Ergod. Th. & Dynam. Sys. (2015), **35**, 64–110 © *D. Calegari & J. Maher 2014.*
doi:10.1017/etds.2013.43 *Published by Cambridge University Press*

Statistics and compression of scl

DANNY CALEGARI† and JOSEPH MAHER‡

† *University of Chicago, Chicago, IL 60637, USA*
(e-mail: dannyc@math.uchicago.edu)
‡ *Department of Mathematics, CUNY College of Staten Island, Staten Island, NY 10314, USA and Department of Mathematics, CUNY Graduate Center, NY 10016, USA*
(e-mail: joseph.maher@csi.cuny.edu)

(*Received* 26 *June* 2012 *and accepted in revised form* 26 *April* 2013)

Abstract. We obtain sharp estimates on the growth rate of stable commutator length on random (geodesic) words, and on random walks, in hyperbolic groups and groups acting non-degenerately on hyperbolic spaces. In either case, we show that with high probability stable commutator length of an element of length n is of order $n/\log n$. This establishes quantitative refinements of qualitative results of Bestvina and Fujiwara and others on the infinite dimensionality of two-dimensional bounded cohomology in groups acting suitably on hyperbolic spaces, in the sense that we can control the geometry of the unit balls in these normed vector spaces (or rather, in random subspaces of their normed duals). As a corollary of our methods, we show that an element obtained by random walk of length n in a mapping class group cannot be written as a product of fewer than $O(n/\log n)$ reducible elements, with probability going to 1 as n goes to infinity. We also show that the translation length on the complex of free factors of a random walk of length n on the outer automorphism group of a free group grows linearly in n.

Contents

1 Introduction 65
1.1 Concentration and hyperbolicity 65
1.2 Bounded cohomology 66
1.3 Stable commutator length as a random variable 67
1.4 The geometry of Q as an obstruction 67
1.5 Random dynamical systems in the continuum limit 68
1.6 Summary of results 68
2 Stable commutator length 72
2.1 Stable commutator length 72
2.2 Quasimorphisms 73
3 Hyperbolic groups 73
3.1 Hyperbolic groups 73

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

3.2 Combings 74
3.3 Markov chains 74
3.4 Ergodicity at infinity 76
3.5 Mixing times and a Chernoff-type estimate 77
3.6 Anti-alignment 78
3.7 Comparing components 79
4 The scl of random geodesics in hyperbolic groups 81
4.1 Almost pairing subwords 81
4.2 Counting quasimorphisms 81
4.3 Homology 83
5 The scl of random walks in groups 85
5.1 Special case: free groups 85
5.2 Universal upper bounds 86
5.3 Template for obtaining lower bounds 87
5.4 Non-degenerate random walks 88
5.5 Proximal points and unfolded walks 90
5.6 Anti-aligned segments and matching 91
5.7 Uniform lower bounds for hyperbolic spaces 93
5.8 Geometry of the scl norm of a random subspace 95
5.9 Concentration versus compression 96
5.10 Random walks on Out(F_n) 97
6 Universal lower bounds 102
6.1 Groups acting on circles 103
6.2 Random turtles in the hyperbolic plane 105
6.3 Central limit theorem for arbitrary quasimorphisms 107
Acknowledgements 109
References 109

1. *Introduction*

The goal of this paper is to use probabilistic and ergodic theoretic methods to obtain bounds and estimates for the growth rate of characteristic norms in certain classes of groups, especially *stable commutator length* in *hyperbolic groups* (and some groups acting on hyperbolic spaces). These growth rates are a *quantitative* obstruction to the existence of non-trivial homomorphisms between groups, which refine known qualitative obstructions due to Bestvina and Fujiwara [**5**] and others. They also illustrate a *probabilistic* connection between the growth rate of geometric and algebraic quantities, which are a manifestation of the phenomenon of Mostow rigidity in a broad context.

1.1. *Concentration and hyperbolicity.*

Let ξ_n be a sequence of random real variables. For many natural sequences ξ_n one can prove a *law of large numbers*, i.e. the almost sure existence of a limit $L := \lim_{n\to\infty} \xi_n/n$ (this is colloquially known as *convergence to the mean*). Such laws of large numbers hold under remarkably broad hypotheses.

It is natural to look for examples in geometric group theory, and it is not hard to find them. If G is a group and $\xi : G \to \mathbb{R}^+$ any subadditive function, one can take ξ_n to be

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

the value of ξ on some Markov process on G. For any group G, random walk is the most obvious and important example of a natural Markov process, but for specific classes of groups, other natural Markov processes exist, arising (for example) from an automatic structure on G. The law of large numbers then follows from an application of Kingman's subadditive ergodic theorem.

For example, when G is hyperbolic, Cannon [**14**] showed how to construct a stationary finite Markov chain which gives a geodesic *combing* of G. And when G is the fundamental group of a closed, negatively curved manifold, Ratner [**42**] showed how to build a Markov coding for the geodesic flow, whose closed orbits are naturally in bijection with conjugacy classes in G. Roughly speaking, the *geometric* hyperbolicity implies *dynamical* hyperbolicity for the geodesic flow, and this dynamical hyperbolicity gives rise to laws of large numbers.

It is rarer to find examples where ξ_n has *sublinear* growth, but nevertheless concentrates at some deterministic scale. Colloquially, we say that a sequence of non-negative random variables ξ_n is *concentrated* if there is some deterministic function $f(n)$ so that $\xi_n/f(n)$ converges in probability to a Dirac mass at 1, and we say that ξ_n is *compressed* if (again for some $f(n)$) every weak limit of $\xi_n/f(n)$ is a probability measure on $\mathbb{R}^+$ with support bounded away from 0 and ∞.

1.2. *Bounded cohomology.* Bounded cohomology, as introduced by Gromov [**26**], is (among other things) a functor from the category of groups and homomorphisms to the category of normed vector spaces and norm-decreasing linear maps. One of the main virtues of this functor is its *monotonicity*: if the invariants associated to a group G are 'smaller' than the invariants associated to a group H, there are no interesting homomorphisms from G to H. As a well-known example, Bestvina and Fujiwara [**5**] used two-dimensional bounded cohomology to show that every homomorphism from a higher rank lattice to a mapping class group factors through a finite group (this fact was known earlier by work of Farb and Masur [**20**], building on work of Kaimanovich and Masur [**34**]).

To date, such tools have generally been used somewhat crudely, because of the enormous difficulty in computing bounded cohomology, or deriving useful invariants from it. Most authors have concentrated on bounded cohomology in dimension two, and have focused almost exclusively on a trichotomous distinction: namely, for a given group G, whether $H_b^2(G)$ (i.e. two-dimensional bounded cohomology with real coefficients) is trivial, non-trivial but finite dimensional, or infinite dimensional.

In a way, this misses the main point, which is that $H_b^2(G)$ is canonically a *Banach space*. Almost nothing is known about the (large scale) geometry of this Banach space in any non-trivial cases. In this paper, we are able to derive strong geometric information about the geometry of these Banach spaces in the important case of hyperbolic groups (and some groups acting suitably on hyperbolic spaces). This is done via the relationship between two-dimensional bounded cohomology, quasimorphisms, and stable commutator length.

For any group G there is an exact sequence of real vector spaces

$$0 \to H^1(G) \to Q(G) \to H_b^2(G) \to H^2(G),$$

where Q denotes the space of *homogeneous quasimorphisms* on G (see §2 for a precise definition, and e.g. [**11**, Theorem 2.50] for a proof). For a finitely presented group, H^1

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

and H^2 are finite dimensional, so H^2_b and Q carry (almost) the same information. Moreover, H^2_b and Q/H^1 are Banach spaces in a functorial way, and the map $Q/H^1 \to H^2_b$ is two-bilipschitz. We focus on Q/H^1 in this paper, although it would be straightforward to reinterpret our results in terms of H^2_b.

Bavard [**2**] interpreted Q/H^1 in terms of an *a priori* algebraic invariant called *stable commutator length* (hereafter scl). If G is a group and $[G, G]$ denotes its commutator subgroup, the *commutator length* of any $g \in [G, G]$ (denoted $\mathrm{cl}(g)$) is the least number of commutators in G whose product is equal to g, and the *stable commutator length* of g is the limit

$$\mathrm{scl}(g) := \lim_{n\to\infty} \mathrm{cl}(g^n)/n.$$

In fact, scl extends in a natural way to a (pseudo-)norm on the vector space $B_1(G)$ of real group one-boundaries, and descends to the quotient $B_1^H(G) := B_1(G)/\langle g^n - ng,\ g - hgh^{-1}\rangle$. In many important cases, scl is a *norm* on B_1^H. In every case it turns out that Q/H^1 is the *dual* of B_1^H with its scl (pseudo-)norm, where Q/H^1 carries the so-called *defect norm* $2D(\cdot)$ (see §2). This statement is usually known as generalized Bavard duality, and in this generality is proved in [**11**].

1.3. *Stable commutator length as a random variable.* In this paper we study the growth rate of scl and the geometry of B_1^H on *random elements* and in *random subspaces* of hyperbolic (and other) groups, with respect to two natural families of probability measures. First, we obtain results for *random geodesics* of length n in hyperbolic groups. Secondly, we obtain results for *random walks* of length n in hyperbolic groups, and in groups acting suitably on hyperbolic spaces. These two senses of random are conceptually related, but in general become mutually singular as $n \to \infty$.

There are technical subtleties in either case which are somewhat complementary. The theory of stable commutator length is well-adapted to the geometry of quasigeodesics in hyperbolic groups, since quasigeodesity can be certified—and stable commutator length estimated—from *local* contributions. A combing determines a geodesic representative of each element in a group, and the Markov process associated to an automatic structure allows one to pick a random element of prescribed word length. But there is no guarantee that the Markov process in question is *ergodic*. Complementarily, the Markov process defining a random walk is always ergodic, but random paths in hyperbolic spaces are typically not quasigeodesic. Nevertheless, in either context we are able to obtain compression results for scl, at the deterministic growth rate of $n/\log n$.

1.4. *The geometry of Q as an obstruction.* If $G \to H$ is a surjective homomorphism, the pullback $Q(H) \to Q(G)$ is injective. Hence, if $Q(G)$ vanishes but $Q(H)$ does not, no such surjective homomorphism exists. More generally, no such surjective homomorphism exists if $Q(G)$ is finite dimensional but $Q(H)$ is infinite dimensional.

In their seminal paper [**5**], Bestvina and Fujiwara showed that if H is a non-elementary subgroup of a mapping class group, then $Q(H)$ is infinite dimensional. On the other hand, $Q(\Gamma)$ is finite dimensional whenever Γ is a lattice in a higher rank Lie group; consequently, every homomorphism from such a lattice to a mapping class group factors through an elementary subgroup, and therefore (by Margulis) through a finite group.

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

Independently, Hamenstädt [**30**] and Bestvina and Feighn [**3**] proved a similar theorem for subgroups of $\mathrm{Out}(F_n)$ (Hamenstädt's construction considers bounded cohomology with twisted coefficients).

This application of bounded cohomology makes use of Q as an $\mathbb{R}$-vector space, but completely neglects its (Banach) geometry. Any homomorphism $G \to H$ is non-increasing for scl. A random walk on G pushes forward under a homomorphism to a random walk on H; the growth rate of scl under random walks is therefore an obstruction to the existence of such a homomorphism. To apply these ideas in practice one must be able to understand the growth rate of scl under random walks in the target group with respect to a measure whose support can be arbitrary; when H is the target group one must be prepared to consider the case that the support of the measure generates a subgroup which is not quasiconvex. Carrying this out for hyperbolic groups, and groups acting in a suitable way on hyperbolic spaces, is one of the main goals of this paper.

1.5. *Random dynamical systems in the continuum limit.* In two dimensions, Pesin theory says that all the entropy of a diffeomorphism is carried by Smale horseshoes, which is to say, by periodic orbits with hyperbolic dynamics. Up to 'zero entropy' therefore, one aims to recover the essential dynamics of a diffeomorphism from an inverse system of combinatorial data, namely the braid type of the diffeomorphism modulo a finite invariant subset.

Ghys [**24**] (building on work of Arnold, Sullivan, Tresser, Fathi, Gambaudo and others) has proposed to understand the group of area-preserving diffeomorphisms of a surface as a kind of 'limit' of braid groups of finer and finer meshes of discrete points in the surface, and has explained how to use quasimorphisms on such braid groups to construct similar quasimorphisms on the transformation groups. This gives highly non-trivial information about the algebraic structure of such groups, but it is difficult in practice to interpret this information dynamically.

One can approach this question from the other direction, by taking families of random walks in the braid group with respect to finer and finer meshes of discrete particles, and trying to obtain a law on random measure-preserving transformations in the limit. In order to do this, one needs to obtain a uniform modulus of continuity on the motion of the discrete particles on every fixed scale, to ensure that the limiting transformations are *continuous*. One way to obtain such estimates would be to show that a random element of a braid or mapping class group has a short factorization as a product of elements with support in subsurfaces with simple topology—i.e. to give an upper bound on its *reducible length*. In fact, our methods give *lower bounds* on reducible length of the same order of magnitude as (stable) commutator length. It therefore becomes a provocative question whether there are complementary upper bounds of the same order of magnitude.

1.6. *Summary of results.* We now summarize the remainder of the paper. Section 2 contains a brief review of the theory of stable commutator length and quasimorphisms that we use in the following. This material is standard.

In §3 we discuss the ergodic theory of geodesic combings on hyperbolic groups, and establish the main technical results necessary to obtain estimates on scl of random

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

geodesics. Roughly speaking, a combing for a hyperbolic group determines a directed graph with edges labeled by generators, so that directed paths of length n in the graph correspond bijectively to elements of the group of word length n (with respect to some fixed generating set). The directed graph can be thought of as a topological Markov chain; assigning transition probabilities to edges determines a stationary Markov process (in the usual sense). There is an assignment of transition probabilities of maximal entropy corresponding to the Patterson–Sullivan measure on the group; this allows us to define what is meant by a 'random element of G of length n' in such a way that we can produce such elements by a Markov process.

The first main technical result in this section is Proposition 3.13. In words, this proposition gives a Chernoff-type estimate for the number of times any given path σ of length $\ell \log n / \log \lambda$ occurs in a random path of length n in an ergodic Markov process, where $\ell < 1$ is fixed, and λ is the entropy of the Markov process. The key subtlety is that the length of the subpaths we are counting is not fixed, but depends on the length of the big random path. The second main technical result in this section is the content of Lemmas 3.18 and 3.19, which say that the subwords of fixed length appearing in different ergodic components of the Patterson–Sullivan Markov process are 'coarsely equivalent' (as isometry classes of geodesic segments in the group) in distribution.

In §4 we apply the estimates obtained in §3 to prove concentration for scl on random geodesics in hyperbolic groups at the scale $n/\log n$.

HYPERBOLIC GEODESIC THEOREM 4.10. *Let G be a hyperbolic group, and S a finite generating set. There are constants C_1, $C_2 > 0$, $C_3 > 0$, $C_4 > 1$, $C_5 > 0$ so that if g is a random element with $|g|_S \in [n - C_1, n + C_1]$ conditioned to lie in $[G, G]$, then*

$$\mathbf{P}(C_2 n/\log n \leq \mathrm{scl}(g) \leq C_3 n/\log n) = 1 - O(C_4^{-n^{C_5}}).$$

In fact, we obtain the stronger result $\mathrm{cl}(g) \leq C_3 n/\log n$, with the same estimate in probability.

The proofs of the upper and lower bounds are somewhat different. The proof of the upper bound is related directly to the definition of (stable) commutator length, and depends on writing a random g efficiently as a product of commutators by directly pairing almost inverse subwords of g of length $O(\log n)$, using the Chernoff-type estimate in Proposition 3.13. The proof of the lower bound is obtained via Bavard duality, by constructing a homogeneous quasimorphism with controlled defect whose value on g is large. The homogeneous quasimorphism is a variant on the *small counting quasimorphisms* considered by Epstein and Fujiwara [**19**], for which, in contrast with the 'big' counting quasimorphisms considered by Brooks [**9**], controlling the defect is easy. On the other hand, while obtaining the correct order of magnitude estimate of the defect is straightforward, obtaining exact estimates is substantially harder; this explains the gap between the bounds $C_2 n/\log n$ and $C_3 n/\log n$ in the hyperbolic geodesic theorem, which we believe to be an artifact of the method of proof, rather than a genuine statistical phenomenon. In fact, the authors of [**13**] conjectured that one should be able to take C_2 and C_3 arbitrarily close to $\log \lambda/6$, where λ is the growth entropy of G with respect to S

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

(i.e. the number such that the ball of radius n in G in the word metric with respect to S has $\Theta(\lambda^n)$ elements).

In §5, we obtain *universal* estimates on scl on random walks in finitely generated groups, and obtain precise order of magnitude estimates for hyperbolic groups and certain groups acting on hyperbolic spaces. For simplicity we consider random walks obtained by repeatedly convolving a symmetric measure of finite support, although our methods could presumably be extended to a more general context. Because of the symmetry of the Cayley graph of a free group with respect to a free generating set, our sharp estimates for scl on random geodesics in free groups gives sharp estimates for scl on random walks in free groups; together with the monotonicity of scl under homomorphisms, this gives the following universal upper bound.

UNIVERSAL UPPER BOUND THEOREM 5.5. *Let G be a group with a finite symmetric generating set S, and let $|S| = 2k$. Let g be obtained by random walk on G (with respect to S) of length n (even), conditioned to lie in $[G, G]$. Then, for any $\epsilon > 0$ there are constants $C_1 > 1$, $C_2 > 0$ so that with probability $1 - O(C_1^{-n^{C_2}})$ there is an inequality*

$$\mathrm{scl}(g) \leq ((k-1)\log(2k-1) + \epsilon)/6k \cdot n/\log n.$$

The remainder of the section is devoted to finding lower bounds of the same order of magnitude for groups acting in a suitable way on hyperbolic spaces. We consider a group G acting simplicially on a δ-hyperbolic simplicial complex Y, and let μ be a symmetric probability measure with support equal to a finite subset S of G (note that we do not assume that S generates G). We say that the triple (G, μ, Y) is *non-degenerate* if it satisfies a short list of axioms, spelled out in Definition 5.7. These axioms are satisfied in the following important cases:

(1) G is a hyperbolic group, Y is the Cayley graph of G, and $\langle S\rangle$ is a non-elementary subgroup;
(2) G is hyperbolic relative to a family of subgroups, Y is the Groves–Manning hyperbolic complex certifying relative hyperbolicity (see [**28**, §3]), and $\langle S\rangle$ is non-elementary;
(3) G is the mapping class group of a surface, Y is the complex of curves, and $\langle S\rangle$ is not reducible or virtually abelian.

A discussion of more general examples (including the action of $\mathrm{Out}(F_n)$ on certain hyperbolic complexes) is deferred to §5.10.

In these terms, the main theorem proved in this section is the hyperbolic lower bound theorem.

HYPERBOLIC LOWER BOUND THEOREM 5.27. *Let Y be a δ-hyperbolic simplicial complex (not assumed to be locally finite), and let G be a finitely generated group that acts simplicially on Y. Let μ be a symmetric probability measure of finite support on G so that (G, μ, Y) is non-degenerate, in the sense of Definition 5.7.*

Let g be obtained by random walk on G (with respect to μ) of length n. Then for any C_1 there is a $C_2 > 0$ and C_3 depending only on δ and G, so that with probability at least $1 - n^{-C_1}$ there is a homogeneous quasimorphism ϕ on G satisfying the following properties:

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

(1) $\phi(g) \geq nC_2/\log n$;
(2) $D(\phi) \leq C_3$;
(3) $|\phi(h)| \leq 2d_Y(q, hq)C_3/\log n$ *for any* $g \in G$ *and any* $q \in Y$.

In particular, for any $C_1 > 0$ *there is a constant* $C > 1$ *so that if we condition on* $g \in [G, G]$ *(for* n *even), then*

$$\mathbf{P}(C^{-1}n/\log n \leq \mathrm{scl}(g) \leq Cn/\log n) \geq 1 - n^{-C_1}.$$

Two significant corollaries follow immediately.

REDUCIBLE BOUND COROLLARY 5.29. *Let* μ *be a symmetric probability measure on the mapping class group of* Σ *of finite support, and suppose the subgroup it generates is not reducible or virtually abelian. Then for any* C_1 *there is* C *so that if* g *is obtained by random walk in* G *of length* n*, then* g *cannot be expressed as a product of fewer than* $Cn/\log n$ *reducible elements, with probability at least* $1 - n^{-C_1}$.

