COMPOSITIO MATHEMATICA

Central limit theorems for counting measures in coarse negative curvature

Ilya Gekhtman, Samuel J. Taylor and Giulio Tiozzo

FOUNDATION
COMPOSITIO
MATHEMATICA

LONDON
MATHEMATICAL
SOCIETY
EST. 1865

CrossMark

Central limit theorems for counting measures in coarse negative curvature

Ilya Gekhtman, Samuel J. Taylor and Giulio Tiozzo

Abstract

We establish central limit theorems for an action of a group G on a hyperbolic space X with respect to the counting measure on a Cayley graph of G. Our techniques allow us to remove the usual assumptions of properness and smoothness of the space, or cocompactness of the action. We provide several applications which require our general framework, including to lengths of geodesics in geometrically finite manifolds and to intersection numbers with submanifolds.

1. Introduction

The goal of this paper is to provide a novel approach to the central limit theorem (CLT) on groups acting on hyperbolic spaces, for sampling with respect to the word length in the group. We shall replace the traditional approach based on thermodynamic formalism with techniques coming from the theory of random walks on groups. This allows us to establish new applications, including CLTs for lengths of geodesics in geometrically finite hyperbolic manifolds, for intersection numbers with submanifolds, and for homomorphisms between hyperbolic groups.

Motivation
The distribution of lengths of closed orbits for smooth flows on manifolds has long been a topic of considerable interest. For instance, Sinai [Sin60] and then Ratner [Rat73] proved a CLT for the geodesic flow on a hyperbolic manifold (see also Lalley [Lal87]). One prominent technique, pioneered by Sinai [Sin72], Bowen [Bow75], Ruelle [Rue04], Parry and Pollicott [PP90], and others, uses Markov partitions to reduce the study of smooth flows to symbolic dynamics to which one can apply tools from thermodynamic formalism. This approach has been successful in a variety of settings, especially applied to Anosov flows and their generalizations.

More recently, there has been renewed interest in the statistical properties of geodesic length and other geometric quantities with respect to a different sampling, namely according to the *counting measure*, that is to say, uniform measure on spheres in a Cayley graph of a finitely generated group G. For instance, Pollicott and Sharp [PS98] considered the ratio between the word length and the geometric length, Calegari and Fujiwara [CF10] compared word lengths with

respect to different generating sets for general hyperbolic groups, and Horsham and Sharp [HS09] established a CLT for quasimorphisms on free groups.

In [GTT19] the authors, building on [GTT18], settled a conjecture of Chas, Li and Maskit [CLM13] about the distribution of hyperbolic lengths of closed geodesics on compact surfaces when sampling with respect to word length. Further, a CLT and statistical laws have been established for cocompact, proper actions of hyperbolic groups on CAT(-1) spaces by Cantrell [Can21].

All of these results are based on a symbolic coding and thermodynamic formalism. Although these techniques are quite powerful, they necessarily impose strong constraints on the actions of interest, usually requiring that the space X is CAT(-1) and that the action $G \curvearrowright X$ is proper cocompact. While this is the case in the classical setting, they are not satisfied for most actions on Gromov hyperbolic spaces.

The goal of this paper is to provide a new approach to the CLT on groups $G \curvearrowright X$ acting on hyperbolic spaces, which will allow us to consider in particular:

(i) groups G which are not necessarily word hyperbolic;
(ii) actions on spaces (X, d) which are δ-hyperbolic, but not necessarily CAT(-1) or proper;
(iii) group actions $G \curvearrowright X$ which need not be free, convex cocompact or even proper.

For the sake of concreteness, we will now present a version of our main theorem (Theorem 1.1) from which we will then derive several applications. Our discussion here will be a special case of the most general theorems (Theorems 7.3 and 7.4) which we will state and prove in §7.

Main results

Let G be a finitely generated group acting by isometries on a δ-hyperbolic metric space (X, d), and fix a finite generating set S. We require that the action is *nonelementary* in the sense that there are two independent loxodromic elements.

Let $S_n := \{g \in G : \|g\| = n\}$ be the sphere of radius n for the word metric with respect to S. We denote by $\mathcal{N}_\sigma$ the Gaussian measure $d\mathcal{N}_\sigma(t) = (1/\sqrt{2\pi}\sigma)e^{-t^2/2\sigma^2} \, dt$ if $\sigma > 0$, and the Dirac measure at 0 if $\sigma = 0$. We require that G admits a *thick geodesic combing* for S and we refer the reader to §2.1 for definitions. We note here that these general conditions are satisfied in a variety of settings; for example, see the applications below and Lemma 8.1.

THEOREM 1.1. *Let G be a group which admits a thick geodesic combing for the generating set S. Let $G \curvearrowright X$ be a nonelementary action by isometries on a δ-hyperbolic space (X, d), and let $o \in X$ be a base point.*

(i) (CLT for displacement) *Then there exist $\ell > 0$, $\sigma \geq 0$ such that for any $a < b$ we have*
$$\lim_{n \to \infty} \frac{1}{\#S_n} \#\left\{g \in S_n : \frac{d(o, go) - \ell n}{\sqrt{n}} \in [a, b]\right\} = \int_a^b d\mathcal{N}_\sigma(t).$$

(ii) (CLT for translation length) *Moreover, if $\tau(g)$ denotes the translation length of g on X, we also have for any $a < b$,*
$$\lim_{n \to \infty} \frac{1}{\#S_n} \#\left\{g \in S_n : \frac{\tau(g) - \ell n}{\sqrt{n}} \in [a, b]\right\} = \int_a^b d\mathcal{N}_\sigma(t).$$

(iii) *Further, $\sigma = 0$ if and only if there exists a constant C such that*
$$|d(o, go) - \ell\|g\|| \leq C$$
for any $g \in G$.

1981

We remark that as a consequence of (iii), if $\sigma = 0$ then the translation length of any $g \in G$ with respect to its action on X is a constant multiple of its translation length in the word metric. Moreover, if the action $G \curvearrowright X$ is not proper, then $\sigma > 0$.

Another way to formulate (i) is to say that we have the convergence in distribution

$$\frac{d(o, go) - \ell n}{\sqrt{n}} \longrightarrow \mathcal{N}_\sigma,$$

hence from now on we will use the above notation as a shorthand.

Applications

There are a number of applications to the above theorems and we summarize a few of them here. For the proofs, see § 8.

Geometrically finite hyperbolic manifolds. First, let us state an extension of our previous work on surfaces [GTT19] to general hyperbolic manifolds, possibly with cusps. If $M = \mathbb{H}^n/\Gamma$ is a hyperbolic manifold and $\gamma \in \Gamma = \pi_1(M)$, then we set $\ell(\gamma)$ to be the length of the geodesic freely homotopic to γ unless γ is peripheral (i.e. homotopic into a cusp), in which case we set $\ell(\gamma) = 0$.

THEOREM 1.2. *Suppose that M is a geometrically finite hyperbolic manifold and let S' be any generating set for $\pi_1(M)$. Then there is a finite generating set $S \supset S'$ and $\ell, \sigma > 0$ such that*

$$\frac{\ell(\gamma) - \ell n}{\sqrt{n}} \longrightarrow \mathcal{N}_\sigma,$$

where γ is chosen uniformly at random in the sphere of radius n with respect to S.

If, moreover, $\pi_1(M)$ is word hyperbolic, then we can take $S = S'$.

The statement includes the cases where M is either finite volume or convex cocompact, and is new even when M is a finite area surface. We remark that when M is either convex cocompact or a surface, the above theorem works for *any* generating set S. In the convex cocompact case, the needed action $\pi_1(M) \curvearrowright \mathbb{H}^n$ is sufficiently tame so that the techniques of thermodynamics may be applicable [PS98, Can21]. However, this is not the case when the manifold M has cusps.

We note that Theorem 1.2 further extends to manifolds of variable negative curvature, as long as the peripheral subgroups are virtually abelian, and the same proof applies.

Geometrically infinite 3-manifolds. In the case of 3-manifolds, the previous result can be strengthened further as follows.

THEOREM 1.3. *Let M be a hyperbolic 3-manifold such that $\pi_1(M)$ is finitely generated and not virtually abelian. Suppose further that M does not have any rank 2 cusps. Then for any finite generating set S of $\pi_1(M)$, there are $\ell, \sigma > 0$ such that*

$$\frac{\ell(\gamma) - \ell n}{\sqrt{n}} \longrightarrow \mathcal{N}_\sigma,$$

where γ is chosen uniformly at random in the sphere of radius n with respect to S.

Moreover, if M has rank 2 cusps, the same statement holds after enlarging the generating set as in Theorem 1.2.

To the authors' knowledge, this is the first CLT for lengths of closed geodesics for possibly geometrically infinite 3-manifolds.

Intersection numbers with a submanifold. For our next application, the required actions are on locally infinite trees, which are nonproper hyperbolic spaces.

Let M be a smooth orientable manifold and Σ a smooth orientable codimension 1 submanifold which is π_1-injective on each component. We say Σ is *fiber-like* if for each boundary component of the cut manifold $M|\Sigma$ its induced subgroup in the fundamental group of the corresponding component of $M|\Sigma$ has index at most 2.

For $\gamma \in \pi_1(M)$, let $i(\gamma, \Sigma)$ denote the minimal intersection number of Σ with loops in M freely homotopic to γ.

