. as in Equation (28) that is also assumed to be independent of the choice $\mathbf{s}$. Lemma 5.4 and Proposition 5.6 together imply the following for each $0 \leqslant k < n$:

$$\mathbb{P}(\#P_{k+1}(\mathbf{s}) \geqslant i + j \mid \#P_k(\mathbf{s}) = i) \geqslant \begin{cases} 1 - \frac{4}{N_0} & \text{if } j = 1, \\ 1 - (\frac{4}{N_0})^{-j+1} & \text{if } j < 0. \end{cases} \quad (i = 0, 1, 2, \ldots) \tag{29}$$

Hence, there exists a non-negative random variable U_k such that $\#P_{k+1} - U_k$ and $\#P_k + X_{k+1}$ have the same distribution.

For each $1 \leqslant k \leqslant n$, we claim that $\mathbb{P}(\#P_k \geqslant i) \geqslant \mathbb{P}(X_1 + \cdots + X_k \geqslant i)$ for each i. For $k = 1$, we have $\#P_{k-1} = 0$ and the claim follows from inequality (29). Given the claim for k, we have

$$\begin{aligned} \mathbb{P}(\#P_{k+1} \geqslant i) \geqslant \mathbb{P}(\#P_k + X_{k+1} \geqslant i) &= \sum_j \mathbb{P}(\#P_k \geqslant j)\mathbb{P}(X_{k+1} = i - j) \\ &\geqslant \sum_j \mathbb{P}(X_1 + \cdots + X_k \geqslant j)\mathbb{P}(X_{k+1} = i - j) \\ &= \mathbb{P}(X_1 + \cdots + X_k + X_{k+1} \geqslant i). \end{aligned}$$

The second claim holds since X_i have finite exponential moment and $\mathbb{E}[X_i] \geqslant 1 - 9/N_0$. □

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

We now describe a simpler situation when $v_i = \mathrm{id}$ for all i, i.e., we study

$$w_0 a_1 b_1 c_1 d_1 w_1 \cdots a_n b_n c_n d_n w_n \cdots .$$

Before defining the pivoting, note that the conditions

$$(\bar{\Gamma}^-(\beta), \Pi(\gamma)o), (o, \Gamma^+(\gamma)) \quad \text{are } K_0\text{-aligned}$$

and are satisfied by every pair of Schottky sequences $(\beta, \gamma) \in S^2$, as we proved in Lemma 5.2. Hence, criterion (A) while defining the set of pivotal times is simplified as follows.

(A') When $(z_{n-1}, \Upsilon(\alpha_n))$ and $(\Upsilon(\delta_n), w_{n+1,2}^- o)$ are K_0-aligned, we set $P_n = P_{n-1} \cup \{n\}$ and $z_n = w_{n,1}^+ o$ (see Figure 4).

Moreover, $\tilde{\mathrm{S}}$ contains all of $\{(\beta, \gamma, \mathrm{id}) : \beta, \gamma \in S\}$. Hence, when $\mathbf{w}$ and $\mathbf{v} = (\mathrm{id})_{i=1}^{\infty}$ are fixed, the previous definition reads as follows: given a choice $\mathbf{s} = (\alpha_1, \beta_1, \ldots, \gamma_n, \delta_n)$ in S^{4n}, we say that $\bar{\mathbf{s}} \in S^{4n}$ is *pivoted from* $\mathbf{s}$ if:

- $\alpha_i = \bar{\alpha}_i$, $\delta_i = \bar{\delta}_i$ for all $i \in \{1, \ldots, n\}$;
- $(\beta_i, \gamma_i) = (\bar{\beta}_i, \bar{\gamma}_i)$ for each $i \in \{1, \ldots, n\} \setminus P_n(\mathbf{s})$.

Therefore, for $\mathbf{s} \in S^{4n}$, $\bar{\mathbf{s}} \in \mathcal{E}_n(\mathbf{s})$ is parameterized by the choices $(\bar{\beta}_i, \bar{\gamma}_i)_{i \in P_n(\mathbf{s})}$ distributed according to the uniform measure on $S^{2\#P_n(\mathbf{s})}$.

5.2 Pivotal times in random walks

In this subsection, we define pivotal times for random walks and prove Proposition 4.2. Let μ be a non-elementary probability measure on G and $S \subseteq (\mathrm{supp}\mu)^{M_0}$ be a large enough K_0-Schottky set with cardinality $N_0 \geqslant 400$. We also fix an integer $n \geq 0$.

Let μ_S be the uniform measure on S. By taking suitably small α, we can decompose μ^{4M_0} as

$$\mu^{4M_0} = \alpha \mu_S^4 + (1 - \alpha)\nu,$$

for some probability measure ν. We then consider Bernoulli RVs $(\rho_i)_i$ with $\mathbb{P}(\rho_i = 1) = \alpha$ and $\mathbb{P}(\rho_i = 0) = 1 - \alpha$, $(\eta_i)_i$ with the law μ_S^4 and $(\nu_i)_i$ with the law ν, all independent, and define

$$(g_{4M_0 k+1}, \ldots, g_{4M_0 k + 4M_0}) = \begin{cases} \nu_k \text{ when } \rho_k = 0, \\ \eta_k \text{ when } \rho_k = 1. \end{cases}$$

Then $(g_i)_{i=1}^{\infty}$ has the law μ^{∞}. Since we need to prove Proposition 4.2 by fixing the choice of $g_1, \ldots, g_{\lfloor n/2 \rfloor + 1}$ and g_{n+1}, we slightly modify ρ_i, namely,

$$\rho_i^{(n)} := \begin{cases} 0, & \text{if } i \leqslant n/8M_0 \text{ or } i = \lfloor n/4M_0 \rfloor, \\ \rho_i, & \text{otherwise.} \end{cases}$$

Let Ω be the ambient probability space on which the above RVs are all measurable. We denote by $\mathscr{B}(k) := \sum_{i=0}^{k} \rho_i^{(n)}$ the number of the Schottky slots until k and by $\vartheta(i) := \min\{j \geqslant 0 : \mathscr{B}(j) = i\}$ the ith Schottky slot. We also set $\vartheta(0) = -1$. Note that $\{j \geqslant 0 : \rho_j^{(n)} = 1\} = \{\vartheta(1) < \vartheta(2) < \cdots\}$.

For each $\omega \in \Omega$ and $i \geqslant 1$ we define

$$\begin{aligned} w_{i-1} &:= g_{4M_0[\vartheta(i-1)+1]+1} \cdots g_{4M_0 \vartheta(i)}, \\ \alpha_i &:= (g_{4M_0\vartheta(i)+1}, \ldots, g_{4M_0\vartheta(i)+M_0}), \\ \beta_i &:= (g_{4M_0\vartheta(i)+M_0+1}, \ldots, g_{4M_0\vartheta(i)+2M_0}), \end{aligned}$$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

$$\gamma_i := (g_{4M_0\vartheta(i)+2M_0+1}, \ldots, g_{4M_0\vartheta(i)+3M_0}),$$
$$\delta_i := (g_{4M_0\vartheta(i)+3M_0+1}, \ldots, g_{4M_0\vartheta(i)+4M_0}).$$

In other words, $\eta_{\vartheta(i)}$ becomes $(\alpha_i, \beta_i, \gamma_i, \delta_i)$ (with M_0 steps each) and w_i is the product of intermediate steps between $\eta_{\vartheta(i-1)}$ and $\eta_{\vartheta(i)}$. As in § 5.1, we write $a_i := \Pi(\alpha_i)$, $b_i := \Pi(\beta_i)$ and so on. We then have

$$\omega_{4M_0\vartheta(l+1)} = w_0 a_1 b_1 v_1 c_1 d_1 w_1 \cdots a_l b_l v_l c_l d_l w_l, \tag{30}$$

for each $l > 0$. Following the discussion in § 5.1, we define

$$\begin{aligned} P_1(\omega) &= P_1((a_1, b_1, c_1, d_1); (w_i)_{i=0}^1, (\mathrm{id})_{i=1}^1), \\ P_2(\omega) &= P_2((a_i, b_i, c_i, d_i)_{i=1}^2; (w_i)_{i=0}^2, (\mathrm{id})_{i=1}^2), \\ &\ \ \vdots \end{aligned} \tag{31}$$

and we finally define

$$\mathcal{P}(\omega) := \{4M_0\vartheta(i) + 2M_0 : i \in \liminf_k P_k(\omega)\}.$$

Recall that P_k is formed from P_{k-1} by adjoining a new element k or taking an initial section of P_{k-1}. Hence, any initial section $\{i(1) < \cdots < i(N)\}$ of $\liminf_k P_k(\omega)$ is an initial section of some $P_m(\omega)$ (in fact, for all sufficiently large m). Proposition 5.1 then tells us the following.

Observation 5.9. Let $\mathcal{P}(\omega) = \{i(1) < i(2) < \cdots\}$. Then

$$(o, \Upsilon(\alpha_{i(1)}), \Upsilon(\beta_{i(1)}), \Upsilon(\gamma_{i(1)}), \Upsilon(\delta_{i(1)}), \ldots, \Upsilon(\alpha_{i(k)}), \Upsilon(\beta_{i(k)}), \Upsilon(\gamma_{i(k)}), \Upsilon(\delta_{i(k)}), \ldots)$$

is D_0-semi-aligned.

Note that $(w_i)_i$ and $(\alpha_i, \beta_i, \gamma_i, \delta_i)_{i>0}$ are independent, the latter being i.i.d. with the uniform distribution on S^4. By Corollary 5.8, P_k linearly increases.

Observation 5.10. There exists $K > 1$ such that

$$\mathbb{P}(\#P_k(\omega) \leqslant k/K \mid w_0, w_1, \ldots) \leqslant K e^{-k/K} \tag{32}$$

for every $k > 0$ and every choice of $(w_i)_i$.

Here, the growth rate is independent of the g_i that are not involved in $(\alpha_i, \beta_i, \gamma_i, \delta_i)$. In particular, it is independent of $g_1, \ldots, g_{\lfloor n/2 \rfloor + 1}$ and g_{n+1}.

To couple the words $w_{k,2}^-$ and the actual random walk Z_n, we need to control $\vartheta(i)$. For each $k \geqslant \lfloor n/4M_0 \rfloor$ and $L > 0$, we have

$$\begin{aligned} \mathbb{P}(\mathscr{B}(k) \leqslant L) \cdot e^{-L} \leqslant \mathbb{E}[e^{-\mathscr{B}(k)}] &= \prod_{i=1}^{k} \mathbb{E}[\exp(-\rho_i^{(n)})] && (\because \text{Markov's inequality}) \\ &= \prod_{n/8M_0 < i \leqslant k, i \neq \lfloor n/4M_0 \rfloor} \mathbb{E}[\exp(-\rho_i^{(n)})] \\ &= (1 - \alpha(1 - e^{-\alpha}))^{k - \lceil n/8M_0 \rceil - 1} \\ &\leqslant (1 - \alpha^2/2)^{n/3M_0 - 4} && (\because e^{-\alpha} \leqslant \alpha/2 \text{ for } 0 \leqslant \alpha \leqslant 1). \end{aligned}$$

By plugging in $L = ((\log(1 + \alpha^2/2))/3M_0)k$, we obtain

$$\mathbb{P}(\mathscr{B}(k) \leqslant k/K') \leqslant K' e^{-k/K'} \quad (k \geqslant n/4M_0), \tag{33}$$

for some $K' = K'(\alpha, M_0) > 1$ (independent of α).

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Let us now combine the ingredients and prove Proposition 4.2. Given the measure μ, integers $k \geqslant n$, and the choices of $g_1, \ldots, g_{\lfloor n/2 \rfloor+1}$ and g_{n+1}, we do the above construction. Then we have

$$\mathbb{P}(A_k := \{\mathscr{B}(\lfloor k/4M_0 \rfloor) \leqslant \lfloor k/4K'M_0 \rfloor\}) \leqslant K' e^{-\lfloor k/4K'M_0 \rfloor}.$$

Let us fix a combination of the values of $\{\rho_i^{(n)} : i > 0\}$ in A_k^c, which determines $\vartheta(i)$. Furthermore, we fix a combination of the values of $\{\eta_i : i > 0\}$. These choices determine

$$\{g_j : j \notin \cup_{i>0}\{4M_0\vartheta(i)+1, \ldots, 4M_0\vartheta(i)+4M_0\}\}, \tag{34}$$

and consequently w_i. Note also that $\{(\alpha_i, \beta_i, \gamma_i, \delta_i) = \eta_{\vartheta(i)} : i > 0\}$ are i.i.d.s distributed according to μ_S^4. Hence, conditioned on choices of $\{\rho_i^{(n)} : i > 0\} \in A_k^c$ and $\{\eta_i : i > 0\}$, we are now reduced to the combinatorial model. From Observation 5.10, we deduce that

$$\begin{aligned} &\mathbb{P}(\#P_l(\omega) \geqslant \lfloor k/4KK'M_0 \rfloor \text{ for all } l \geqslant \lfloor k/4K'M_0 \rfloor \mid w_0, w_1, \ldots) \\ &\geqslant 1 - K \sum_{l \geqslant \lfloor k/4K'M_0 \rfloor} e^{-l/K} \geqslant 1 - \frac{K}{1-e^{-1/K}} e^{-\lfloor k/4KK'M_0 \rfloor}. \end{aligned}$$

In other words, except for probability $(K/(1-e^{-1/K}))e^{-\lfloor k/4KK'M_0 \rfloor}$ (under the conditioning), the initial $\lfloor k/4KK'M_0 \rfloor$-sections of $P_{\lfloor k/4K'M_0 \rfloor}(\omega)$ remain the same in $P_l(\omega)$ for $l \geqslant \lfloor k/4M_0 \rfloor$. Hence, it becomes an initial section of $\liminf_l P_l(\omega)$. This means that

$$\#(\mathcal{P}(\omega) \cap \{4M_0\vartheta(i)+2M_0 : i \in P_{\lfloor k/4K'M_0 \rfloor}(\omega)\}) \geqslant \lfloor k/4KK'M_0 \rfloor. \tag{35}$$

Meanwhile, since $\{\rho_i^{(n)} : i > 0\}$ is determined in A_k^c, we have $\mathscr{B}(\lfloor k/4M_0 \rfloor) > \lfloor k/4K'M_0 \rfloor$ and

$$P_{\lfloor k/4K'M_0 \rfloor}(\omega) \subseteq \{\vartheta(1), \ldots, \vartheta(\lfloor k/4K'M_0 \rfloor)\} \subseteq \{1, \ldots, \lfloor k/4M_0 \rfloor - 1\}.$$

This implies $\{4M_0\vartheta(i)+2M_0 : i \in P_{\lfloor k/4K'M_0 \rfloor}(\omega)\} \subseteq \{1, \ldots, k-2M_0\}$. Combined with display (35), this implies

$$\#(\mathcal{P}(\omega) \cap \{1, \ldots, k\}) \geqslant \lfloor k/4KK'M_0 \rfloor. \tag{36}$$

Summing up the conditional probabilities, we have

$$\mathbb{P}(\#(\mathcal{P}(\omega) \cap \{1, \ldots, k\}) \geqslant \lfloor k/4KK'M_0 \rfloor \,|A_k^c) \geqslant 1 - \frac{K}{1-e^{-1/K}} e^{-\lfloor k/4KK'M_0 \rfloor}.$$

Since $\mathbb{P}(A_k)$ decays exponentially, we conclude inequality (7).