GROWTH OBSTRUCTION COROLLARY 5.30. *Let* G *be a group and* μ *a symmetric probability measure with finite support which generates* G. *Suppose that for any* $\epsilon > 0$ *there is a* $\delta > 0$ *so that if* g *is obtained by random walk on* G *of length* n *(even) conditioned to lie in* $[G, G]$*, we have*

$$\mathbf{P}(\mathrm{scl}(g) < \epsilon \cdot n/\log n) > \delta.$$

Then every homomorphism from G *to a hyperbolic group or to a mapping class group has virtually abelian image.*

The growth obstruction corollary is the promised quantitative strengthening of [**5**] alluded to in §1.4.

In §5.8 we spell out the analogs of these theorems for parametric families of elements, in terms of the geometry of the unit ball in the scl norm of a random subspace of $B_1^H(G)$ for G as above. In words: we show that the unit ball in a random subspace has uniformly bounded geometry (up to rescaling by a deterministic factor), and conjecture that it is C^0 close to a cross-polytope (i.e. to a unit ball in an L^1 norm). This conjecture is only known to be true for free groups, by [**13**].

In §5.10 we consider random walks on non-locally compact Gromov hyperbolic spaces in some generality, and prove (under mild conditions) that all the axioms from Definition 5.7 are satisfied except possibly the condition of acylindricity. A key intermediate step is to show that random walks converge to the (Gromov) boundary almost surely. Although this does not have immediate implications for scl, it enables us to obtain linear lower bounds on translation length of elements obtained by random walk. By 'translation length' $\tau(g)$ of an isometry g on a metric space Y we mean here *asymptotic translation length*, i.e. $\lim_{n\to\infty} d_Y(y, g^n(y))/n$ for arbitrary $y \in Y$. Precisely, we show the following theorem.

LINEAR TRANSLATION LENGTH THEOREM 5.35. *Let* G *be a group of isometries of a Gromov hyperbolic space* Y*, which is not necessarily locally compact. Let* μ *be a probability distribution with finite support on* G*, such that the support of* μ *generates a non-elementary subgroup of* G. *Then there are constants* $L > 0$ *and* $c < 1$ *such that*

$$\mathbf{P}(\tau(w_n) \leq Ln) \leq O(c^n),$$

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

where w_n is the group element obtained by a random walk of length n, and $\tau(w_n)$ is the translation length of w_n acting on Y.

This theorem applies (for example) to the action of $\mathrm{Out}(F_n)$ on the free splitting complex, and on the complex of free factors, both of which are (Gromov) hyperbolic; see Bestvina and Feighn [**3, 4**] and Handel and Mosher [**31**] for details.

Finally, in §6 we survey what is known for arbitrary groups, and deduce universal lower bounds on scl (known to be sharp in certain cases) from the work of Björklund and Hartnick [**6**]. Burger and Monod [**10**] showed that any quasimorphism differs by a bounded amount from a *harmonic function*. In particular, the expectation of this (adjusted) quasimorphism on a random walk is constant, i.e. the value of this function is a martingale. By applying the martingale central limit theorem, Björklund and Hartnick were able to obtain a central limit theorem for the distribution of values of a quasimorphism under a random walk. Their result applies in great generality, but if one specializes to finitely generated groups G with $Q(G)$ finite dimensional and $H_1(G)$ torsion for simplicity, Bavard duality plus the main theorem of [**6**] implies that for any ϵ there are positive constants a, b (depending only on ϵ) so that $\mathbf{P}(a < \mathrm{scl}_n/\sqrt{n} < b) \geq 1 - \epsilon$, where scl_n denotes the value of scl under a random walk of length n. Colloquially, one could say that scl has growth rate of order $\sqrt{n}$ in this case. We obtained a special case of the theorem of Björklund and Hartnick before their work was announced; because the method of proof is more geometric, we decided it was worth including in §6.1. One nice geometric corollary we obtain is as follows.

AREA AND WINDING THEOREM 6.5. *Fix some angle α and length ℓ. Let P_n be a random polygon in the hyperbolic plane with (cyclic) vertices $p_0, p_1, \ldots, p_n$, where $d(p_i, p_{i+1}) = \ell$ for each $0 \leq i \leq n-1$ and an angle of $\pm\alpha$ at each p_i with $0 < i < n$, with signs independently and uniformly chosen from ± 1. Let A_n be the algebraic area enclosed by P_n, and W_n the winding number of ∂P_n. Then A_n and W_n both satisfy a central limit theorem with mean 0.*

2. *Stable commutator length*

We recall some standard definitions and facts for the convenience of the reader. A basic reference for the material in this section is [**11**].

2.1. *Stable commutator length.*

Definition 2.1. Let G be a group, and G' its commutator subgroup. Given $g \in G'$, the *commutator length* of g, denoted $\mathrm{cl}(g)$, is the least number of commutators in G whose product is g, and the *stable commutator length*, denoted $\mathrm{scl}(g)$, is the limit

$$\mathrm{scl}(g) = \lim_{n\to\infty} \frac{\mathrm{cl}(g^n)}{n}.$$

The following estimates are elementary.

LEMMA 2.2. *Let G be a group, and g, h elements of G'. Then there is an inequality*

$$\mathrm{scl}(gh) \leq \mathrm{scl}(g) + \mathrm{scl}(h) + 1/2.$$

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

LEMMA 2.3. *Let G be a group with finite symmetric generating set S. Let $g \in G'$. Then there is a constant C_1 depending on G and S so that*

$$\mathrm{scl}(g) \leq C_1 |g|_S,$$

where $|\cdot|_S$ denotes word length with respect to S.

2.2. *Quasimorphisms.* There is a duality between stable commutator length and certain functions on G called *homogeneous quasimorphisms*.

Definition 2.4. Let G be a group. A function $\phi : G \to \mathbb{R}$ is a *quasimorphism* if there is some least non-negative number $D(\phi)$ called the *defect*, so that for all $g, h \in G$, there is an inequality

$$|\phi(gh) - \phi(g) - \phi(h)| \leq D(\phi).$$

A quasimorphism is *homogeneous* if, further, it satisfies $\phi(g^n) = n\phi(g)$ for all $g \in G$ and all $n \in \mathbb{Z}$.

Denote the vector space of all quasimorphisms on G by $\widehat{Q}(G)$, and the subspace of homogeneous quasimorphisms by $Q(G)$.

LEMMA 2.5. [**11**, Lemmas 2.21 and 2.58] *Given any $\psi \in \widehat{Q}$, the* homogenization $\overline{\psi}$, *defined by*

$$\overline{\psi}(g) = \lim_{n\to\infty} \psi(g^n)/n,$$

exists and satisfies $|\overline{\psi} - \psi| \leq D(\psi)$. Moreover, $\overline{\psi}$ is a homogeneous quasimorphism with $D(\overline{\psi}) \leq 2D(\psi)$.

Bavard duality is the following theorem.

THEOREM 2.6. (Bavard [**2**]; [**11**, Theorem 2.70]) *If G is a group and $g \in G'$ then*

$$\mathrm{scl}(g) = \sup_{\phi \in Q(G)} \frac{\phi(g)}{2D(\phi)}.$$

We use Bavard duality to obtain lower bounds on stable commutator length.

3. *Hyperbolic groups*

A standard introduction to the theory of hyperbolic groups is [**27**]. More specialized references for the material in this section are [**12, 14**] and [**16**].

3.1. *Hyperbolic groups.* Let G be a group with a generating set S. Let $C_S(G)$ (or just C for brevity) denote the Cayley graph of G with respect to the generating set S. The Cayley graph $C_S(G)$ can be made into a geodesic metric space, by giving each edge length 1. With this metric, $d(\mathrm{id}, g) = |g|_S$.

Definition 3.1. A group G is *hyperbolic* if $C_S(G)$ is hyperbolic as a metric space. That is, if there is some δ so that if pqr is a geodesic triangle, any point on the geodesic pq is within distance δ of $qr \cup rp$.

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

We are casual about identifying G with the vertices of $C_S(G)$, and in this way think of G as a metric space. If S is given, we say G is δ-*hyperbolic* for any δ as above.

There is an equivalence relation on proper geodesic rays in $C_S(G)$, where two rays are equivalent if and only if they are a finite Hausdorff distance apart. The set of equivalence classes is itself a compact space called the *Gromov boundary* of G, and denoted $\partial_\infty G$. The left action of G on itself (or on $C_S(G)$) extends to an action on $\partial_\infty G$ by homeomorphisms.

3.2. *Combings.* For an introduction to combings, regular languages etc, see [**18**].

Let S^* denote the set of finite words in the generating set S, let $|\cdot|$ denote word length in S^*, and let $\text{eval} : S^* \to G$ denote the evaluation map. A word $w \in S^*$ is a *geodesic* if $|\text{eval}(w)|_S = |w|$. Under evaluation of prefixes, a word $w \in S^*$ determines a *directed path* in the Cayley graph $C_S(G)$ from id to $\text{eval}(w)$. We denote this path by $\text{path}(w)$.

Suppose X is a finite directed graph (hereafter digraph) with a distinguished initial vertex, and with edges labeled by elements of S in such a way that there is at most one outgoing edge from each vertex with a given label.

Let Γ denote the set of finite directed simplicial paths in X, and Γ_0 the subset starting at the initial vertex. There is an injective map $\text{word} : \Gamma_0 \to S^*$ which takes a path γ to the string of edge labels on the edges it traverses, in order. The composition $\text{path} \circ \text{word}$ takes paths in Γ_0 to paths in $C_S(G)$ starting at id. When the meaning is clear from context, we denote $\text{path} \circ \text{word}(\gamma)$ simply by γ.

A subset of the form $\text{word}(\Gamma_0) \subset S^*$ is necessarily prefix-closed, since Γ_0 is closed under taking initial subpaths. Furthermore, $\text{word}(\Gamma_0)$ is a regular language; in fact, a subset $L \subset S^*$ is a prefix-closed regular language if and only if there is some X with $\text{word}(\Gamma_0) = L$.

Definition 3.2. A *combing* of G (with respect to a generating set S) is a subset $L \subset S^*$ for which there is some labeled digraph X as above with $\text{word}(\Gamma_0) = L$ such that:

(1) the evaluation map $\text{eval} : L \to G$ is a bijection; and

(2) the words of L are geodesics.

We say the digraph X *parameterizes* the combing.

Note that L is prefix-closed by our conventions. Note also that for a combing, the image of $\gamma \in \Gamma_0$ under $\text{path} \circ \text{word}$ is a geodesic in $C_S(G)$.

Remark 3.3. Many conflicting definitions of combings appear in the literature. Our definition is by no means standard.

THEOREM 3.4. (Cannon [**14**]) *Let G be a hyperbolic group, and let S be a generating set. Then there is a combing for G with respect to S.*

Example 3.5. An ordering of S determines a lexicographic order (i.e. dictionary order) on S^*. In any hyperbolic group G, the language L of *lexicographically first geodesic representatives* is an example of a combing in the sense of Definition 3.2.

3.3. *Markov chains.* The digraph X is a topological Markov chain. Associated to X is the *transition matrix* M whose M_{ij} entry counts the number of directed edges from vertex i

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

to vertex j. Note that M has a non-negative real eigenvalue λ so that $|\xi| \leq \lambda$ for every other eigenvalue ξ of M.

Two vertices in a digraph are said to be *communicating* if there are directed paths between the vertices in both directions. This defines an equivalence relation on X, and we define a *component* to be a maximal subgraph whose vertices are a communicating class. There is a natural digraph $C(X)$ obtained as a quotient of X, whose vertices are the communicating classes of X. From the definition it is clear that the digraph $C(X)$ has no directed loops.

If C is a component of X, the adjacency matrix M_C of C has biggest real eigenvalue $\lambda_C \leq \lambda$.

Definition 3.6. A component C is *maximal* if $\lambda_C = \lambda$.

In general, there are no constraints on $C(X)$ other than the fact that it has no directed loops. However, for G a hyperbolic group, and X a digraph parameterizing a combing, we have the following, due (implicitly) to Coornaert.

THEOREM 3.7. (Coornaert [**16**]; [**12**, Lemma 4.15]) *Let G be a hyperbolic group, and let X be a digraph parameterizing a combing. Then each directed path in $C(X)$ contains at most one maximal component. Equivalently, if λ is the maximal real eigenvalue of M, there are positive constants C_1, C_2 so that*

$$C_1\lambda^n \leq |G_n| \leq C_2\lambda^n.$$

Remark 3.8. Actually, Coornaert does not use the language of combings, and only proves the *a priori* weaker fact that $C_1\lambda^n \leq |G_{\leq n}| \leq C_2\lambda^n$. The stronger fact with $G_{\leq n}$ replaced by G_n follows immediately once one realizes that G_n counts the number of elements of length n in a prefix-closed regular language. As far as we know, this sharper observation appears for the first time in print in [**12**].

There is a natural stationary Markov chain with states the vertices of X, for which the positive transition probabilities are a subset of the (directed) edges of X. This Markov chain is described in detail in [**12**, §4]. We let N_{ij} denote the probability of a transition from state i to state j (note $N_{ij} = 0$ if $M_{ij} = 0$) and μ a certain stationary measure for N of maximal entropy.

LEMMA 3.9. [**12**, Lemmas 4.9–4.10] *The matrix N is a stochastic matrix, and μ is stationary for N. The support of μ is the union of the maximal components of X.*

For each n, let Γ^n denote the subset of Γ consisting of paths of length n, and let Γ^n_0 denote the subset starting at the initial vertex. Note that Γ^n_0 bijects with G_n under eval $\circ$ word. There is a probability measure on Γ^n which we denote μ by abuse of notation, defined by

$$\mu(\gamma) = \mu(\gamma(0))N_{\gamma(0)\gamma(1)}N_{\gamma(1)\gamma(2)} \cdots N_{\gamma(n-1)\gamma(n)}.$$

Thinking of Γ^n as the cylinder sets in the space Γ^∞ of infinite paths, there is an associated probability measure on Γ^∞ which is invariant under the shift map s which takes a path to the suffix obtained by omitting the first vertex.

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

We can also define a measure ν on Γ_0^n by

$$\nu(\gamma) = N_{0,\gamma(1)} N_{\gamma(1)\gamma(2)} \cdots N_{\gamma(n-1)\gamma(n)}$$

(since $\gamma(0) = 0$ by definition for $\gamma \in \Gamma_0^n$), and extend to a measure ν on Γ_0^∞. The shift map s takes Γ_0^∞ into Γ^∞, and we have $\mu = \lim_{n\to\infty}(1/n)\sum_{i=0}^{n-1} s_*^i \nu$. Note that ν is just μ conditioned on $\Gamma_0 \subset \Gamma$ (or on $\Gamma_0^n \subset \Gamma^n$ for each n). Note too that if the maximal components of X (the support of μ) are all aperiodic, then $\mu = \lim_{n\to\infty} s_*^n \nu$, and in general $\mu = \lim_{n\to\infty}(1/m)\sum_{i=0}^{m-1} s_*^{i+n}\nu$, where m is the lcm of the periods of the maximal components.

3.4. *Ergodicity at infinity.* For each n, let ν_n be the probability measure on G defined by

$$\nu_n = \frac{\sum_{|g|_S \le n} \lambda^{-|g|_S} \delta_g}{\sum_{|g|_S \le n} \lambda^{-|g|_S}},$$

where δ_g is the Dirac measure supported at g. This extends trivially to a probability measure on $G \cup \partial_\infty G$. It turns out that the limit $\nu = \lim_{n\to\infty} \nu_n$ exists and is supported on $\partial_\infty G$, where it is known as a *Patterson–Sullivan* measure.

THEOREM 3.10. (Coornaert [**16**, Theorem 7.7]) *Let ν be the Patterson–Sullivan measure. The action of G on $\partial_\infty G$ preserves the measure class of ν, and is ergodic.*

The relationship between the measures ν_n and the measures ν on Γ_0^n is as follows. Let $\gamma \in \Gamma_0^n$ and let $g = \mathrm{eval}(\mathrm{word}(\gamma))$. Define $\mathrm{cone}(g)$ to be the image in G of the union of the $\gamma' \in \Gamma_0^m$ extending γ. Then $\nu(\gamma) = \nu(\mathrm{cone}(g)) := \lim_{n\to\infty} \nu_n(\mathrm{cone}(g))$. Let $\gamma' \in \Gamma_0^{n+1}$ be obtained by extending γ, and let $h = \mathrm{eval}(\mathrm{word}(\gamma'))$ so that $h = gs$ for some generator s. Let i be the terminal vertex of γ, and j the terminal vertex of γ'. Then $N_{ij} = \nu(\mathrm{cone}(h))/\nu(\mathrm{cone}(g))$. See [**12**, §4.3] for details.

The bijection $\Gamma_0^n \to G_n$ lets us compare the ν-measure on Γ_0^n with the uniform measure on G_n. From the relationship between ν and N_{ij}, the ν-measure of $g \in G_n$ (i.e. $\nu(\mathrm{cone}(g))$) depends only on cone type of g, or equivalently on the terminal vertex of the corresponding path in Γ_0^n. Since there are only finitely many cone types, and since every $g \in G$ is a uniformly bounded distance from some g' whose cone has growth $\Theta(\lambda^n)$ (and therefore has $\nu(g')$ bounded away from 0) we derive the following proposition.

PROPOSITION 3.11. [**12**, §4] *There are positive constants C_1, C_2, C_3 with the following property. Let $g \in G_n$, and let B denote the intersection of the ball of radius C_1 about g (in $C_S(G)$) with G_n. Let B_Γ denote the preimage of B under the natural bijection $\Gamma_0^n \to G_n$. Then there is an inequality*

$$C_2 \nu(B_\Gamma)/\nu(\Gamma_0^n) \le |B|/|G_n| \le C_3 \nu(B_\Gamma)/\nu(\Gamma_0^n).$$

Finally, we state a proposition which describes the typical behavior of a ν-random element of Γ_0^n.

PROPOSITION 3.12. [**12**, Lemma 4.10] *Let γ be a ν-random element of Γ_0^n. Then as $n \to \infty$, with probability $1 - O(C_1^{-C_2 n})$, apart from a prefix of size $C_3 \log n$, the path γ is entirely contained in a single maximal component C, and it enters (and stays in) a given maximal component C with probability $\mu(C)$.*

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

The collective significance of these propositions can be summarized in the following way: every $g \in G_n$ is a uniformly bounded distance away from some $h \in G_n$ corresponding to a path $\gamma \in \Gamma_0^n$, which is generated (apart from a prefix of length $O(\log n)$) by one of finitely many ergodic stationary Markov processes, corresponding to one of the maximal components of X.

We will establish properties which hold for typical sequences in each of these stationary Markov processes. Sequences in different Markov processes can be compared using Theorem 3.10. This will let us establish properties which hold for ν-typical $g \in G_n$, and we will deduce that the properties also hold for typical $g \in G_n$ (with respect to the uniform measure).

3.5. *Mixing times and a Chernoff-type estimate.* Let X be a finite stationary ergodic Markov chain with transition matrix N, and μ the stationary measure. Define measures μ on Γ^n and Γ^∞ as before. In our applications, X will be a maximal component $C(X)$ of X as above, and μ will be μ conditioned on $C(X)$.

If $\sigma, \gamma \in \Gamma$, define $C_\sigma(\gamma)$ to be the number of times σ appears as a subsequence of γ. Now if γ_n is a μ-random element of Γ_n, then ergodicity of X implies that $(1/n)C_\sigma(\gamma_n) \to \mu(\sigma)$ in probability. A *Chernoff-type estimate* says that the probability that the deviation $|C_\sigma(\gamma_n) - n\mu(\sigma)|$ is of size $\Theta(n^\delta)$ decays exponentially in n, for δ in a suitable range (the critical exponent is half the exponent of $n\mu(\sigma)$, corresponding to the distribution in the central limit theorem). For our purposes, it is necessary to estimate the probability when $|\sigma| = O(\log n)$.

Let λ be the entropy of X (as above), and let $m = \log_\lambda n = \log n / \log \lambda$ so that there are $O(n^\ell)$ elements of Γ of length ℓm.