THEOREM 1.4. *Suppose that M is a closed orientable hyperbolic manifold and let S be any generating set for $\pi_1(M)$. Let Σ be a smooth orientable codimension 1 submanifold that is π_1-injective but not fiber-like. Then there are $\ell, \sigma > 0$ such that*

$$\frac{i(\gamma, \Sigma) - \ell n}{\sqrt{n}} \longrightarrow \mathcal{N}_\sigma,$$

where γ is chosen uniformly at random in the sphere of radius n with respect to S.

The theorem is new even for surfaces; in that context, Chas and Lalley [CL12] proved a CLT for self-intersection numbers of curves with respect to word length. Following Theorem 1.2, a similar result could be formulated for more general hyperbolic manifolds.

Homomorphisms between hyperbolic groups. Our next application is to homomorphisms between hyperbolic groups. Interestingly, the condition for nonzero variance can be recast in terms of the induced Patterson–Sullivan measures.

THEOREM 1.5. *Suppose that $\phi: G \to G'$ is a homomorphism between hyperbolic groups such that the image of ϕ is not virtually cyclic. For any fixed generating sets S and S' of G and G', respectively, there are $\ell > 0$ and $\sigma \geq 0$ such that*

$$\frac{\|\phi(g)\|_{S'} - \ell\|g\|_S}{\sqrt{\|g\|_S}} \longrightarrow \mathcal{N}_\sigma,$$

for $g \in G$ chosen uniformly at random in the sphere of radius n with respect to S.

Moreover, $\sigma = 0$ (i.e. the Gaussian is degenerate) if and only if ϕ has finite kernel and there is an induced embedding $\partial\phi: \partial G \to \partial G'$ that pushes the Patterson–Sullivan measure class for (G, d_S) to the Patterson–Sullivan measure class for $(\phi(G), d_{S'})$.

The above result generalizes a result due to Calegari and Fujiwara [CF10, Corollary 4.27]. It is also related to [Can21, Theorem 1.6], who proved a CLT where ϕ is the abelianization homomorphism, which in turn generalizes work of Rivin [Riv10] for free groups.

Hyperplanes crossed in right-angled Artin and Coxeter groups. Our final application is to a collection of groups that is not necessarily relatively hyperbolic.

Suppose that G is a right-angled Artin group or right-angled Coxeter group that is not a direct product. Let V be its set of vertex generators. For each $v \in V$, define a function $\#_v: G \to \mathbb{Z}$ that counts the number of occurrences of $v^{\pm 1}$ in a shortest spelling of $g \in G$ with respect to V. Equivalently, $\#_v(g)$ is the number of hyperplanes labeled by v separating o and go in the cube complex associated to G.

THEOREM 1.6. *For G as above, there are $\ell, \sigma > 0$ such that for any vertex v,*

$$\frac{\#_v(g) - \ell n}{\sqrt{n}} \longrightarrow \mathcal{N}_\sigma,$$

where g is chosen uniformly at random in the sphere of radius n with respect to the vertex generators.

We conclude by noting that our methods are sufficiently general to apply beyond the case of 'nonpositively curved' groups. Moreover, we do *not* need to assume that our counting measures are associated to geodesic combings. See Theorems 7.3 and 7.4 for the most general result. For example, by using the standard graph structure associated to the language of geodesics for a free group, we obtain a CLT for nonbacktracking random walks on *any* group with a nonelementary action on a hyperbolic space X.

From thermodynamics to random walks

Most CLTs for counting measures established so far use a coding for geodesics with finite paths, and then apply classical results in thermodynamic formalism, like the existence and uniqueness of Gibbs measures for shifts of finite type. There, the observable is assumed to be Hölder continuous with respect to the standard metric on the shift space.

In this paper, instead, we *do not* assume any good geometric property on the action. Let us recall that displacement is not a quasimorphism, is in general not *weakly combable* (in the language of [CF10]) if the action is not convex cocompact, and is not a *Hölder weight function* in the sense of [PS98, Proposition 1] if X is not CAT(-1). Thus, the observable need not be Hölder and the thermodynamic approach does not appear to work. We also do not use transfer operators or ζ-functions (e.g. as in [PP90, Can21]).

Rather, our general strategy is as follows.

 (i) We start with a *graph structure*, that is, a graph whose paths parameterize the group elements we want to count. We first consider a vertex v of this graph, and consider a random walk on the semigroup Γ_v of loops based at this vertex. Here, we apply the CLT for cocycles for groups acting on hyperbolic spaces, as devised by Benoist and Quint [BQ16] and generalized by Horbez [Hor18] to actions on nonproper spaces.

 (ii) Then we consider the set of paths in a maximal component for the graph as a suspension on the space of loops at v, and we apply results of Melbourne and Török [MT18] to 'lift' the CLT to the suspended transformation. To be precise, we need to consider a skew product over the shift space.

(iii) Now we note that a thick graph structure is *almost semisimple*, hence there exists a power p for which the transition matrix M^p is semisimple. We use this to prove that the counting measure starting at an initial vertex converges to a convex combination of stationary measures for the Markov chains on the maximal components.

(iv) Using a coherence condition on the graph structure, we show that all the CLTs for all Markov chains have the same mean and variance. This implies a CLT for the counting measure on the set of paths starting at any vertex in a semisimple structure.

 (v) Finally, for a general thick structure of period p we condition on the first prefix of length r; since all these distributions for the conditional measures converge to the same law (by (iv) above), the CLT for the entire sequence holds.

We conclude by mentioning some natural generalizations of this work, which we do not attempt here. A Berry–Esseen type estimate of the error term appears within reach, using the corresponding result for random walks and the bounds we work out here comparing the counting and random walk measures. It would also be very interesting to obtain a local limit theorem, for which, however, much further work is needed.

2. Background

2.1 Graph structures for countable groups

Given a countable group G, we define a *graph structure* on G as a triple $(\Gamma, v_0, \mathrm{ev})$, where Γ is a finite, directed graph, v_0 is a vertex of Γ which we call its *initial vertex*, and $\mathrm{ev}\colon E(\Gamma) \to G$ is a map that labels the edges of Γ with group elements. Given this data, we extend the map ev by defining for each finite path $g = g_1 \cdots g_n$ the group element $\mathrm{ev}(g) = \mathrm{ev}(g_1) \cdots \mathrm{ev}(g_n)$. We remark that here and throughout, a path in Γ will always mean a directed path. To simplify notation, we will use $\overline{g} = \mathrm{ev}(g)$ to denote the group element associated to the path g. We denote by $\|g\|$ the length of the path g. The graph structure is *bounded* if there is a uniform bound on the number of paths in Γ mapping to any single element of G.

For a graph structure Γ, we define Ω to be the set of all infinite paths starting at any vertex of Γ and $\sigma\colon \Omega \to \Omega$ to be the shift map. Given a path $\omega = (g_1, \ldots, g_n, \ldots)$, we denote by $w_n := g_1 \cdots g_n$ its prefix of length n. The set of all finite paths starting at any vertex of Γ will be denoted by Ω^*.

We define two vertices v_i, v_j to be *equivalent* if there is a path from v_i to v_j and a path from v_j to v_i, and the *components* of Γ as the equivalence classes for this relation.

We will denote by M the transition matrix for Γ. By Perron–Frobenius, M has a real eigenvalue of largest modulus, which we will denote by λ. Moreover, such a matrix is *almost semisimple* if for any eigenvalue of maximal modulus, its geometric and algebraic multiplicity agree. Furthermore, such a matrix is *semisimple* if its only eigenvalue of maximal modulus is real positive. We call a graph structure *(almost) semisimple* if its associated transition matrix is.

Let Γ be almost semisimple, and let λ be the leading eigenvalue of M. Then we define a vertex v to be of *large growth* if

$$\lim_{n \to \infty} \frac{1}{n} \log \#\{\text{paths of length } n \text{ starting at } v\} = \lambda$$

and of *small growth* otherwise (in which case, the limit above is less than λ). Furthermore, a component C is *maximal* if

$$\lim_{n \to \infty} \frac{1}{n} \log \#\{\text{paths of length } n \text{ inside } C\} = \lambda.$$

As discussed in [GTT20], the global structure of Γ is as follows: there is no path between maximal components, and vertices of large growth are precisely the ones which have a path to a maximal component.

Given a vertex v, we denote by Γ_v the *loop semigroup* of v, that is, the set of all finite paths from v to itself. This is a semigroup under concatenation, and all its elements lie entirely in the component of v. We denote by $\overline{\Gamma}_v$ the image of Γ_v in G under the evaluation map.

DEFINITION 2.1 (Thick graph structure). A graph structure Γ is *thick* if, for any vertex v in a maximal component, there exists a finite set $B \subseteq G$ such that

$$G = B \cdot \overline{\Gamma}_v \cdot B,$$

where the equality is *in the group G*.

In what follows, we often make the evaluation map implicit in our notation. In particular, if G acts on a metric space (X, d), $o \in X$ is a base point, and g is a finite path in Γ, we will often write go to mean the point $\overline{g}o \in X$.