It remains to partition the probability space Ω into pivotal equivalence classes that satisfy Definition 4.1, with $\mathcal{P}(\omega)$ as the set of pivotal times. We say that $\bar{\omega} \in \Omega$ is pivoted from ω if they only differ in the value of β_i for $i \in \liminf_l P_l(\omega)$. Then, being pivoted from each other is an equivalence relation. On an equivalence class $\mathcal{E}$, all random paths have the same set of pivotal times $\mathcal{P}(\mathcal{E}) = \{j(1) < j(2) < \cdots\} \subseteq M_0\mathbb{Z}$ that avoids $1, \ldots, \lfloor n/2 \rfloor$ and n. Moreover, the steps g_i are uniform across $\mathcal{E}$ except for

$$s_k := (g_{j(k)-M_0+1}, g_{j(k)-M_0+2}, \ldots, g_{j(k)}) \quad (k = 1, 2, \ldots),$$

which are i.i.d.s chosen from S according to μ_S. Lastly, observe that

$$\begin{aligned} \Upsilon(\beta_{i(k)}) &= (Z_{(4M_0+M')\vartheta(i(k))+M_0}o, \ldots, Z_{(4M_0+M')\vartheta(i(k))+2M_0}o) \\ &= (Z_{j(k)-M_0}o, \ldots, Z_{j(k)}o) = \mathbf{Y}_{j(k)}. \end{aligned}$$

By Observation 5.9, $(o, \mathbf{Y}_{j(1)}, \mathbf{Y}_{j(2)}, \ldots)$ is always D_0-semi-aligned. Proposition 4.2 is now proved.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

6. Large deviation principles

In this section, we consider a more delicate pivoting that leads to the large deviation principle. Definition 6.1 and Proposition 6.2 rephrases Gouëzel's result in [Gou22, Section 5A] in terms of strongly contracting isometries.

Definition 6.1. Let μ and ν be non-elementary probability measures on G and $(\Omega, \mathbb{P})$ be a probability space for μ. Let $0 < \epsilon < 1$, let $K_0, N > 0$ and let $S \subseteq (\mathrm{supp}\mu)^{M_0}$ be a long enough and large K_0-Schottky set for μ.

A subset $\mathcal{E}$ of Ω is called an (n, N, ϵ, ν)*-pivotal equivalence class for* μ, associated with the *set of pivotal times*

$$\mathcal{P}^{(n,N,\epsilon,\nu)}(\mathcal{E}) = \{j(1) < j'(1) < \cdots < j(\#\mathcal{P}/2) < j'(\#\mathcal{P}/2)\} \subseteq M_0\mathbb{Z}_{>0},$$

if the following hold:

(i) for each $\omega \in \mathcal{E}$ and $k \geqslant 1$,

$$\begin{aligned} s_k(\omega) &:= (g_{j(k)-M_0+1}(\omega), g_{j(k)-M_0+2}(\omega), \ldots, g_{j(k)}(\omega)),\\ s'_k(\omega) &:= (g_{j'(k)-M_0+1}(\omega), g_{j'(k)-M_0+2}(\omega), \ldots, g_{j'(k)}(\omega)), \end{aligned}$$

are Schottky sequences;

(ii) for each $\omega \in \mathcal{E}$, $(o, \mathbf{Y}_{j(1)}, \mathbf{Y}_{j'(1)}, \ldots, \mathbf{Y}_{j(\#\mathcal{P}/2)}, \mathbf{Y}_{j'(\#\mathcal{P}/2)}, Z_n o)$ is D_0-semi-aligned;

(iii) for the RV defined as

$$r_k := g_{j(k)+1}g_{j(k)+2}\cdots g_{j'(k)-M_0},$$

$(s_k, s'_k, r_k)_{k>0}$ on $\mathcal{E}$ are i.i.d.s and r_k are distributed almost according to $\mu^{*2M_0N} * \nu^{*((j'(k)-j(k))/2M_0)-N-0.5}$ in the sense that the following holds for every $g \in G$:

$$\begin{aligned} &(1-\epsilon)(\mu^{*2M_0N} * \nu^{*((j'(k)-j(k))/2M_0)-N-0.5})(g)\\ &\quad \leqslant \mathbb{P}(r_k = g) \leqslant (1+\epsilon)(\mu^{*2M_0N} * \nu^{*((j'(k)-j(k))/2M_0)-N-0.5})(g) \end{aligned}$$

for each $g \in G$.

Proposition 6.2. *Let $M_0 > 0$, μ be a non-elementary probability measure on G, let $0 < \epsilon < 1$ and let $S \subseteq (\mathrm{supp}\mu)^{M_0}$ be a long enough and large Schottky set for μ with cardinality greater than $100/\epsilon$. Then there exists a non-elementary probability measure ν on G such that the following holds.*

For each sufficiently large integer N, there exists $K > 0$ such that for each n we have a probability space $(\Omega, \mathbb{P})$ for μ and its measurable partition $\mathscr{P}_{n,N,\epsilon,\nu} = \{\mathcal{E}_\alpha\}_\alpha$ into (n, N, ϵ, ν)-pivotal equivalence classes that satisfies

$$\mathbb{P}\left(\omega : \frac{1}{2}\#\mathcal{P}^{(n,N,\epsilon,\nu)}(\omega) \leqslant (1-\epsilon)\frac{n}{2M_0N}\right) \leqslant Ke^{-n/K}. \tag{37}$$

We will in fact prove a statement that is more explicit than Proposition 6.2.

Proposition 6.3. *Let $0 < \epsilon < 1$, let $K_0, M_0 > 0$ and let $S \subseteq G^{M_0}$ be a long enough and large K_0-Schottky set with $\#S \geqslant 100/\epsilon$. Let μ be a probability measure on G such that $m := \min\{\mu^{M_0}(s) : s \in S\}$ is positive. Let $N > 40/m^2\epsilon$ and let ν be the measure defined by*

$$\nu = \frac{1}{1-0.5m^2}(\mu^{*2M_0} - 0.5m^2 \cdot (\text{uniform measure on } \{\Pi(s)\Pi(s') : s, s' \in S\})).$$

Then ν is a non-elementary probability measure. Moreover, there exists $K > 0$ depending only on S, m, N and ϵ (but not on μ) such that, for each n, we have a probability space $(\Omega, \mathbb{P})$ for μ and

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

its measurable partition $\mathscr{P}_{n,N,\epsilon,\nu} = \{\mathcal{E}_\alpha\}_\alpha$ *into* (n, N, ϵ, ν)*-pivotal equivalence classes, associated with the set of pivotal times* $\mathcal{P}^{(n,N,\epsilon,\nu)}$*, that satisfies*

$$\mathbb{P}\left(\omega : \frac{1}{2}\#\mathcal{P}^{(n,N,\epsilon,\nu)}(\omega) \leqslant (1-\epsilon)\frac{n}{2M_0N}\right) \leqslant Ke^{-n/K} \quad (\forall n \in \mathbb{Z}_{>0}).$$

Gouëzel proved Proposition 6.2 for random walks on a Gromov hyperbolic space in [Gou22, Section 5C]. We adapt his proof to our setting here.

Proof. Let us denote the uniform measure on S by μ_S. In this proof, when a probability measure τ on G^k is given, we denote by τ^* the pushforward measure by convolution

$$\tau^*(g) := \sum_{(g_1,\ldots,g_k)\in G^k,\ g_1\cdots g_k = g} \tau(g_1, \ldots, g_k).$$

Let $N_0 = \#S$ be the cardinality of S. Note that $10/N_0 \leqslant \epsilon/10$. Consider the decomposition

$$\mu^{2M_0} = 0.5m^2\mu_S^2 + (1-0.5m^2)\tau, \tag{38}$$

where τ is a probability measure on G^{2M_0} with $\tau^* = \nu$. Recall that S is a long enough and large K_0-Schottky set, so there exists $a, b \in S$ such that $\Pi(a)$ and $\Pi(b)$ are independent strongly contracting isometries. Since τ has the same support with μ^{2M_0}, ν puts nonzero weights on a^2 and b^2. Hence, ν is non-elementary.

Given the decomposition as in Equation (38), we consider Bernoulli RVs $(\rho_i)_{i\geq 0}$ with $\mathbb{P}(\rho_i = 1) = 0.5m^2$ and $\mathbb{P}(\rho_i = 0) = 1 - 0.5m^2$, $(\eta_i)_{i\geq 0}$ with the law of μ_S^2, $(\tau_i)_i$ with the law of τ, and $(\xi_i)_{i\geqslant 0}$ with the law of μ^{2M_0}, all independent. We define RVs $\{t_j, t'_j\}_{j=1}^{\infty}$. First, t_1 is the smallest $i > 0$ with $\rho_i = 1$, and $t'_1 := \min\{i > t_1 + N : \rho_i = 1\}$. Inductively, we define

$$t_k := \min\{i > t'_{k-1} : \rho_i = 1\}, \quad t'_k := \min\{i > t_k + N : \rho_i = 1\}.$$

For convenience, we set $t'_0 := 0$. We then define

$$(g_{2M_0(k-1)+1}, \ldots, g_{2M_0(k-1)+2M_0}) := \begin{cases} \eta_k, & \text{when } k \in \{t_j, t'_j\}_{j=1}^{\infty}, \\ \xi_k, & \text{when } t_j + 1 \leqslant k \leqslant t_j + N \text{ for some } j, \\ \tau_k, & \text{otherwise.} \end{cases}$$

Then $(g_i)_{i=1}^{\infty}$ is distributed according to the product measure μ^∞ [Gou22, Claim 5.11]. We let $\mathscr{B}(k) := \#\{j \geqslant 1 : t'_j < k\}$. Now define

$$\begin{aligned} w_{i-1} &:= g_{2M_0t'_{i-1}+1}\cdots g_{2M_0(t_i-1)}, \\ \alpha_i &:= (g_{2M_0t_i-2M_0+1}, \ldots, g_{2M_0t_i-M_0}), \\ \beta_i &:= (g_{2M_0t_i-M_0+1}, \ldots, g_{2M_0t_i}), \\ v_i &:= g_{2M_0t_i+1}\cdots g_{2M_0t'_i-2M_0}, \\ \gamma_i &:= (g_{2M_0t'_i-2M_0+1}, \ldots, g_{2M_0t'_i-M_0}), \\ \delta_i &:= (g_{2M_0t'_i-M_0+1}, \ldots, g_{2M_0t'_i}), \end{aligned}$$

for $i = 1, \ldots, \mathscr{B}(\lfloor n/2M_0 \rfloor)$ and define $w_{\mathscr{B}(\lfloor n/2M_0\rfloor)} = g_{2M_0t'_{\mathscr{B}(\lfloor n/2M_0\rfloor)}+1}\cdots g_n$. Using these data, we define the set of pivotal times

$$P_{\mathscr{B}(\lfloor n/2M_0\rfloor)}(\omega) = P_{\mathscr{B}(\lfloor n/2M_0\rfloor)}((\alpha_i, \beta_i, \gamma_i, \delta_i)_{i=1}^{\mathscr{B}(\lfloor n/2M_0\rfloor)}; (w_i)_{i=0}^{\mathscr{B}(\lfloor n/2M_0\rfloor)}, (v_i)_{i=1}^{\mathscr{B}(\lfloor n/2M_0\rfloor)})$$

as in § 5.1.