PROPOSITION 3.13. *Suppose σ has length ℓm with $\ell < 1$. Then for any $\epsilon > 0$ there are constants $C_1 > 1$ and $C_2 > 0$ (depending on X and ℓ but not on n or σ) so that*

$$\mathbf{P}(|C_\sigma(\gamma_n) - n\mu(\sigma)| < n^{\epsilon+(1-\ell)/2}) = 1 - O(C_1^{-n^{C_2}}).$$

In fact, there is a fixed integer L so that for each residue $i \bmod Lm$, if $\gamma_{i,k}$ is the subpath of γ of length ℓm starting at $i + kLm$, and $C_{\sigma,i}(\gamma_n)$ is the number of copies of σ amongst the $\gamma_{i,k}$, then

$$\mathbf{P}\left(\left|C_{\sigma,i}(\gamma_n) - \frac{n}{Lm}\mu(\sigma)\right| < n^{\epsilon+(1-\ell)/2}\right) = 1 - O(C_1^{-n^{C_2}}).$$

Proof. This proposition is essentially a special case of a theorem of Dinwoodie [**17**], and we sketch the proof, referring to [**17**] for details (also compare [**13**, §2]). For simplicity, we assume X is aperiodic (the general case is very similar).

In fact, the first estimate follows from the second, so we prove the second. Choose some big fixed integer L, and let i be a residue mod Lm. Let $\gamma_{i,k}$ be the subpath of γ of length ℓm starting at $i + kLm$. For fixed i, the successive $\gamma_{i,k}$ are not independent, but their correlation is extremely small. Explicitly, if $\lambda_1 < 1$ is the second largest eigenvalue of the transition matrix N, then, conditional on any given value of $\gamma_{i,k}$, the probability that the first vertex of $\gamma_{i,k+1}$ is equal to some j differs from $\mu(j)$ by $\lambda_1^{(L-\ell)m}$. So, conditioned on any given value of $\gamma_{i,k}$, the probability that $\gamma_{i,k+1}$ is equal to σ is bounded

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

above by $\mu(\sigma)(1+\lambda_1^{(L-\ell)m}/\mu(\sigma(0)))$ and below by $\mu(\sigma)(1-\lambda_1^{(L-\ell)m}/\mu(\sigma(0)))$, where $\sigma(0)$ is the initial vertex of σ. If L is sufficiently large, these bounds are approximately $\mu(\sigma)(1\pm n^{-C_3})$, where C_3 may be taken to be as large as we like.

Let $C_{\sigma,i}(\gamma_n)$ be the number of copies of σ amongst the $\gamma_{i,k}$. By the argument above, $C_{\sigma,i}(\gamma_n)$ may be estimated from the Chernoff bound for *independent* Bernoulli variables; see, e.g., [**47**, Theorem 1.3.13]. Explicitly, we obtain an estimate of the form

$$\mathbf{P}\left(\left|C_{\sigma,i}(\gamma_n)-\frac{n}{Lm}\mu(\sigma)(1\pm n^{-C_3})\right|<n^{\epsilon+(1-\ell)/2}\right)=1-O(C_1^{-n^{C_2}}).$$

Taking C_3 sufficiently large, we may absorb it into the $n^{\epsilon+(1-\ell)/2}$ term, at the cost of increasing ϵ an arbitrarily small amount.

Summing over i we get

$$\mathbf{P}(|C_\sigma(\gamma_n)-\mu(\sigma)|<Lm\cdot n^{\epsilon+(1-\ell)/2})=1-O(Lm\cdot C_1^{-n^{C_2}}).$$

Since $Lm=O(\log n)$, by adjusting constants the result holds. □

Remark 3.14. It is worth spelling out the meaning of λ, Γ^n, μ and so on for a simple example. For example, if F is the free group of rank k, after fixing a free generating set, there is a combing of F consisting of the regular language of all reduced words in the generators. Then Γ^n consists of the set of reduced words of length n, the entropy λ is equal to $2k-1$ since $|\Gamma^n|=2k\cdot(2k-1)^{n-1}$, and μ is the uniform probability measure on Γ^n for each n.

Remark 3.15. The meaning of Proposition 3.13 is somewhat hard to glean directly from the statement. It can be explained informally as follows. We would like to prove a strong equidistribution theorem for the set of subpaths of γ of length ℓm. If these subpaths were independent, the Chernoff bound would give us what we want. But the subpaths are not independent—in fact, two subpaths with a big overlap are highly dependent on each other. However, the correlations between subpaths decay *exponentially quickly*, and therefore a collection of subpaths with the property that the distance (in γ) between distinct subpaths is at least Lm for some fixed $L\gg 1$ are 'independent enough' for a modified Chernoff bound to apply. The set of all subpaths of length ℓm can be partitioned into Lm subsets so that each subset satisfies the desired separation property, and we get an equidistribution result that holds on each subset with high probability. Combining the subsets together merely has the effect of multiplying the error by $Lm=O(\log(n))$, which can be absorbed into one of the constants.

3.6. *Anti-alignment.*

Definition 3.16. Two oriented geodesic segments γ_1, γ_2 in $C_S(G)$ are K *anti-aligned* if there is $g\in G$ so that the terminal point of $g(\gamma_1)$ is within distance K of the initial point of γ_2, and *vice versa*.

By the defining property of δ-hyperbolicity, the translate of γ_1 is contained in the $K+2\delta$-neighborhood of the translate of γ_2, and *vice versa*. Note that the property of being K anti-aligned is a property of *orbit classes* of geodesic segments in $C_S(G)$ under

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

the action of G. Given any path $\gamma \in \Gamma$, the orbit class of γ in $C_S(G)$ is well defined, so it makes sense to ask if a pair of elements of Γ are K anti-aligned.

PROPOSITION 3.17. *Fix $L > 2$ and some constant K. Let γ be a ν-random element of Γ_0^n, and let γ_i denote the successive subpaths of γ of length Lm. Then there is some $\epsilon > 0$ and constants C_1, C_2 so that*

$$\mathbf{P}(\#\{i : \gamma_i \text{ } K \text{ anti-aligns some } \gamma' \subset \gamma\} < n^{1-\epsilon}) = 1 - O(C_1^{-n^{C_2}}).$$

Proof. By Proposition 3.12, we can assume that apart from a prefix of size $O(\log n)$, γ is contained in some maximal component C.

Now, if γ_i, γ' subpaths of length Lm are K anti-aligned, then there are *disjoint* $\gamma_i' \subset \gamma_i$ and $\gamma'' \subset \gamma'$ of length at least $Lm/2$ which are $K + 2\delta$ anti-aligned. Replace K by $K + 2\delta$ and L by $L/2$, and by abuse of notation let γ_i denote successive subpaths of length Lm. Then we need only show that

$$\mathbf{P}(\#\{i : \gamma_i \text{ } K \text{ anti-aligns some } \gamma' \text{ disjoint from } \gamma_i\} < n^{1-\epsilon}) = 1 - O(C_1^{-n^{C_2}}).$$

For any given location of γ' in γ, there are $O(n^L)$ distinct elements of Γ_0^n which agree with γ outside γ', all with approximately the same μ-measure. If γ' anti-aligns some γ_i, then γ' is determined by γ_i, by $h_1, h_2 \in G$ of length $\leq K$, and by the initial vertex of γ' in X. Hence, for each γ_i there are only C_3 possible choices for γ', and therefore only $O(n)$ choices, where we range over all i. In particular, conditioning on γ outside γ', the probability that a given γ' K anti-aligns some γ_i that it is disjoint from is at most $O(n^{1-L})$. The proposition follows. □

3.7. *Comparing components.* Proposition 3.13 lets us very accurately estimate the number of copies of a subword σ in a random word γ_n of length n in a maximal component C of X, at least when $|\sigma| = \ell m$ with $\ell < 1$ for $m = \log_\lambda n$.

However, it is necessary to compare such distributions for *different* maximal components. We introduce some notation.

For each maximal component C, let $\mu|_C$ be the conditional probability measure with support in C. That is, define $\mu(\cdot)|_C = \mu(\cdot)/\mu(C)$ in C, and $\mu(\cdot)|_C = 0$ on the complement of C. By abuse of notation, this defines a probability measure $\mu|_C$ on Γ, defined on cylinder sets by

$$\mu|_C(\gamma) = \mu|_C(\gamma(0))N_{\gamma(0)\gamma(1)} \cdots N_{\gamma(n-1)\gamma(n)}.$$

Proposition 3.13 says that for γ_n a ν-random path in X conditioned to enter a specific component C, we have $|C_\sigma(\gamma_n) - \mu|_C(\sigma)| < n^{\epsilon+(1-\ell)/2}$ with exponentially few exceptions. Under the evaluation map, for each M the measure $\mu|_C$ on Γ^M pushes forward to a probability measure $\mu|_C$ on G_M. We would like to compare the measures $\mu|_C$ on G_M for different maximal components C.

It is too much to hope that these measures will be *equal*. However, it turns out that each $\mu|_{C_i}$ can be obtained from $\mu|_{C_j}$ by (roughly speaking) a random convolution process, as described in the following lemma.

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

LEMMA 3.18. *Suppose the Cayley graph $C_S(G)$ is δ-hyperbolic. Then the following is true: for each pair of big components C_i, C_j of X, and for each M, there is a map $f_{i,j} : G_M \to \mathrm{Prob}(G_M)$ satisfying:*

(1) *for all g and all h in the support of $f_{i,j}(g)$ there is $a \in G$ with $|a| \leq \delta$ so that $d(g, aha^{-1}) \leq 2\delta$; and*

(2) $\mu|_{C_j}(h) = \sum_g \mu|_{C_i}(g) f_{i,j}(g)(h)$.

We express point (2) *by saying that $\mu|_{C_j}$ is obtained by* convolving *$\mu|_{C_i}$ with $f_{i,j}$.*

Proof. In fact, the proof of this lemma is a very simple trick, which is a variation on the main trick of [**12**]. By Kakutani's random ergodic theorem (see [**35**]; alternately, this follows from Proposition 3.13) and the definition of $\mu_{i,m}$, almost every infinite path in X that enters C_i is composed of subpaths of length M that are distributed in G_M according to $\mu|_{C_i}$. Call an infinite path (i.e. an element of Γ) *C_i-typical* if its subpaths have this property. Let γ_i be a C_i-typical infinite path, and by abuse of notation, let γ_i denote the corresponding infinite geodesic path in G starting at id, limiting to some point in $\partial_\infty G$. By ergodicity of the action of G on $\partial_\infty G$ (i.e. Theorem 3.10), there are typical γ_i and γ_j and some $a \in G$ so that $a \cdot \gamma_i$ and γ_j have the same endpoint, and therefore, outside some compact subset, they synchronously fellow-travel. It follows that we can subdivide $a \cdot \gamma_i$ and γ_j into subpaths of length M, throwing away finitely many at the start, so that corresponding subpaths are each contained in the δ-neighborhood of each other. If two associated subpaths evaluate to g and h in G_M respectively, then $g = a_1 h a_2$, where each of a_1, a_2 has length at most δ. Hence, $d(g, a_1 h a_1^{-1}) \leq 2\delta$. So define $f_{i,j}(g)(h)$ to be the probability that a subpath of γ_i evaluating to g fellow-travels (as above) a subpath of γ_j evaluating to h. □

We conclude that each individual measure $\mu|_{C_i}$ can be convolved (in the sense above) by a weighted average $\sum p_j f_{i,j}$ to obtain a single probability measure μ' on G_M. Now, though it might not be true that $\mu'(v) = \mu'(v^{-1})$ for all $v \in G_M$, this is *approximately* true, up to convolution in the sense of Lemma 3.18.

LEMMA 3.19. *Suppose the Cayley graph $C_S(G)$ is δ-hyperbolic. Then the following is true: for each M there is a map $f : G_M \to \mathrm{Prob}(G_M)$ satisfying:*

(1) *for all g and all h in the support of $f(g)$ there is $a \in G$ with $|a| \leq \delta$ so that $d(g, ah^{-1}a^{-1}) \leq 2\delta$; and*

(2) $\mu'(h) = \sum_g \mu'(g) f(g)(h)$.

Proof. Choose some $N \gg M$, so that with probability $1 - O(C_1^{-n^{C_2}})$, a random path γ in Γ^N satisfies $|C_\sigma(\gamma) - N\mu|_{C_i}(\sigma)| < N^{1-C_3}$ for every $\sigma \in \Gamma^M$, where C_i is the big component that γ enters (the existence of such an N given M follows from Proposition 3.13).

We call a γ for which such an estimate holds (for some i) *almost typical*. By Proposition 3.11, there are almost typical paths γ and γ' evaluating to g and g' respectively in G_N, and satisfying $d(g^{-1}, g') \leq C_4$.

Let γ^{-1} denote the path in G from id to g^{-1} obtained by reversing and translating γ (note that γ^{-1} will not typically be an element of Γ). Since subpaths of γ^{-1} are in bijection with subpaths of γ but oriented oppositely, $C_\sigma(\gamma) = C_{\sigma^{-1}}(\gamma^{-1})$ for all σ of length M.

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

Since γ^{-1} and γ' synchronously fellow-travel, after throwing away suffixes of uniformly bounded length (depending only on C_4), we can pair subpaths of γ^{-1} and γ' of length M so that corresponding subpaths are each contained in the δ-neighborhood of each other. Taking $N \to \infty$ (for fixed M) the proof follows. □

4. *The scl of random geodesics in hyperbolic groups*

We now have nearly all the necessary tools to estimate scl on random elements in hyperbolic groups. There is one significant additional complication for hyperbolic groups G for which $H_1(G)$ has positive rank, namely that scl is only defined on elements in the commutator subgroup.

We will show in §4.3 that the relative proportion of $[G, G]$ in the set of elements of length $[n - C_1, n + C_1]$ (for a suitable constant C_1) is of size $O(n^{-k/2})$, where k is the rank of $H_1(G)$; in particular, it is *polynomial* in n. It follows that properties of words that hold with probability $1 - O(C_2^{-n^{C_3}})$ in G_n will also hold with a similar estimate in probability for words conditioned to lie in $[G, G]$.

4.1. *Almost pairing subwords.*

PROPOSITION 4.1. *Let G be a hyperbolic group, and let S be a generating set. Let λ be the growth rate for G (i.e. the number such that $|G_n| = \Theta(\lambda^n)$). Then for any $\epsilon > 0$, there are constants $C_1 > 1$, $C_2 > 0$ so that if g is a random element of G_n, with probability $1 - O(C_1^{-n^{C_2}})$ the following holds:*

(1) *there is some h with $|h|_S \leq 8\delta \cdot n(\log \lambda + \epsilon)/\log n + o(n/\log n)$; and*
(2) *the product gh^{-1} is in $[G, G]$; and*
(3) *there is an estimate $\mathrm{cl}(gh^{-1}) \leq n(\log \lambda + \epsilon)/2 \log n$.*

Proof. The proof follows by assembling the results of the previous section. Let $\gamma \in \Gamma_0^n$ correspond to g. By Proposition 3.11, it suffices to let γ be a ν-random element of Γ_0^n. With the desired probability, we can assume, apart from a prefix and suffix of size $O(\log n)$, that γ is contained in a single maximal component C. Fix some $\ell < 1$, and consider the set of successive subpaths γ_i of γ of length ℓm, where $m = \log n/\log \lambda$. By Proposition 3.13, the distribution of the γ_i is very close to $\mu|_C$. By Lemmas 3.18 and 3.19 we can pair most of the γ_i in such a way that the evaluation of each pair is approximately inverse, i.e. if g_i, g_j are the values of the pair of segments in G, there are a_1, a_2 with $|a_i| \leq 4\delta$ and $g_i = a_1 g_j^{-1} a_2$. We can therefore cancel each g_i with g_j at the cost of a commutator plus a word of length at most 8δ. The product of these 8δ words is h. All but $o(n/\log n)$ can be paired in this way; the remainder can be absorbed into h at the cost of adjusting ϵ an arbitrarily small amount (for big n). The proof follows. □

4.2. *Counting quasimorphisms.* We obtain lower bounds on scl via Bavard duality, by constructing explicit quasimorphisms with uniformly bounded defect, and value $O(n/\log n)$ on random g. The quasimorphisms in question are a variant on the *small counting quasimorphisms* of Epstein and Fujiwara [**19**]. In fact, it is useful to work in the generality of a group acting on a δ-hyperbolic complex, following [**21**] (also see [**11**, §3.5]).

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

Let Y be a δ-hyperbolic simplicial complex (*not* necessarily locally finite) and let G act on Y simplicially. If σ is a finite oriented simplicial path in Y, let σ^{-1} denote the same path with the opposite orientation. A *copy* of σ is a translate $g \cdot \sigma$ for some $g \in G$.

Definition 4.2. Let Y be a δ-hyperbolic simplicial complex, and let $p \in Y$ be a base point. Let Σ be a (possibly infinite) collection of oriented simplicial paths in Y, and let Σ^{-1} denote the collection obtained by reversing the orientations on all $\sigma \in \Sigma$. For any oriented simplicial path γ in Y define

$$c_\Sigma(\gamma) = \text{maximal number of } \textit{disjoint} \text{ copies of } \sigma \in \Sigma \text{ contained in } \gamma$$

and then for $g \in G$ define

$$c_\Sigma(g) = d(p, g(p)) - \inf_\gamma(\text{length}(\gamma) - c_\Sigma(\gamma)),$$

where the infimum is taken over *all* oriented simplicial paths γ from p to $g(p)$. Define the *small counting quasimorphism* h_Σ by the formula

$$h_\Sigma(g) := c_\Sigma(g) - c_{\Sigma^{-1}}(g).$$

A path γ from p to $g(p)$ realizing the infimum of $\text{length}(\gamma) - c_\Sigma(\gamma)$ is called a *realizing path* for γ. Since the values of this function are integers, a realizing path exists.

LEMMA 4.3. (Fujiwara [**21**, Lemma 3.3]) *Suppose the length of every $\sigma \in \Sigma$ is ≥ 2. Then any realizing path is a $(2, 4)$-quasigeodesic.*

It follows that any realizing path is within distance C_1 of a geodesic, where C_1 depends only on δ.

LEMMA 4.4. (Fujiwara [**21**, Proposition 3.10]) *Suppose the length of every $\sigma \in \Sigma$ is ≥ 2. Then there is a constant C_2 depending only on δ such that $D(h_\Sigma) \leq C_2$.*

Remark 4.5. Actually, Epstein and Fujiwara only consider small counting functions for a single σ, but the bounds on the geometry of realizing paths and on the defect of $D(h_\Sigma)$ are valid for arbitrary collections Σ as above. This is by contrast with the 'big' counting functions introduced by Brooks [**9**], in which one counts *all* (possibly overlapping) copies of σ, not just a maximal disjoint collection.

Remark 4.6. As defined, the quasimorphism c_Σ depends on the choice of basepoint p. However, different choices of points p give rise to quasimorphisms with the same homogenization. Since it is the homogenization we really care about, we gloss over this detail.

We apply this construction to the case that $Y = C_S(G)$ and $p = \text{id}$ so that $d(p, g(p)) = |g|_S$.

PROPOSITION 4.7. *Let G be a hyperbolic group, and let S be a generating set. Let λ be the growth rate for G. Then for any $\epsilon > 0$ there are constants $C_1 > 1$, $C_2 > 0$ and C_3 depending only on δ so that if g is a random element of G_n, with probability $1 - O(C_1^{-n^{C_2}})$, there is a homogeneous quasimorphism ϕ on G satisfying the following properties:*

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

(1) $\phi(g) \geq n \log \lambda/(2+\epsilon) \log n$;
(2) $D(\phi) \leq C_3$;
(3) $|\phi(h)| \leq 2|h|_S \log \lambda/(2-\epsilon) \log n$ *for any* $h \in G$.

Proof. By Proposition 3.11 it suffices to prove the theorem for a ν-random element g. Fix $L = 2+\epsilon$, and recall the notation $m = \log n/\log \lambda$. Let $\gamma \in \Gamma_0^n$ correspond to g, and let γ_i denote the successive subpaths of γ of length Lm. For a suitable constant K (to be determined shortly) let Σ be the subset of the γ_i which do not K anti-align any $\gamma' \subset \gamma$. If c_Σ denotes the small counting function associated to the set Σ (thought of as a collection of orbit classes of subpaths of the Cayley graph $C_S(G)$) then by Proposition 3.17 we obtain an estimate $c_\Sigma(g) \geq n \log \lambda/(2+\epsilon) \log n$.

Now, there is a constant K so that any $(2, 4)$-quasigeodesic in a δ-hyperbolic space stays within distance K of a genuine geodesic. If we choose K with this property, then by Lemma 4.3 and the definition of Σ we have $c_{\Sigma^{-1}}(g) = 0$, and therefore $h_\Sigma(g) \geq n \log \lambda/(2+\epsilon) \log n$. On the other hand, $D(\phi) \leq C_3$ for some constant C_3 depending only on δ, by Lemma 4.4.