Geodesic combings. For particular applications, it is also useful to define the notion of a geodesic graph structure. A graph structure Γ is *geodesic* if the length $\|g\|$ of any path g is equal

1985

the word length of $\overline{g}$ in the subgroup generated by the edge labels, using edge labels as the (finite) generating set. A geodesic graph structure is called a *geodesic combing* if, in addition, there is a directed path from v_0 to any other vertex of Γ and the evaluation map is a bijection from the set of finite paths starting at v_0 to the set of elements of G. Note that a geodesic combing is automatically a bounded graph structure. We say that Γ is a geodesic combing *associated to* a finite generating set S if, up to adding inverses, S is the set of edge labels for the graph structure. In this case, $\|g\|$ is equal to the word length of $\overline{g}$ with respect to S.

We emphasize that although the geodesic condition is used in the applications of our main theorem (as in Theorems 1.1–1.6), it is not required for the most general results; see Theorems 7.3 and 7.4. There we use the following more general notion.

DEFINITION 2.2 (Coherent graph structure). A graph structure Γ for G is *coherent* if it is bounded and if for any finite set $B \subseteq G$ there exists a constant $\mathcal{B} \geq 0$ such that if g and h are finite length paths in Γ, and $\overline{g} = b_1 \overline{h} b_2$ in G, then

$$\left| \|g\| - \|h\| \right| \leq \mathcal{B}.$$

The idea of coherence is simple; informally, if two group elements are close in G then a coherent graph structure Γ codes them with paths that are roughly the same length, and the coding is uniformly bounded-to-one.

Our first lemma summarizes some properties of thick, coherent graph structures that we will need in the sequel.

LEMMA 2.3. *A thick, coherent graph structure Γ is almost semisimple.*

Moreover, if $G \curvearrowright X$ is a nonelementary action on a hyperbolic space, then the actions of both semigroups Γ_v and Γ_v^{-1} on X are also nonelementary, for each vertex v contained in a maximal component.

Proof. Suppose the graph structure Γ is thick and coherent. Since a coherent graph structure is bounded, there is a $\mathfrak{b} \geq 0$ so that the evaluation map $\mathrm{ev} \colon \Omega^* \to G$ is at most $\mathfrak{b}$-to-one. Let $B \subset G$ be the finite set given by thickness of Γ (Definition 2.1) and let $\mathcal{B}$ be the resulting constant from Definition 2.2. By thickness and coherence we have

$$\overline{S_n} \subseteq \bigcup_{b_1, b_2 \in B} \bigcup_{|k| \leq \mathcal{B}} b_1 (\overline{\Gamma_v \cap S_{n+k}}) b_2,$$

where $S_n \subset \Omega^*$ is the set of all length n paths starting at any vertex. Then the definition of $\mathfrak{b}$ implies that

$$\#S_n \leq \mathfrak{b} |B|^2 \sum_{|k| \leq \mathcal{B}} \#(\Gamma_v \cap S_{n+k}).$$

This, in turn, is bounded by a constant times λ^n since it is no more than the growth of paths in the maximal component of Γ containing v, and k is uniformly bounded. On the other hand, if the transition matrix M of Γ is not almost semisimple, then M has a Jordan block for an eigenvalue of modulus λ of size $k \geq 2$ (see [GTT18, § 2]). In particular, the growth of paths in Γ is at least a constant times $n^{k-1} \lambda^n$. We conclude that Γ is almost semisimple.

The statement that the action of Γ_v (and hence Γ_v^{-1}) on X is nonelementary is proven in [GTT20, Proposition 6.3]. $\square$

The following lemma is immediate from the definitions.

LEMMA 2.4. *A geodesic combing for G is coherent.*

Other sources of (not necessarily geodesic) coherent graph structures come from biautomatic groups, in the sense of [Mos97] or [EPC$^+$92, Lemma 2.5.5].

2.2 Cocycles and horofunctions

Let (X, d) be a metric space, and let $o \in X$ be a base point. Given $z \in X$, we define the *Busemann function* $\rho_z : X \to \mathbb{R}$ as

$$\rho_z(x) := d(x, z) - d(o, z).$$

Thus, setting

$$\Phi(z) := \rho_z$$

defines a map

$$\Phi : X \to \mathrm{Lip}_o^1(X)$$

where $\mathrm{Lip}_o^1(X)$ is the space of 1-Lipschitz functions on X which vanish at o.

We define the *horofunction compactification* $\overline{X}^h$ as the closure of $\Phi(X)$ in $\mathrm{Lip}_o^1(X)$, with respect to the topology of pointwise convergence. Elements of $\overline{X}^h$ will be called *horofunctions*. We denote by $\overline{X}_\infty^h$ the space of *infinite horofunctions*, that is, the set of $h \in \overline{X}^h$ such that $\inf_{x \in X} h(x) = -\infty$.

For any $\xi \in \overline{X}^h$, the *Busemann cocycle* is defined as

$$\beta_\xi(x, y) := \lim_{z_n \to \xi}[d(y, z_n) - d(x, z_n)]$$

$$= h_\xi(y) - h_\xi(x),$$

where h_ξ is the horofunction associated to ξ. This has the usual cocycle property $\beta_\xi(x, z) = \beta_\xi(x, y) + \beta_\xi(y, z)$.

Remark 2.5. Benoist and Quint [BQ16] and Horbez [Hor18] define $B \colon G \times \overline{X}^h \to \mathbb{R}$ by

$$B(g, \xi) = h_\xi(g^{-1}o).$$

To compare their definition with ours:

$$B(g, \xi) = h_\xi(g^{-1}o) = \lim_{z_n \to \xi}[d(g^{-1}o, z_n) - d(o, z_n)] = \beta_\xi(o, g^{-1}o).$$

3. Central limit theorem for random walks on the loop semigroups

Let Γ be a graph structure for G, and let v be a vertex in a maximal component. Recall that a (directed) loop is *prime* if it is not itself a product of nontrivial loops; since every loop can be decomposed uniquely as a concatenation of prime loops, prime loops freely generate Γ_v as a semigroup.

Given a probability measure μ on the set of edges of Γ, one defines the *first return measure* μ_v on Γ_v as follows. If $l = g_1 \cdots g_n$ is a prime loop in Γ_v, then we set

$$\mu_v(l) := \mu(g_1) \cdots \mu(g_n).$$

We set $\mu_v(l) = 0$ for all other loops. Note that inversion defines a map $\Gamma_v \to \Gamma_v^{-1}$ and we define the measure $\check{\mu}$ on Γ_v^{-1} by $\check{\mu}(l) = \mu(l^{-1})$. These measures push forward to measures on the group G under the evaluation map. We say that μ_v is *nondegenerate* if it gives positive measure to any prime loop of Γ_v.

Let $\mathcal{M}$ be a metric space on which G acts by homeomorphisms. A measure ν on $\mathcal{M}$ is μ-*stationary* if $\nu = \int_G g_\star \nu \, d\mu(g)$, and μ-*ergodic* if it cannot be written as a nontrivial convex combination of μ-stationary measures.

3.1 Central limit theorems for cocycles

Recall that a *cocycle* is a function $\sigma\colon G \times \mathcal{M} \to \mathbb{R}$ such that

$$\sigma(gh, x) = \sigma(g, hx) + \sigma(h, x), \quad \forall\, g, h \in G, \ \forall\, x \in \mathcal{M}.$$

A cocycle $\sigma\colon G \times \mathcal{M} \to \mathbb{R}$ has *constant drift* λ if there exists $\lambda \in \mathbb{R}$ such that

$$\int_G \sigma(g, x) \, d\mu(g) = \lambda$$

for any $x \in \mathcal{M}$. A cocycle $\sigma\colon G \times \mathcal{M} \to \mathbb{R}$ is *centerable* if it can be written as

$$\sigma(g, x) = \sigma_0(g, x) + \psi(x) - \psi(g \cdot x),$$

where σ_0 is a cocycle with constant drift and where $\psi\colon \mathcal{M} \to \mathbb{R}$ is a bounded, measurable function. In this case, we say that σ_0 is the *centering* of σ; note that $\lambda = \int_{G \times \mathcal{M}} \sigma(g, x) \, d\mu(g) \, d\nu(x)$ for any μ-stationary ν. We say that the cocycle σ has *finite second moment* with respect to a measure μ on G if

$$\int_G \sup_{x \in \mathcal{M}} |\sigma(g, x)|^2 \, d\mu(g) < +\infty.$$

We now use the following CLT for centerable cocycles; as remarked in [Hor18, Remark 1.7], the proof is exactly the same as the proof of [BQ16, Theorem 4.7].

THEOREM 3.1 (Central limit theorem for cocycles). *Let G be a discrete group, $\mathcal{M}$ be a compact metrizable G-space and μ a probability measure on G. Let ν be a μ-ergodic, μ-stationary probability measure on $\mathcal{M}$, and let $\mathcal{M}_0$ be a G-invariant subset of $\mathcal{M}$ of full ν-measure. Let $\sigma\colon G \times \mathcal{M}_0 \to \mathbb{R}$ be a centerable cocycle with drift λ and finite second moment. Then there exist $\sigma \geq 0$ such that for any continuous $F\colon \mathbb{R} \to \mathbb{R}$ with compact support, we have for ν-almost every (a.e.) $x \in \mathcal{M}$,*

$$\lim_{n \to \infty} \int_G F\left(\frac{\sigma(g, x) - \lambda n}{\sqrt{n}} \right) d\mu^{*n}(g) = \int_{\mathbb{R}} F(t) \, d\mathcal{N}_\sigma(t).$$

We now apply this result to the loop semigroup. Let F_v be the group freely generated by the prime loops in Γ_v.