We first determine the values of ρ_j. Observe that $\mathscr{B}(\lfloor n/2M_0\rfloor)$ and $\{t_j, t'_j\}_j$ depend solely on $\{\rho_j\}_j$ and counts the renewal times in $[0, n/2M_0]$ formed with a geometric distribution after a delay N. More explicitly, if we 'omit' ρ_{t_k+i} for $k > 0$ and $i = 1, \ldots, N$ and define

$$(\rho'_1, \rho'_2, \rho'_3, \ldots) := (\rho_1, \ldots, \rho_{t_1}, \rho_{t_1+N+1}, \rho_{t_1+N+2}, \ldots, \rho_{t_2}, \rho_{t_2+N+1}, \ldots),$$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

then $\{\rho'_i\}_i$ are i.i.d. Bernoulli RVs and $t'_k = kN + \min\{j : \sum_{i=1}^{j} \rho'_i = 2k\}$. Hence, we have

$$\begin{aligned}\mathbb{P}\bigg(\mathscr{B}(\lfloor n/2M_0\rfloor) < (1-\epsilon/10)\frac{n}{2M_0N}\bigg) &= \mathbb{P}(t'_{\lceil(1-\epsilon/10)n/2M_0N\rceil} > \lfloor n/2M_0\rfloor) \\ &= \mathbb{P}\bigg(\bigg\lceil(1-\epsilon/10)\frac{n}{2M_0}\bigg\rceil + \min\bigg\{j : \sum_{i=1}^{j}\rho'_i = \bigg\lceil(1-\epsilon/10)\frac{n}{M_0N}\bigg\rceil\bigg\} \geq \bigg\lfloor\frac{n}{2M_0}\bigg\rfloor\bigg) \\ &\leqslant \mathbb{P}\bigg(\sum_{i=1}^{\lceil\epsilon n/20M_0\rceil+3}\rho'_i < (1-\epsilon/10)\frac{n}{M_0N}\bigg),\end{aligned}$$

which decays exponentially because $\mathbb{E}[\rho'_i] = 0.5m^2 > 20/\epsilon N$. Hence, there exists $K_1 > 0$ that depends on m, ϵ and N such that

$$\mathbb{P}\bigg(\mathscr{B}(\lfloor n/2M_0\rfloor) < (1-\epsilon/10)\frac{n}{2M_0N}\bigg) \leqslant K_1 e^{-n/K_1}. \tag{39}$$

Let us fix the choices of $(\rho_i)_{i\geqslant 0}$. This determines $(t_i, t'_i)_{i>0}$ and $\mathscr{B}(\lfloor n/2M_0\rfloor)$. We then fix the data $(\tau_i, \xi_i)_{i>0}$ and $\{\eta_i : i > t'_{\mathscr{B}(\lfloor n/2M_0\rfloor)}\}$. These in turn determine $(w_i)_{i=0}^{\mathscr{B}(\lfloor n/2M_0\rfloor)}$ and $(v_i)_{i=1}^{\mathscr{B}(\lfloor n/2M_0\rfloor)}$. Furthermore

$$(\alpha_i, \beta_i)_{i=1}^{\mathscr{B}(\lfloor n/2M_0\rfloor)} = (\eta_{t_i})_{i=1}^{\mathscr{B}(\lfloor n/2M_0\rfloor)}, \quad (\gamma_i, \delta_i)_{i=1}^{\mathscr{B}(\lfloor n/2M_0\rfloor)} = (\eta_{t'_i})_{i=1}^{\mathscr{B}(\lfloor n/2M_0\rfloor)}$$

are all independent and identically distributed according to μ_S^2. Hence, the situation is reduced to the combinatorial model in § 4. Corollary 5.8 asserts the following for some $K_2 > 0$:

$$\mathbb{P}(\#P_{\mathscr{B}(\lfloor n/2M_0\rfloor)} \leqslant (1-10/N_0)\mathscr{B}(\lfloor n/2M_0\rfloor) \mid (\rho_i, \tau_i, \xi_i)_{i\geqslant 0}) \leqslant K_2 e^{-\mathscr{B}(\lfloor n/2M_0\rfloor)/2M_0K_2}. \tag{40}$$

Combining inequalities (39) and (40), we can conclude that $\mathbb{P}(\#P_{\mathscr{B}(\lfloor n/2M_0\rfloor)} \leqslant (1-\epsilon)n/2M_0N)$ decays exponentially.

Now, given $\omega \in \Omega$ with $P_{\mathscr{B}(\lfloor n/2M_0\rfloor)(\omega)}(\omega) = \{i(1) < i(2) < \cdots\}$, we define

$$\begin{aligned}\mathcal{P}^{(n,N,\epsilon,\nu)}(\omega) &= \{j(1) < j'(1) < j(2) < j'(2) < \cdots\} \\ &:= \{2M_0t_{i(1)}, 2M_0t'_{i(1)} - M_0, 2M_0t_{i(2)}, 2M_0t'_{i(2)} - M_0, \ldots\}.\end{aligned}$$

We just established the estimate is display (37) for this $\mathcal{P}^{(n,N,\epsilon,\nu)}$. Furthermore, note that

$$\begin{aligned}\Upsilon(\beta_{i(k)}) &= (Z_{2M_0t_{i(k)}-M_0}o, \ldots, Z_{2M_0t_{i(k)}}o) = (Z_{j(k)-M_0}o, \ldots, Z_{j(k)}o) = \mathbf{Y}_{j(k)}, \\ \Upsilon(\gamma_{i(k)}) &= (Z_{2M_0t'_{i(k)}-2M_0}o, \ldots, Z_{2M_0t'_{i(k)}-M_0}o) = (Z_{j'(k)-M_0}o, \ldots, Z_{j'(k)}o) = \mathbf{Y}_{j'(k)},\end{aligned}$$

are Schottky axes, and that $w^-_{\mathscr{B}(\lfloor n/2M_0\rfloor)+1,2} = Z_n$. Proposition 5.1 tells us that $(o, \mathbf{Y}_{j(1)}, \mathbf{Y}_{j'(1)}, \mathbf{Y}_{j(2)}, \mathbf{Y}_{j'(2)}, \ldots, Z_n o)$ is always D_0-semi-aligned. This settles items (i) and (ii) in Definition 6.1.

It remains to realize the partition as in Definition 6.1 and check item (iii) in Definition 6.1. We declare the equivalence by pivoting. More precisely, given $\omega \in \Omega$ with $P_{\mathscr{B}(\lfloor n/2M_0\rfloor}(\omega) = \{i(1) < i(2) < \cdots\}$, we declare that another element $\omega' \in \Omega$ is equivalent to ω if it has the same values of $(\rho_i)_{i\geqslant 0}$ (hence the same values of $(t_i, t'_i)_{i>0}$) as ω , and if it has the same values of $(\eta_i, \tau_i, \xi_i)_{i\geqslant 0}$ as ω, possibly except for

$$\{\eta_i : i \in \cup_k\{t_{i(k)}, t'_{i(k)}\}\}, \quad \{\xi_i : i \in \cup_k[t_{i(k)}+1, t_{i(k)}+N]\}, \quad \{\tau_i : i \in \cup_k[t_{i(k)}+N, t'_{i(k+1)}-1]\}.$$

Further, we require that $(\alpha_{i(k)}(\omega'), \delta_{i(k)}(\omega')) = (\alpha_{i(k)}(\omega), \delta_{i(k)}(\omega))$ for each k and

$$(\beta_{i(l)}(\omega'), \gamma_{i(l)}(\omega'), v_{i(l)}(\omega')) \in \tilde{S} \quad (l = 1, \ldots, \#P_{\mathscr{B}(\lfloor n/2M_0\rfloor}).$$

Note that, under this requirement, ω' has the same values of $(w_i)_{i=0}^{\mathscr{B}(\lfloor n/2M_0\rfloor)}$ and $\{v_i : i \neq i(1), \ldots, i(\#P_{\mathscr{B}(\lfloor n/2M_0\rfloor)})\}$ as ω. By Lemma 5.5, we have $P_{\mathscr{B}(\lfloor n/2M_0\rfloor)}(\omega) = P_{\mathscr{B}(\lfloor n/2M_0\rfloor)}(\omega')$, and the above relation becomes an equivalence relation.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Recall that, conditioned on the data $(\rho_i)_{i\geq 0}$, (β_j, γ_j, v_j) is distributed according to $\mu_S^2 \times (\mu^{*2M_0} * (\tau^*)^{*(t_j'-t_j-N-0.5)}) = \mu_S^2 \times (\mu^{*2M_0} * \nu^{*(t_j'-t_j-N-0.5)})$. Now let $\mathcal{E}$ be a pivotal equivalence class that has pivotal times $P_{\mathscr{B}(\lfloor n/2M_0 \rfloor)} = \{i(1) < \cdots < i(m)\}$. Then $(\beta_{i(l)}, \gamma_{i(l)}, v_{i(l)})$ are independent and distributed according to the restriction of $\mu_S^2 \times (\mu^{*2M_0 N} * \nu^{*(t_j'-t_j-N-1)})$ onto the set of 'legitimate choices' $\tilde{\mathrm{S}}$. To describe this, let us define a (not necessarily probability) measure

$$\mu^{(1)}(s', s'', r) := \begin{cases} \mu_S(s')\mu_S(s'')(\mu^{*2M_0 N} * \nu^{*(t_j'-t_j-N-1)})(r), & \text{if } (s', s'', r) \in \tilde{\mathrm{S}}, \\ 0, & \text{otherwise.} \end{cases}$$

Then $(\beta_{i(l)}, \gamma_{i(l)}, v_{i(l)})$ is distributed according to the normalized version $\mu^{(0)}$ of $\mu^{(1)}$, namely, $\mu^{(0)}(A) := (1/(\mu^{(1)}(S^2 \times G)))\mu^{(1)}(A)$ for each $A \subseteq S^2 \times G$.

For each $r \in G$, among N_0^2 choices of s' and s'' in S at least $N_0^2 - 2N_0$ choices qualify the criterion and make $(s', s'', r) \in \tilde{\mathrm{S}}$ by the Schottky property. (See the discussion in displays (20) and (21).) This implies the bound for each $r \in G$,

$$\left(1 - \frac{2}{N_0}\right)(\mu^{*2M_0 N} * \nu^{*(t_j'-t_j-N-1)})(r) \leqslant \mu^{(1)}(S^2 \times \{r\}) \leqslant (\mu^{*2M_0 N} * \nu^{*(t_j'-t_j-N-1)})(r).$$

Summing this up for all $r \in G$, we obtain $1 - 2/N_0 \leqslant \mu^{(1)}(S^2 \times G) \leqslant 1$. Combining these two estimates, we conclude the following for every $g \in G$:

$$\begin{aligned} \left(1 - \frac{2}{N_0}\right)(\mu^{*2M_0 N} * \nu^{*(t_j'-t_j-N-1)})(r) &\leqslant \mathbb{P}(v_{i(l)} = r) = \mu^{(0)}(S^2 \times \{r\}) \\ &\leqslant \left(1 + \frac{3}{N_0}\right)(\mu^{*2M_0 N} * \nu^{*(t_j'-t_j-N-1)})(r). \end{aligned}$$

This settles item (iii) in Definition 6.1, as desired. □

We now establish the large deviation principle for random walks.

Theorem 6.4. *Let (X, G, o) be as in Convention 2.11 and let $(Z_n)_n$ be the random walk generated by a non-elementary probability measure μ on G. Let $\lambda(\mu) = \lim_n (1/n)\mathbb{E}[d(o, Z_n o)]$ be the drift of μ. Then for each $0 < L < \lambda(\mu)$, the probability $\mathbb{P}(d(o, Z_n o) \leqslant Ln)$ decays exponentially as n goes to infinity.*

Recall that $\lambda(\mu) = +\infty$ when μ has infinite first moment, by Corollary 4.12.

Proof. Due to the subadditivity, we have $\mathbb{E}_{\mu^{*N}}[d(o, go)] \geqslant \lambda(\mu)N$ for each $N > 0$. Since L is smaller than $\lambda(\mu)$, there exists $\epsilon > 0$ such that

$$(1-\epsilon)^3 \lambda(\mu) > L + \epsilon.$$

For this $\epsilon > 0$, let S be a long enough Schottky set for μ with cardinality greater than $100/\epsilon$. By Proposition 6.2, there exists a non-elementary probability measure ν, and for each sufficiently large N, a partition $\mathscr{P}_{n,N,\epsilon}$ into (n, N, ϵ, ν)-pivotal equivalence classes for each n such that

$$\mathbb{P}\left(\omega : \frac{1}{2}\#\mathcal{P}^{(n,N,\epsilon,\nu)}(\omega) \leqslant (1-\epsilon)\frac{n}{2M_0 N}\right)$$

decays exponentially in n. Let $C > 0$ be a constant for ν provided by Corollary 4.11: we have

$$\mathbb{P}_{\nu^{*m}}(h : d(o, gho) \geqslant d(o, go) - C) \geqslant 1 - \epsilon/2,$$

for each $g \in G$ and each $m > 0$. We now fix an N such that $N > C/2M_0\lambda(\mu)\epsilon$.

Let $\mathcal{E}$ be an equivalence class such that $\frac{1}{2}\#\mathcal{P}^{(n,N,\epsilon,\nu)}(\mathcal{E}) \geqslant (1-\epsilon)(n/2M_0 N)$. Then for each $\omega \in \mathcal{E}$, $(o, \mathbf{Y}_{j(1)}, \ldots, \mathbf{Y}_{j'(\#\mathcal{P}/2)}, Z_n o)$ is D_0-semi-aligned. The second inequality in item (ii) of

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Lemma 3.18 tells us that

$$d(o, Z_n o) \geqslant \sum_{i=1}^{\#\mathcal{P}/2} d(Z_{j(i)}o, Z_{j'(i)-M_0}o) = \sum_{i=1}^{\#\mathcal{P}/2} d(o, r_i o) \quad (r_i := g_{j(i)+1} \cdots g_{j'(i)-M_0}).$$

Since r_i are non-negative i.i.d. with

$$\mathbb{E}[d(o, r_i o)] \geqslant (1-\epsilon)\mathbb{E}_{\mu^{*2M_0N}}[d(o, go) - C] \geqslant (1-\epsilon)^2 \cdot 2M_0 N \lambda(\mu),$$

we can apply the classical theory of large deviation. As a result, there exists $K' > 0$ such that

$$\mathbb{P}(d(o, Z_n o) \leq (1-\epsilon)^3 \lambda(\mu) n \,|\, \mathcal{E}) \leqslant K' e^{-n/K'} \quad (\forall n > 0).$$

Summing up this conditional probability, we obtain the desired exponential bound. □

We now connect Theorem 6.4 with the large deviation principle. In [BMSS22, Proposition 2.3, Theorem 2.8], Boulanger, Mathieu, Sert, and Sisto presented a general theory of large deviation principles on metric spaces with Schottky sets. Combining their result with Theorem 1.1, we establish the large deviation principle for random walks on the mapping class group.