Finally, if h is any element of word length $|h|_S$, any realizing path for c_Σ or $c_{\Sigma^{-1}}$ is a $(2, 4)$-quasigeodesic, which therefore has length at most $2|h|_S + 4$ and thus contains at most $(2|h|_S + 4)/Lm$ disjoint paths of length Lm. □

In particular, if $g \in [G, G]$ then $\mathrm{scl}(g) \geq C_1 n/\log n$ with very high probability, and if there is h with $|h|_S = O(n/\log n)$ and $gh^{-1} \in [G, G]$ then $\mathrm{scl}(gh^{-1}) \geq C_2 n/\log n$.

4.3. *Homology.* Propositions 4.1 and 4.7 together give upper and lower bounds on $\mathrm{scl}(gh^{-1})$ for some h with $|h|_S = O(n/\log n)$ and $gh^{-1} \in [G, G]$. If g is in $[G, G]$ then so is h, and $\mathrm{scl}(h)$ can be estimated from its length (see Lemma 4.8). We will see in this section that the relative proportion of homologically trivial g in G_n is polynomial of bounded degree, and therefore a random element of G_n *conditioned to lie in* $[G, G]$ will have two-sided bounds on scl, with high probability.

The following lemma is elementary.

LEMMA 4.8. *Let G be a group and S a finite generating set. There is a constant C_1 so that if h is an element of $[G, G]$ then* $\mathrm{cl}(h) \leq C_1|h|_S$.

Proof. For simplicity, we suppose $H_1(G)$ is torsion-free; the general case is not significantly harder.

Assume without loss of generality that $S = S^{-1}$, and let L be the free abelian semigroup generated by S as a set. We think of L as the intersection of the positive orthant in $\mathbb{R}^{|S|}$ with $\mathbb{Z}^{|S|}$. The map from S to $H_1(G; \mathbb{R})$ extends by linearity to $\mathbb{R}^{|S|}$, and the kernel K is a rational subspace. It follows that $K \cap L$ is finitely generated as an abelian semigroup (such a generating set is known as a *Hilbert basis*); see, e.g., [**1**]. It follows that if we write h as a (geodesic) word in the elements of S, there is a constant C_1 so that there is a subset of letters of cardinality at most C_1 whose image in homology is trivial. We move these letters to the right, at the cost of at most C_1 commutators. In other words, we can write $h = h_1h_2h_3$, where each h_i is in $[G, G]$, where $|h_1|_S < |h|_S$, where $|h_2|_S \leq C_1$, and where $\mathrm{cl}(h_3) \leq C_1$. Since there are only finitely many elements of G with $|\cdot|_S \leq C_1$, there is a

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

bound on the commutator length of the homologically trivial ones. Hence, $|h_2|_S \leq C_2$ and therefore, by induction, $\mathrm{cl}(h) \leq (C_1 + C_2) \cdot |h|_S$ as required. □

The main proposition of this section estimates the relative density of $[G, G]$ in G_n.

PROPOSITION 4.9. *Let G be a hyperbolic group with finite generating set S, and suppose the rank of $H_1(G)$ is k. There are constants C_1, C_2 so that if g is a random element of $[n - C_1, n + C_1]$ (with the uniform probability), then $\mathbf{P}(g \in [G, G]) \geq C_2 \cdot n^{-k/2}$.*

Proof. The case that G is free and S is a free generating set is due to Sharp [**45**], with $C_1 = 1$. However, the proof does not use very particular properties of free groups, and generalizes to hyperbolic groups. The key probabilistic component is a (standard) local limit theorem for random sums in ergodic (finite, stationary) Markov chains.

For general hyperbolic groups, one cannot apply such local limit theorems directly because X might have more than one maximal component; however, we can apply local limit theorems to each maximal component *individually*. To do this we need to know that the *expected* value in $H_1(G)$ of (the evaluation of) a random walk conditioned to lie in a maximal component C_i of X is zero. This follows from the main theorem of Calegari and Fujiwara [**12**], since a homomorphism to $\mathbb{Z}$ is an example of a bicombable function. It follows that the Markov chains associated to each component C_i satisfy the conditions in [**45**, Theorem 2] and we can obtain sharp estimates of the desired form on the probability that the evaluation of a random path in C_i has trivial abelianization.

Now, a ν-random $\gamma \in \Gamma_0^n$ is of the form $\gamma_1\gamma_2$, where γ_2 is contained in some maximal component. It is not true that we can bound the length of γ_1, but it *is* true that there is a constant C_1 so that $\mathbf{P}(|\gamma_1| \leq C_1) \geq 1/2$. If α denotes the abelianization map $\alpha : \Gamma^n \to H_1(G)$, then $|\gamma_1| \leq C_1$ implies a uniform estimate $|\alpha(\gamma_1)| \leq C_2$.

Because of the Markov property, we get an estimate for g a ν-random word of length n of the form

$$\mathbf{P}(|\alpha(g)| \leq C_2) \geq C_3 \cdot n^{-k/2}.$$

By Proposition 3.11, we get a similar estimate (but with different constants) for g a random element of G_n with the uniform probability. Now, there is a constant C_4 so that for every g with $|\alpha(g)| \leq C_2$ there is h with $|h|_S \leq C_2C_4$, so that $gh \in [G, G]$. The map $g \to gh$ is bounded-to-one, so the cardinality of the intersection of $[G, G]$ with the set of words of length in the interval $[n - C_2C_4, n + C_2C_4]$ is at least $C_5 n^{-k/2} \cdot |G_n|$. The proof follows. □

Putting this all together, we obtain the main theorem in this section.

THEOREM 4.10. (Hyperbolic geodesic theorem) *Let G be a hyperbolic group, and S a finite generating set. There are constants C_1, $C_2 > 0$, $C_3 > 0$, $C_4 > 1$, $C_5 > 0$ so that if g is a random element with $|g|_S \in [n - C_1, n + C_1]$ conditioned to lie in $[G, G]$, then*

$$\mathbf{P}(C_2 n/\log n \leq \mathrm{scl}(g) \leq C_3 n/\log n) = 1 - O(C_4^{-n^{C_5}}).$$

In fact, we obtain the stronger result $\mathrm{cl}(g) \leq C_3 n/\log n$, with the same estimate in probability.

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

Proof. The estimates in Propositions 4.7 and 4.1 hold with probability $1 - O(C^{-n^c})$, and therefore they still hold with the same order of probability (with different constants) conditioned on $g \in [G, G]$, by Proposition 4.9. Bavard duality (Theorem 2.6) therefore gives the lower bound.

Similarly, for random g conditioned to lie in $[G, G]$, Proposition 4.1 says we can write $g = gh^{-1}h$, where $\mathrm{cl}(gh^{-1}) \leq Cn/\log n$, and where $|h|_S \leq Cn/\log n$. Since $g \in [G, G]$ and $gh^{-1} \in [G, G]$, we have $h \in [G, G]$. So Lemma 4.8 gives $\mathrm{cl}(h) \leq Cn/\log n$. Putting this together gives the upper bound on $\mathrm{cl}(g)$, with the desired estimate in probability. □

As remarked in the introduction, the gap between the upper and lower bounds is presumably an artifact of the method of proof; in fact, in [**13**] the authors conjectured that there should be concentration for the random variable $\mathrm{scl}(g) \log n/n$ at $\log \lambda/6$, where λ is the growth entropy of G with respect to the generating set S. The main theorem of [**13**] proves this for a free group with respect to a free generating set (see Theorem 5.1 below for a precise statement).

5. *The scl of random walks in groups*

In this section we obtain estimates on the value of $\mathrm{scl}(g)$, where g is obtained by a *random walk* in a group G, providing G satisfies certain hypotheses. Even in a hyperbolic group G with a fixed generating set S, the probability distributions defined by random geodesics and by random walks are not usually uniformly comparable, and typically become mutually singular at infinity. However, there is one very important special case in which the two probability distributions can be compared very precisely, namely the case of free groups with a free generating set. It follows that we obtain upper bounds on scl in random words in free groups, with high probability.

The significance of this is not that we are interested in free groups *per se*, but rather that scl is *monotone* under homomorphisms. If G is any group, and S any (symmetric) generating set, then there is a surjective homomorphism $F_S \to G$, where F_S is the free group on S, and (simple) random walk on F_S pushes forward to random walk on G. It follows that any upper bound on scl on random walks in free groups gives a universal upper bound on scl on random walks in *any* group G.

These upper bounds are complemented by lower bounds for hyperbolic groups, and for certain groups acting on hyperbolic spaces (e.g. braid groups, mapping class groups). Universal lower bounds, valid for all groups, are obtained by a quite different method, and discussed in the next section.

5.1. *Special case: free groups.*

For F a free group of rank k, Calegari and Walker obtained a sharpening of Theorem 4.10.

THEOREM 5.1. (Calegari and Walker [**13**, Theorem 4.1]) *Let F be a free group of rank k, and let g be a random element of length n in a free generating set, where n is even, and g is conditioned to lie in $[F, F]$. Then for every $\epsilon > 0$ and every C there is an estimate*

$$\mathbf{P}(|\mathrm{scl}(g) \log n/n - \log(2k-1)/6| \leq \epsilon) = 1 - O(n^{-C}).$$

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

In fact, the upper bound (i.e. $\mathrm{scl}(g) \log n/n - \log(2k-1)/6) \leq \epsilon$) is proved to hold with probability $1 - O(C^{-n^c})$ for some $C > 1$, $c > 0$; see [**13**, Proposition 4.2]. Theorem 5.1 is derived from a proposition, valid for g random of length n, and then conditioning on $g \in [F, F]$. The following proposition is implicit in [**13**]; for completeness, we indicate how it follows immediately from §§4.1–4.3 of that paper.

PROPOSITION 5.2. *Let F be a free group of rank k. Then for any $\epsilon > 0$, there are constants $C_1 > 1$, $C_2 > 0$ and $\delta > 0$ so that if g is a random element of F_n, the following holds with probability $1 - O(C_1^{-n^{C_2}})$:*

(1) *there is some h with $|h|_S \leq O(n^{1/2+\epsilon})$; and*
(2) *the product gh^{-1} is in $[F, F]$; and*
(3) *there is an estimate $\mathrm{scl}(gh^{-1}) \leq n(\log(2k-1) + \epsilon)/6 \log n$.*

Proof. It is convenient to use the (well-known) extended definition of scl as a norm on homologically trivial formal real group 1-boundaries; see [**11**, §2.6] for details. Lemma 4.7 from [**13**] says that there is some formal 1-chain Γ with $|\Gamma|_S = O(n^{1-\delta})$ for some $\delta > 0$ so that $\mathrm{scl}(g - \Gamma) \leq n(\log(2k-1) + \epsilon)/6 \log n$. By an estimate of Rivin (see, e.g., [**44**]) we can assume that the L^1 norm of $[g] \in H_1(F) = \mathbb{Z}^k$ is $O(n^{1/2+\epsilon})$ for any ϵ, with probability $1 - O(C_1^{-n^{C_2}})$, and therefore there is h with $[g] = [h]$ and $|h|_S \leq O(n^{1/2+\epsilon})$. We estimate

$$\begin{aligned} \mathrm{scl}(gh^{-1}) &\leq \mathrm{scl}(g + h^{-1}) + 1/2 \\ &\leq \mathrm{scl}(g - \Gamma) + \mathrm{scl}(\Gamma + h^{-1}) + 1/2 \\ &\leq n(\log(2k-1) + \epsilon)/6 \log n + O(n^{1-\delta}) \end{aligned}$$

and the $O(n^{1-\delta})$ may be absorbed into the ϵ. □

Remark 5.3. The reader who is uncomfortable with the proof of Proposition 5.2 can safely use Proposition 4.1 instead in the following, after observing that $\delta = 0$ in a free group with a free generating set. The only cost is that the constant in Proposition 4.1 is worse by a factor of 6, whereas the constant in Proposition 5.2 is sharp.

5.2. *Universal upper bounds.* We now compare random words with random walks. The Cayley graph of F with respect to a free generating set is a regular $2k$-valent tree. The group of simplicial automorphisms of this tree, fixing the origin, acts transitively on the set of vertices at distance m, for any m. Let μ be the uniform probability measure on the generators, and let μ^{*n} denote the n-fold convolution, i.e. the probability measure associated to a random walk of length n. Then μ^{*n} is a weighted sum of uniform measures on the sets F_m for $m \leq n$. The generating function for the weights can be determined explicitly (see, e.g., [**49**, Lemma 1.24]), and a straightforward calculation shows that for any ϵ, all but $O(C_1^{-n^{C_2}})$ of the mass of μ^{*n} is concentrated on the set of F_m with $m/n \in [(k-1)/k - \epsilon, (k-1)/k + \epsilon]$. We therefore obtain the following proposition.

PROPOSITION 5.4. *Let F be a free group of rank k. Then for any $\epsilon > 0$, there are constants $C_1 > 1$, $C_2 > 0$ and $\delta > 0$ so that if g is obtained by random walk on F (in a free generating set) of length n, the following holds with probability $1 - O(C_1^{-n^{C_2}})$:*

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

(1) *there is some* h *with* $|h|_S \leq O(n^{1/2+\epsilon})$*; and*
(2) *the product* gh^{-1} *is in* $[F, F]$*; and*
(3) *there is an estimate* $\mathrm{scl}(gh^{-1}) \leq ((k-1)\log(2k-1)+\epsilon)/6k \cdot n/\log n$.

Remarkably, from this elementary estimate, we obtain a *universal sharp* upper bound on scl for random walks in *arbitrary* finitely generated groups.

THEOREM 5.5. (Universal upper bound) *Let* G *be a group with a finite symmetric generating set* S*, and let* $|S| = 2k$*. Let* g *be obtained by random walk on* G *(with respect to* S*) of length* n *(even), conditioned to lie in* $[G, G]$*. Then for any* $\epsilon > 0$ *there are constants* $C_1 > 1$*,* $C_2 > 0$ *so that with probability* $1 - O(C_1^{-n^{C_2}})$ *there is an inequality*

$$\mathrm{scl}(g) \leq ((k-1)\log(2k-1)+\epsilon)/6k \cdot n/\log n.$$

Proof. Let $\phi : F_k \to G$ take a free symmetric generating set for F_k to S. Then random walk in F_k (with respect to the standard generating set) pushes forward to random walk in G with respect to S. Since scl is monotone non-increasing under homomorphisms, the theorem follows from Proposition 5.4 and Lemma 4.8, together with the fact that a random walk of length n has probability at least $\Theta(n^{-k/2})$ of being homologically trivial (for n even). □

5.3. *Template for obtaining lower bounds.* We will obtain lower bounds, complementing Theorem 5.5, for random walks in hyperbolic groups, and certain groups acting on hyperbolic spaces; the most important example of the latter will be (not necessarily quasiconvex) finitely generated subgroups of hyperbolic groups. The lower bounds are obtained from the counting quasimorphism construction, described in §4.2; however, the argument is complicated by the fact that a random walk in a hyperbolic group (or in a hyperbolic graph) is almost certainly not quasigeodesic.

The abstract template for obtaining lower bounds is the following somewhat technical proposition, which is basically just a restatement of some properties of small counting quasimorphisms.

PROPOSITION 5.6. *Let* G *be a group acting by isometries on a* δ*-hyperbolic simplicial complex* Y *(*not *assumed to be locally finite) with a basepoint* p*. Let* $g \in G$ *be given, and let* $\gamma(g)$ *(or* γ *for short) be a geodesic in* Y *from* p *to* $g(p)$*. We set* $n = \mathrm{length}(\gamma)$*. Fix* $C_1 > 0$*, and let* γ_i *be the successive (non-overlapping) subpaths of* γ *of length* $C_1 \log n$*. Let* $K(\delta)$ *(or* K *for short) be such that any* $(2, 4)$*-quasigeodesic in a* δ*-hyperbolic space stays within distance* K *of a genuine geodesic. Suppose that there is an* $\epsilon > 0$ *for which the following inequality holds:*

$$\#\{i : \gamma_i \ K \text{ anti-aligns some } \gamma' \subset \gamma\} < n^{1-\epsilon}.$$

Then there are constants $C_2 > 0$ *and* C_3 *depending only on* C_1*,* ϵ *and* δ*, and* C_4 *depending only on the action, and a homogeneous quasimorphism* ϕ *on* G *satisfying the following properties:*

(1) $\phi(g) \geq nC_2/\log n$*;*
(2) $D(\phi) \leq C_3$*;*
(3) $|\phi(h)| \leq 2|h|_S C_4/\log n$ *for any* $h \in G$.

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

Proof. Let K be such that any $(2, 4)$-quasigeodesic in a δ-hyperbolic space stays within distance K of a genuine geodesic. Choose K with this property, and let Σ be the set of γ_i that do not K anti-align some $\gamma' \subset \gamma$. Let c_Σ denote the small counting function associated to the set Σ, and similarly $c_{\Sigma^{-1}}$. Then by Lemma 4.3 and the definition of Σ we have $c_\Sigma(g) \geq nC_2/\log n$ and $c_{\Sigma^{-1}}(g) = 0$. Then let ϕ be the homogenization of h_Σ.

To obtain the last bullet point, observe that $d(p, h(p)) \leq C_5|h|_S$ for any $h \in G$, for some constant C_5. □

Obtaining lower bounds on scl for random walks thus reduces to showing that for certain G, the condition in Proposition 5.6 holds with high probability for g the result of a random walk of length n/L (where L is the drift). The main technical issue is to relate subwalks of a random walk with subpaths of the geodesic joining the endpoints. We address this point in what follows.

5.4. *Non-degenerate random walks.* We fix a group G and probability measure μ supported on a finite subset S of G so that $S = S^{-1}$, and $\mu(s) = \mu(s^{-1})$ for each $s \in S$ (such a measure is said to be *symmetric*). We also fix a simplicial action of G on a δ-hyperbolic simplicial complex Y (not assumed to be locally finite) with a basepoint p.

The subset S generates some subgroup $\langle S \rangle$ of G; we say μ is *non-elementary* if $\langle S \rangle$ does not fix any finite subset of $\partial_\infty Y$.

A random sequence $\mathrm{id} = g_0, g_1, g_2, \ldots$ (finite or infinite) is a *μ-random walk in G* if the successive differences $s_i := g_{i-1}^{-1} g_i$ are independent random elements of S each with the distribution μ. For such a walk, define the associated *μ-random walk in Y*, namely the sequence $p = p_0, p_1, p_2, \ldots$, where $p_i = g_i p$. Since S is finite, the set of distances $d_Y(p_i, p_{i+1})$ is uniformly bounded; we say such a walk has *bounded increments*.

For any measure μ of finite first moment on any group acting isometrically on any metric space, Kingman's subadditive ergodic theorem implies that there is a constant $L \geq 0$ called the *drift* so that $L = \lim_{n\to\infty} d(p, p_n)/n$ almost surely.

Recall that for any points p, q in a hyperbolic space Y and any constant K, the *shadow* $S_p(q, K)$ is defined to be the set of all points $r \in Y$ so that every geodesic from p to r comes within distance K of q. This may be expressed equivalently (up to slightly adjusting the constant R) in terms of the 'Gromov product' $(x \cdot y)_p := 1/2(d_Y(p, x) + d_Y(p, y) - d_Y(x, y))$ as the set of points with $(q \cdot r)_p \geq d_Y(p, q) - R$.

We will make the following assumptions about G, Y, μ.

Definition 5.7. Let G be a group acting simplicially on a δ-hyperbolic simplicial complex Y, and let μ be a symmetric probability measure with support equal to some set $S \subset G$. A triple (G, μ, Y) as above is *non-degenerate* if it satisfies the following conditions:

(1) *(non-elementary:)* $\langle S \rangle$ does not fix any finite subset of $\partial_\infty Y$;
(2) *(positive drift:)* the drift L is positive;
(3) *(acylindricity:)* for any G-orbit $Gp \subset Y$ and for any $C_1 > 0$ there are constants C_2, C_3 so that if q, r are points in Gp with $d_Y(q, r) \geq C_2$, there are at most C_3 elements $g \in G$ with $d_Y(q, gq) \leq C_1$ and $d_Y(r, gr) \leq C_1$;
(4) *(linear progress:)* there are constants $C_1 > 1$, $C_2 > 0$ so that

$$\mathbf{P}(d_Y(p_0, p_n) \in [LC_1^{-1}n, LC_1 n]) \geq 1 - e^{-nC_2}.$$

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

(5) *(exponential decay:)* there are constants C_1 and $C_2 > 0$ such that for any $y \in Y$ and any K, the probability that the result of a random walk of length n lies in the shadow $S_{p_0}(y, K)$ decays exponentially in the distance to the shadow, i.e.

$$\mathbf{P}(p_n \in S_{p_0}(y, K)) \leq C_1 e^{-C_2(d_Y(p_0,y)-K)}.$$

Remark 5.8. We do not claim that every condition in this list is logically necessary; rather it reflects the ingredients that go into our proof of Theorem 5.27. It is natural to wonder whether the condition of acylindricity could be replaced by Bestvina and Fujiwara's *weakly properly discontinuous* condition (see [**5**, p. 76]), since the latter condition is known to hold for a wider class of group actions; but our arguments do not seem to easily allow it.