Let $N\colon \Gamma_v \to \mathbb{Z}$ be the semigroup homomorphism $N(g) := -\|g\|$, where $\|g\|$ is the length in Γ of the loop g. There is a natural inclusion $\Gamma_v \to F_v$ as a subsemigroup and we can extend the semigroup homomorphism above to a group homomorphism $N\colon F_v \to \mathbb{Z}$. Moreover, we also extend the natural semigroup homomorphism $\Gamma_v \to G$, induced by evaluation, to a group homomorphism $e\colon F_v \to G$. Now, using the homomorphism $e\colon F_v \to G$, the free group F_v has a nonelementary action on X, and, moreover, μ_v^{*n} is supported on $\Gamma_v \subseteq F_v$ for all $n \geq 1$.

Finally, for some $\ell \in \mathbb{R}$ to be specified below, we define $\eta\colon F_v \times \overline{X}^h \to \mathbb{R}$ as

$$\eta(g, \xi) := \beta_\xi(o, g^{-1}o) - \ell N(g).$$

LEMMA 3.2. *Suppose that the action of Γ_v on X is nonelementary and μ_v is nondegenerate. Then, for any $\ell \in \mathbb{R}$, the restriction of $\eta\colon F_v \times \overline{X}^h \to \mathbb{R}$ to $F_v \times \overline{X}_\infty^h$ is a centerable cocycle.*

Proof. We have

$$
\begin{aligned}
\eta(gh, \xi) &= \beta_\xi(o, h^{-1}g^{-1}o) - \ell N(gh) \\
&= \beta_\xi(o, h^{-1}o) + \beta_\xi(h^{-1}o, h^{-1}g^{-1}o) - \ell N(g) - \ell N(h) \\
&= \beta_\xi(o, h^{-1}o) + \beta_{h\xi}(o, g^{-1}o) - \ell N(g) - \ell N(h) \\
&= \eta(h, \xi) + \eta(g, h\xi),
\end{aligned}
$$

hence η is a cocycle. Moreover, by [Hor18, Proposition 1.5], using [Hor18, Corollary 2.7] and [Hor18, Proposition 2.8], the cocycle $B(g, \xi) = \beta_\xi(o, g^{-1}o)$ is centerable on $F_v \times \overline{X}_\infty^h$. Then, since $\eta(g, \xi) - B(g, \xi) = \ell N(g)$ is a homomorphism and depends only on g, we have that $\eta(g, \xi)$ is also centerable on $F_v \times \overline{X}_\infty^h$. $\qquad\square$

Thus, we obtain the following consequence of Theorem 3.1.

COROLLARY 3.3. *Let Γ be a thick structure, let v be a vertex in a maximal component of Γ. Suppose that the first return measure μ_v is nondegenerate, and let ν_v be a $\check{\mu}_v$-ergodic, $\check{\mu}_v$-stationary measure on $\overline{X}^h$. Then there exist $\ell, \sigma \geq 0$ such that for any continuous $F \colon \mathbb{R} \to \mathbb{R}$ with compact support, we have for ν_v-a.e. ξ,*

$$
\lim_{n \to \infty} \int_G F\left(\frac{\beta_\xi(o, go) - \ell\|g\|}{\sqrt{n}} \right) d\mu_v^{*n}(g) = \int_{\mathbb{R}} F(t) \, d\mathcal{N}_\sigma(t).
$$

Proof. We apply Theorem 3.1 to the measure $\check{\mu}_v$, supported on Γ_v^{-1}, where ℓ is chosen so that $\lambda = \int_{F_v \times \overline{X}^h} \eta(g, \xi) \, d\check{\mu}_v(g) \, d\nu_v(\xi) = 0$. Note that, by [MT18, Proposition 4.4] and the fact that Γ_v^{-1} is nonelementary (Lemma 2.3), we have $\nu_v(\overline{X}_\infty^h) = 1$. Moreover, for any $g \in \Gamma_v$ we have

$$
\eta(g^{-1}, \xi) = \beta_\xi(o, go) - \ell\|g\|.
$$
$\qquad\square$

3.2 Skew products and invariance on the loop semigroup

Let $\mathcal{M}$ be a compact metric space with a continuous G-action. We define the *skew product* $T \colon \Omega \times \mathcal{M} \to \Omega \times \mathcal{M}$ as

$$
T(\omega, \xi) := (\sigma(\omega), g_1^{-1}\xi)
$$

where $\omega = (g_1, g_2, \ldots)$.

A graph structure Γ is *primitive* if its associated transition matrix M is primitive, that is, has a positive power. Now let Γ be a primitive graph structure, let v be a vertex of Γ, let Γ_v be the loop semigroup, and let μ_v be the first return measure.

Finally, let $\Omega_v = (\Gamma_v)^{\mathbb{N}}$ with shift map σ_v. To highlight the difference, we denote the elements of $\Gamma_v^{\mathbb{N}}$ by $(l_1, l_2, \ldots)$, since each element of the sequence is a loop, while the elements of Ω will be denoted by $\omega = (g_1, g_2, \ldots)$, since its elements are edges. Let us define the map $T_v \colon \Omega_v \times \mathcal{M} \to \Omega_v \times \mathcal{M}$ as

$$
T_v(\omega, \xi) = (\sigma_v(\omega), l_1^{-1}\xi).
$$

LEMMA 3.4. *A measure ν on $\mathcal{M}$ is $\check{\mu}_v$-stationary if and only if $\mu_v^{\mathbb{N}} \otimes \nu$ is T_v-invariant.*

Proof. Fix $C \subset \Omega_v$ measurable and let $C_l \subset \Omega_v$ be the subset consisting of sequences beginning with $l \in \Gamma_v$ such that $\sigma_v(C_l) = C$. Then, for any $A \subset \mathcal{M}$ measurable,

$$
T_v^{-1}(C \times A) = \bigcup_l C_l \times lA.
$$

Since

$$\mu_v^{\mathbb{N}} \otimes \nu \left(\bigcup_l C_l \times lA \right) = \mu(C) \sum_l \mu(l)\nu(lA)$$

$$= \mu(C) \sum_l \check{\mu}(l) l_* \nu(A),$$

the lemma follows. □

LEMMA 3.5. *There exists an ergodic $\check{\mu}_v$-stationary measure ν_v on $\mathcal{M}$ such that the product measure $\mu_v^{\mathbb{N}} \otimes \nu_v$ is T_v-invariant and ergodic.*

Proof. Since $\mathcal{M}$ is a compact metric space, there exists a $\check{\mu}_v$-stationary measure ν_1 on $\mathcal{M}$ [Fur63, Lemma 1.2]. Then by Lemma 3.4 the measure $\lambda_1 := \mu_v^{\mathbb{N}} \otimes \nu_1$ is T_v-invariant. If λ_1 is not ergodic, let us consider its ergodic decomposition, and take one of its ergodic components λ_v. By definition, $\lambda_v \ll \lambda_1$ and λ_v is T_v-invariant and ergodic. Then, by [Mor88, Corollary 3.1], λ_v is of the form $\lambda_v = \mu_v^{\mathbb{N}} \otimes \nu_v$ for some measure ν_v on $\mathcal{M}$. Finally, again by Lemma 3.4, the measure ν_v is $\check{\mu}_v$-stationary. □

LEMMA 3.6. *Consider the function $f \colon \Omega \times \overline{X}^h \to \mathbb{R}$ defined as*

$$f(\omega, \xi) := \beta_\xi(o, g_1 o).$$

Then for any n we have

$$\sum_{j=0}^{n-1} f(T^j(\omega, \xi)) = \beta_\xi(o, w_n o). \tag{1}$$

Proof. The cocycle property implies

$$\beta_\xi(o, w_n o) = \sum_{j=0}^{n-1} \beta_\xi(w_j o, w_{j+1} o) = \sum_{j=0}^{n-1} \beta_{w_j^{-1}\xi}(o, g_{j+1} o)$$

for any $\xi \in \overline{X}^h$. Moreover, by definition and G-equivariance we have

$$f(T^j(\omega, \xi)) = \beta_{w_j^{-1}\xi}(o, g_{j+1} o)$$

and the claim follows. □

An analogous statement holds by replacing T, Ω by T_v, Ω_v.

4. Central limit theorem for Markov chains of primitive graph structures

We begin by recalling the following. If Γ is a directed graph whose transition matrix M is primitive with leading eigenvalue λ, then

$$\lim_{n\to\infty} \frac{M^n}{\lambda^n} = \rho u^T,$$

where $M\rho = \lambda\rho$, $u^T M = \lambda u^T$, and $u^T \rho = 1$ [Wal00, Theorem 0.17]. We will consider the following measures associated to Γ. The stationary measure on vertices for the corresponding Markov chain is given by setting the starting probability at vertex v_i to be $\pi_i = \rho_i u_i$. Assigning the probability $\rho_j/\lambda\rho_i$ to an edge e from v_i to v_j then determines the transition probability p_{ij} from v_i to v_j to be $p_{ij} = M_{ij}\rho_j/\lambda\rho_i$ (see, for example, [Wal00, §6.6 (7)]). The corresponding Markov measure $\mathbb{P}$ on the path space Ω determined by these probabilities gives the measure of maximal entropy

for the shift on Ω (which in this section is unique since the graph structure is assumed to be primitive) and is known as the *Parry measure* [Par64]. For a vertex v of Γ, we use $\mathbb{P}_v$ to denote the measure on the space of paths $\Omega_v \subset \Omega$ starting at v obtained by beginning the Markov chain at v and using the above transition probabilities. From now on we use these edge probabilities to define the first return measure μ_v as in §3.