Corollary 6.5 (Large deviation principle). *Let (X, G, o) be as in Convention 2.11 and let $(Z_n)_{n \geq 0}$ be the random walk generated by a non-elementary probability measure μ on G. Then there exists a proper convex function $I : \mathbb{R} \to [0, +\infty]$, vanishing only at the drift $\lambda(\mu)$, such that*

$$-\inf_{x \in \mathrm{int}(E)} I(x) \leqslant \liminf_{n \to \infty} \frac{1}{n} \log \mathbb{P}\Big(\frac{1}{n} d(\mathrm{id}, Z_n) \in E\Big),$$
$$-\inf_{x \in \bar{E}} I(x) \geq \limsup_{n \to \infty} \frac{1}{n} \log \mathbb{P}\Big(\frac{1}{n} d(\mathrm{id}, Z_n) \in E\Big),$$

holds for every measurable set $E \subseteq \mathbb{R}$.

We note the work of Corso [Cor21], who proved that the rate function exists and is proper for random walks involving strongly contracting isometries. Our Corollary 6.5 strengthens Corso's result by showing that $I(x) \neq 0$ for $x \in [0, \lambda(\mu))$, which is a consequence of Theorem 1.1.

Part II. Random walks with weakly contracting isometries

In this part, we deal with groups acting on a space X and another space $\tilde{X}$ equivariantly, where the action on X involves strong contraction and the action on $\tilde{X}$ involves weak contraction; see Convention 7.2. After studying alignment of weakly contracting directions in § 8, we establish limit theorems for mapping class groups in § 9.

7. Mapping class groups and HHGs

Let Σ be a finite-type hyperbolic surface, let $(\tilde{X}, \tilde{d})$ be the Cayley graph of the mapping class group $G = \mathrm{Mod}(\Sigma)$ of Σ, and let (X, d) be the curve complex of Σ or the Teichmüller space of Σ. The action of G on (X, d) satisfies Convention 2.11: G contains independent pseudo-Anosov mapping classes that have strongly contracting orbits on X ([Min96, Contraction Theorem], [MM99, Proposition 4.6]).

Let $\mathrm{Pr} : \tilde{X} \to X$ be the orbit map: $\mathrm{Pr}(g) = go$, where $o \in X$ is the basepoint. Since G is finite generated and acts on X by isometries, the map Pr is coarsely Lipschitz and is G-equivariant.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

We will denote by $\tilde{A}$ the object 'in the upper space' corresponding to an object A 'in the lower space'. For example, we fix basepoints $\tilde{o} = \mathrm{id} \in \tilde{X}$ and $o \in X$ that satisfy $\mathrm{Pr}(\tilde{o}) = o$.

For each subset $\tilde{A} \subseteq \tilde{X}$, we define the projection $\tilde{\pi}_{\tilde{A}}$ from $\tilde{X}$ onto $\tilde{A}$ by referring to the closest point projection at the lower space X. Namely, for $\tilde{x} \in \tilde{X}$ and its projection $x := \mathrm{Pr}(\tilde{x})$, we define $\tilde{\pi}_{\tilde{A}}(\tilde{x}) := \mathrm{Pr}^{-1} \circ \pi_A \circ \mathrm{Pr}$ by $\tilde{a} \in \tilde{\pi}_{\tilde{A}}(\tilde{x}) \;\Leftrightarrow\; a \in \pi_A(x)$.

Lemma 7.1. *For each $C > 1$ there exists $D > 1$ such that if a C-quasigeodesic $\tilde{\gamma} : I \to \tilde{X}$ on $\tilde{X}$ has projection γ onto X that is a C-contracting axis, then $\tilde{\gamma}$ is D-weakly contracting with respect to the map $\tilde{\pi}_{\tilde{\gamma}} := Pr^{-1} \circ \pi_{\gamma} \circ Pr$.*

Proof. Let us first consider the case that $\tilde{X}$ is the Cayley graph of $\mathrm{Mod}(\Sigma)$ and (X, d) is the curve complex $\mathcal{C}(\Sigma)$ of Σ. Recall that there are coarsely Lipschitz projections $\mathrm{Pr}_U : \tilde{X} \to \{\text{uniformly bounded subsets of } \mathcal{C}U\}$ from $\tilde{X}$ to the curve complex $\mathcal{C}U$ of subsurfaces $U \subseteq \Sigma$, and $\rho_U^V : \mathcal{C}U \to \{\text{uniformly bounded subsets of } \mathcal{C}V\}$ for every pair of nested subsurfaces $V \subseteq U \subseteq \Sigma$. Further, Pr_U and $\rho_\Sigma^U \circ \mathrm{Pr}$ are uniformly coarsely equivalent.

Since $\tilde{\gamma}$ and $\gamma = \mathrm{Pr} \circ \tilde{\gamma}$ are C-quasigeodesics, $\{\mathrm{Pr}_U(\tilde{\gamma}) : U \subsetneq \Sigma\}$ have uniformly bounded diameter (depending on C). This is due to the bounded geodesic image property. Namely, given a proper subsurface $U \subsetneq \Sigma$, there exists a uniformly bounded neighborhood N of $\partial U \subseteq \mathcal{C}(\Sigma)$ such that $\gamma \setminus N$ is uniformly close to a geodesic on $\mathcal{C}(\Sigma)$ that is disjoint from ∂U. By [MM00, Theorem 3.1], $\rho_\Sigma^U(\gamma \setminus N)$ has bounded diameter. Since N is bounded, $\rho_\Sigma^U(\gamma \cap N)$ is also bounded.

Given the uniform boundedness of $\mathrm{Pr}_U(\tilde{\gamma})$, i.e., the coboundedness of $\tilde{\gamma}$, the weakly contracting property of $\tilde{\gamma}$ follows from [DR09, Theorem 4.2] (cf. [Beh06, Lemma 5.6]). More explicitly, [DR09, Theorem 4.2] guarantees a constant $E = E(C)$ such that, for each $\tilde{x} \in \tilde{X}$, we have

$$\mathrm{diam}_X \left(\pi_\gamma \circ \mathrm{Pr}\Big(\Big\{ \tilde{p} \in \tilde{X} : d(\tilde{x}, \tilde{p}) < \frac{1}{E} \tilde{d}(\tilde{x}, \tilde{\gamma}) \Big\} \Big) \right) < E.$$

Since the d-diameter along γ and $\tilde{d}$-diameter along $\tilde{\gamma}$ are coarsely equivalent (as γ is a quasigeodesic), we conclude that $\tilde{\gamma}$ is weakly contracting with respect to $\tilde{\pi}_{\tilde{\gamma}}$.

When (X, d) is the Teichmüller space of Σ, the strongly contracting property of γ implies that the Teichmüller geodesics $[\gamma(t), \gamma(s)]$ for $t < s$ are contained in a uniform neighborhood of γ (Corollary 3.4) and are hence uniformly thick. This in turn implies that $\mathrm{Pr}_U(\tilde{\gamma}(t), \tilde{\gamma}(s))$ for $t < s$ and proper subsurfaces $U \subsetneq \Sigma$ are uniformly bounded ([Raf05, Theorem 1.1], [Raf14, Theorem 5.5], [RS09, Theorem 4.1] and [DT15, Lemma 5.1]). Then we similarly deduce the weakly contracting property of $\tilde{\gamma}$ by [DR09, Theorem 4.2]. □

In general, Lemma 7.1 can be generalized to the setting where G is a HHG, $(\tilde{X}, \tilde{d})$ is its Cayley graph and (X, d) is the top curve graph for G. This follows from [ABD21, Corollary 6.2] ((3) ⇒ (2)) and [ABD21, Theorem 4.4]. Note that, even though Corollary 6.2 and Theorem 4.4 assume the unbounded products of the HHG structure for G, which is not granted in general, the directions we need do not require such an assumption.

In particular, the pseudo-Anosov axes on $\tilde{X}$ are weakly contracting. Hence, our setting is as follows.

Convention 7.2. *We fix $B > 0$ and assume that:*

(i) *$(\tilde{X}, \tilde{d})$, (X, d) are geodesic metric spaces;*
(ii) *$Pr : \tilde{X} \to X$ is a coarsely Lipschitz map, i.e., for all $\tilde{x}, \tilde{y} \in \tilde{X}$*

$$d(Pr(\tilde{x}), Pr(\tilde{y})) \leqslant B\tilde{d}(\tilde{x}, \tilde{y}) + B;$$

(iii) *G is a countable group of isometries acting on $\tilde{X}$ and X equivariantly;*
(iv) *$\tilde{o} \in \tilde{X}$ and $o \in X$ are basepoints that satisfy $Pr(\tilde{o}) = o$;*

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

(v) *for each $C > 1$ there exists $D > 1$ such that a path $\tilde{\gamma}$ on $\tilde{X}$ is D-weakly contracting with respect to $\tilde{\pi}_{\tilde{\gamma}}$ whenever its projection $Pr(\tilde{\gamma})$ is a C-contracting axis;*
(vi) *G contains two independent strongly contracting isometries of X.*

For each object $\tilde{A} \subseteq \tilde{X}$, we denote by A its projection $Pr(\tilde{A}) \subseteq X$.

When item (iii) is replaced with the coarse equivariance condition, this setting also covers HHGs acting on the top curve graph. For simplicity, we denote the word norm of $g \in G$ by $|g|$. In the general case, one can replace $|g|$ with $\tilde{d}(\tilde{o}, g\tilde{o})$.

8. Alignment II: weakly contracting axes

Throughout, we adopt Convention 7.2. We define the alignment among paths $\tilde{\kappa}_1, \ldots, \tilde{\kappa}_n$ on $\tilde{X}$ based on Definition 3.6 *with respect to the projections $\tilde{\pi}_{\tilde{\kappa}_i} := Pr^{-1} \circ \pi_{\kappa_i} \circ Pr$.*

Lemma 8.1. *For each $K > 1$ there exists $K' > K$ such that the following hold. Let $\tilde{x}, \tilde{y} \in \tilde{X}$ and let κ be a path on $\tilde{X}$ whose projection κ on X is a K-contracting axis. Then $\tilde{\kappa}$ is a K'-quasigeodesic that is K'-weakly contracting with respect to $\tilde{\pi}_{\tilde{\kappa}}$. Moreover, for each $C > 1$ we have the implication*

$$(\tilde{x}, \tilde{\kappa}, \tilde{y}) \text{ is } C\text{-aligned} \Rightarrow (x, \kappa, y) \text{ is } K'C\text{-aligned},$$
$$(x, \kappa, y) \text{ is } C\text{-aligned} \Rightarrow (\tilde{x}, \tilde{\kappa}, \tilde{y}) \text{ is } K'C\text{-aligned}.$$

Proof. Let $\tilde{\kappa} : I \to \tilde{X}$ and $\kappa := \mathrm{Pr} \circ \tilde{\kappa} : I \to X$. The weakly contracting property of $\tilde{\kappa}$ is given by Lemma 7.1. If we denote by F the coarse inverse of κ, Pr and $\tilde{\kappa} \circ F$ are maps between κ and $\tilde{\kappa}$, and are coarse inverses of each other. This implies the coarse comparison

$$\frac{1}{K''}\tilde{d}(\tilde{p}, \tilde{q}) - K'' \leqslant d(p, q) \leqslant K''\tilde{d}(\tilde{p}, \tilde{q}) + K'',$$

for all points $\tilde{p}, \tilde{q}$ on $\tilde{\kappa}$, for some $K'' = K''(K)$. This implies the remaining items. □

We now prove the main proposition of this section.

Proposition 8.2. *For each $K, D > 1$, there exist $E, L' > K, D$ such that the following holds.*

Let $L \geqslant L'$, let $\tilde{x}, \tilde{y} \in \tilde{X}$ and let $\tilde{\kappa}_1, \ldots, \tilde{\kappa}_n$ be paths on $\tilde{X}$ whose domains are longer than L and such that their projections are K-contracting axes. Suppose that $(x, \kappa_1, \ldots, \kappa_n, y)$ is D-aligned. Then there exist points $\tilde{p}_1, \ldots, \tilde{p}_n$ on $[\tilde{x}, \tilde{y}]$, in order from left to right, such that

$$\tilde{d}(\tilde{p}_i, \tilde{\kappa}_i) \leqslant \sum_{j=1}^{n+1} e^{-|j-i-0.5|L/E} \operatorname{diam}_{\tilde{X}}(\tilde{\kappa}_{j-1} \cup \tilde{\kappa}_j) + E. \tag{41}$$

Here, we plug in $\tilde{\kappa}_0 = \tilde{x}$ and $\tilde{\kappa}_{n+1} = \tilde{y}$.

Proof. Let B be the coarse Lipschitzness constant for Pr, and define the constants:

- let $K_2 = K'(K)$ be as in Lemma 8.1, which is larger than $K > 1$;
- let $K_4 = K'(K)$ be as in Lemma 2.10, which is larger than $K > 1$;
- let $E_1 = E(K, D)$ be as in Lemma 3.9, which is larger than $K > 1$;
- let $E_2 = E(K, D)$, $L_0 = L(K, D)$ be as in Proposition 3.10.

Now we define constants

$$E = 16KK_2K_4(1 + \log 2K_4) + E_1 + E_2 + B,$$
$$L' = L_0 + 4KK_2(B + 5K + K_2E).$$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Then, the following hold for all $L \geqslant L'$:

$$\frac{L}{K} - E - K \geqslant L/2K \geqslant E_1 + 2E_2 + B + 4K,$$
$$L \geqslant 4KK_2 \left(\frac{B+5K}{K_2} + K_2 E \right),$$
$$\frac{1}{2} e^{-2L/E} \geqslant K_4 e^{-L/4KK_2K_4}.$$

Note the following.