Remark 5.9. The condition that we call 'acylindricity' in Definition 5.7 is weaker than what is usually called acylindricity for an action, in that the constants C_2, C_3 are allowed to depend on both C_1 and the choice of G-orbit G_p. If a distinction needs to be made, we refer to our weaker notion as *orbit acylindricity*. If the action of G on Y is cocompact, both notions of acylindricity are equivalent.

The following proposition is largely obtained by assembling known results.

PROPOSITION 5.10. *In each of the following cases, (G, μ, Y) is non-degenerate:*

(1) *G is a hyperbolic group, Y is the Cayley graph of G with respect to some finite generating set, and $\langle S \rangle$ is non-elementary;*
(2) *G is a (strongly) relatively hyperbolic group, Y is the Groves–Manning space associated to G, and $\langle S \rangle$ is non-elementary;*
(3) *for some surface Σ, the group G is the mapping class group, Y is the complex of curves, and $\langle S \rangle$ is not reducible or virtually abelian.*

Proof. The non-elementary axiom follows by hypothesis in all three cases. Exponential decay in case (3) follows from Maher [**40**], and again the arguments go through verbatim in the setting of a non-elementary action of a group on a proper Gromov hyperbolic space, so cases (1) and (2) also follow from this, though presumably the result is standard in case (1). We now verify the other properties.

In case (1), the Cayley graph Y is a proper Gromov hyperbolic space, and acylindricity follows from the fact that the action of G on itself is properly discontinuous. Positive drift holds in cases (1) by Kaimanovich [**33**, Theorem 7.3], and linear progress follows from Kesten's estimate [**49**, Lemma 8.1b] for a random walk on a non-amenable group.

In case (2), given a (strongly) relatively hyperbolic group G, Groves and Manning [**28**] construct a *proper* Gromov hyperbolic space Y (called the *cusped space*—see [**28**, §3]) on which G acts by isometries properly discontinuously, but not cocompactly; and orbit acylindricity follows from the fact that the space Y is proper, and the action of G on Y is properly discontinuous. Positive drift holds by Kaimanovich [**33**, Theorem 7.3], and linear progress follows from Maher [**40**]. Although the results of [**40**] are stated in terms of the action of the mapping class group on the complex of curves, the results hold in the (simpler) case of an action of a non-elementary group on a proper Gromov hyperbolic space Y.

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

In case (3), the complex of curves is a locally infinite Gromov hyperbolic simplicial complex on which the mapping class group acts discontinuously by simplicial isometries, and acylindricity is a theorem of Bowditch [**7**, Theorem 1.3]. Positive drift and linear progress follow from Maher [**38**, Theorem 1.1]. □

5.5. *Proximal points and unfolded walks.* Since S is finite, the length of successive steps $d(p_i, p_{i+1})$ is uniformly bounded by a constant, and therefore we can think of the random walk as a (coarse) path in Y of length $\leq C_1 n$. We would like to use this estimate to show that with very high probability, 'most' of the p_i are within a bounded distance of the geodesic from p_0 to p_n. Actually, it turns out to be easier (and just as useful) to show that most points on the geodesic from p_0 to p_n are within a bounded distance of some p_i, and moreover this fact can be deduced directly from linear progress (see Definition 5.7) and elementary hyperbolic geometry.

Definition 5.11. Let $p_0, p_1, \ldots, p_n$ be a walk on a δ-hyperbolic space Y. If for all i there is an inequality $d(p_i, p_{i+1}) \leq C_1$ we say p_i has C_1-*bounded increments*.

Definition 5.12. Let $p_0, p_1, \ldots, p_n$ be a walk on a δ-hyperbolic space Y, and let γ be a geodesic from p_0 to p_n. A point $q \in \gamma$ is K-*proximal* if $d_Y(q, p_i) \leq K$ for some p_i. We denote the K-proximal subset of γ by γ_K.

LEMMA 5.13. *Let $p_0, p_1, \ldots, p_n$ be a walk on a δ-hyperbolic space Y with C_1-bounded increments. Let γ be a geodesic from p_0 to p_n. Suppose* $\text{length}(\gamma) \geq C_2 n$. *Then for any $\epsilon > 0$ there is a constant $K(\epsilon, C_1, C_2)$ so that if γ_K denotes the K-proximal subset of γ, there is an estimate*

$$\text{length}(\gamma_K)/\text{length}(\gamma) \geq 1 - \epsilon.$$

Proof. We assume $K \gg C \gg \delta$ for convenience.

There is a nearest point projection π from the p_i to γ so that $d(\pi(p_i), \pi(p_{i+1})) \leq C + O(\delta)$ and consequently every point in γ is within distance $C/2 + O(\delta)$ from some $\pi(p_i)$. If $p \in \gamma$ is not K-proximal, then $d(p_i, \gamma) \geq K - O(C)$ and therefore $d(p_j, \gamma) \geq 3K/4 - O(C)$ for $|j - i| \leq K/4C$. On the other hand, $d(p_i, p_j) \leq |j - i|C$ so any geodesic from p_j to p_i does not come within distance $K/2$ of γ. By δ-thinness, we can conclude that $d(\pi(p_j), \pi(p_i)) = O(\delta)$ for $|j - i| \leq K/4C$.

If the set of non K-proximal points has length at least $\epsilon \cdot \text{length}(\gamma)$, there are at least $\epsilon \cdot \text{length}(\gamma)/O(\delta)$ such points whose mutual pairwise distances is at least $O(\delta)$. To each such point we can associate a sequence of $K/2C$ points p_i whose projections to γ are within $O(\delta)$ of it, and therefore these collections of points are *disjoint*. The total number of p_i in these collections is at least $(K/2C)\epsilon C^{-1} n/O(\delta) = nK\epsilon C^{-2}/O(\delta)$, so if $K > \epsilon^{-1} C^2 O(\delta)$ we get a contradiction, as desired. □

From the linear progress axiom, we deduce the following.

LEMMA 5.14. (Proximal) *Let (G, μ, Y) be non-degenerate. Let $p_0, \ldots, p_n$ be a random walk, and γ a geodesic from p_0 to p_n. There are constants $C_1, C_2 > 0$ so that for any $\epsilon > 0$ there is a further constant $K(C_1, \epsilon)$ so that*

$$\mathbf{P}((\text{length}(\gamma) \geq C_1 n) \wedge (\text{length}(\gamma_K)/\text{length}(\gamma) \geq 1 - \epsilon)) \geq 1 - e^{-nC_2}.$$

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

Note that the first condition implies the second by Lemma 5.13, but it is convenient to state both conditions explicitly.

Definition 5.15. We say γ as above is C_1-*long* (or just *long* if C_1 is understood) if $\text{length}(\gamma) \geq C_1 n$.

In the following, we use the convention that oriented geodesics are parameterized proportional to arclength. We also use the convention that γ is oriented from p_0 to p_n, so we write $\gamma(0) = p_0$ and $\gamma(1) = p_n$.

Now, by definition, for every s, t with $\gamma(s), \gamma(t) \in \gamma_K$ there are indices i, j so that $d_Y(p_i, \gamma(s)) \leq K$ and $d_Y(p_j, \gamma(t)) \leq K$. It is not necessarily true, however, that $s < t$ implies $i < j$. Nevertheless, this should be true whenever $d_Y(\gamma(s), \gamma(t))$ is sufficiently large, with big probability. We quantify this.

Definition 5.16. Fix some big constant $M \gg 0$. We say a walk $p_0, \ldots, p_n$ is $(K, M \log n)$-*unfolded* (or just *unfolded* if K and M are understood) if for every $j > i$ with $j - i > M \log n$ and every geodesic γ_i from p_0 to p_i, we have $d_Y(p_j, \gamma_i) > K$.

LEMMA 5.17. (Unfolded) *Let (G, μ, Y) be non-degenerate. Let $p_0, \ldots, p_n$ be a random walk. For any K there is $C_1 > 0$ so that*

$$\mathbf{P}(\textit{walk is } (K, M \log n)\textit{-unfolded}) \geq 1 - n^{2-C_1 M}.$$

Proof. Pick indices i, j so that $j > i$ and $j - i > M \log n$. By the linear progress property, the walk $p_i, p_{i+1}, \ldots, p_j$ satisfies $d_Y(p_i, p_j) > C_1 M \log n$ with probability $1 - n^{-C_2 M}$ for suitable C_1, C_2. The subset of the K-neighborhood of γ_i which lies outside of $B(p_i, C_1 M \log n)$ is contained in a shadow $S_{p_i}(y, K + C_3)$, where y is the point distance $C_1 M \log n$ from p_i along γ_i, and C_3 only depends on δ, the constant of hyperbolicity. By linear progress and the exponential decay property for shadows, the probability that p_j is within distance K of γ_i is at most $n^{-C_4 M}$, for some constant C_4. Hence

$$\mathbf{P}((d_Y(p_i, p_j) > C_1 M \log n) \wedge (d_Y(p_j, \gamma_i') > K)) > 1 - n^{-C_5 M}.$$

There are fewer than n^2 indices i, j as above; the lemma follows. □

Remark 5.18. The condition that a walk should be $(K, M \log n)$-unfolded is very strong, probably much stronger than we need. But it does simplify the proofs to come. If we just insist that *most* pairs of indices i, j with $j - i > M \log n$ are 'unfolded' in the obvious sense, then the probability will be at least $1 - e^{-nC_1}$, for sufficiently big (but fixed) M. We do not use this fact in the following.

5.6. *Anti-aligned segments and matching.* Let $p_0, p_1, \ldots, p_n$ as above be a walk in Y with bounded increments, and γ the geodesic from p_0 to p_n. We assume γ is long, and let $N = \text{length}(\gamma)$. We fix some constants $M \gg 0$ and $K \gg 0$ (to be determined later).

Definition 5.19. An R-*match* (or just a *match* if R is understood) is a triple (α, β, h), where α, β are geodesic segments of γ of length R, and $h \in G$ so that $h\alpha$ *anti-aligns* β, i.e. $d(h\alpha(0), \beta(1)) \leq K$ and $d(h\alpha(1), \beta(0)) \leq K$.

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

Ultimately we will be concerned with R-matches where $R = M \log n$. It is convenient for our R-matches not to be too close to each other; that can be achieved by the following 'cut in half' lemma.

LEMMA 5.20. (Cut in half) *Let* (α, β, h) *be an* R*-match, and let* $\gamma_K \subset \gamma$ *so that* $\text{length}(\gamma_K \cap \alpha) \geq (1-\epsilon)R$ *and* $\text{length}(\gamma_K \cap \beta) \geq (1-\epsilon)R$. *Then there are subsegments* $\alpha' \subset \alpha$ *and* $\beta' \subset \beta$ *with endpoints in* $\gamma_K \cap \alpha$ *and* $\gamma_K \cap \beta$ *respectively, so that* (α', β', h) *is an* R'*-match and* $d_Y(\alpha', \beta') \geq R/2$, *for some* $R'/R \in [1/4 - 4\epsilon, 1/4 + 4\epsilon]$. *Furthermore,* $\text{length}(\gamma_K \cap \alpha') \geq (1-5\epsilon)R'$ *and similarly for* β'.

Proof. If ν is an oriented geodesic, let ν^- denote the initial half of ν, and ν^+ the terminal half. One of the pairs (α^+, β^-) and (α^-, β^+) must be disjoint, and therefore one of the pairs $(\alpha^{++}, \beta^{--}), \ldots, (\alpha^{--}, \beta^{++})$ must have segments separated from each other by distance at least $R/2$. Each of these segments has length $R/4$, so the length of their intersections with γ_K are both at least $R/4 - \epsilon R$, and there must be further subsets in each of length at least $R/4 - 2\epsilon R$ matched by h. Let α', β' be maximal subsegments with endpoints in these subsets. □

Definition 5.21. Let γ be a geodesic in Y of length N. Let α_i be successive subpaths of γ of length $M \log N$, so that there are $N/M \log N$ of them. We say γ is *well-matched* if for *every* subset I of indices with $|I| > N/10M \log N$ there are at least $9|I|/10$ disjoint geodesics β_j in γ and elements $h_j \in G$ and indices $i(j) \in I$ so that $(\alpha_{i(j)}, \beta_j, h_j)$ is a match.

Note that the definition of a match and well-matched implicitly depend on K and M. If we need to specify them we use the terminology $(K, M \log N)$ *well-matched*.

LEMMA 5.22. *Let* G *act simplicially on a* δ*-hyperbolic simplicial space* Y. *Let* p *be a basepoint, and let* γ *be a geodesic from* p *to* $g(p)$ *for some* $g \in G$. *Then there are constants* K *and* C_1 *depending only on* δ, *so that if* γ *is not* $(K, M \log N)$ *well-matched, there is a counting quasimorphism* ϕ *on* G, *supported on words of length* $M \log N$, *with* $D(\phi) \leq C_1$ *and* $\phi(g) \geq N/(100M \log N)$.

Proof. Choose a set of indices I as above, witnessing that γ fails to be K well-matched. Then no realizing path for g can contain more than $9|I|/10$ disjoint anti-aligned copies of K. Define c_I to be the counting function which counts *disjoint* copies of any of the α_i with $i \in I$ in a realizing path. We can take γ itself as a realizing path for c_I, so that $c_I(g) \geq |I|$. On the other hand, any realizing path for g^{-1} for c_I contains at most $9|I|/10$ disjoint copies of α_i, so $\phi(g) := c_I(g) - c_I(g^{-1}) \geq |I|/10$. □

Remark 5.23. We have implicitly used the fact that disjoint copies of α_i^{-1} on any realizing path for g^{-1} are close to geodesic segments of γ which overlap only in segments of length $\leq 2K + O(\delta)$; trimming these overlaps, we can assume the nearby segments in γ are disjoint.

It remains to understand when γ associated to a random walk is $(K, M \log N)$ well-matched, for suitable constants K, M.

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

The next lemma shows that any well-matching can be 'relativized' to any subset γ_K with $\text{length}(\gamma_K)/\text{length}(\gamma) \geq 1 - \epsilon$.

LEMMA 5.24. *Let γ be a geodesic of length N, and let β_j be a collection of at least $9N/(100M \log N)$ disjoint subpaths, each of length $M \log N$. Let γ_K be a subset of γ with $\text{length}(\gamma_K)/\text{length}(\gamma) \geq 1 - \epsilon$. Then at least $(1 - \sqrt{\epsilon})$ of the β_j satisfy*

$$\text{length}(\gamma_K \cap \beta_j)/\text{length}(\beta_j) \geq 1 - 12\sqrt{\epsilon}.$$

Proof. This is essentially just Chebyshev's inequality. The total length of the β_j is at least $9N/100$. For each β_j which fails to satisfy the desired inequality, $\text{length}(\gamma_K^C \cap \beta_j) \geq 12\sqrt{\epsilon} \cdot \text{length}(\beta_j)$ and therefore

$$\epsilon N \geq \text{length}(\gamma_K^C) \geq 12\sqrt{\epsilon}\delta 9N/100,$$

where δ is the proportion of 'failing' β_j. We conclude $\delta \leq \sqrt{\epsilon}$, as claimed. □

Combining Lemma 5.24 with Lemma 5.20 we deduce the following.

PROPOSITION 5.25. (Relative well-matching) *Let γ be a geodesic in Y of length N. Suppose γ is $(K, M \log N)$ well-matched, and let γ_K be a subset of γ with $\text{length}(\gamma_K)/\text{length}(\gamma) \geq 1 - \epsilon$.*

If ϵ is sufficiently small, there are at least $8N/(10M \log N)$ segments α_j in γ, and $8N/(10M \log N)$ disjoint *segments β_j in γ, each of length at least $M \log N/5$, so that for each β_j there is a match (α_j, β_j, h_j) satisfying:*

(1) *the endpoints of β_j are in γ_K and similarly for $\alpha_{i(j)}$;*
(2) *there is an inequality $\text{length}(\gamma_K \cap \beta_j)/\text{length}(\beta_j) \geq 1 - 100\sqrt{\epsilon}$ and similarly for $\alpha_{i(j)}$; and*
(3) *$d_Y(\beta_j, \alpha_{i(j)}) \geq M \log N/2$.*

Proof. Of all the successive segments α_i of γ of length $M \log N$, let I' be the set of indices for which $\text{length}(\gamma_K \cap \alpha_i)/\text{length}(\alpha_i) \geq 1 - 12\sqrt{\epsilon}$. Then $|I'| \geq (1 - \sqrt{\epsilon})N/(M \log N)$ by Lemma 5.24. Since γ is well-matched, we can find at least $9|I'|/10$ disjoint geodesics β_j in γ and matches $(\alpha_{i(j)}, \beta_j, h_j)$ with $i(j) \in I'$. Applying Lemma 5.24 again, there is a subset $I \subset I'$ with $|I| \geq (1 - \sqrt{\epsilon})^2 N/(M \log N)$ so that, further, $\text{length}(\gamma_K \cap \beta_j)/\text{length}(\beta_j) \geq 1 - 12\sqrt{\epsilon}$.

For each β_j, choose some $\alpha_{i(j)}$ as above, and relabel it as α_j. Note that although the β_j are distinct for different j, we do *not* assume the α_j are distinct for different j (for all we know, they might all be the same!). Applying Lemma 5.20 to each (α_j, β_j, h_j) gives rise to new $(\alpha'_j, \beta'_j, h_j)$ of at least $1/4 - 100\sqrt{\epsilon}$ times the length, satisfying the desired properties. Notice now that even if $\alpha_i = \alpha_j$ we might have $\alpha'_i \neq \alpha'_j$. □

5.7. *Uniform lower bounds for hyperbolic spaces.* We are now ready to return to probability.

LEMMA 5.26. *Let (G, μ, Y) be non-degenerate, and let $p_0, \ldots, p_n$ be a random walk of length n. Fix C_1 and K. Then there is a constant C_2 so that for any M the following holds. Consider the collection of indices $a < a' < b < c < c' < d$ for which there are geodesics α from p_a to p_b and β from p_c to p_d with the following properties:*

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

(1) $\text{length}(\alpha) \geq MC_1 \log n$ *and similarly for* β*;*
(2) *there is* $t \in [0.1, 0.2]$ *so that* $d_Y(p_{a'}, \alpha(1-t)) \leq K$ *and* $d_Y(p_{c'}, \beta(t)) \leq K$*;*
(3) *there is some* $h \in G$ *so that* $d_Y(h\alpha(0), \beta(1)) \leq K$ *and* $d_Y(h\alpha(1), \beta(0)) \leq K$.
The probability that this collection of indices is non-empty is at most n^{6-C_2M}.

Proof. Fix a collection of indices $a < a' < b < c < c' < d$. The key point is that we don't care about α and β *per se*, but only their G-orbits $G\alpha$ and $G\beta$, since β is only compared with a translate $h\alpha$. Speaking somewhat loosely, since $b < c$, the random variables $G\alpha$ and $G\beta$ are *independent*. In fact, if we let β'' be a geodesic from p_c to $p_{c'}$, and β' a geodesic from $p_{c'}$ to p_d, then the orbit $G\beta'$ is independent of $G\alpha$ and $G\beta''$.

By orbit acylindricity (this is the only place in the entire argument that acylindricity is used), there are only *boundedly* many $h \in G$ with $d(h\alpha(1), \beta(0)) \leq K$ and $d(h\alpha(1-t), \beta(t)) \leq 3K$, providing n is bigger than some universal constant. So we can suppose that we are given a *fixed* geodesic α' of length $\geq (1-t)MC_1 \log n$ starting within distance $3K$ of $p_{c'}$, and we want to estimate the probability that a random walk from $p_{c'}$ to p_d ends up within distance K of the endpoint $h\alpha(0)$. This can be bounded from above by the probability that a random walk started at $p_{c'}$ lies in the finite collection of shadows $S_{p_{c'}}(h\alpha(0), K + C_2)$, for some constant C_2 which only depends on δ, the constant of hyperbolicity. By the exponential decay property for shadows, this probability is at most n^{-C_3M} for some C_3.

Summing over all possible choices of indices gives probability at most n^{6-C_3M}. □

We deduce our main theorem on non-degenerate random walks on hyperbolic spaces.