We also define μ_n as the distribution on the space of paths of length n induced by the Markov measure $\mathbb{P}$; that is, μ_n is the pushforward of $\mathbb{P}$ to Ω^* under the map which sends an infinite path to its prefix of length n. In this section we prove the following result.

THEOREM 4.1. *Suppose that Γ is a primitive graph structure and let μ_n be the associated distribution on the space of paths of length n. There are constants ℓ and σ such that, for any continuous function $F \colon \mathbb{R} \to \mathbb{R}$ with compact support, we have*

$$\lim_{n \to \infty} \int_{\Omega^*} F\left(\frac{d(o, go) - \ell n}{\sqrt{n}}\right) d\mu_n(g) = \int_{\mathbb{R}} F(t)\, d\mathcal{N}_\sigma(t).$$

The main technique to obtain the CLT for the Markov chain as above from that from the random walk on the loop semigroup is using a *suspension flow*, adapting the approach of Melbourne and Török [MT04] for dynamical systems.

4.1 Suspension flows

Let $S \colon (\mathcal{X}, \lambda) \to (\mathcal{X}, \lambda)$ be a measure-preserving dynamical system, and let $r \colon \mathcal{X} \to \mathbb{N}$ be a measurable, integrable function, which we call the *roof function*. Then the *discrete suspension flow* of S with roof function r is the dynamical system given by the map $\widehat{S} \colon \widehat{\mathcal{X}} \to \widehat{\mathcal{X}}$ where

$$\widehat{\mathcal{X}} := \{(x, n) \in \mathcal{X} \times \mathbb{N} : 0 \leq n \leq r(x) - 1\}$$

with measure $\widehat{\lambda} := (1/\overline{r})(\lambda \otimes \delta)$, where δ is the counting measure on $\mathbb{N}$ and $\overline{r} := \int_{\mathcal{X}} r\, d\lambda$. Then the map $\widehat{S}$ is defined as

$$\widehat{S}(x, n) = \begin{cases} (x, n+1) & \text{if } n \leq r(x) - 2, \\ (S(x), 0) & \text{if } n = r(x) - 1. \end{cases}$$

Since in this case the system has discrete time, the above construction is also called a *Kakutani skyscraper*.

We now state the main theorem of Melbourne and Török [MT04, Theorem 1.1].

THEOREM 4.2. *Let $S \colon (\mathcal{X}, \lambda) \to (\mathcal{X}, \lambda)$ be an ergodic, measure-preserving transformation, and let $\widehat{S} \colon (\widehat{\mathcal{X}}, \widehat{\lambda}) \to (\widehat{\mathcal{X}}, \widehat{\lambda})$ be the suspension flow with roof function r. Let $\phi \colon \widehat{\mathcal{X}} \to \mathbb{R}$ be such that $\int \phi\, d\widehat{\lambda} = 0$, and define $\Phi(x) := \sum_{k=0}^{r(x)-1} \phi(x, k)$. Let $\phi \in L^b(\widehat{\mathcal{X}})$ and let $r \in L^a(\mathcal{X})$ be the roof function, with $(1 - 1/a)(1 - 1/b) \geq 1/2$. Suppose that Φ and r satisfy a CLT. Then ϕ satisfies a CLT.*

Moreover, if the CLT for Φ has variance σ_1^2, then the CLT for ϕ has variance $\sigma^2 = \sigma_1^2/\overline{r}$.

4.2 Invariant measure on the suspended space

For any $\omega \in \Omega_v$, let $r(\omega)$ be the length in Γ of $l_1(\omega)$. This is the first return time for the loop determined by ω. Let us define the suspension of the skew product

$$\Omega^{(s)} := \{(\omega, k, \xi) \in \Omega_v \times \mathbb{N} \times \mathcal{M} : 0 \leq k \leq r(\omega) - 1\}$$

and

$$\widehat{T}(\omega, k, \xi) = \begin{cases} (\omega, k+1, \xi) & \text{if } k \leq r(\omega) - 2, \\ (\sigma_v(\omega), 0, l_1^{-1}\xi) & \text{if } k = r(\omega) - 1. \end{cases}$$

1991

Let us now denote $R := \int_{\Gamma_v} \|g\| \, d\mu_v(g) = \int r(\omega) \, d\mathbb{P}_v(\omega)$ and define the probability measure $\nu^{(s)} := (1/R)(\mu_v^{\mathbb{N}} \otimes \delta \otimes \nu_v)$ on $\Omega^{(s)}$.

Lemma 4.3. *Let ν_v be the $\check{\mu}_v$-stationary measure constructed in Lemma 3.5. Then $\nu^{(s)}$ on $\Omega^{(s)}$ is $\widehat{T}$-invariant and ergodic.*

Proof. It suffices to check invariance of the measure using cylinder sets $C_{l_1,\dots,l_n}$ consisting of loops beginning with $l_1 \cdots l_n$. We have

$$\widehat{T}^{-1}(C_{l_1,\dots,l_n} \times \{k\} \times A) = \begin{cases} C_{l_1,\dots,l_n} \times \{k-1\} \times A & \text{if } k > 0, \\ \displaystyle\bigsqcup_{l \in P_v} C_{l,l_1,\dots,l_n} \times \{\|l\|-1\} \times lA & \text{if } k = 0, \end{cases}$$

where $P_v \subseteq \Gamma_v$ is the set of prime loops. Hence, in the first case, the equality

$$\nu^{(s)}(\widehat{T}^{-1}(C_{l_1,\dots,l_n} \times \{k\} \times A)) = \nu^{(s)}(C_{l_1,\dots,l_n} \times \{k\} \times A)$$

is obvious. In the second case,

$$\begin{aligned} \nu^{(s)}(\widehat{T}^{-1}(C_{l_1,\dots,l_n} \times \{k\} \times A)) &= \frac{1}{R} \sum_{l \in P_v} \mu_v(l) \mu_v(l_1) \cdots \mu_v(l_n) \nu_v(lA) \\ &= \frac{1}{R} \mu_v(l_1) \cdots \mu_v(l_n) \sum_{l \in P_v} \mu_v(l) \nu_v(lA) \\ &= \frac{1}{R} \mu_v(l_1) \cdots \mu_v(l_n) \nu_v(A) \\ &= \nu^{(s)}(C_{l_1,\dots,l_n} \times \{k\} \times A), \end{aligned}$$

hence $\nu^{(s)}$ is $\widehat{T}$-invariant. Moreover, the suspension of an ergodic measure is ergodic (see, for example, [Sar20, Proposition 1.11]). $\square$

4.3 Pushforward of the $\widehat{T}$-invariant measure to $\Omega \times \mathcal{M}$

Recall that Ω is the space of all infinite sample paths in Γ starting at any vertex. Let us define the projection $\pi : \Omega^{(s)} \to \Omega \times \mathcal{M}$ as

$$\pi(\omega, k, \xi) = (\sigma^k(\omega), (g_1 \cdots g_k)^{-1}\xi)$$

and recall that the skew product $T : \Omega \times \mathcal{M} \to \Omega \times \mathcal{M}$ is

$$T(\omega, \xi) := (\sigma(\omega), g_1^{-1}\xi).$$

Lemma 4.4. *The following diagram commutes.*

$$\begin{array}{ccc} & \overset{\widehat{T}}{\curvearrowright} & \overset{T}{\curvearrowright} \\ (\Omega^{(s)}, \nu^{(s)}) & \overset{\pi}{\longrightarrow} & \Omega \times \mathcal{M} \\ \downarrow & & \\ (\Omega_v \times \mathcal{M}, \mu_v^{\mathbb{N}} \otimes \nu_v) & & \end{array}$$

As a consequence, in the hypotheses of the previous lemmas, the measure $\overline{\nu} := \pi_\star \nu^{(s)}$ is T-invariant and ergodic.

Proof. We show that the horizontal arrow is equivariant for the shifts. This follows from the fact that if we write $l_1(\omega)$ for the first return loop of ω then $l_1(\omega) = g_1(\omega) \ldots g_{r(\omega)}(\omega)$. Hence,

$$\pi \circ \widehat{T}(\omega, r(\omega) - 1, \xi) = \pi((\sigma_v(\omega), 0, l_1^{-1}\xi))$$
$$= (\sigma^{r(\omega)}(\omega), (g_1 \cdots g_{r(\omega)})^{-1}\xi),$$

which is equal to $T \circ \pi((\omega, r(\omega) - 1, \xi))$. The other cases being trivial, this proves the first statement.

Finally, since $\overline{\nu}$ is the pushforward of an ergodic measure, it is ergodic. $\qquad\square$

4.4 Return times and invariant measures for the Markov chain

Recall that in the previous section we produced a measure ν_v on $\mathcal{M}$ which is $\check{\mu}_v$-stationary and such that the product measure $\mu_v^{\mathbb{N}} \otimes \nu_v$ is T_v-invariant and ergodic. Then, by lifting it to the suspension and pushing it forward to $\Omega \times \mathcal{M}$, we have an ergodic, T-invariant measure $\overline{\nu}$ on $\Omega \times \mathcal{M}$.