Fact 8.3. Let $C > 0$, let $x \in X$ and let κ be a K-contracting axis on X with L-long domain. If (x, κ) is C-aligned, then (κ, x) is not $((L/K) - C - K)$-aligned.

Let $L \geqslant L'$, let $\tilde{x}, \tilde{y} \in \tilde{X}$ and let $\tilde{\kappa}_1, \ldots, \tilde{\kappa}_n$ be paths on $\tilde{X}$ whose domains are longer than L and whose projections κ_i onto X are K-contracting axes. Recall our convention that, whenever we define $\tilde{A} \subseteq \tilde{X}$, we use the notation $A := \Pr(\tilde{A})$.

Step 1. We prove the following for $n \geqslant 2$: We induct on the number n of the contracting axes. First, Proposition 3.10 implies that (x, κ_i, y) is E_2-aligned for each i. In view of Fact 8.3, (κ_1, x) is not $(E_1 + E_2 + B + 4K)$-aligned but (κ_1, y) is E_2-aligned. Now, note that:

- π_{κ_1} is $(1, 4K)$-coarsely Lipschitz (Lemma 2.2);
- Pr is B-coarsely Lipschitz and hence $\pi_{\kappa_1} \circ \Pr$ is $(B, B + 4K)$-coarsely Lipschitz; and
- the geodesic $[\tilde{x}, \tilde{y}]$ is connected.

Pick the rightmost point $\tilde{z}_2 \in [\tilde{x}, \tilde{y}]$ such that (κ_1, z_2) is not $(E_1 + E_2)$-aligned. Then (κ_1, z_2) is $(E_1 + E_2 + B + 4K)$-aligned, and hence E-aligned, since $\pi_{\kappa_1} \circ \Pr$ is B-coarsely Lipschitz. Since (κ_1, κ_2) is D-aligned and (κ_1, z_2) is not E_1-aligned, Lemma 3.9 implies that (z_2, κ_1) is E_1-aligned.

When $n = 2$, the proof ends here. Otherwise, note that (z_2, κ_2) is E-aligned and $(\kappa_2, \ldots, \kappa_n, y)$ is D-aligned. By the induction hypothesis, there exist $\tilde{z}_3, \ldots, \tilde{z}_n$ on $[\tilde{z}_2, \tilde{y}]$, in order from left to right, such that (κ_{i-1}, z_i) and (z_i, κ_i) are E-aligned for each $i \geqslant 3$. The claim now follows.

Step 2: *Construction of* $\tilde{p}_j$. We now assume that $(x, \kappa_1, \ldots, \kappa_n, y)$ is D-aligned. By Step 1, we obtain points $\tilde{z}_2, \ldots, \tilde{z}_n$ on $[\tilde{x}, \tilde{y}]$, in order from left to right. We let $\tilde{z}_1 := \tilde{x}$ and $\tilde{z}_{n+1} := \tilde{y}$.

Pick $j \in \{1, \ldots, n\}$. Then (z_j, κ_j) is E-aligned, and hence $L/2K$-aligned. Meanwhile, (κ_j, z_{j+1}) is E-aligned, so (z_{j+1}, κ_j) is not $(L/K - E - K)$-aligned, and hence not $L/2K$-aligned. Now let $\tilde{p}_j$ to be the rightmost point on $[\tilde{z}_j, \tilde{z}_{j+1}]$ such that (p_j, κ_j) is not $L/2K$-aligned. Then by the $(B, B+4K)$-Lipschitzness of $\pi_{\kappa_j} \circ \Pr$, we have that

$$L/2K - (B + 4K) \leqslant \operatorname{diam}_X(\text{beginning point of } \kappa_j \cup \pi_{\kappa_j}(p_j)) \leqslant L/2K + (B + 4K).$$

In particular, (p_j, κ_j) is not $(L/2K - (B + 4K))$-aligned. Moreover, (κ_j, p_j) is not $(L/2K - (B + 5K))$-aligned by Fact 8.3. Denoting the beginning point of $\tilde{\kappa}_j$ by q_j, Lemma 8.1 implies

$$\begin{aligned} \operatorname{diam}_{\tilde{X}}(\tilde{\pi}_{\tilde{\kappa}_j}(\tilde{z}_j) \cup \tilde{\pi}_{\tilde{\kappa}_j}(\tilde{p}_j)) &\geqslant \operatorname{diam}_{\tilde{X}}(\tilde{\pi}_{\tilde{\kappa}_j}(\tilde{z}_j) \cup q_j) - \operatorname{diam}_{\tilde{X}}(q_j \cup \tilde{\pi}_{\tilde{\kappa}_j}(\tilde{p}_j)) \\ &\geqslant \frac{1}{K_2}(\frac{L}{2K} - B - 4K) - K_2 E \geqslant \frac{L}{4KK_2}. \end{aligned}$$

For a similar reason, $\operatorname{diam}_{\tilde{X}}(\tilde{\pi}_{\tilde{\kappa}_j}(\tilde{z}_{j+1}) \cup \tilde{\pi}_{\tilde{\kappa}_j}(\tilde{p}_j))$ is at least $(1/K_2)((L/2K) - B - 5K) - K_2 E \geqslant L/4KK_2$. Now Lemma 2.10 implies

$$\begin{aligned} \tilde{d}(\tilde{p}_j, \tilde{\kappa}_j) &\leqslant K_4 e^{-L/4KK_2K_4} \tilde{d}(\tilde{z}_j, \tilde{\kappa}_j) + K_4 e^{-L/4KK_2K_4} \tilde{d}(\tilde{z}_{j+1}, \tilde{\kappa}_j) + K_4 \\ &\leqslant \tfrac{1}{2} e^{-2L/E} \tilde{d}(\tilde{z}_j, \tilde{\kappa}_j) + \tfrac{1}{2} e^{-2L/E} \tilde{d}(\tilde{z}_{j+1}, \tilde{\kappa}_j) + K_4. \end{aligned} \tag{42}$$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Step 3: *Estimating* $\tilde{d}(\tilde{p}_j, \tilde{\kappa}_j)$. Given inequality (42), it now suffices to prove

$$\tilde{d}(\tilde{z}_i, \tilde{\kappa}_{i-1}) + \tilde{d}(\tilde{z}_i, \tilde{\kappa}_i) \leqslant \sum_{j=1}^{n+1} 2e^{-|j-i|L/E} \operatorname{diam}_{\tilde{X}}(\tilde{\kappa}_{j-1} \cup \tilde{\kappa}_j) + E \tag{43}$$

for $i = 2, \ldots, n$. To prove this, we collect indices i that violate inequality (43). Let $I = \{m, m+1, \ldots, m'\}$ be a maximal one-connected set of such indices. We aim to show that I is empty.

Suppose on the contrary that I is nonempty. Note first that $\tilde{z}_1$ and $\tilde{z}_{n+1}$ satisfy inequality (43), i.e., $1, n+1 \notin I$; hence, $m \geqslant 2$ and $m' \leqslant n$. We now compute $\tilde{d}(\tilde{z}_{m-1}, \tilde{z}_{m'+1})$ in two different ways. First, using inequality (42) we deduce

$$\begin{aligned}
\tilde{d}(\tilde{z}_{m-1}, \tilde{z}_{m'+1}) &= \sum_{j=m}^{m'+1} \tilde{d}(\tilde{z}_{j-1}, \tilde{z}_j) \geqslant \sum_{j=m}^{m'+1} (\tilde{d}(\tilde{z}_{j-1}, \tilde{p}_{j-1}) + \tilde{d}(\tilde{p}_{j-1}, \tilde{z}_j)) \\
&\geqslant \sum_{j=m}^{m'+1} (\tilde{d}(\tilde{z}_{j-1}, \tilde{\kappa}_{j-1}) + \tilde{d}(\tilde{\kappa}_{j-1}, \tilde{z}_j) - 2\tilde{d}(\tilde{p}_{j-1}, \tilde{\kappa}_{j-1})) \\
&\geqslant \sum_{j=m}^{m'+1} ((1 - e^{-2L/E})(\tilde{d}(\tilde{z}_{j-1}, \tilde{\kappa}_{j-1}) + \tilde{d}(\tilde{\kappa}_{j-1}, \tilde{z}_j)) - 2K_4).
\end{aligned}$$

Recall that $\tilde{d}(\tilde{\kappa}_{j-1}, \tilde{z}_j) + \tilde{d}(\tilde{z}_j, \tilde{\kappa}_j) \geqslant \sum_k 2e^{-|k-j|L/E} \operatorname{diam}_{\tilde{X}}(\tilde{\kappa}_{k-1} \cup \tilde{\kappa}_k) + E$ holds for $m \leqslant j \leqslant m'$. Moreover, $E \cdot (\#I) \geqslant 2K_4 \cdot (\#I + 1) + 0.5E$ because $\#I \geqslant 1$ and $E \geqslant 8K_4$. Hence, we obtain

$$\begin{aligned}
\tilde{d}(\tilde{z}_{m-1}, \tilde{z}_{m'+1}) \geqslant\ & (1 - e^{-2L/E})(\tilde{d}(\tilde{z}_{m-1}, \tilde{\kappa}_{m-1}) + \tilde{d}(\tilde{\kappa}_{m'}, \tilde{z}_{m'+1})) \\
&+ (1 - e^{-2L/E}) \sum_{j=m}^{m'} \sum_{k=1}^{N+1} 2e^{-|k-j|L/E} \operatorname{diam}_{\tilde{X}}(\tilde{\kappa}_{k-1} \cup \tilde{\kappa}_k) + (1 - e^{-2L/E}) \cdot 0.5E.
\end{aligned}$$

If we rearrange the double summation with respect to k, the right-hand side is at least

$$\begin{aligned}
&(1 - e^{-2L/E})(\tilde{d}(\tilde{z}_{m-1}, \tilde{\kappa}_{m-1}) + \tilde{d}(\tilde{z}_{m'+1}, \tilde{\kappa}_{m'}) + 0.5E) + 2(1 - e^{-2L/E}) \sum_{j=m}^{m'} \operatorname{diam}(\tilde{\kappa}_{j-1} \cup \tilde{\kappa}_j) \\
&\quad + 2(1 - e^{-2L/E}) \Bigg(\sum_{1 \leqslant k < m} e^{-(m-k)L/E} \operatorname{diam}_{\tilde{X}}(\tilde{\kappa}_{k-1} \cup \tilde{\kappa}_k) \\
&\quad + \sum_{m' < k \leqslant N+1} e^{-(k-m')L/E} \operatorname{diam}_{\tilde{X}}(\tilde{\kappa}_{k-1} \cup \tilde{\kappa}_k) \Bigg).
\end{aligned}$$

Next, we will obtain an upper bound of $\tilde{d}(\tilde{z}_{m-1}, \tilde{z}_{m'+1})$

$$\begin{aligned}
\tilde{d}(\tilde{z}_{m-1}, \tilde{z}_{m'+1}) &\leqslant \tilde{d}(\tilde{z}_{m-1}, \tilde{\kappa}_{m-1}) + \sum_{j=m}^{m'} \operatorname{diam}(\tilde{\kappa}_{j-1} \cup \tilde{\kappa}_j) + \tilde{d}(\tilde{z}_{m'+1}, \tilde{\kappa}_{m'}) \\
&\leqslant (1 - e^{-2L/E})(\tilde{d}(\tilde{z}_{m-1}, \tilde{\kappa}_{m-1}) + \tilde{d}(\tilde{z}_{m'+1}, \tilde{\kappa}_{m'})) + \sum_{j=m}^{m'} \operatorname{diam}(\tilde{\kappa}_{j-1} \cup \tilde{\kappa}_j) \\
&\quad + 2e^{-2L/E} \cdot E + e^{-2L/E}(\tilde{d}(\tilde{\kappa}_{m-2}, \tilde{z}_{m-1}) + \tilde{d}(\tilde{z}_{m-1}, \tilde{\kappa}_{m-1}) - E) \\
&\quad + e^{-2L/E}(\tilde{d}(\tilde{\kappa}_{m'}, \tilde{z}_{m'+1}) + \tilde{d}(\tilde{z}_{m'+1}, \tilde{\kappa}_{m'+1}) - E).
\end{aligned}$$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

We know that $6e^{-L/E} \leqslant 1$ because $L > 8E$. Having this in mind, we now make use of the fact that $m-1$ and $m'+1$ are not contained in I: $\tilde{d}(\tilde{z}_{m-1}, \tilde{z}_{m'+1})$ is bounded from above by

$$\begin{aligned}
&(1-e^{-2L/E})(\tilde{d}(\tilde{z}_{m-1}, \tilde{\kappa}_{m-1}) + \tilde{d}(\tilde{z}_{m'+1}, \tilde{\kappa}_{m'}) + 0.5E) + \sum_{j=m}^{m'} \operatorname{diam}_{\tilde{X}}(\tilde{\kappa}_{j-1} \cup \tilde{\kappa}_j) \\
&\quad + e^{-2L/E}\left(2\sum_{k=1}^{N+1}(e^{-|k-m+1|L/E} + e^{-|k-m'-1|L/E}) \operatorname{diam}_{\tilde{X}}(\tilde{\kappa}_{k-1} \cup \tilde{\kappa}_k)\right) \\
&\leqslant (1-e^{-2L/E})(\tilde{d}(\tilde{z}_{m-1}, \tilde{\kappa}_{m-1}) + \tilde{d}(\tilde{z}_{m'+1}, \tilde{\kappa}_{m'}) + 0.5E) \\
&\quad + \sum_{j=m}^{m'}(1+4e^{-2L/E}) \operatorname{diam}_{\tilde{X}}(\tilde{\kappa}_{j-1} \cup \tilde{\kappa}_j) + \sum_{1 \leqslant k < m} 4e^{-L/E} \cdot e^{-(m-k)L/E} \operatorname{diam}_{\tilde{X}}(\tilde{\kappa}_{k-1} \cup \tilde{\kappa}_k) \\
&\quad + \sum_{m' < k \leqslant N+1} 4e^{-L/E} \cdot e^{-(k-m')L/E} \operatorname{diam}_{\tilde{X}}(\tilde{\kappa}_{k-1} \cup \tilde{\kappa}_k),
\end{aligned}$$

which is a contradiction. Hence, $I = \emptyset$ and inequality (43) is established. □

9. Limit laws for mapping class groups

We continue to employ the notion of Schottky sets defined in Definition 3.15. Once a Schottky set S and its element s is understood, the translates of $\Gamma^{\pm}(s)$ are now called *Schottky axes on* X, whereas the translates of $\tilde{\Gamma}^{\pm}(s)$ are called *Schottky axes on* $\tilde{X}$.