THEOREM 5.27. (Hyperbolic lower bound) *Let* Y *be a* δ*-hyperbolic simplicial complex (not assumed to be locally finite), and let* G *be a finitely generated group that acts simplicially on* Y*. Let* μ *be a symmetric probability measure of finite support on* G *so that* (G, μ, Y) *is non-degenerate, in the sense of Definition 5.7.*

Let g *be obtained by random walk on* G *(with respect to* μ*) of length* n*. Then for any* C_1 *there is a* $C_2 > 0$ *and* C_3 *depending only on* δ *and* G*, so that with probability at least* $1 - n^{-C_1}$ *there is a homogeneous quasimorphism* ϕ *on* G *satisfying the following properties:*
(1) $\phi(g) \geq nC_2/\log n$*;*
(2) $D(\phi) \leq C_3$*;*
(3) $|\phi(h)| \leq 2d_Y(q, hq)C_3/\log n$ *for any* $g \in G$ *and any* $q \in Y$.
In particular, for any $C_1 > 0$ *there is a constant* $C > 1$ *so that if we condition on* $g \in [G, G]$ *(for* n *even), then*

$$\mathbf{P}(C^{-1}n/\log n \leq \text{scl}(g) \leq Cn/\log n) \geq 1 - n^{-C_1}.$$

Proof. We use the notation K, M as above.

Let $p_0, \ldots, p_n$ be the associated walk from the basepoint p to gp, and let γ be the geodesic from p to gp. We assume that γ is long (see Definition 5.15), that the walk is unfolded (see Definition 5.16), and that the conclusions of Lemma 5.26 hold. By Lemmas 5.14, 5.17 and 5.26, the probability of all three of these things happening is at least $1 - n^{6-C_1M}$ for some C_1, where M is as above.

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

Since γ is long, $\text{length}(\gamma) = Cn$ for some constant C bounded above and below. If γ is not $(K, M \log(Cn))$ well-matched, the desired conclusion follows from Lemma 5.22, so we suppose γ is well-matched.

By Proposition 5.25 we can find a pair of segments (in fact, many such pairs) α, β of length $M \log(Cn)/5$, with endpoints in γ_K with $\text{length}(\gamma_K \cap \alpha)/\text{length}(\alpha) \geq 1 - 100\sqrt{\epsilon}$ (and similarly for β), and which are distance at least $\geq M \log(Cn)/2$ apart in γ. Pick $t \in [0.1, 0.2]$ with $\alpha(1-t)$ and $\beta(t)$ in γ_K. Since γ is unfolded, the points $\alpha(0)$, $\alpha(1-t)$, $\alpha(1)$, $\beta(0)$, $\beta(t)$, $\beta(1)$ are within distance K of points p_a, $p_{a'}$, p_b, p_c, $p_{c'}$, p_d on the walk so that the indices appear in the same order that the corresponding points appear in γ.

But the conclusion of Lemma 5.26 says that no such pairs α, β can exist, so we get a contradiction. It follows that γ is not well-matched after all, and the desired quasimorphism exists. □

From Proposition 5.10 we obtain the following corollaries.

COROLLARY 5.28. *Let G be a hyperbolic group (respectively the mapping class group of a surface Σ), and let μ be a symmetric probability measure on G with finite support generating a non-elementary subgroup (respectively a subgroup which is not reducible or virtually abelian). Then for any C_1 there is C so that if g is obtained by random walk in G of length n (even) and conditioned to lie in $[G, G]$ we have*

$$\mathbf{P}(C^{-1} n/\log n \leq \text{scl}(g) \leq Cn/\log n) \geq 1 - n^{-C_1}.$$

Since reducible elements in mapping class groups fix points in the complex of curves, we obtain the following corollary.

COROLLARY 5.29. *Let μ be a symmetric probability measure on the mapping class group of Σ of finite support, and suppose the subgroup it generates is not reducible or virtually abelian. Then for any C_1 there is C so that if g is obtained by random walk in G of length n, then g cannot be expressed as a product of fewer than $Cn/\log n$ reducible elements, with probability at least $1 - n^{-C_1}$.*

From monotonicity of scl under homomorphisms, we obtain the following corollary.

COROLLARY 5.30. *Let G be a group and μ a symmetric probability measure with finite support which generates G. Suppose that for any $\epsilon > 0$ there is a $\delta > 0$ so that if g is obtained by random walk on G of length n (even) conditioned to lie in $[G, G]$, we have*

$$\mathbf{P}(\text{scl}(g) < \epsilon \cdot n/\log n) > \delta.$$

Then every homomorphism from G to a hyperbolic group or to a mapping class group has virtually abelian image.

Proof. This just depends on the observation that reducible subgroups of mapping class groups are themselves mapping class groups of simpler surfaces. □

5.8. *Geometry of the scl norm of a random subspace.* As remarked earlier, stable commutator length is not merely a function, but actually defines a pseudo-norm on the space $B_1(G)$ of (real) group 1-boundaries, i.e. formal real linear combinations $\sum t_i g_i$

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

with $t_i \in \mathbb{R}$ and $g_i \in G$, representing 0 in $H_1(G; \mathbb{R})$. In fact, scl descends to a pseudo-norm on the 'homogenized' quotient $B_1^H(G) := B_1(G)/\langle g - hgh^{-1}, g^n - ng\rangle$—see [**11**, §2.6] for details.

A nice corollary of Theorem 5.27 is to obtain *a priori* geometric control over the geometry of a random subspace of $B_1^H(G)$ of fixed dimension.

THEOREM 5.31. (Random norm ball theorem) *Let Y be a δ-hyperbolic simplicial complex (not assumed to be locally finite), and let G be a finitely generated group that acts simplicially on Y. Let μ be a symmetric probability measure of finite support on G so that (G, μ, Y) is non-degenerate, in the sense of Definition 5.7.*

Fix k and let $g_1, g_2, \ldots, g_k$ be obtained by random walk on G (with respect to μ) of (even) lengths $\ell_1 n, \ell_2 n, \ldots, \ell_k n$, all conditioned to lie in $[G, G]$. Then for any C_1 there is a $C_2 > 0$ and C_3 depending only on δ and G, so that with probability at least $1 - n^{-C_1}$, for any formal sum $\sum t_i g_i$ there is an estimate

$$\mathbf{P}\left(C^{-1}\left(\sum_i t_i \ell_i n\right)\Big/ \log n \leq \mathrm{scl}\left(\sum t_i g_i\right) \leq C\left(\sum_i t_i \ell_i n\right)\Big/ \log n\right) \geq 1 - n^{-C_1}.$$

Proof. We give the argument in the case $k = 2$ and $\ell_1 = \ell_2 = 1$. The general case follows by a minor modification of the argument.

We know that $\mathrm{scl}(g_1)$ and $\mathrm{scl}(g_2)$ are between $C^{-1}n/\log n$ and $Cn/\log n$ with high probability. By the definition of a norm, $\mathrm{scl}(g_1 + g_2) \leq C2n/\log n$. By the argument of Theorem 5.27, again with high probability we can find subsets I_i $(i = 1, 2)$ of the indices associated to the quasimorphisms ϕ_i certifying the lower bound for $\mathrm{scl}(g_i)$ so that $|I_i|/|I| \geq (1 - \epsilon)$ and such that if $\phi_{1,2}$ is the small counting quasimorphism associated to the union of the I_i segments, then $\phi_{1,2}(g_i) \geq (1 - \epsilon)\phi_i(g_i)$. Note that $D(\phi_{1,2}) \leq C_3$ for the same constant as in Theorem 5.27, since this constant is universal for *any* small counting quasimorphism. Moreover, $\phi_{1,2}(t_1 g_1 + t_2 g_2) = t_1\phi_{1,2}(g_1) + t_2\phi_{1,2}(g_2) \geq (t_1 + t_2)C^{-1}n/\log n$. This gives the desired estimate in the orthant where the t_i are both positive. In an orthant in which t_i is negative, use g_i^{-1} in place of g_i together with the observation that $t_i g_i = -t_i g_i^{-1}$ in B_1^H. □

5.9. *Concentration versus compression.* It is natural, in view of Theorem 5.1, to believe that the estimates in Theorems 5.27 and 5.31 can be sharpened. We make the following conjecture.

CONJECTURE 5.32. (Concentration conjecture) *Let (G, μ, Y) be non-degenerate in the sense of Definition 5.7. Then there is a constant $C > 0$ and $C_1 > 0$ so that for any $\epsilon > 0$, if g is obtained by random walk on G of length n (even), conditioned to lie in $[G, G]$, then*

$$\mathbf{P}(|\mathrm{scl}(g)\log n/n - C| \leq \epsilon) \geq 1 - e^{-n^{C_1}}.$$

Conjecture 5.32 would imply the following conjecture about the geometry of the norm in a random subspace.

CONJECTURE 5.33. (Norm conjecture) *Let (G, μ, Y) be non-degenerate in the sense of Definition 5.7. Then there is a constant $C > 0$ and $C_1 > 0$ so that for any finite integer k*

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

and any $\epsilon > 0$, if $g_1, \ldots, g_k$ are obtained by independent random walks on G of length $\ell_1 n, \ldots, \ell_k n$ (even), conditioned to lie in $[G, G]$, then

$$\mathbf{P}\left(\left|\mathrm{scl}\left(\sum t_i g_i\right) \log n \Big/ n - C \sum t_i \ell_i\right| \leq \epsilon\right) \geq 1 - e^{-n^{C_1}}.$$

It is tempting to conjecture further that $C = \lambda/6$, where λ is the entropy of μ, but this might be premature without first understanding a wider range of examples.

5.10. *Random walks on Out(F_n).* In this section we summarize our current knowledge of the behavior of random walks on $\mathrm{Out}(F_n)$. In this context, it is natural to consider random walks on the isometry group of a non-locally compact Gromov hyperbolic simplicial complex. In this section we show that if μ is a probability distribution with finite support on the isometry group of a non-locally compact Gromov hyperbolic simplicial complex Y, whose support generates a discrete non-elementary subgroup, then all of the axioms from Definition 5.7 are satisfied, except possibly for acylindricity. A key step is to show that random walks converge to the boundary.

THEOREM 5.34. *Let μ be a probability measure with finite support on the isometry group of a (not-necessarily proper) Gromov hyperbolic simplicial complex Y with basepoint y_0, whose support generates a discrete non-elementary subgroup. Then almost every sample path $\{g_n y_0\}$ converges to the Gromov boundary ∂Y, and the hitting measure ν is non-atomic, and is the unique μ-stationary measure on ∂Y.*

The arguments from [**38**] and [**40**] then go through in this case, to give linear progress, positive drift and exponential decay. We can then show the following theorem.

THEOREM 5.35. (Linear translation length) *Let G be a group of isometries of a simplicial Gromov hyperbolic space Y, which is not necessarily locally compact. Let μ be a probability distribution with finite support on G, such that the support of μ generates a non-elementary subgroup of G. Then there are constants $L > 0$ and $c < 1$ such that*

$$\mathbf{P}(\tau(w_n) \leq Ln) \leq O(c^n),$$

where w_n is the group element obtained by a random walk of length n, and $\tau(w_n)$ is the translation length of w_n acting on Y.

The group $\mathrm{Out}(F_n)$ acts on a (locally infinite) simplicial complex, the complex of free factors, which Bestvina and Feighn [**3**] have shown is Gromov hyperbolic, so the results above apply in this case. If the action of $\mathrm{Out}(F_n)$ on the complex of free factors is acylindrical, then this would imply that the growth rate of scl for random walks in $\mathrm{Out}(F_n)$ is $n/\log n$. However, the acylindricity of the action of $\mathrm{Out}(F_n)$ is still unknown. If an element of $\mathrm{Out}(F_n)$ acts on the complex of free factors with positive translation length, then it is fully irreducible with fully irreducible powers (some use the terminology *iwip*) and so this generalizes results of Rivin [**43**] and Kowalski [**36**]. Finally, we remark that $\mathrm{Out}(F_n)$ also acts on the complex of free splittings, which Handel and Mosher [**31**] have shown is Gromov hyperbolic, but in this case the action is definitely not acylindrical.

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

Remark 5.36. In a recent preprint, Sisto [**46**] introduces the notion of a *weakly contracting* element in a group (with respect to a so-called *path system*), and shows (for example) that an element of $\mathrm{Out}(F_n)$ obtained by (suitable) random walk will be weakly contracting (and therefore iwip) with probability going to 1 exponentially fast with the length of the walk. Sisto proves many other interesting results in his preprint, with methods that do not seem to overlap much with ours.

Remark 5.37. Since this paper was first posted, more facts have been established about the action of $\mathrm{Out}(F_n)$ on various hyperbolic complexes. We refer the reader to [**4, 31**] for details. Theorem 5.35 applies to all known actions. However, since in no case the action is known to be acylindrical, Theorem 5.27 does not apply (as far as we know).

We now prove Theorem 5.34. Kaimanovich [**33**] showed that a random walk on a non-elementary subgroup of isometries of a proper Gromov hyperbolic space Y converges to the boundary ∂Y almost surely. If Y is proper, then $Y \cup \partial Y$ is compact. However, the only place where this is used is to show that there is a weak limit of the convolution measures μ^{*n}, which is a probability measure on $Y \cup \partial Y$. Our initial task is therefore to show that for a random walk on a non-elementary subgroup of isometries of a non-proper Gromov hyperbolic space, the convolution measures μ^{*n} converge to a μ-invariant probability measure supported on the Gromov boundary ∂Y.

LEMMA 5.38. *Let G be the isometry group of a (not necessarily proper) Gromov hyperbolic simplicial complex Y, and let μ be a probability distribution with finite support which generates a non-elementary subgroup of G. Then there is a probability distribution ν on the Gromov boundary ∂Y which is a weak limit of the convolution measures μ^{*n}.*

As the Gromov boundary is not compact, it will be convenient to consider an alternative compactification, namely the horofunction compactification, which we now describe. Let Y be a Gromov hyperbolic simplicial space, which is not necessarily locally compact, and let $C(Y, \mathbb{R})$ be the space of continuous functions on Y, with the compact-open topology, which in this case is equivalent to the topology of uniform convergence on compact sets. As $\mathbb{R}$ is Hausdorff, $C(Y, \mathbb{R})$ is also Hausdorff. Let $y_0 \in Y$ be a basepoint. There is a map from Y to $C(Y, \mathbb{R})$, defined by sending y to the corresponding horofunction, $h_y(z) = d(z, y) - d(y_0, y)$. Let Y_h be the closure of Y in $C(Y, \mathbb{R})$, which is sequentially compact, and hence compact as $C(Y, \mathbb{R})$ is Hausdorff. The space Y_h is known as the horofunction compactification of Y, and $Y_h \backslash Y$ is called the horofunction boundary.

We now define a 'local minimum' map $\phi\colon Y_h \to Y \cup \partial Y$. Given a function $h \in Y_h$, consider $\inf(h) = \inf_{y \in Y} h(y)$, which takes values in $[-\infty, 0]$. If a horofunction corresponds to a point in Y, i.e. $h = h_y$ for some $y \in Y$, then $\inf(h_y) = -d(y_0, y)$. We now consider the two cases depending on whether $\inf(h) > -\infty$, or $\inf(h) = -\infty$. If $\inf(h) > -\infty$, then there is a point y such that $h(y) \leq \inf(h) + 1$, and we shall set $\phi(h) = y$. Note that if h is equal to h_y for some $y \in Y$ then we may choose $\phi(h) = y$. If $\inf(h) = -\infty$, then choose a sequence y_n with $h(y_n) \to -\infty$, and set $\phi(h)$ equal to the limit of y_n in ∂Y. Note that we can choose this sequence y_n to be quasigeodesic, and such that $h(y_n) = -n$.

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

PROPOSITION 5.39. *The local minimum map* $\phi\colon Y_h \to Y \cup \partial Y$ *is coarsely well defined if* $\inf(h) > -\infty$*, and well defined if* $\inf(h) = -\infty$.

Proof. We start with some preliminary observations about horofunctions $h = h_y$ corresponding to points $y \in Y$, i.e. $h_y(z) = d(z, y) - d(y_0, y)$. In this case, h_y achieves its minimum value of $-d(y_0, y)$ at the point y. Furthermore, for any geodesic γ, the restriction of h_y to γ has a coarsely well-defined minimum a bounded distance away from the closest point projection p of y to γ. The value of h_y at p is equal to $d(p, y) - d(y_0, y)$, up to bounded error depending only on the constant of hyperbolicity δ, and for any other point $q \in \gamma$, the value of $h_y(q)$ is equal to $d(p, q) + d(p, y) - d(y_0, y)$, up to bounded error depending only on δ, i.e.

$$h_y(q) - K \leq d(p, q) + d(p, y) - d(y_0, y) \leq h_y(q) + K, \tag{5.10.1}$$

for some constant K, depending only on δ.

Let y_1 and y_2 be two points in Y with $h(y_i) \leq \inf(h) + \epsilon$, and let γ be a geodesic connecting them. Let y_n be a sequence of points in Y such that the corresponding horofunctions h_{y_n} converge to h. As γ is compact, for any number $\epsilon > 0$ there is an N such that for all $n \geq N$, and for all points $y \in \gamma$, $|h(y) - h_{x_n}(y)| \leq \epsilon$. Let p_n be the nearest point projection of y_n to γ, then by equation (5.10.1),

$$h_{y_n}(p_n) \leq \inf(h) + 2\epsilon - d(y_1, y_2)/2 + K,$$

where K depends only on δ, and so $d(y_1, y_2)$ is bounded by a constant which only depends on δ.

If $\inf(h) = -\infty$, there is a sequence of points $\{y_n\}$ such that $h(y_n) \to -\infty$. We now show that this sequence $\{y_n\}$ converges to a point in the Gromov boundary, and we shall map h to this point. Recall that a sequence $\{y_n\}$ converges to the boundary if for every number B there is a constant N such that the Gromov product $(y_m \cdot y_n)_{y_0} \geq B$ for all $m \geq N$ and $n \geq N$. So if the sequence $\{y_n\}$ does not converge, then there is a constant B such that for all N there are points in the sequence y_n and y_m with $m \geq N$ and $n \geq N$ such that $(y_m \cdot y_n)_{y_0} \leq B$.

Let γ be a geodesic from y_n to y_m, and let p be the closest point on γ to the basepoint y_0. The distance from p to y_0 is equal to the Gromov product, up to an error which only depends on δ, so $d(y_0, p) \leq B + K$, where K only depends on δ. In particular, this implies that the value of $h(p)$ is bounded below by $-B - K$. However, the value of h at any point on γ is bounded above by the value of h at the endpoints, up to an error which only depends on δ, and so as $h(y_n) \to -\infty$, this implies that $h(p) \to -\infty$, which is a contradiction. Therefore, the sequence y_n converges to a point in the Gromov boundary. □

We now consider the G-equivariance of the local minimum map ϕ.

PROPOSITION 5.40. *The local minimum map* $\phi\colon Y_h \to Y \cup \partial Y$ *is coarsely* G*-equivariant if* $\inf(h) > -\infty$*, and* G*-equivariant if* $\inf(h) = -\infty$.

Proof. First suppose that $\inf(h) > -\infty$. From the definition of a horofunction, $gh(z) = h(g^{-1}z) - h(g^{-1}y_0)$, so $\inf(gh) = \inf(h) - h(g^{-1}y_0)$, so the action of G preserves the set

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

of horofunctions with $\inf(h) > -\infty$. Furthermore,

$$\begin{aligned} gh(g\phi(h)) &= h(\phi(h)) - h(g^{-1}y_0) \\ &\leq \inf(gh) + 1. \end{aligned}$$

Therefore, both $g\phi(h)$ and $\phi(gh)$ have the property that their values under gh are within 1 of $\inf(gh)$, and so they are a bounded distance apart. Therefore, ϕ is coarsely G-equivariant on the set of elements with $\inf(h) > -\infty$.

Now suppose that $\inf(h) = -\infty$, and let y_n be a sequence in Y such that $h_{y_n} \to h$. Recall that $\phi(h_{y_n}) = y_n$, and $\phi(h)$ is the limit point of y_n in the Gromov boundary. The group G acts by homeomorphisms on $Y \cup \partial Y$, so $g\phi(h_{y_n}) = gy_n$, which converges to $g\phi(h)$. As G also acts by homeomorphisms on Y_h, the sequence gh_{y_n} converges to gh. Then $\phi(gh)$ is equal to the limit of $\phi(gh_{y_n}) = gy_n$, and this sequence converges to $g\phi(h)$, as required. Therefore, ϕ is G-equivariant on the set of h such that $\inf(h) = -\infty$. □

We now show that the map ϕ is measurable on the set of Y_h with $\inf(h) = -\infty$. In fact, it is continuous.