Now, for any vertex w other than v we define the measure ν_w on $\mathcal{M}$ as

$$\nu_w(A) = \sum_{\gamma \in \Gamma_{v,w}} \mu(\gamma)\nu_v(\gamma A) \tag{2}$$

where the sum is over the set $\Gamma_{v,w}$ of all paths γ from v to w which do not pass through v in their middle, and $\mu(\gamma)$ is the product of the measures of the edges of γ. Recall also we denote by $\mathbb{P}_w$ the Markov measure on the space of infinite sample paths starting at w.

LEMMA 4.5. *We have*

$$\overline{\nu} := \frac{1}{R} \sum_w \mathbb{P}_w \otimes \nu_w.$$

Proof. Let w be a vertex, and let $g_1, g_2, \ldots, g_n$ be a finite path starting from w. We have for any measurable $A \subseteq \mathcal{M}$

$$\pi^{-1}(C_{g_1,\ldots,g_n} \times A) = \begin{cases} C_{g_1,\ldots,g_n} \times \{0\} \times A & \text{if } w = v, \\ \bigsqcup_{\gamma \in \Gamma_{v,w}} C_{\gamma,g_1,\ldots,g_n} \times \{|\gamma|\} \times \gamma A & \text{if } w \neq v, \end{cases}$$

where the union is over the set $\Gamma_{v,w}$ of all paths γ from v to the initial vertex w of g_1 which do not pass through v in their middle. Thus we have

$$\overline{\nu}(C_{g_1,\ldots,g_n} \times A) = \frac{1}{R} \sum_{\gamma \in \Gamma_{v,w}} \mu(\gamma)\mu(g_1) \cdots \mu(g_n)\nu_v(\gamma A)$$
$$= \frac{1}{R}\mu(g_1) \cdots \mu(g_n)\nu_w(A)$$
$$= \frac{1}{R}\mathbb{P}_w(C_{g_1,\ldots,g_n})\nu_w(A)$$

which proves the claim, since both measures agree on all rectangles. $\qquad\square$

Recall that $R = \int r(\omega) \, d\mathbb{P}_v(\omega)$, and set $n_w = \nu_w(\mathcal{M})$. Here we show the following lemma.

LEMMA 4.6. *We have the identities*

(i) $R = 1/\pi_v$, *and*
(ii) $\pi_w = n_w/R$ *for any vertex w of Γ.*

Note that if we replace Γ with the graph $\overline{\Gamma}$ obtained by reversing the direction of each edge, then the transition matrix for $\overline{\Gamma}$ is M^T and so we have that ρ and u switch roles. In particular, the new transition probability from v_i to v_j is (in terms of the quantities defined in §4) $\overline{p}_{ij} = M_{ji}u_j/\lambda u_i$ but the stationary measure on vertices is unchanged.

Proof of Lemma 4.6. Assertion (i) is the well-known *Kac lemma* [Kac47, Theorem 2']. To prove (ii), recall that $\Gamma_{v,w}$ is the set of all paths γ from v to w which do not pass through v in their middle. Hence, if we reverse all the paths in this set, we obtain $\overline{\Gamma}_{v,w}$, the set of all paths $\overline{\gamma}$ from w to v which do not pass through v in their middle. Note that since almost every path starting at w passes through v,

$$1 = \sum_{\overline{\gamma} \in \overline{\Gamma}_{v,w}} \overline{\mu}(\overline{\gamma})$$

$$= \frac{u_v}{u_w} \sum_{\overline{\gamma} \in \overline{\Gamma}_{v,w}} \lambda^{-|\overline{\gamma}|},$$

where $\overline{\mu}(\overline{\gamma})$ is the product of the measures of the edges of $\overline{\gamma}$ with respect to $\overline{p}$ and we have used our previous observation about $\overline{p}$.

Using this and the fact that

$$\sum_{\overline{\gamma} \in \overline{\Gamma}_{v,w}} \lambda^{-|\overline{\gamma}|} = \sum_{\gamma \in \Gamma_{v,w}} \lambda^{-|\gamma|},$$

we compute

$$n_w = \nu_w(\mathcal{M})$$

$$= \sum_{\gamma \in \Gamma_{v,w}} \mu(\gamma) = \frac{\rho_w}{\rho_v} \sum_{\gamma \in \Gamma_{v,w}} \lambda^{-|\gamma|}$$

$$= \frac{\rho_w}{\rho_v} \cdot \frac{u_w}{u_v} = \frac{\pi_w}{\pi_v}.$$

Hence, the lemma follows from (i). □

4.5 The central limit theorem for the Markov chain

We are now in a position to prove Theorem 4.1. By Melbourne and Török [MT04, Theorem 1.1], we have the following proposition.

PROPOSITION 4.7. *Let $\phi\colon \Omega \times \overline{X}^h \to \mathbb{R}$ belong to $L^b(\Omega \times \overline{X}^h, \overline{\nu})$ for some $b > 2$, and let $m := \int \phi \, d\overline{\nu}$. Define $\Phi\colon \Omega_v \times \overline{X}^h \to \mathbb{R}$ as $\Phi(\omega, \xi) := \sum_{k=0}^{r(\omega)-1} \phi(T^k(\omega, \xi)) - mr(\omega)$, and suppose that*

$$\frac{\sum_{j=0}^{n-1} \Phi \circ T_v^j}{\sqrt{n}}$$

converges to a normal distribution in probability on $(\Omega_v \times \overline{X}^h, \mu_v^{\mathbb{N}} \otimes \nu_v)$. Then the sequence

$$\frac{\sum_{j=0}^{n-1} \phi \circ T^j - nm}{\sqrt{n}}$$

converges to a normal distribution in probability on $(\Omega \times \overline{X}^h, \overline{\nu})$.

Proof. Note that since r has exponential tail (see, for example, [GTT18, Lemma 3.2 (2)]), it belongs to $L^a(\Omega_v)$ for any $a \geq 1$. Then the condition $(1 - 1/a)(1 - 1/b) \geq 1/2$ is satisfied as

1994

long as $b > 2$. Moreover, $(r \circ T_v^n(\omega))_n$ is a sequence of independent random variables and so it satisfies a CLT. Hence, we can apply Theorem 4.2 to obtain a central limit theorem for the observable $\phi \circ \pi - m$ and the system $\widehat{T}$, with measure $\nu^{(s)}$. Moreover, since $\phi \circ \pi \circ \widehat{T}^n = \phi \circ T^n \circ \pi$ by Lemma 4.4, this is equivalent to a CLT for the observable ϕ on the system T with the measure $\pi_*(\nu^{(s)}) = \overline{\nu}$. $\qquad\square$

PROPOSITION 4.8. *There exist ℓ, σ such that for any continuous, compactly supported $F \colon \mathbb{R} \to \mathbb{R}$ one has*

$$\int_{\Omega \times \overline{X}^h} F\left(\frac{\beta_\xi(o, g_1 \cdots g_n o) - \ell n}{\sqrt{n}} \right) d\overline{\nu}(\omega, \xi) \to \int_{\mathbb{R}} F(t) \, d\mathcal{N}_\sigma(t),$$

as $n \to \infty$.

Proof. Let us apply the previous proposition with $\phi = f$ where $f \colon \Omega \times \overline{X}^h \to \mathbb{R}$ is defined as $f(\omega, \xi) := \beta_\xi(o, g_1 o)$. Then by definition of Φ and f, Lemma 3.6 gives that for every $\omega \in \Omega_v$,

$$\Phi(\omega, \xi) = \sum_{k=0}^{r(\omega)-1} f(T^k(\omega, \xi)) - \ell r(\omega)$$

$$= \beta_\xi(o, w_{r(\omega)} o) - \ell r(\omega) =: f_v(\omega, \xi),$$

where

$$\ell = m = \frac{\int \beta_\xi(o, go) \, d\mu_v(g) \, d\nu_v(\xi)}{\int \|g\| \, d\mu_v(g)}.$$

Now, by Corollary 3.3, integrating in $d\nu_v$, we have for some $\sigma_1 \geq 0$,

$$\int_{G \times \overline{X}^h} F\left(\frac{\beta_\xi(o, go) - \ell\|g\|}{\sqrt{n}} \right) d\mu_v^{*n}(g) \, d\nu_v(\xi) \to \int_{\mathbb{R}} F(t) \, d\mathcal{N}_{\sigma_1}(t).$$

Note, moreover, that $\beta_\xi(o, l_1 \cdots l_n o) - \ell\|l_1 \cdots l_n\| = \sum_{j=0}^{n-1} f_v(T_v^j(\omega, \xi))$, hence we can rewrite the above equation as

$$\int_{\Omega_v \times \overline{X}^h} F\left(\frac{\sum_{j=0}^{n-1} f_v(T_v^j(\omega, \xi))}{\sqrt{n}} \right) d(\mu_v^{\mathbb{N}} \otimes \nu_v)(\omega, \xi) \to \int_{\mathbb{R}} F(t) \, d\mathcal{N}_{\sigma_1}(t).$$

Thus, by Proposition 4.7 and the above calculation, we also have (for some different σ)

$$\int_{\Omega \times \overline{X}^h} F\left(\frac{\sum_{j=0}^{n-1} f \circ T^j(\omega, \xi) - \ell n}{\sqrt{n}} \right) d\overline{\nu}(\omega, \xi) \to \int_{\mathbb{R}} F(t) \, d\mathcal{N}_\sigma(t).$$

The claim follows by again using that, by Lemma 3.6, we have $\sum_{j=0}^{n-1} f \circ T^j(\omega, \xi) = \beta_\xi(o, g_1 \cdots g_n o)$. $\qquad\square$

We will now need to go from the CLT for the Busemann cocycle to that for displacement. To do so, we use the following variation of [BQ16, Proposition 3.3].