Definition 9.1. Given a constant $K_0 > 0$, we define:

- $K_1 = K'(K_0)$ as in Lemma 8.1;
- $D_0 = D(K_0, K_0)$ as in Lemma 3.8;
- $E_0 = E(K_0, D_0)$, $L_0 = L(K_0, D_0)$ as in Proposition 3.12;
- $E_1 = K'(K_0, E_0)$, $L_1 = L'(K_0, E_0)$ as in Proposition 8.2.

Let $0 < \epsilon < 1$. If a K_0-Schottky set $S \subseteq G^{M_0}$ consists of sequences of length

$$M_0 > \max(L_0, L_1, 2K_1E_1, (-\log(\epsilon^2/4)) \cdot E_1),$$

then we call S an *ϵ-constricting K_0-Schottky set.*

Thanks to Proposition 3.19, for every non-elementary probability measure μ on G and $N, \epsilon > 0$, there exists an ϵ-long enough Schottky set for μ with cardinality N. We are ready to state the following proposition.

Proposition 9.2. *Let μ be a non-elementary probability measure on the mapping class group G and $((\check{Z}_n)_n, (Z_n)_n)$ be the (bi-directional) random walk generated by μ, with step sequences $((\check{g}_n)_n, (g_n)_n)$. Then there exists $K' > 0$ such that*

$$\mathbb{P}(d(\mathrm{id}, [\check{Z}_m, Z_n]) \leqslant K'D_k \quad \text{for all } n, m \geqslant 0 \,|\, \check{g}_{k+1}, g_{k+1}) \leqslant K'e^{-k/K'},$$

holds for all k, where

$$D_k := \sum_{i=1}^{k} |g_i| + \sum_{i=1}^{k} |\check{g}_i| + \sum_{i=1}^{\infty} e^{-i/K'}|g_i| + \sum_{i=1}^{\infty} e^{-i/K'}|\check{g}_i| + 1. \tag{44}$$

Proof. Let S and $\check{S}$ be long enough, large, and $(2/e)$-constricting K_0-Schottky sets for μ and $\check{\mu}$, respectively, for some $K_0 > 0$. Proposition 4.2 determines a constant $K > 0$ (not depending on

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

k but only on μ), a probability space $(\Omega, \mathbb{P})$ for μ and a partition of Ω into pivotal equivalence classes that is independent of the backward steps $(\check{g}_n)_{n>0}$ and such that

$$\mathbb{P}(\#\mathcal{P}(\omega) \cap \{1, \ldots, n\} \geqslant n/K \mid g_{k+1}) \geqslant 1 - Ke^{-k/K} \quad (n \geqslant k),$$

and also another partition into (backward) pivotal equivalence classes that is independent of the forward steps $(g_n)_{n>0}$ and such that

$$\mathbb{P}(\#\mathcal{P}(\check{\omega}) \cap \{1, \ldots, n\} \geqslant n/K \mid \check{g}_{k+1}) \geqslant 1 - Ke^{-k/K} \quad (n \geqslant k).$$

We enumerate $\mathcal{P}(\omega)$ by $\{j(1) < j(2) < \cdots\}$ and $\mathcal{P}(\check{\omega})$ by $\{\check{j}(1) < \check{j}(2) < \cdots\}$. Let us now define the event B_k in Ω; $(\check{\omega}, \omega) \in B_k$ if:

(i) $\#\mathcal{P}(\omega) \cap \{1, \ldots, n\} \geqslant n/K$ for all $n \geqslant k/3$;
(ii) $\#\mathcal{P}(\check{\omega}) \cap \{1, \ldots, n\} \geqslant n/K$ for all $n \geqslant k/3$;
(iii) for each $n \geqslant k$ and $m \geqslant k$, the following are D_0-semi-aligned:

$$(o, \mathbf{Y}_{j(1)}(\omega), \mathbf{Y}_{j(2)}(\omega), \ldots, \mathbf{Y}_{j(\lfloor 2n/3K \rfloor)}(\omega), Z_n o),$$
$$(o, \mathbf{Y}_{\check{j}(1)}(\check{\omega}), \mathbf{Y}_{\check{j}(2)}(\check{\omega}), \ldots, \mathbf{Y}_{\check{j}(\lfloor 2m/3K \rfloor)}(\check{\omega}), \check{Z}_m o);$$

(iv) $(\bar{\mathbf{Y}}_{\check{j}(i)}(\check{\omega}), \mathbf{Y}_{j(i)}(\omega))$ is D_0-aligned for some $i \leqslant k/3K$.

In the proof of Lemma 4.10 we proved that $\mathbb{P}(B_k)$ decays exponentially in k. It remains to prove that $d(\mathrm{id}, [\check{Z}_m, Z_n]) \leqslant K' D_k$ for any $n, m > 0$ and $(\check{\omega}, \omega) \in B_k$, where we set $K' \geqslant 8 + 1.5K + E_1$. From now on, we fix k. When $n \leqslant k$, we automatically have

$$d(\mathrm{id}, [\check{Z}_m, Z_n]) \leqslant d(\mathrm{id}, Z_n) \leqslant \sum_{i=0}^{k} |g_i| \leqslant K' D_k.$$

Similarly, the desired inequality holds when $m \leqslant k$. Now assume $n, m \geqslant k$. The sequence

$$(\check{Z}_m o, \overline{\mathbf{Y}}_{\check{j}(\lfloor 2m/3K \rfloor)}(\check{\omega}), \ldots, \overline{\mathbf{Y}}_{\check{j}(\lfloor k/3K \rfloor)}(\check{\omega}), \mathbf{Y}_{j(\lfloor k/3K \rfloor)}(\omega), \ldots, \mathbf{Y}_{j(\lfloor 2n/3K \rfloor)}(\omega), Z_n o)$$

is D_0-semi-aligned, and hence E_0-aligned by Proposition 3.10. Here, recall that the involved Schottky set is $(2/e)$-long enough and that $-\log(4/e^2 \cdot 1/4) = 2$. Hence, $M_0/E_1 \geqslant 2$. By Proposition 8.2, there exists $\tilde{p} \in [\check{Z}_m, Z_n]$ whose distance to $\tilde{\mathbf{Y}}_{j(\lfloor k/3K \rfloor)}(\omega)$ is at most

$$\begin{aligned}
&\sum_{l=\lfloor k/3K \rfloor + 1}^{\lfloor 2n/3K \rfloor} e^{-l-\lfloor k/3K \rfloor} \operatorname{diam}(\tilde{\mathbf{Y}}_{j(l-1)}(\omega) \cup \tilde{\mathbf{Y}}_{j(l)}(\omega)) \\
&+ \sum_{l=\lfloor k/3K \rfloor + 1}^{\lfloor 2m/3K \rfloor} e^{-l-\lfloor k/3K \rfloor} \operatorname{diam}(\tilde{\mathbf{Y}}_{\check{j}(l-1)}(\check{\omega}) \cup \tilde{\mathbf{Y}}_{\check{j}(l)}(\check{\omega})) \\
&+ e^{-(\lfloor 2n/3K \rfloor - \lfloor k/3K \rfloor)} \operatorname{diam}(\tilde{\mathbf{Y}}_{j(\lfloor 2n/3K \rfloor)}(\omega) \cup Z_n) \\
&+ e^{-(\lfloor 2m/3K \rfloor - \lfloor k/3K \rfloor)} \operatorname{diam}(\tilde{\mathbf{Y}}_{j(\lfloor 2m/3K \rfloor)}(\check{\omega}) \cup \check{Z}_m) \\
&+ e^{-1} \operatorname{diam}(\tilde{\mathbf{Y}}_{j(\lfloor k/3K \rfloor)}(\omega) \cup \tilde{\mathbf{Y}}_{\check{j}(\lfloor k/3K \rfloor)}(\check{\omega})) + E_1.
\end{aligned}$$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Here, note that

$$\operatorname{diam}(\tilde{\mathbf{Y}}_{j(k-1)}(\omega) \cup \tilde{\mathbf{Y}}_{j(k)})(\omega)) \leqslant \sum_{i=j(k-1)-M_0+1}^{j(k)} |g_i|,$$

$$\operatorname{diam}(\tilde{\mathbf{Y}}_{j(\lfloor 2n/3K \rfloor)}(\omega) \cup Z_n) \leqslant \sum_{i=j(\lfloor 2n/3K \rfloor)-M_0+1}^{n} |g_i|.$$

Note also that $l - \lfloor k/3K \rfloor \geqslant \frac{1}{2}l$ for $l > \lfloor 2k/3K \rfloor$. Using these, we deduce

$$\begin{aligned}
\tilde{d}(\tilde{p}, \tilde{\mathbf{Y}}_{j(\lfloor k/3K \rfloor)}(\omega)) \leqslant & \sum_{i=j(\lfloor 2k/3K \rfloor)+1}^{j(\lfloor 2n/3K \rfloor)} 2e^{-\frac{1}{2}\min\{l>0: j(l) \geqslant i\}} |g_i| + \sum_{i=\check{j}(\lfloor 2k/3K \rfloor)+1}^{\check{j}(\lfloor 2m/3K \rfloor)} 2e^{-\frac{1}{2}\min\{l>0: \check{j}(l) \geqslant i\}} |\check{g}_i| \\
& + \sum_{i=j(\lfloor 2n/3K \rfloor)+1}^{n} 2e^{-\frac{1}{2}\lfloor 2n/3K \rfloor} |g_i| + \sum_{i=j(\lfloor 2m/3K \rfloor)+1}^{m} 2e^{-\frac{1}{2}\lfloor 2m/3K \rfloor} |\check{g}_i| \\
& + \sum_{i=1}^{j(\lfloor 2k/3K \rfloor)} 2|g_i| + \sum_{i=1}^{\check{j}(\lfloor 2k/3K \rfloor)} 2|\check{g}_i| + E_1.
\end{aligned}$$

Since we have $j(\lceil i/K \rceil) \leqslant i$ for each $i \geqslant k/3$, this is bounded by

$$\begin{aligned}
& 2\sum_{i=1}^{k} |g_i| + 2\sum_{i=1}^{k} |\check{g}_i| + 2\sum_{i=k+1}^{\lfloor 2n/3 \rfloor} e^{-i/2K} |g_i| + 2e^{-\lfloor n/3K \rfloor} \sum_{i=\lfloor 2n/3 \rfloor+1}^{n} |g_i| \\
& + 2\sum_{i=k+1}^{\lfloor 2m/3 \rfloor} e^{-i/2K} |\check{g}_i| + 2e^{-\lfloor m/3K \rfloor} \sum_{i=\lfloor m/3 \rfloor+1}^{m} |\check{g}_i| + E_1.
\end{aligned}$$

Moreover, since $\operatorname{diam}(\mathrm{id} \cup \tilde{\mathbf{Y}}_{j(\lfloor k/3K \rfloor)})$ is bounded by $\sum_{i=1}^{j(\lfloor k/3K \rfloor)} |g_i| \leqslant \sum_{i=1}^{k} |g_i|$, we conclude

$$\tilde{d}(\mathrm{id}, \tilde{p}) \leqslant 4\sum_{i=1}^{k} (|g_i| + |\check{g}|_i) + 2\sum_{i=1}^{\infty} e^{-\frac{1}{3K}i} (|g_i| + |\check{g}|_i) + E_1. \qquad \square$$

PROPOSITION 9.3. *Let $p > 0$ and let $((\check{Z}_n)_n, (Z_n)_n)$ be the (bi-directional) random walk generated by a non-elementary probability measure μ on G with finite pth moment. Then there exists $K > 0$ such that*

$$\mathbb{E}_\mu \Big[\sup_{n,m} \check{d}(\mathrm{id}, [\check{Z}_m, Z_n])^p \Big] < K.$$

In particular, for almost every sample path $(\check{\omega}, \omega)$, every geodesic in $\{[\check{Z}_m, Z_n] : m, n > 0\}$ intersects the same bounded metric ball centered at id.