PROPOSITION 5.41. *The map $\phi\colon Y_h \to \partial Y$ is continuous on the set of $h \in Y_h$ with $\inf(h) = -\infty$.*

Proof. Let h_n be a sequence in Y_h, with $\inf(h_n) = \infty$, which converges to a point $h \in Y_h$, also with $\inf(h) = -\infty$. It suffices to show that $\phi(h_n)$ converges to $\phi(h)$.

For each h_n, there is a quasigeodesic sequence $(y_{n,m})$ with $h_n(y_{n,m}) = -m$, and also for h there is a quasigeodesic sequence y_m with $h(y_m) = -m$. By the definition of ϕ, the sequence $\phi(h_n)$ is equal to the limit of $(y_{n,m})$ and $\phi(h)$ is equal to the limit of y_m.

We now show that the Gromov product $(\phi(h_n) \cdot \phi(h))_{y_0} \to \infty$. It suffices to show that the quasigeodesics $(y_{n,m})_{m\in\mathbb{N}}$ and (y_m) fellow-travel on longer and longer initial subsequences as $n \to \infty$. Suppose that this is not the case, then no sequence $(y_{n,m})$ fellow-travels with (y_m) past some given point, y_p say. Consider the point y_q for $q > p$. The function h_n is decreasing along the quasigeodesic $(y_{n,m})$, and has no maximum on any geodesic. By thin triangles, any geodesic from $y_{n,m}$ to y_q passes close to y_q, and as it is increasing from $y_{n,m}$ to y_q, it must also be increasing from y_p to y_q. In particular, this implies that $h_n(y_q) \geq h_n(y_p)$ for all n. As $h_n \to h$, the functions h_n must in particular converge pointwise at p and q, so $h_n(y_p) \to h(y_p)$ and $h_n(y_q) \to h(y_q)$. This gives a contradiction, as h is decreasing along y_m, and $h(y_q) < h(y_p)$. Therefore, $\phi(x_n) \to \phi(x)$, so ϕ is continuous on those $h \in Y_h$ with $\inf(h) = -\infty$, as required. □

Consider the convolution measures μ^{*n} on Y_h. As Y_h is compact, there is a weak limit ν, which is a μ-stationary probability measure. The measure ν pushes forward to a μ-stationary probability measure on $Y \cup \partial Y$, which by abuse of notation we shall also refer to as ν. Actually, this requires some elaboration: the map from Y_h to $Y \cup \partial Y$ is only coarsely defined when the image is in Y, but is well defined when the image is in ∂Y. On the other hand, it will turn out as a consequence of the next lemma that the part of Y_h having image in Y has zero ν-measure, and therefore the pushforward of ν to ∂Y is well defined and μ-stationary, as claimed.

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

LEMMA 5.42. [**39**, Lemma 3.5] *Consider a random walk generated by a symmetric finitely supported probability distribution μ on the isometry group of a (not necessarily proper) Gromov hyperbolic simplicial complex Y, such that the group G generated by the support of μ is non-elementary. Let ν be a μ-stationary probability measure on $Y \cup \partial Y$, and let $|\cdot|_G$ be any proper metric on G.*

Let X be a set with the property that there is a sequence $\{k_i\}_{i\in\mathbb{N}}$ such that for any translate gX of X there is a sequence $\{w_i\}_{i\in\mathbb{N}}$ of elements of G, such that the translates gX, w_1gX, w_2gX, ... are all disjoint, and $|w_n|_G \leq k_n$ for all n. Then $\nu(X) = 0$.

We shall choose $X = B(y_0, r)$, the ball of radius r in Y. As G is non-elementary it contains a hyperbolic isometry f with translation length greater than $r + 2\delta$. Then for any translate $gB(y_0, r)$, the translates $f^n gB(y_0, r)$ are all disjoint. Choose $|\cdot|_G$ to be word length with respect to the generating set consisting of the support of μ. As $|f^n|_G \leq n|f|_G$, we may choose $k_n = n|f|_G$, and then Lemma 5.42 implies that $\nu(B(y_0, r)) = 0$. As this holds for every r, this implies that $\nu(Y) = 0$, as required. This completes the proof of Lemma 5.38, showing that there is weak limit of the convolution measures supported on the Gromov boundary.

We now prove Theorem 5.34, convergence to the boundary, following the arguments of Kaimanovich [**33**]. The action of G on $Y \cup \partial Y$ satisfies the following two properties.

(CP) If the sequence $g_n y_0$ converges to a point in ∂Y, then the sequence $g_n h y_0$ converges to the same point, for any $h \in G$.

(CS) The boundary ∂Y consists of at least three points, and there is a G-equivariant Borel map S assigning to pairs of distinct points b_1, b_2 in ∂Y subsets (strips) $S(b_1, b_2) \subset Y$, such that for any three pairwise distinct points $b_i \in \partial Y$, $i = 0, 1, 2$, there are neighborhoods $b_0 \in U_0 \subset Y$, and $b_i \in U_i \subset \partial Y$, $i = 1, 2$ with the property that $S(b_1, b_2) \cap U_0 = \varnothing$ for both $b_i \in U_i$, $i = 1, 2$.

We have shown that $(\partial Y, \nu)$ is a μ-boundary for G. Furthermore, the action of G on $Y \cup \partial Y$ satisfies properties (CP) and (CS) from Kaimanovich [**33**], and so all hypotheses of [**33**, Theorem 2.4] are satisfied, except compactness. However, in the proof of Theorem 2.4 compactness is only used to guarantee the existence of a μ-stationary probability measure, and we have shown how to construct such a measure above. Therefore, the conclusion of Theorem 2.4 holds in this setting, and this completes the proof of Theorem 5.34.

The properties of exponential decay and linear progress now follow from [**38**], as these arguments only use δ-hyperbolicity and convergence to the boundary, which we have now established. We may now complete the proof of Theorem 5.35.

Proof of Theorem 5.35. We briefly review some properties of quasigeodesics from [**8**, §III.H], see also Fujiwara [**22**, §1.2]. Let α be a path which is locally quasigeodesic, i.e. there are constants D, K and c such that every subpath of length at most D is a (K, c)-quasigeodesic. Then there is a constant D_0, which depends only on the constant of hyperbolicity δ, and the quasigeodesic constants K and c, such that for all $D \geq D_0$, the locally quasigeodesic path α is globally a (K', c')-quasigeodesic, where K' and c' depend only on δ. A local quasigeodesic may be constructed by concatenating geodesic segments with bounded overlaps. More precisely, let α and β be two geodesics in Y such that the

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

final point of α is equal to the initial point of β, and let γ be a geodesic from the initial point of α to the final point of β. We define the *overlap* $\mathcal{O}(\alpha, \beta)$ of α and β to be the largest distance from any point on $\alpha \cup \beta$ to γ. There is a constant Q, which only depends on δ, such that if α is a path consisting of a union of geodesics, each of length at least Q, such that each successive pair overlaps by at most 2δ, then α is a quasigeodesic, with quasigeodesic constants depending only on δ.

Let γ be a geodesic from y_0 to gy_0, and let $\overline{\gamma}$ be γ with the reverse orientation. Consider the path formed from the union of the geodesic segments $g^k\gamma$. Let m be the midpoint of γ, and let α be a geodesic from m to gm, then the union of the geodesic segments $g^k\alpha$ also forms a path in Y. Let B be the overlap of γ and $g\gamma$. By thin triangles, there is a constant K, which only depends on δ, such that if the length of γ is at least $2B + Q + K$, then the length of α is at least Q, and so the union of the $g^k\alpha$ forms a (K', c')-quasigeodesic, where K' and c' only depend on δ. Furthermore, the distance between m and gm is equal to $\text{length}(\gamma) - 2B$, up to an additive error which depends only on δ, and this in turn is equal to the translation length of g, again up to an additive error which only depends on δ.

If the size of the overlap between γ and $g\gamma$ is at least B, then γ has an initial segment of length B which fellow-travels with a final segment of $g^{-1}\gamma$, which is the initial segment of $\overline{\gamma}$ from y_0 to $g^{-1}y_0$. In particular, this means that $g^{-1}y_0 \in S_{y_0}(gy_0, B + K)$, for some constant K which only depends on δ.

We now estimate the probability that this occurs for a random walk of length n. Recall that by exponential decay, there are constants K and $c_1 < 1$ such that for any $g \in G$

$$\mathbf{P}(w_n y_0 \in S_{y_0}(gy_0, r)) \leq K c_1^{d(y_0, gy_0) - r}.$$

As μ is symmetric, and choosing $r = 3Ln/4$, we obtain

$$\mathbf{P}(w_n^{-1} y_0 \in S_{y_0}(w_n y_0, 3/4Ln)) \leq K c_1^{d(y_0, w_n y_0) - 3/4Ln}.$$

By linear progress, the probability that $d(y_0, w_n y_0) \leq Ln$ decays exponentially in n, which implies

$$\mathbf{P}(w_n^{-1} y_0 \in S_{y_0}(w_n y_0, 3/4Ln)) \leq K c_1^{1/4Ln} + O(c_2^n),$$

where $c_2 < 1$ is the exponential decay constant from linear progress. Therefore, the probability that the overlap of γ and $w_n\gamma$ is at most $Ln/4$ decays exponentially in n. As we have shown that the translation length of w_n is equal to $\text{length}(\gamma) - 2\mathcal{O}(\gamma, \overline{\gamma})$, up to additive error depending only on δ, this implies that the probability that the translation length of w_n is at least $Ln/2$ tends to 1 exponentially fast, as required. □

6. *Universal lower bounds*

In this section we shift our focus abruptly, and concentrate on obtaining uniform *lower bounds* on scl for random walks in arbitrary finitely generated groups. If G is any group, then Bavard duality implies that either scl vanishes identically on $[G, G]$, or else there is a homogeneous quasimorphism ϕ on G. In the former case, there is nothing to say. In the latter case, one obtains uniform lower bounds on scl from an estimate on the distribution of values of ϕ.

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

It turns out that there is a *central limit theorem* for ϕ, proved by Björklund and Hartnick [**6**]. This theorem allows us to obtain uniform lower bounds on scl of order $O(\sqrt{n})$. An important special case of this theorem concerns quasimorphisms obtained from actions of groups on circles (we obtained this special case independently of Björklund and Hartnick, though the method of proof is similar). In the next section we prove the central limit theorem for circle actions, and derive some geometric applications, of independent interest.

6.1. *Groups acting on circles.* Let $\mathrm{Homeo}^+(S^1)$ denote the group of orientation-preserving homeomorphisms of the circle. This group has a universal central extension consisting of the group of homeomorphisms of $\mathbb{R}$ that commute with integer translation; we denote this group $\mathrm{Homeo}^+(\mathbb{R})^{\mathbb{Z}}$.

Poincaré defined a function $\mathrm{rot} : \mathrm{Homeo}^+(\mathbb{R})^{\mathbb{Z}} \to \mathbb{R}$, called *rotation number*, by the formula

$$\mathrm{rot}(h) = \lim_{n\to\infty} \frac{h(n)}{n}.$$

This descends to a function $\mathrm{rot} : \mathrm{Homeo}^+(S^1) \to \mathbb{R}/\mathbb{Z}$. The function rot is a homogeneous quasimorphism on $\mathrm{Homeo}^+(\mathbb{R})^{\mathbb{Z}}$ with defect $D(\mathrm{rot}) = 1$.

Now let G be any group. Suppose G acts on the circle by orientation-preserving homeomorphisms, i.e. suppose we have $G \to \mathrm{Homeo}^+(S^1)$. The preimage of G in $\mathrm{Homeo}^+(\mathbb{R})^{\mathbb{Z}}$ is a (possibly split) central extension $\overline{G}$ of G and we obtain a homogeneous quasimorphism rot on $\overline{G}$. The defect of rot on $\overline{G}$ is usually equal to 1, but might be smaller, for instance, if the centralizer of $\overline{G}$ in $\mathrm{Homeo}^+(\mathbb{R})^{\mathbb{Z}}$ is bigger than the center $\mathbb{Z}$.

Let $\overline{G} \to \mathrm{Homeo}^+(\mathbb{R})^{\mathbb{Z}}$ be some representation as above, and let $\overline{S}$ be a finite generating set for $\overline{G}$ (note that $\overline{G}$ is finitely generated if and only if G is). Let μ be the uniform probability measure on $\overline{S}$, and let μ_n be the n-fold convolution of μ. That is, μ_n is the probability measure associated to a random walk of length n on G in the generators $\overline{S}$. Note that we do *not* require S to be symmetric or to generate G as a group.

THEOREM 6.1. (Central limit theorem) *Let $\overline{G}$ be a subgroup of* $\mathrm{Homeo}^+(\mathbb{R})^{\mathbb{Z}}$*, and let $\overline{S}$ be a finite generating set. Let $g_0, g_1, \ldots$ be a Markov process on G, where $g_0 = \mathrm{id}$, and where each g_n is obtained from g_{n-1} by right multiplication by a random element $s \in \overline{S}$ (in the uniform measure). Then there is a central limit theorem for* rot*, i.e. there are constants E and σ so that $n^{-1/2}(\mathrm{rot}(g_n) - En)$ converges in probability to the Gaussian measure $N(0, \sigma)$.*

Proof. Let G be the image of $\overline{G}$ in $\mathrm{Homeo}^+(S^1)$, and let S be the image of $\overline{S}$. Further, let F be the semigroup generated by S (which might be smaller than G). Let $M(S^1)$ denote the space of probability measures on S^1, and let $\Delta_S : M(S^1) \to M(S^1)$ be defined by

$$\Delta_S(\nu) = \frac{1}{|S|} \sum_{s\in S} s_*\nu.$$

A fixed point of Δ_S is called a *stationary* (or *harmonic*) measure on S^1. We let ν be an ergodic stationary measure. By construction, the action of F on S^1 is absolutely continuous with respect to ν.

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

If ν contains atoms, then there is an atom of biggest measure, supported at some point p. Since $\nu(p)$ is equal to the average of $\nu(sp)$, it follows that the measure of every $\nu(sp)$ is equal to the measure of $\nu(p)$. It follows that every point in Fp has the same (atomic) measure, and therefore this set is finite and F-invariant. Since the action is by homeomorphisms, this set is invariant under G, and therefore G preserves a probability measure on S^1. In this case it is well known that rot is a *homomorphism* from $\overline{G}$ to $\mathbb{R}$, and the ordinary central limit theorem applies. So we assume in the following that ν contains no atoms.

Since ν is stationary, for any measurable $I \subset S^1$ (in particular, for every interval I), there is an equality

$$\frac{1}{|S|}\sum_{s\in S}\nu(S(I)) = \nu(I).$$

Let $\overline{\nu}$ be the Radon measure on $\mathbb{R}$ obtained by identifying $\mathbb{R}$ locally with S^1. Observe that $\overline{\nu}$ is invariant under integer translation, and furthermore it satisfies $\overline{\nu}([t, t+1]) = 1$ for all $t \in \mathbb{R}$. Furthermore, $\overline{\nu}$ is evidently stationary for $\overline{S}$, i.e. $(1/|S|)\sum_{s\in\overline{S}}\overline{\nu}(s(I)) = \overline{\nu}(I)$ for all measurable $I \subset \mathbb{R}$.

Define $F : \mathbb{R} \to \mathbb{R}$ as follows. For each t choose $T \ll 0$ and $T \ll t$, and define

$$F(t) = \overline{\nu}[T, t] - \overline{\nu}[T, 0]$$

(evidently, F does not depend on the choice of sufficiently negative T). For any $t \in \mathbb{R}$, define

$$f(t) = \left(\frac{1}{|S|}\sum_{s} F(s(t))\right) - F(t).$$

If u is arbitrary, and I is the interval with extremal points t and u, then

$$f(u) = f(t) + \left(\frac{1}{|S|}\sum_{s}\overline{\nu}(s(I))\right) - \overline{\nu}(I) = f(t).$$

Hence $f(\cdot)$ is *constant*, and equal to some fixed E, which we call the *drift* of S.

It follows from this that the function $F(g_n(0)) - nE$ is a (bounded) *martingale*; that is, the expected value of $F(g_n(0)) - nE$ given g_{n-1} is $F(g_{n-1}(0)) - (n-1)E$.

Now, if X_i is any martingale with bounded increments, if σ_i^2 is the expectation of $(X_{i+1} - X_i)^2$ given $X_1, X_2, \ldots, X_i$, and if τ_n is the minimum n so that $\sum_{i=1}^{n}\sigma_i^2 \geq n$, then the Martingale central limit theorem (see, e.g., [**29**], especially Theorem 3.2 on p. 58) says that $X_{\tau_n}/\sqrt{n}$ converges in probability to a normal distribution $N(0, 1)$. In our particular case, it turns out that τ_n/n converges in probability to σ^2 for some constant σ. The expectation of $(F(g_n) - F(g_{n-1}) - E)^2$ given g_{n-1} depends only on $g_{n-1}(0)$ mod $\mathbb{Z}$, and since by hypothesis the measure ν is ergodic for the action of Δ_S, the random ergodic theorem (see [**23**, Theorem 3.1]) implies such convergence in probability (even in L^1).

Now, $|\mathrm{rot}(g) - g(0)| \leq 1$ for any $g \in \overline{G}$, and moreover $|F(t) - t| \leq 1$. Therefore, a central limit theorem for the function $F(g_n(0)) - nE$ implies one for $\mathrm{rot}(g_n) - nE$ and the theorem is proved. □

Remark 6.2. If $S = S^{-1}$ (symmetric random walk in a group), then the random process is invariant under taking inverses, and therefore $E = 0$. Similarly, if S is conjugated to itself (in $\mathrm{Homeo}^+(\mathbb{R})^{\mathbb{Z}}$) by some reflection $t \to 2C - t$, then $E = -E = 0$ by symmetry.

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

COROLLARY 6.3. *Let G be a finitely generated subgroup of* $\mathrm{Homeo}^+(S^1)$*, and let $\overline{G}$ be the preimage in* $\mathrm{Homeo}^+(\mathbb{R})^{\mathbb{Z}}$*. Suppose further that* scl *vanishes identically on G. Then if* scl_n *denotes the value of* scl *on a random walk in $\overline{G}$ (in some finite symmetric generating set), there is some σ for which the following is true:*

$$\lim_{n\to\infty} \mathbf{P}(a < \mathrm{scl}_n/\sigma\sqrt{n} < b) = \frac{2}{2\pi}\int_a^b \mathbf{1}_{[0,\infty)} e^{-x^2/2}\,dx.$$

Proof. For any group G there is a short exact sequence

$$0 \to H^1(G) \to Q(G) \to H_b^2(G) \to H^2(G),$$

where H_b^2 denotes bounded cohomology. Since scl vanishes identically on G by hypothesis, $Q(G)/H^1(G) = 0$ by Bavard duality, i.e. Theorem 2.6. However, the coboundary $\delta\mathrm{rot}$ exists as an element of $H_b^2(G)$ whose image in $H^2(G)$ is non-trivial, and equal to the familiar *Euler class*.

Since $\overline{G}$ is the central extension associated to the Euler class, it follows that $Q(\overline{G})/H^1(\overline{G})$ is one-dimensional, and spanned by rot. So $\mathrm{scl}(g) = |\mathrm{rot}(g)|/2D(\mathrm{rot})$ on the commutator subgroup of $\overline{G}$, and the conclusion follows. □

Example 6.4. Corollary 6.3 applies to many naturally occurring families of groups, including Hilbert modular groups $\mathrm{SL}(2, \mathcal{O}(n))$, where $\mathcal{O}(n)$ is the ring of integers in $\mathbb{Q}(\sqrt{n})$ for n square-free, $\mathrm{SL}(2, \mathbb{Z}[1/2])$, Thompson's circle group F and certain generalized Stein–Thompson groups, and many others. The fact that scl vanishes identically on these groups follows from the stronger property that they are *boundedly generated by commutators*.

For Hilbert modular groups, this is a consequence of a deep theorem of Carter, Keller and Paige, namely [**48**, Theorem 6.1], which says that if A is the ring of integers in a number field K containing infinitely many units, and T is an element of $\mathrm{SL}(2, A)$ which is not a scalar matrix, then $\mathrm{SL}(2, A)$ has a finite index normal subgroup which is boundedly generated by conjugates of T.

The case $\mathrm{SL}(2, \mathbb{Z}[1/2])$ is due to Liehl [**37**], who proves that the group is boundedly generated by elementary matrices (which are themselves products of commutators of bounded length).

The case of Thompson's group is due to Ghys and Sergiescu [**25**], and some generalizations are due to Zhuang [**50**]. For an introduction to Thompson's groups and their properties, see [**15**].