By [MT18, Proposition 4.4] and the fact that Γ_v^{-1} is nonelementary (Lemma 2.3), we have $\nu_v(\overline{X}_\infty^h) = 1$ for any vertex v.

LEMMA 4.9. *For any $\epsilon > 0$ there exists T such that for all vertices w in Γ, all $\xi \in \overline{X}_\infty^h$ and all $n \geq 1$ we have*

$$\mathbb{P}_w(\omega : |d(o, g_1 \cdots g_n o) - \beta_\xi(o, g_1 \cdots g_n o)| \leq T) \geq 1 - \epsilon.$$

1995

Proof. Recall that by [MT18, § 3.3] there exists a G-equivariant map $\pi\colon \overline{X}_\infty^h \to \partial X$, where ∂X is the Gromov boundary. Then, by definition of Gromov product and δ-hyperbolicity, we have

$$d(o, go) - \beta_\xi(o, go) = 2(go, \pi(\xi))_o + O(\delta) \tag{3}$$

for any $\xi \in \overline{X}_\infty^h$ (see, for example, [Hor18, Lemma 2.4]). Now, since the pushforward of the stationary measure $\mathbb{P}_w$ for the Markov chain starting at w to the Gromov boundary of X is not atomic [GTT20, Lemma 4.2], we have that for every $\epsilon > 0$ there exists T such that

$$\mathbb{P}_w\Big(\omega \in \Omega_w : \sup_{n \geq 1}(w_n o, \pi(\xi))_o \leq T \Big) \geq 1 - \epsilon$$

for all $\xi \in \overline{X}_\infty^h$ and for all w. This, combined with (3), yields the desired estimate. □

Proof of Theorem 4.1. Let $F\colon \mathbb{R} \to \mathbb{R}$ be continuous with compact support. Since F is uniformly continuous and by Lemma 4.9, for any $\eta > 0$ there exists n_0 such that for any $n \geq n_0$, any w and any $\xi \in \overline{X}_\infty^h$ one has

$$\left| F\left(\frac{d(o, w_n o) - \ell n}{\sqrt{n}} \right) - F\left(\frac{\beta_\xi(o, w_n o) - \ell n}{\sqrt{n}} \right) \right| < \eta$$

with probability $\mathbb{P}_w$ at least $1 - \epsilon$. Thus, since $\overline{\nu} = (1/R) \sum_w \mathbb{P}_w \otimes \nu_w$, for any $\eta > 0$ there exists n_0 such that for any $n \geq n_0$ we have

$$\left| F\left(\frac{d(o, g_1 \cdots g_n o) - \ell n}{\sqrt{n}} \right) - F\left(\frac{\beta_\xi(o, g_1 \cdots g_n o) - \ell n}{\sqrt{n}} \right) \right| < \eta \tag{4}$$

on a subset of $\Omega \times \overline{X}^h$ of $\overline{\nu}$-measure $\geq 1 - \epsilon$. On the other hand, by Proposition 4.8, we have

$$\int_{\Omega \times \overline{X}^h} F\left(\frac{\beta_\xi(o, g_1 \cdots g_n o) - \ell n}{\sqrt{n}} \right) d\overline{\nu}(\omega, \xi) \to \int_\mathbb{R} F(t) \, d\mathcal{N}_\sigma(t).$$

Hence,

$$\int_{\Omega \times \overline{X}^h} F\left(\frac{d(o, g_1 \cdots g_n o) - \ell n}{\sqrt{n}} \right) d\overline{\nu}(\omega, \xi) \to \int_\mathbb{R} F(t) \, d\mathcal{N}_\sigma(t).$$

Since the integrand does not depend on ξ, then we also have

$$\int_\Omega F\left(\frac{d(o, g_1 \cdots g_n o) - \ell n}{\sqrt{n}} \right) dP(\omega) \to \int_\mathbb{R} F(t) \, d\mathcal{N}_\sigma(t),$$

where P is the pushforward of $\overline{\nu}$ to Ω. Finally, since $\overline{\nu} = (1/R) \sum_w \mathbb{P}_w \otimes \nu_w$, the pushforward of $\overline{\nu}$ to Ω equals $P = \sum_w \frac{n_w}{R} \mathbb{P}_w$, where $n_w = \nu_w(M)$. Hence, Lemma 4.6 implies that $P = \sum_w \pi_w \mathbb{P}_w = \mathbb{P}$, thus we also have

$$\int_\Omega F\left(\frac{d(o, g_1 \cdots g_n o) - \ell n}{\sqrt{n}} \right) d\mathbb{P}(\omega) \to \int_\mathbb{R} F(t) \, d\mathcal{N}_\sigma(t)$$

as required. □

5. Uniqueness of drift and variance

Now suppose that Γ is a semisimple graph structure. In particular, each maximal component C_i of Γ gives a primitive graph structure (without an initial vertex) on G to which the results of the previous section (in particular, Theorem 4.1) apply. Hence, for each maximal component C_i of Γ, Theorem 4.1 gives constants ℓ_i and σ_i for the associated CLT.

In this section we show that the CLTs for the recurrent components of Γ are compatible in the sense that they have the same drift and variance. This is the primary place where we will use thickness and coherence of Γ. For this, first recall that Ω^* is the set of finite length paths in Γ starting at any vertex and that the graph structure is *bounded* if there exists a constant $\mathfrak{b} \geq 1$ so that the evaluation map $\mathrm{ev} \colon \Omega^* \to G$ is at worst $\mathfrak{b}$-to-one, that is, $\#(\mathrm{ev}^{-1}(g)) \leq \mathfrak{b}$ for all $g \in G$. Also recall the definition of coherence from Definition 2.2.

Remark 5.1. Our standing assumption until §7 is that Γ is a semisimple graph structure on G. This implies that the transition matrix for each component of maximal growth is primitive.

5.1 Uniformly bicontinuous functions

Let us begin by introducing a class of functions that are well behaved under bounded perturbations in the group.

DEFINITION 5.2. A function $f \colon \Omega^* \to \mathbb{R}$ is *uniformly bicontinuous* if, for any finite set $B \subseteq G$ and any $\eta > 0$, there exists $N \geq 0$ such that if $\|g\| \geq N$ and $b_1 \overline{g} b_2 = \overline{h}$ in G for some $b_1, b_2 \in B$, then

$$|f(g) - f(h)| < \eta.$$

In our application, the uniformly bicontinuous property will be a consequence of the fact that displacement is Lipschitz in both the right and left word metric on G. This definition is inspired by [CF10] (in particular, the proof of [CF10, Lemma 4.24]), where the bi-Lipschitz property is extensively used.

We next introduce the primary functions of interest used throughout this section. Define the following functions on Ω^*: for any $\ell \in \mathbb{R}$,

$$\varphi(g) := \frac{d(o, go) - \ell\|g\|}{\sqrt{\|g\|}}, \quad \psi(g) := \frac{d(o, go)}{\|g\|}.$$

We remark that the constant ℓ will be chosen once and for all after the proof of Lemma 5.6.

LEMMA 5.3. *If the graph structure is coherent, then the functions ψ and φ defined above are uniformly bicontinuous for any ℓ.*

Proof. Fix a finite set $B \subset G$, and let $\mathcal{B}$ be the resulting constant from Definition 2.2. Suppose that $\overline{h} = b_1 \overline{g} b_2$ for some $b_1, b_2 \in B$. Then $|\|g\| - \|h\|| \leq \mathcal{B}$ and by the triangle inequality $|d(o, go) - d(o, ho)| \leq \mathcal{B}_1$ where $\mathcal{B}_1 := 2\max_{b \in B} d(o, bo)$. Finally, denote by $\mathcal{L}$ the Lipschitz constant so that $d(o, go) \leq \mathcal{L}\|g\|$ for any $g \in \Omega^*$.

(1) By the above estimates,

$$
\begin{aligned}
\left| \frac{d(o, ho)}{\|h\|} - \frac{d(o, go)}{\|g\|} \right| &= \frac{|(d(o, ho) - d(o, go))\|g\| + d(o, go)(\|g\| - \|h\|)|}{\|g\|\|h\|} \\
&\leq \frac{|d(o, ho) - d(o, go)|}{\|h\|} + \frac{d(o, go)}{\|g\|} \cdot \frac{|\|g\| - \|h\||}{\|h\|} \\
&\leq \frac{\mathcal{B}_1 + \mathcal{L}\mathcal{B}}{\|g\| - \mathcal{B}},
\end{aligned}
$$

and the right-hand side tends to 0 as $\|g\| \to \infty$.

(2) We can write

$$|\varphi(g) - \varphi(h)| = \frac{x}{\sqrt{n}} - \frac{x+y}{\sqrt{n+d}}$$

where $x = d(o, go) - \ell\|g\|$, $y = d(o, ho) - \ell\|h\| - d(o, go) + \ell\|g\|$, $n = \|g\|$, and $d = \|h\| - \|g\|$.