Proof. Let $K' > 0$ be as in Proposition 9.2. Let D_k be as defined by Equation (44) and let

$$A_k := \{(\check{\omega}, \omega) : d(\mathrm{id}, [\check{Z}_n, Z_m]) \leqslant D_k \text{ for all } n, m \geqslant 0\}.$$

Then, we have

$$\frac{1}{K'} \sup_{n,m} \check{d}(\mathrm{id}, [\check{Z}_m, Z_n]) \leqslant \sum_{k=1}^{\min\{m : (\check{\omega}, \omega) \in A_m\}} (|g_k| + |\check{g}_k|) + \sum_{k=1}^{\infty} e^{-k/K'} |g_k| + \sum_{k=1}^{\infty} e^{-k/K'} |\check{g}_k| + 1. \quad (45)$$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Noting that $|x+y|^p \leqslant (2\max(|x|,|y|))^p \leqslant |2x|^p + |2y|^p$ for $x,y>0$, it suffices to bound $\mathbb{E}[I_i^p]$ for

$$I_1 := \sum_{k=1}^{\min\{m:(\check\omega,\omega)\in A_m\}} |g_k|, \quad I_2 := \sum_{k=1}^{\min\{m:(\check\omega,\omega)\in A_m\}} |\check g_k|,$$
$$I_3 := \sum_{k=1}^{\infty} e^{-k/K'}|g_k|, \qquad I_4 := \sum_{k=1}^{\infty} e^{-k/K'}|\check g_k|.$$

Observe the following: when $|g_k|e^{-k/2K'}$ is bounded by M for all k, we have

$$I_3 = \sum_{k=1}^{\infty} e^{-k/K'}|g_k| \leqslant M\sum_{k=1}^{\infty} e^{-k/2K'} \leqslant MC,$$

for $C=(1-e^{-1/2K'})^{-1}$. This means

$$\begin{aligned}\mathbb{E}[I_3^p] &\leqslant C^p\mathbb{E}\Big[(\max_k e^{-k/2K'}|g_k|)^p\Big] \leqslant C^p\mathbb{E}\Big[\sum_k (e^{-k/2K'}|g_k|)^p\Big]\\ &= C^p\mathbb{E}_\mu|g|^p\cdot\sum_k e^{-kp/2K'} < +\infty.\end{aligned}$$

For I_1, recall the inequality $|t^p - s^p| \leqslant 2^p(|t-s|^p + s^{p-n_p}|t-s|^{n_p})$ for each $t,s\geqslant 0$ and $p>0$, where $n_p = p$ for $0\leqslant p\leqslant 1$ and $n_p=1$ otherwise. From this, we have

$$\mathbb{E}[I_1^p] \leqslant 2^p\sum_{k=0}^{\infty}\mathbb{E}\left[\left(|g_{k+1}|^p + \left(\sum_{i=1}^k |g_i|\right)^{p-n_p}|g_{k+1}|^{n_p}\right)1_{A_k^c}\right].$$

Since $\mathbb{P}(A_k^c \mid g_{k+1}) \leqslant K'e^{-k/K'}$ by Proposition 9.2, $\mathbb{E}(|g_{k+1}|^p 1_{A_k^c}) \le (\mathbb{E}_\mu|g|^p)\cdot K'e^{-k/K'}$ is summable. Moreover,

$$\begin{aligned}\mathbb{E}\left[\left(\sum_{i=1}^k |g_i|\right)^{p-n_p}|g_{k+1}|^{n_p}1_{A_k^c}\right] &\leqslant \mathbb{E}\left[\left(\left(\sum_{i=1}^k |g_i|\right)^p c^{-n_p} + c^{p-n_p}\right)|g_{k+1}|^{n_p}1_{A_k^c}\right]\\ &\leqslant c^{-n_p}\cdot k^{p+1}(\mathbb{E}_\mu|g|^p)^2 + K'c^{p-n_p}e^{-k/K'}\mathbb{E}_\mu|g|^p\end{aligned}$$

holds for $c = e^{-k/2pK'}$, which is summable for k. Hence, $\mathbb{E}[I_1^p]$ is finite. The remaining terms $\mathbb{E}[I_2^p]$ and $\mathbb{E}[I_4^p]$ can be handled in a similar way. □

We obtain an analogous estimate for random walks with finite exponential moment. Because it follows from the proof of Corollary 4.16 given inequality (45), we omit the proof.

PROPOSITION 9.4. *Let $((\check Z_n)_{n>0},(Z_n)_{n>0})$ be the (bi-directional) random walk generated by a non-elementary probability measure μ on G with finite exponential moment. Then there exists $K>0$ such that*

$$\mathbb{E}_\mu\left[\exp\left(\frac{1}{K}\sup_{n,m}\check d(\mathrm{id},[\check Z_m, Z_n])\right)\right] < K.$$

Using Proposition 9.3, we obtain the uniform second-moment deviation inequality for non-elementary probability measures on the mapping class group. Now employing Theorems 4.1 and 4.2 of [MS20] and the proof of Theorem 4.20, we establish Theorem 1.2.

To prove Theorem 1.3, for each $k\geqslant 0$ and $(\check\omega,\omega)\in\check\Omega\times\Omega$ we define the infinite geodesic $\Gamma_k(\check\omega,\omega)$ to be a subsequential limit of $\{[\check Z_{n+k}, Z_{n-k}] : n = 1,2,\ldots\}$ (which exists by the Arzelá–Ascoli theorem and the second result in Proposition 9.3). Note that $\tilde d(Z_k, \Gamma_0(\check\omega,\omega))$ are identically

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

distributed with $\tilde{d}(\mathrm{id}, \Gamma_k(\check{\omega}, \omega))$, which are all dominated by $\tilde{d}(\mathrm{id}, \sup_{n,m}[\check{Z}_n, Z_m])$. It follows that

$$\mathbb{P}(\tilde{d}(Z_k, \Gamma_0(\check{\omega}, \omega)) \geqslant g(k)),$$

is summable for some $o(k^{1/p})$-function $g(k)$ ($K \log k$ for some $K > 0$, respectively) when the underlying measure has finite pth moment (finite exponential moment, respectively). By the Borel–Cantelli lemma, we deduce

$$\lim_n \frac{1}{n^{1/p}} \tilde{d}(Z_n, \Gamma_0) = 0 \quad \left(\limsup_n \frac{1}{\log n} \tilde{d}(Z_n, \Gamma_0) < K, \text{ respectively}\right).$$

We conclude this paper by establishing Theorem 1.1. Recall that Proposition 4.2 guaranteed the alignment of Schottky axes along $[o, Z_n o]$ *on* X, which led to the reverse triangle inequality for distances on X (Lemma 3.18). We now record the corresponding result for distances on $\tilde{X}$.

Lemma 9.5. *Let $0 < \epsilon < 1$ and let S be a long enough, ϵ-constricting K_0-Schottky set. Let $\tilde{x}, \tilde{y} \in \tilde{X}$ and let $\tilde{\kappa}_1, \ldots, \tilde{\kappa}_N$ be Schottky axes on $\tilde{X}$. If $(x, \kappa_1, \ldots, \kappa_N, y)$ is D_0-semi-aligned, then we have*

$$\tilde{d}(\tilde{x}, \tilde{y}) \geqslant (1-\epsilon)\Big(\mathrm{diam}_{\tilde{X}}(\tilde{x}, \tilde{\kappa}_1) + \sum_{i=2}^n \mathrm{diam}_{\tilde{X}}(\tilde{\kappa}_{i-1}, \tilde{\kappa}_i) + \mathrm{diam}_{\tilde{X}}(\tilde{\kappa}_n, \tilde{y})\Big) - 4\sum_{i=1}^n \mathrm{diam}_{\tilde{X}}(\tilde{\kappa}_i).$$

Proof. Let M_0 be such that $S \subseteq G^{M_0}$. Note that $4e^{-M_0/2E_0} \leqslant \epsilon < 1$, which implies

$$\sum_{j \in \mathbb{Z}} e^{-|j-0.5|M_0/E_0} \leqslant \frac{\epsilon}{4} \cdot \frac{1}{1-(\epsilon/4)^2} \leqslant \frac{\epsilon}{2}.$$

Now let $\tilde{\kappa}$ be an arbitrary Schottky axis on $\tilde{X}$. Because $M_0 > 2K_1E_1$ for E_1 as in Definition 9.1 and $\tilde{\kappa}$ is a K_1-quasigeodesics by Lemma 8.1, we have $\mathrm{diam}_{\tilde{X}}(\tilde{\kappa}) > E_1$.

Since $(x, \kappa_1, \ldots, \kappa_N, y)$ are D_0-semi-aligned, they are E_0-aligned by Proposition 3.12. Consequently, $(\tilde{x}, \tilde{\kappa}_1, \ldots, \tilde{\kappa}_N, \tilde{y})$ is K_1E_0-aligned by Lemma 8.1. Since the domains of κ_i are longer than $M_0 \geq L_1$, we can obtain the points $\tilde{p}_i$ on $[\tilde{x}, \tilde{y}]$ as described in Proposition 8.2.

We now have

$$\begin{aligned}
\tilde{d}(\tilde{x}, \tilde{y}) &= \tilde{d}(\tilde{x}, \tilde{p}_1) + \sum_{i=2}^n \tilde{d}(\tilde{p}_{i-1}, \tilde{p}_i) + \tilde{d}(\tilde{p}_n, \tilde{y}) \\
&\geqslant (\mathrm{diam}_{\tilde{X}}(\tilde{x}, \tilde{\kappa}_1) - \mathrm{diam}_{\tilde{X}}(\tilde{\kappa}_1) - \tilde{d}(\tilde{\kappa}_1, \tilde{p}_1)) + (\mathrm{diam}_{\tilde{X}}(\tilde{y}, \tilde{\kappa}_n) - \mathrm{diam}_{\tilde{X}}(\tilde{\kappa}_n) - \tilde{d}(\tilde{\kappa}_n, \tilde{p}_n)) \\
&\quad + \sum_{i=2}^n (\mathrm{diam}_{\tilde{X}}(\tilde{\kappa}_{i-1} \cup \tilde{\kappa}_i) - \mathrm{diam}_{\tilde{X}}(\tilde{\kappa}_{i-1}) - \mathrm{diam}_{\tilde{X}}(\tilde{\kappa}_i) - \tilde{d}(\tilde{\kappa}_{i-1}, \tilde{p}_{i-1}) - \tilde{d}(\tilde{\kappa}_i, \tilde{p}_i)).
\end{aligned}$$

Here, Proposition 8.2 tells us that

$$\begin{aligned}
\sum_{i=1}^n \tilde{d}(\tilde{\kappa}_i, \tilde{p}_i) &\leqslant \Big(\mathrm{diam}_{\tilde{X}}(\tilde{x}, \tilde{\kappa}_1) + \sum_{i=2}^n \mathrm{diam}_{\tilde{X}}(\tilde{\kappa}_{i-1} \cup \tilde{\kappa}_i) + \mathrm{diam}_{\tilde{X}}(\tilde{\kappa}_n, \tilde{y})\Big) \cdot \sum_{j \in \mathbb{Z}} e^{-|j-0.5|M_0/E_1} + E_1 n \\
&\leqslant \frac{\epsilon}{2}\Big(\mathrm{diam}_{\tilde{X}}(\tilde{x}, \tilde{\kappa}_1) + \sum_{i=2}^n \mathrm{diam}_{\tilde{X}}(\tilde{\kappa}_{i-1} \cup \tilde{\kappa}_i) + \mathrm{diam}_{\tilde{X}}(\tilde{\kappa}_n, \tilde{y})\Big) + \sum_{i=1}^n \mathrm{diam}_{\tilde{X}}(\tilde{\kappa}_i).
\end{aligned}$$

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Using this, we conclude

$$\tilde{d}(\tilde{x},\tilde{y}) \geqslant (1-\epsilon)\Big(\operatorname{diam}_{\tilde{X}}(\tilde{x},\tilde{\kappa}_1) + \sum_{i=2}^{n} \operatorname{diam}_{\tilde{X}}(\tilde{\kappa}_{i-1}\cup\tilde{\kappa}_i) + \operatorname{diam}_{\tilde{X}}(\tilde{\kappa}_n,\tilde{y})\Big) - 4\sum_{i=1}^{n}\operatorname{diam}_{\tilde{X}}(\tilde{\kappa}_i).$$

□

COROLLARY 9.6 [Gou22, Lemma 4.14]. *Let ν be a non-elementary probability measure on G and let $(Z_n)_n$ be the random walk generated by ν. Then, for each $\epsilon > 0$, there exists $C > 0$ such that*

$$\mathbb{P}(|gZ_n| \geqslant (1-\epsilon)|g| - C \text{ for all } n \geqslant 0) \geqslant 1-\epsilon/2 \quad (\forall g \in G).$$

Proof. Let S be a large, long enough and ϵ-constricting K_0-Schottky set for μ in G^{M_0}, for some suitable $K_0, M_0 > 0$. (This determines the constants $K_1, D_0, \ldots$ as in Definition 9.1.)