6.2. *Random turtles in the hyperbolic plane.* In this subsection we give a geometric application of Theorem 6.1 of independent interest. Consider the following random process. A turtle starts at the origin in the hyperbolic plane, and moves by alternately moving forward some fixed distance ℓ, and by turning either left or right through some fixed angle α. Let $p_0, p_1, \ldots$ denote the locations of the turtle after each successive move forward. Note p_0 is the origin, and $d(p_i, p_{i+1}) = \ell$ for each i. We think of the p_i as the vertices of a random polygonal path.

For each n, let P_n be the polygon with $n+1$ (cyclic) vertices $p_0, p_1, \ldots, p_n$. There are (at least) two natural geometric quantities to associate to P_n. If $\gamma : S^1 \to \mathbb{R}^2$ is

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

a C^1 immersion, the *winding number* is the degree of the Gauss map $\theta \to \gamma'(\theta)/|\gamma'(\theta)|$. Moreover, the *algebraic area* enclosed by γ is the integral $\int_{\mathbb{R}^2} \mathrm{wind}(\gamma, p) d\mathrm{area}(p)$, where the *local winding number* $\mathrm{wind}(\gamma, p)$ is the degree of the map $\theta \to (\gamma(\theta) - p)/|(\gamma(\theta) - p)|$.

Although ∂P_n is only piecewise linear, it can be smoothed canonically by rotating the tangent vector left or right through an angle α at each vertex p_i with $0 < i < n$ (according to the behavior of the turtle), and then in an arbitrary way at p_0 and p_n. Thus, we can assign to the sequence p_n two geometric quantities W_n, the winding number of ∂P_n, and A_n, the algebraic area enclosed by ∂P_n.

Let α_i be the signed turning angle of P_n at the vertex p_i. Then the Gauss–Bonnet theorem for immersed polygons says there is an equality

$$2\pi W_n - \sum_i \alpha_i = -A_n.$$

THEOREM 6.5. (Area and winding theorem) *Fix some angle α and length ℓ. Let P_n be a random polygon in the hyperbolic plane with (cyclic) vertices $p_0, p_1, \ldots, p_n$, where $d(p_i, p_{i+1}) = \ell$ for each $0 \le i \le n-1$ and an angle of $\pm\alpha$ at each p_i with $0 < i < n$, with signs independently and uniformly chosen from ± 1. Let A_n be (as above) the algebraic area enclosed by P_n, and W_n the winding number of ∂P_n. Then A_n and W_n both satisfy a central limit theorem with mean 0.*

Proof. Let L and R be the hyperbolic isometries which translate the origin a distance ℓ and then rotate either left or right through angle α. As matrices in $\mathrm{PSL}(2, \mathbb{R})$, we can take

$$R = \begin{pmatrix} e^{\ell/2}\cos(\alpha/2) & e^{\ell/2}\sin(\alpha/2) \\ -e^{-\ell/2}\sin(\alpha/2) & e^{-\ell/2}\cos(\alpha/2) \end{pmatrix}, \quad L = \begin{pmatrix} e^{\ell/2}\cos(\alpha/2) & -e^{\ell/2}\sin(\alpha/2) \\ e^{-\ell/2}\sin(\alpha/2) & e^{-\ell/2}\cos(\alpha/2) \end{pmatrix}.$$

The successive p_i are obtained from p_0 by right multiplication by a random sequence of matrices R and L. We think of $p_i = w_i p_0$, where w_i is an element of the free semigroup generated by R and L. By exponentiation, the unit tangent circle at each point in $\mathbb{H}^2$ is canonically identified with the ideal circle S^1_∞ (this defines a canonical flat projective connection on the unit tangent bundle). Therefore, the winding number can be computed from the (lifted) action on the ideal circle. We think of R and L as elements of $\mathrm{Homeo}^+(\mathbb{R})^{\mathbb{Z}}$, normalized to move some basepoint distance < 1 in the positive and negative directions respectively. Then $|W_i - \mathrm{rot}(w_i)| < 1$, so Theorem 6.1 shows that W_n satisfies a central limit theorem. Note that $E = 0$, by the left–right symmetry (see Remark 6.2).

With notation as in the proof of Theorem 6.1, the central limit theorem for W_n follows from the martingale property of the function $F(w_n(0))$. However, the function $F(w_n(0)) - \sum_{i=1}^n \alpha_i/2\pi$ is also a martingale, so the same proof gives a central limit theorem for A_n, as claimed. □

Remark 6.6. If α is small enough compared to ℓ, the polygonal path $p_0, p_1, p_2, \ldots$ is uniformly quasigeodesic. In this case, the areas of successive triangles p_0, p_n, p_{n+1} define a Hölder continuous function on the one-sided shift space on the alphabet $\{L, R\}$. The Hölder continuity amounts to the observation that for four points a, b, c, d in $\mathbb{H}^2$

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

with $d(a, b) = d(c, d) = \ell$, the difference of area(a, c, d) and area(b, c, d) is bounded by a constant that decays exponentially fast in the distance from b to c. The central limit theorem for Hölder functions on shift spaces gives a different proof in this case. This kind of argument is implicit in [**41**] and a related argument is pursued in [**12, 32**].

Remark 6.7. There is nothing very special (apart from its charm) about the particular random model we chose for the polygons P_n. We could just as easily fix some finite subset $S \subset \mathrm{PSL}(2, \mathbb{R})$ and define a random sequence p_i by $p_i = sp_{i-1}$ for some random $s \in S$. The winding number and algebraic area of P_n satisfy a central limit theorem in this case too, and with essentially the same proof. Note that in this generality, the drift might be non-zero.

Remark 6.8. It is interesting to study how the statistical quantities associated to A_n and W_n vary as a function of the parameters. Let us fix α and let ℓ vary, and consider the random winding number W_n. By the left–right symmetry, the drift $E(\ell)$ is zero for all ℓ. But the standard deviation $\sigma(\ell)$ undergoes a phase transition: it is zero for $\ell \in [2 \cosh^{-1}(1/\sin(\alpha/2)), \infty)$, and increases monotonically to α as ℓ decreases to 0. In this case $\sigma(\ell)$ is real analytic (as a function of ℓ) on $[0, 2 \cosh^{-1}(1/\sin(\alpha/2)))$.

6.3. *Central limit theorem for arbitrary quasimorphisms.* In fact, very shortly after proving Theorem 6.1, we learned that a completely general statement has independently been obtained by Björklund and Hartnick [**6**]. They prove the following theorem (in fact, their results hold in considerably greater generality).

THEOREM 6.9. (Björklund and Hartnick [**6**, Theorem 1.1]) *Let G be a finitely generated group, and S a finite generating set. Let $f : G \to \mathbb{R}$ be a quasimorphism, and $X_n := s_n \cdots s_1$ an i.i.d. left-random walk on G in the generating set S. Then f satisfies a central limit theorem with respect to X_n. Moreover, if the homogenization of f is non-zero, then the central limit is non-degenerate.*

By Bavard duality we immediately conclude the following.

COROLLARY 6.10. *Let G be any finitely generated group, and suppose $H_1(G)$ is finite. Let S be a finite symmetric generating set. Suppose $Q(G)$ is finite dimensional but non-zero. Let g_n be obtained by random walk of length n with respect to the uniform measure on S. Then for any ϵ there are positive constants a, b depending on ϵ such that $\mathbf{P}(a < \mathrm{scl}(g_n)/\sqrt{n} < b) \geq 1 - \epsilon$ for $n \gg 0$.*

Theorem 6.9 generalizes Theorem 6.1 considerably, but the proof turns out not to be too much harder. Under very general conditions, Björklund and Hartnick show that a quasimorphism f on a group G with a probability measure μ (satisfying some conditions) has a *(bi-)harmonic representative*, i.e. there is a function f' that differs from f by a bounded amount, and with the property that f' is invariant under convolution with μ (the existence of such a harmonic representative was also proved in quite a different way by Burger and Monod [**10**]). It is this harmonic feature of f' that lets one prove a central limit theorem using the martingale central limit theorem, as in the proof of Theorem 6.1.

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

Remark 6.11. We would like to point out that in the first version of their paper, Björklund and Hartnick required that the probability measure on S be *symmetric*, whereas we never required this as a hypothesis.

Remark 6.12. Burger and Monod and Björklund and Hartnick both work with individual quasimorphisms. A finite collection of quasimorphisms $\phi_1, \phi_2, \ldots, \phi_m$ can be put together into a single function $\Phi : G \to \mathbb{R}^m$ whose coordinates are the ϕ_i. It makes sense to ask for some harmonic representative Φ' with $|\Phi' - \Phi| < \infty$ and a central limit theorem for Φ' (and therefore also for Φ). In fact, the arguments in [**6, 10**] easily generalize to this situation. The existence of a harmonic representative is essentially elementary. The quasimorphism property implies by definition that there is a constant C so that for any g, h we have $|1/2(\Phi(gh) + \Phi(gh^{-1})) - \Phi(g)| \leq C$. It follows that for μ symmetric, the n-fold convolutions $\Phi_n(g) := \int_G \Phi(gh)\, d\mu^{*n}(h)$ lie in a precompact family, and therefore we obtain in the usual way a fixed point for convolution with μ, i.e. a harmonic representative. Following Björklund and Hartnick we deduce a central limit theorem for Φ.

The condition in Corollary 6.10 that $H_1(G)$ is finite is an annoying technical restriction. The problem is that one cannot perform the 'homological correction trick' that we used earlier for two reasons. Firstly, there is no compression of the support of a Gaussian measure, and therefore conditioning on a rare event can change the order of magnitude of the distribution. Secondly, the order of magnitude of the error term (i.e. the size of the image in homology) is the same as the term we are trying to control, so it is impossible to naively obtain lower bounds, since the two terms might cancel. Nevertheless, the following conjectures seem reasonable. The first just asserts that the hypothesis that $H_1(G)$ is finite in Corollary 6.10 can be removed.

CONJECTURE 6.13. (Local limit conjecture) *Let G be any finitely generated group. Let S be a finite symmetric generating set. Suppose $Q(G)$ is finite dimensional but non-zero. Let g_n be obtained by random walk of length n with respect to the uniform measure on S, conditioned to lie in $[G, G]$. Then for any ϵ there are positive constants a, b depending on ϵ such that $\mathbf{P}(a < \mathrm{scl}(g_n)/\sqrt{n} < b) \geq 1 - \epsilon$ for $n \gg 0$.*

The second conjecture is more ambitious, and asserts that the $O(\sqrt{n})$ growth rate should characterize (finitely presented) groups with $Q(G)$ finite dimensional.

CONJECTURE 6.14. (Finite dimensionality) *Let G be a finitely presented group, and let g_n be obtained by random walk of length n conditioned to lie in $[G, G]$. Suppose that for any ϵ there is a positive constant b so that*

$$\mathbf{P}(\mathrm{scl}(g_n)/\sqrt{n} < b) \geq 1 - \epsilon$$

for $n \gg 0$. Then $Q(G)$ is finite dimensional.

Finally, we ask for finitely presented groups for which scl has *intermediate growth*.

QUESTION 6.15. *Is there a finitely presented group G so that if g_n is obtained by random walk of length n conditioned to lie in $[G, G]$, then for any $C > 0$ and $\epsilon > 0$, there is an estimate*

$$\mathbf{P}(\mathrm{scl}(g_n) < C\sqrt{n} \text{ or } \mathrm{scl}(g_n) > C^{-1}n/\log n) < \epsilon$$

for $n \gg 0$?

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

We believe that it should be possible to produce finitely *generated* groups with the property sought by Question 6.15, by a careful small cancellation argument. It would also be very interesting to find examples of (uncountable) *transformation groups G* containing finitely generated subgroups Γ so that the growth rate of scl_G on random walk in Γ is intermediate between $\sqrt{n}$ and $n/\log n$.

QUESTION 6.16. *Let Σ be a closed surface with $\chi(\Sigma) < 0$, equipped with a smooth area form. Let G be the group of Hamiltonian diffeomorphisms of Σ. Is there a well-defined growth rate of scl_G on random walk in Γ for Γ a 'generic' finitely generated subgroup of G? Does it grow like $n/\log n$?*

Acknowledgements. Danny Calegari was supported by NSF grants DMS 0707130 and DMS 1005246. Joseph Maher was supported by NSF grant DMS 0706764 and PSC-CUNY Award 60019-43-20, and would like to thank the Hausdorff Research Institute for Mathematics for its hospitality during work on this paper. We would like to thank Jason Behrstock, Michael Björklund, Sean Cleary, David Fisher, Alex Furman, Bradley Groff, Anders Karlsson, Nikolai Makarov, Curt McMullen, Andrés Navas, Jay Rosen, Richard Sharp, Brian Simanek, Alessandro Sisto, Alden Walker and the anonymous referee for some useful conversations about this material.

REFERENCES

[1] A. Barvinok. *Integer Points in Polyhedra (Zurich Lectures in Advanced Mathematics)*. European Mathematical Society, Zürich, 2008.
[2] C. Bavard. Longueur stable des commutateurs. *Enseign. Math.* (2) **37** (1991), 109–150.
[3] M. Bestvina and M. Feighn. A hyperbolic Out(F_n)-complex. *Groups Geom. Dyn.* **4**(1) (2010), 31–58.
[4] M. Bestvina and M. Feighn. Hyperbolicity of the complex of free factors. *Preprint*, 2001, arXiv:1107.3308.
[5] M. Bestvina and K. Fujiwara. Bounded cohomology of subgroups of mapping class groups. *Geom. Topol.* **6** (2002), 69–89.
[6] M. Björklund and T. Hartnick. Biharmonic functions on groups and limit theorems for quasimorphisms along random walks. *Geom. Topol.* **15**(1) (2011), 123–143.
[7] B. Bowditch. Tight geodesics in the curve complex. *Invent. Math.* **171**(2) (2008), 281–300.
[8] M. Bridson and A. Haefliger. *Metric Spaces of Non-positive Curvature (Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 319)*. Springer, Berlin, 1999.
[9] R. Brooks. Some remarks on bounded cohomology. *Riemann Surfaces and Related topics: Proceedings of the 1978 Stony Brook Conference (SUNY, Stony Brook, NY, 1978) (Annals of Mathematics Studies, 97)*. Princeton University Press, Princeton, NJ, 1981, pp. 53–63.
[10] M. Burger and N. Monod. Bounded cohomology of lattices in higher rank Lie groups. *J. Eur. Math. Soc.* **1**(2) (1999), 199–235.
[11] D. Calegari. *scl (MSJ Memoirs, 20)*. Mathematical Society of Japan, Tokyo, 2009.
[12] D. Calegari and K. Fujiwara. Combable functions, quasimorphisms and the central limit theorem. *Ergod. Th. & Dynam. Sys.* **30**(5) (2010), 1343–1369.
[13] D. Calegari and A. Walker. Random rigidity in the free group. *Preprint*, arXiv:1104.1768, *Geom. Topol.*, to appear.
[14] J. Cannon. The combinatorial structure of cocompact discrete hyperbolic groups. *Geom. Dedicata* **16**(2) (1984), 123–148.
[15] J. Cannon, W. Floyd and W. Parry. Introductory notes on Richard Thompson's groups. *Enseign. Math. (2)* **42**(3–4) (1996), 215–256.
[16] M. Coornaert. Mesures de Patterson–Sullivan sur le bord d'un espace hyperbolique au sens de Gromov. *Pacific J. Math.* **159**(2) (1993), 241–270.

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press

[17] I. Dinwoodie. Expectations for nonreversible Markov chains. *J. Math. Anal. Appl.* **220** (1998), 585–596.
[18] D. Epstein, J. Cannon, D. Holt, S. Levy, M. Paterson and W. Thurston. *Word Processing in Groups*. Jones and Bartlett, Boston, MA, 1992.
[19] D. Epstein and K. Fujiwara. The second bounded cohomology of word-hyperbolic groups. *Topology* **36**(6) (1997), 1275–1289.
[20] B. Farb and H. Masur. Superrigidity and mapping class groups. *Topology* **37**(6) (1998), 1169–1176.
[21] K. Fujiwara. The second bounded cohomology of a group acting on a Gromov-hyperbolic space. *Proc. Lond. Math. Soc. (3)* **76**(1) (1998), 70–94.
[22] K. Fujiwara. Subgroups generated by two pseudo-Anosov elements in a mapping class group. I. Uniform exponential growth. *Groups of Diffeomorphisms (Advanced Studies in Pure Mathematics, 52)*. Mathematical Society of Japan, Tokyo, 2008, pp. 283–296.
[23] A. Furman. *Random Walks on Groups and Random Transformations (Handbook of Dynamical Systems, 1a)*. North-Holland, Amsterdam, 2002, pp. 931–1014.
[24] É. Ghys. *Knots and Dynamics (International Congress of Mathematicians. 1)*. European Mathematical Society, Zürich, 2007, pp. 247–277.
[25] É. Ghys and V. Sergiescu. Sur un groupe remarquable de difféomorphismes du cercle. *Comment. Math. Helv.* **62**(2) (1987), 185–239.
[26] M. Gromov. Volume and bounded cohomology. *Publ. Math. Inst. Hautes Études Sci.* **56** (1982), 5–99.
[27] M. Gromov. *Hyperbolic Groups (Essays in Group Theory, MSRI Publication 8)*. Springer, New York, 1987, pp. 75–263.
[28] D. Groves and J. Manning. Dehn filling in relatively hyperbolic groups. *Israel J. Math.* **168** (2008), 317–429.
[29] P. Hall and C. Heyde. *Martingale Limit Theory and Its Application*. Academic Press, New York, 1980.
[30] U. Hamenstädt. Lines of minima in outer space. *Preprint*, arXiv:0911.3620, *Duke Math. J.*, to appear.
[31] M. Handel and L. Mosher. The free splitting complex of a free group I: hyperbolicity. *Preprint*, 2011, arXiv:1111.1994.
[32] M. Horsham and R. Sharp. Lengths, quasi-morphisms and statistics for free groups. *Spectral Analysis in Geometry and Number Theory (Contemporary Mathematics, 484)*. American Mathematical Society, Providence, RI, 2009, pp. 219–237.
[33] V. Kaimanovich. The Poisson formula for groups with hyperbolic properties. *Ann. of Math. (2)* **152** (2000), 659–692.
[34] V. Kaimanovich and H. Masur. The Poisson boundary of the mapping class group. *Invent. Math.* **125**(2) (1996), 221–264.
[35] S. Kakutani. Random ergodic theorems and Markoff processes with a stable distribution. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950*. University of California Press, Berkeley, CA, 1951, pp. 247–261.
[36] E. Kowalski. *The Large Sieve and its Applications: Arithmetic Geometry, Random Walks and Discrete Groups (Cambridge Tracts in Mathematics, 175)*. Cambridge University Press, Cambridge, 2008.
[37] B. Liehl. Beschränkte Wortlänge in SL_2. *Math. Z.* **186**(4) (1984), 509–524.
[38] J. Maher. Linear progress in the complex of curves. *Trans. Amer. Math. Soc.* **362**(6) (2010), 2963–2991.
[39] J. Maher. Random Heegaard splittings. *J. Topol.* **3**(4) (2010), 997–1025.
[40] J. Maher. Exponential decay in the mapping class group. *J. Lond. Math. Soc.* (2) **86**(2) (2012), 366–386.
[41] J.-C. Picaud. Cohomologie bornée des surfaces et courants géodésiques. *Bull. Soc. Math. France* **125**(1) (1997), 115–142.
[42] M. Ratner. Markov partitions for Anosov flows on n-dimensional manifolds. *Israel J. Math.* **15** (1973), 92–114.
[43] I. Rivin. Walks on groups, counting reducible matrices, polynomials, and surface and free group automorphisms. *Duke Math. J.* **142** (2008), 353–379.
[44] I. Rivin. Growth in free groups (and other stories)—twelve years later. *Illinois J. Math.* **54**(1) (2010), 327–370.
[45] R. Sharp. Local limit theorems for free groups. *Math. Ann.* **321** (2001), 889–904.
[46] A. Sisto. Contracting elements and random walks. *Preprint*, 2011, arXiv:1112.2666.
[47] D. Stroock. *Probability Theory, an Analytic View*. Cambridge University Press, Cambridge, 1993.
[48] D. Witte-Morris. Bounded generation of $\mathrm{SL}(n, A)$ (after D. Carter, G. Keller, and E. Paige). *New York J. Math.* **13** (2007), 383–421.
[49] W. Woess. *Random Walks on Infinite Graphs and Groups (Cambridge Tracts in Mathematics, 138)*. Cambridge University Press, Cambridge, 2000.
[50] D. Zhuang. Irrational stable commutator length in finitely presented groups. *J. Mod. Dyn.* **2**(3) (2008), 499–507.

https://doi.org/10.1017/etds.2013.43 Published online by Cambridge University Press