As in the proof of Corollary 4.11, there exists $N > 0$ independent of g such that

$$\mathbb{P}\begin{pmatrix}\text{there exists } i < N \text{ such that } \mathbf{Y}_i \text{ is a Schottky axis and} \\ (g^{-1}o, \mathbf{Y}_i(\omega), Z_n o) \text{ is } D_0\text{-semi-aligned for each } n \geqslant N\end{pmatrix} \geqslant 1-\epsilon/4.$$

Also, when $(g^{-1}o, \mathbf{Y}_i, Z_n o)$ is D_0-semi-aligned, Lemma 9.5 implies that

$$\begin{aligned}\tilde{d}(g^{-1}, Z_n) &\geqslant (1-\epsilon)(\operatorname{diam}_{\tilde{X}}(g^{-1}\cup\tilde{\mathbf{Y}}_i) + \operatorname{diam}_{\tilde{X}}(\tilde{\mathbf{Y}}_i \cup Z_n)) - 4\operatorname{diam}_{\tilde{X}}(\tilde{\mathbf{Y}}_i)\\ &\geqslant (1-\epsilon)\tilde{d}(g^{-1}, \tilde{Z}_{i-M_0}) - 4\cdot(K_1M_0 + K_1)\\ &\geqslant (1-\epsilon)|g| - (1-\epsilon)\sum_{j=1}^{i-M_0}|g_j| - C'' \geqslant (1-\epsilon)|g| - \sum_{j=1}^{N}|g_j| - C'',\end{aligned}$$

where $C'' = 4(K_1M_0 + K_1)$. This bound also holds for $n \leqslant N$

$$\tilde{d}(g^{-1}, Z_n) \geqslant |g^{-1}| - |Z_n| \geqslant |g| - \sum_{j=1}^{N}|g_j|.$$

Given these, the proof ends by taking large enough $C' > 0$ such that

$$\mathbb{P}\Bigg(\sum_{j=1}^{N}|g_j| \geqslant C' - C''\Bigg) \leqslant \epsilon/4.$$

□

Proof of Theorem 1.1. As in the proof of Theorem 6.4, we first take $\epsilon > 0$ such that

$$(1-\epsilon)^4\lambda(\mu) > L + \epsilon.$$

Let S be a large, long enough and ϵ-constricting Schottky set for μ with cardinality greater than $10/\epsilon$, and let

$$C'' = 4\max_{s\in S}\operatorname{diam}_{\tilde{X}}\tilde{\Gamma}^+(s).$$

By Proposition 6.2, there exists a non-elementary probability measure ν, and for each sufficiently large N, a partition $\mathscr{P}_{n,N,\epsilon}$ into (n, N, ϵ, ν)-pivotal equivalence classes for each n such that

$$\mathbb{P}\bigg(\omega : \frac{1}{2}\#\mathcal{P}^{(n,N,\epsilon)}(\omega) \leqslant (1-\epsilon)\frac{n}{2M_0N}\bigg)$$

decays exponentially in n. Let $C > 0$ be a constant for ν provided by Corollary 9.6: we have

$$\mathbb{P}_{\nu^{*m}}(h : |gh| \geqslant (1-\epsilon)|g| - C) \geqslant 1-\epsilon/2$$

for each $g \in G$ and each $m > 0$. We now fix an N such that $N > C + C''/M_0\lambda(\mu)\epsilon$.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Let $\mathcal{E}$ be an equivalence class such that $\frac{1}{2}\#\mathcal{P}^{(n,N,\epsilon)}(\mathcal{E}) \geqslant (1-\epsilon)(n/2M_0N)$. For each $\omega \in \mathcal{E}$ $(o, \mathbf{Y}_{j(1)}, \ldots, \mathbf{Y}_{j'(\#\mathcal{P}/2)}, Z_n o)$ is D_0-semi-aligned. Lemma 9.5 tells us that

$$|Z_n| \geqslant \sum_{i=1}^{\#\mathcal{P}/2} ((1-\epsilon)\tilde{d}(Z_{j(i)}, Z_{j'(i)-M_0}) - C'') = \sum_{i=1}^{\#\mathcal{P}/2} (|r_i| - C'') \quad (r_i := g_{j(i)+1} \cdots g_{j'(i)-M_0}).$$

Since r_i are i.i.d. with

$$\mathbb{E}[|r_i| - M] \geqslant (1-\epsilon)\mathbb{E}_{\mu^{*2M_0N}}[(1-\epsilon)|g| - C] - C'' \geqslant (1-\epsilon)^3 \cdot 2M_0 N \lambda(\mu),$$

we can find $K' > 0$ not depending on n such that

$$\mathbb{P}(|Z_n| \leq (1-\epsilon)^4 \lambda n \,|\, \mathcal{E}) \leqslant K' e^{-n/K'} \quad (\forall n > 0).$$

Summing up this conditional probability, we obtain the desired exponential bound. $\square$

Acknowledgements

The author thanks H. Baik, K. Chawla, I. Gekhtman, V. He, S.-h. Kim, J. Maher, H. Masai, Y. Qing, K. Rafi, S. Taylor, G. Tiozzo and W. Yang for helpful discussions. The author is indebted to the anonymous referees' helpful and careful comments. The author is also grateful to the American Institute of Mathematics and the organizers and the participants of the workshop 'Random walks beyond hyperbolic groups' in April 2022 for helpful and inspiring discussions.

Conflicts of interest

None.

Financial support

The author was supported by Samsung Science & Technology Foundation (SSTF-BA1702-01 and SSTF-BA1301-51) and by a KIAS Individual Grant (SG091901). Part of the revision was conducted while the author was supported by Mid-Career Researcher Program (RS-2023-00278510) through the NRF funded by the government of Korea. This work constitutes part of the author's PhD thesis.

Journal information

Compositio Mathematica is owned by the Foundation Compositio Mathematica and published by the London Mathematical Society in partnership with Cambridge University Press. All surplus income from the publication of *Compositio Mathematica* is returned to mathematics and higher education through the charitable activities of the Foundation, the London Mathematical Society and Cambridge University Press.

References

ABD21 C. Abbott, J. Behrstock and M. G. Durham, *Largest acylindrical actions and stability in hierarchically hyperbolic groups*, Trans. Amer. Math. Soc. Ser. B **8** (2021), 66–104, with an appendix by Daniel Berlyne and Jacob Russell.

ACGH17 G. N. Arzhantseva, C. H. Cashen, D. Gruber and D. Hume, *Characterizations of Morse quasi-geodesics via superlinear divergence and sublinear contraction*, Doc. Math. **22** (2017), 1193–1224.

ACGH19 G. N. Arzhantseva, C. H. Cashen, D. Gruber and D. Hume, *Negative curvature in graphical small cancellation groups*, Groups Geom. Dyn. **13** (2019), 579–632.

ACT15 G. N. Arzhantseva, C. H. Cashen and J. Tao, *Growth tight actions*, Pacific J. Math. **278** (2015), 1–49.

BCK21 H. Baik, I. Choi and D. Kim, *Linear growth of translation lengths of random isometries on Gromov hyperbolic spaces and Teichmüller spaces*, Preprint (2021), arXiv:2103.13616.

Beh06 J. A. Behrstock, *Asymptotic geometry of the mapping class group and Teichmüller space*, Geom. Topol. **10** (2006), 1523–1578.

BQ16 Y. Benoist and J.-F. Quint, *Central limit theorem on hyperbolic groups*, Izv. Ross. Akad. Nauk Ser. Mat. **80** (2016), 5–26.

BF02 M. Bestvina and K. Fujiwara, *Bounded cohomology of subgroups of mapping class groups*, Geom. Topol. **6** (2002), 69–89.

BMSS22 A. Bounlanger, P. Mathieu, Ç. Sert and A. Sisto, *Large deviations for random walks on hyperbolic spaces*, Ann. Sci. Éc. Norm. Supér. (4) (2022), 885–944.

CW21 M. Calvez and B. Wiest, *Morse elements in Garside groups are strongly contracting*, Preprint (2021), arXiv:2106.14826.

CCT25 K. Chawla, I. Choi and G. Tiozzo, *Genericity of contracting geodesics in groups.* Groups Geom. Dyn., to appear. Preprint (2025), https://ems.press/journals/ggd/articles/14298985.

Cho22a I. Choi, *Limit laws on Outer space, Teichmüller space, and CAT(0) spaces*, Preprint (2022a), arXiv:2207.06597v1.

Cho22b I. Choi, *Random walks and contracting elements III: Outer space and outer automorphism group*, Preprint (2022b), arXiv:2212.12122.

Cho23 I. Choi, *Central limit theorem and geodesic tracking on hyperbolic spaces and teichmüller spaces*, Adv. Math. **431** (2023), 109236.

Cho24 I. Choi, *Pseudo-Anosovs are exponentially generic in mapping class groups*, Geom. Topol. **28** (2024), 1923–1955.

Cor21 E. Corso, *Large deviations for irreducible random walks on relatively hyperbolic groups*, Preprint (2021), arXiv:2110.14592.

Cou22 R. Coulon, *Patterson-Sullivan theory for groups with a strongly contracting element*, Preprint (2022), arXiv:2206.07361.

DH18 F. Dahmani and C. Horbez, *Spectral theorems for random walks on mapping class groups and $Out(F_N)$*, Int. Math. Res. Not. IMRN **2008** (2018), 2693–2744.

DR09 M. Duchin and K. Rafi, *Divergence of geodesics in Teichmüller space and the mapping class group*, Geom. Funct. Anal. **19** (2009), 722–742.

DT15 M. G. Durham and S. J. Taylor, *Convex cocompactness and stability in mapping class groups*, Algebr. Geom. Topol. **15** (2015), 2839–2859.

Fer18 T. Fernós, *The Furstenberg-Poisson boundary and CAT(0) cube complexes*, Ergodic Theory Dynam. Syst. **38** (2018), 2180–2223.

FLM24 T. Fernós, J. Lécureux and F. Mathéus, *Contact graphs, boundaries, and a central limit theorem for CAT(0) cubical complexes*, Groups Geom. Dyn. **18** (2024), 677–704.

GGPY21 I. Gekhtman, V. Gerasimov, L. Potyagailo and W. Yang, *Martin boundary covers Floyd boundary*, Invent. Math. **223** (2021), 759–809.

GP13 V. Gerasimov and L. Potyagailo, *Quasi-isometric maps and Floyd boundaries of relatively hyperbolic groups*, J. Eur. Math. Soc. (JEMS) **15** (2013), 2115–2137.

GS21 A. Goldsborough and A. Sisto, *Markov chains on hyperbolic-like groups and quasi-isometries*, Preprint (2021), arXiv:2111.09837.

Gou22 S. Gouëzel, *Exponential bounds for random walks on hyperbolic spaces without moment conditions*, Tunis. J. Math. **4** (2022), 635–671.
Hor18 C. Horbez, *Central limit theorems for mapping class groups and Out*(F_N), Geom. Topol. **22** (2018), 105–156.
KM96 V. A. Kaimanovich and H. Masur, *The Poisson boundary of the mapping class group*, Invent. Math. **125** (1996), 221–264.
Kar03 A. Karlsson, *Free subgroups of groups with nontrivial Floyd boundary*, Comm. Algebra **31** (2003), 5361–5376.
KL06 A. Karlsson and F. Ledrappier, *On laws of large numbers for random walks*, Ann. Probab. **34** (2006), 1693–1706.
KM99 A. Karlsson and G. A. Margulis, *A multiplicative ergodic theorem and nonpositively curved spaces*, Comm. Math. Phys. **208** (1999), 107–123.
Kes59 H. Kesten, *Full Banach mean values on countable groups*, Math. Scand. **7** (1959), 146–156.
LB22a C. Le Bars, *Central limit theorem on cat(0) spaces with contracting isometries*, Preprint (2022a), arXiv:2209.11648.
LB22b C. Le Bars, *Random walks and rank-one isometries on CAT(0) spaces*, Preprint (2022b), arXiv:2205.07594.
Leg22 X. Legaspi, *Constricting elements and the growth of quasi-convex subgroups*, Preprint (2022), arXiv:2206.06749.
Mah10a J. Maher, *Linear progress in the complex of curves*, Trans. Amer. Math. Soc. **362** (2010a), 2963–2991.
Mah10b J. Maher, *Random Heegaard splittings*, J. Topol. **3** (2010b), 997–1025.
Mah11 J. Maher, *Random walks on the mapping class group*, Duke Math. J. **156** (2011), 429–468.
Mah12 J. Maher, *Exponential decay in the mapping class group*, J. Lond. Math. Soc. (2) **86** (2012), 366–386.
MT18 J. Maher and G. Tiozzo, *Random walks on weakly hyperbolic groups*, J. Reine Angew. Math. **742** (2018), 187–239.
MM99 H. A. Masur and Y. N. Minsky, *Geometry of the complex of curves. I. Hyperbolicity*, Invent. Math. **138** (1999), 103–149.
MM00 H. A. Masur and Y. N. Minsky, *Geometry of the complex of curves. II. Hierarchical structure*, Geom. Funct. Anal. **10** (2000), 902–974.
MS20 P. Mathieu and A. Sisto, *Deviation inequalities for random walks*, Duke Math. J. **169** (2020), 961–1036.
Min96 Y. N. Minsky, *Quasi-projections in Teichmüller space*, J. Reine Angew. Math. **473** (1996), 121–136.
PSZ24 H. Petyt, D. Spriano and A. Zalloum, *Hyperbolic models for CAT*(0) *spaces*, Adv. Math. **450** (2024), 109742.
QRT20 Y. Qing, K. Rafi and G. Tiozzo, *Sublinearly Morse boundary II: Proper geodesic spaces*, Preprint (2020), arXiv:2011.03481.
Raf05 K. Rafi, *A characterization of short curves of a Teichmüller geodesic*, Geom. Topol. **9** (2005), 179–202.
Raf14 K. Rafi, *Hyperbolicity in Teichmüller space*, Geom. Topol. **18** (2014), 3025–3053.
RS09 K. Rafi and S. Schleimer, Covers and the curve complex, Geom. Topol. **13** (2009), 2141–2162.
San06 P. Sankaran, *On homeomorphisms and quasi-isometries of the real line*, Proc. Amer. Math. Soc. **134** (2006), 1875–1880.

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press

Sis17 A. Sisto, *Tracking rates of random walks*, Israel J. Math. **220** (2017), 1–28.

Sis18 A. Sisto, *Contracting elements and random walks*, J. Reine Angew. Math. **742** (2018), 79–114.

Sun20 M. Sunderland, *Linear progress with exponential decay in weakly hyperbolic groups*, Groups Geom. Dyn. **14** (2020), 539–566.

Yan14 W.-y. Yang, *Growth tightness for groups with contracting elements*, Math. Proc. Cambridge Philos. Soc. **157** (2014), 297–319.

Yan19 W.-y. Yang, *Statistically convex-cocompact actions of groups with contracting elements*, Int. Math. Res. Not. IMRN **2019** (2019), 7259–7323.

Yan20 W.-y. Yang, *Genericity of contracting elements in groups*, Math. Ann. **376** (2020), 823–861.

Inhyeok Choi inhyeokchoi48@gmail.com
June E Huh Center for Mathematical Challenges, KIAS, Seoul 02455, Republic of Korea
Cornell University, Ithaca, NY 14850, USA

https://doi.org/10.1112/S0010437X25007511 Published online by Cambridge University Press