Recall that by the above inequalities $|d| \le \mathcal{B}$, hence also $|y| \le \mathcal{B}_1 + \ell\mathcal{B}$ and $|x| \le (\mathcal{L} + \ell)\|g\|$. Thus,

$$
\begin{aligned}
\left| \frac{x}{\sqrt{n}} - \frac{x+y}{\sqrt{n+d}} \right| &= \left| \frac{x(\sqrt{n+d} - \sqrt{n})}{\sqrt{n(n+d)}} - \frac{y}{\sqrt{n+d}} \right| \\
&\le \frac{|x|}{n} \frac{n}{\sqrt{n(n+d)}} |\sqrt{n+d} - \sqrt{n}| + \frac{|y|}{\sqrt{n+d}} \\
&\le (\mathcal{L} + \ell) \frac{n}{\sqrt{n(n-\mathcal{B})}} (\sqrt{n+\mathcal{B}} - \sqrt{n}) + \frac{\mathcal{B}_1 + \ell\mathcal{B}}{\sqrt{n-\mathcal{B}}}
\end{aligned}
$$

and the right-hand side tends to 0 uniformly in n. $\qquad\square$

Remark 5.4 (Logarithmic perturbations). As a consequence of the proof that φ is uniformly bicontinuous, we observe that for any $\eta > 0$ there is an N such that if $\|g\| \ge N$ then for any decomposition $g = g_0 g_1 g_2$ with $\|g_0\|, \|g_2\| \le \log N$ we have

$$|\varphi(g) - \varphi(g_1)| < \eta.$$

The main reason why we introduce the bicontinuous functions is the following property. For each i, we denote by $\mu_n^{(i)}$ the distribution on the space of paths of length n induced by the Markov measure associated to the maximal component C_i, as defined in §4. In particular, $\mu_n^{(i)}$ is supported on length n paths of Γ that are contained in C_i.

LEMMA 5.5. *Suppose that the graph structure Γ is thick and coherent. Let $f \colon \Omega^* \to \mathbb{R}$ be a uniformly bicontinuous function, and suppose that for each maximal component C_i there is a finite measure $\mathcal{D}_i$ on $\mathbb{R}$ so that*

$$f(w_n) \to \mathcal{D}_i$$

in distribution with respect to the Markov measure on C_i. If each $\mathcal{D}_i$ is nonatomic, then there exists $a \ge 1$ such that for any interval $I \subset \mathbb{R}$,

$$\frac{1}{a} \mathcal{D}_i(I) \le \mathcal{D}_j(I) \le a\mathcal{D}_i(I).$$

Similarly, if each $\mathcal{D}_i$ is a Dirac measure, then $\mathcal{D}_i = \mathcal{D}_j$.

Proof. Let v be a vertex of C_i, and let C_j be another maximal component. Let $B \subset G$ be the finite set given by thickness of Γ (Definition 2.1) and let $\mathcal{B}$ be the resulting constant from Definition 2.2. Similar to the proof of Lemma 2.3, by thickness and coherence we have

$$\overline{S_n^{(j)}} \subseteq \bigcup_{b_1, b_2 \in B} \bigcup_{|k| \le \mathcal{B}} b_1(\overline{\Gamma_v \cap S_{n+k}})b_2$$

where $S_n^{(j)}$ is the set of paths of length n which lie entirely in C_j. Hence, by uniform continuity, for any $\epsilon > 0$ there exists n_0 such that for all $n \ge n_0$,

$$\#\{g \in S_n^{(j)} : f(g) \in [x, y]\} \le \mathfrak{b}|B|^2 \sum_{|k| \le \mathcal{B}} \#\{g \in \Gamma_v \cap S_{n+k} : f(g) \in [x - \epsilon, y + \epsilon]\},$$

1998

for any interval $I = [x, y]$. Here, we have used that the evaluation map is at most $\mathfrak{b}$-to-one. Now, note that there exists $C > 0$ such that

$$C^{-1} \frac{\#(A \cap S_n^{(i)})}{\lambda^n} \leq \mu_n^{(i)}(A) \leq C \frac{\#(A \cap S_n^{(i)})}{\lambda^n} \tag{5}$$

for any i, any n and any set A. Hence, by noting that $\Gamma_v \cap S_{n+k} \subseteq S_{n+k}^{(i)}$,

$$\mu_n^{(j)}(g : f(g) \in [x, y]) \leq \mathfrak{b}|B|^2 C^2 \sum_{|k| \leq \mathcal{B}} \mu_{n+k}^{(i)}(g : f(g) \in [x - \epsilon, y + \epsilon]).$$

Now taking limits and using the portmanteau theorem, we obtain that

$$\mathcal{D}_j((x, y)) \leq 2\mathcal{B}\mathfrak{b}|B|^2 C^2 \mathcal{D}_i([x - \epsilon, y + \epsilon]),$$

for any $\epsilon > 0$. So if $c = 2\mathcal{B}\mathfrak{b}|B|^2 C^2$, we have that $\mathcal{D}_j((x, y)) \leq c\mathcal{D}_i([x, y])$ because $\mathcal{D}_i([x, y]) = \bigcap_{\epsilon > 0} \mathcal{D}_i([x - \epsilon, y + \epsilon])$ for any finite measure.

Now if each $\mathcal{D}_i$ is nonatomic, we get $\mathcal{D}_j([x, y]) \leq c\mathcal{D}_i([x, y])$, and reversing the roles of i and j completes the proof in this case. Similarly, if $\mathcal{D}_i$ and $\mathcal{D}_j$ were distinct Dirac measures, we would obtain a contradiction by letting (x, y) be a small interval about the atom for $\mathcal{D}_j$ and taking ϵ small enough so that $[x - \epsilon, y + \epsilon]$ does not contain the atom for $\mathcal{D}_i$. This completes the proof. $\square$

5.2 Uniqueness of drift

We now show that all maximal components of Γ determine the same drift.

LEMMA 5.6. *If the graph structure Γ is thick and coherent, then all ℓ_i for each maximal component C_i are the same.*

Proof. By the subadditive ergodic theorem, for each maximal component C_i we have

$$\lim_{n \to \infty} \frac{d(o, w_n o)}{n} = \ell_i$$

almost surely (and hence in distribution) with respect to the measure $\mu_n^{(i)}$. Since $\psi(g) = d(o, go)/\|g\|$ is uniformly bicontinuous by Lemma 5.3, the claim then follows by Lemma 5.5. $\square$

By the above lemma, we now define φ using $\ell = \ell_i$ for any (equivalently all) i.

5.3 Uniqueness of variance

To show that all maximal components of Γ determine the same variance, we use the simple fact that distinct normal distributions with mean 0 can be distinguished by the decay of their tails.

LEMMA 5.7. *If the graph structure Γ is thick and coherent, then for any two maximal components C_i and C_j we have $\sigma_i = \sigma_j$.*

Proof. We have already shown that for ℓ fixed as above and any maximal component C_i of Γ,

$$\phi(w_n) \to \mathcal{N}_{\sigma_i}$$

in distribution with respect to the Markov measure on C_i (Theorem 4.1). Since ϕ is uniformly bicontinuous by Lemma 5.3, Lemma 5.5 states that there exists an $a \geq 1$ so that for any interval I in $\mathbb{R}$,

$$\frac{1}{a} \int_I d\mathcal{N}_{\sigma_i} \leq \int_I d\mathcal{N}_{\sigma_j} \leq a \int_I d\mathcal{N}_{\sigma_i}.$$

But since this holds for every interval, Lemma 5.8 implies that $\sigma_i = \sigma_j$ as required. $\square$

1999

Lemma 5.8. *If there exists $a \geq 1$ such that*

$$\frac{1}{a} \int_I d\mathcal{N}_{\sigma_i} \leq \int_I d\mathcal{N}_{\sigma_j} \leq a \int_I d\mathcal{N}_{\sigma_i} \tag{6}$$

for any interval I, then $\sigma_i = \sigma_j$.

Proof. If we let

$$I_{n,\sigma} := \int_n^{n+1} d\mathcal{N}_\sigma = \frac{1}{\sqrt{2\pi}\sigma} \int_n^{n+1} e^{-(t^2/2\sigma^2)} \, dt$$

we get

$$\frac{1}{\sqrt{2\pi}\sigma} e^{-((n+1)^2/2\sigma^2)} \leq I_{n,\sigma} \leq \frac{1}{\sqrt{2\pi}\sigma} e^{-(n^2/2\sigma^2)}$$

which yields

$$\lim_{n\to\infty} \frac{\log I_{n,\sigma}}{n^2} = -\frac{1}{2\sigma^2}. \tag{7}$$

Then, if (6) holds, then

$$\frac{1}{a} \leq \frac{I_{n,\sigma_i}}{I_{n,\sigma_j}} \leq a$$

hence, by (7)

$$-\frac{1}{2\sigma_i^2} = \lim_{n\to\infty} \frac{\log I_{n,\sigma_i}}{n^2} = \lim_{n\to\infty} \frac{\log I_{n,\sigma_j}}{n^2} = -\frac{1}{2\sigma_j^2}$$

which yields $\sigma_i = \sigma_j$. □

6. The semisimple case

In this section we prove our main theorem for semisimple graph structures. This is completed in Theorem 6.3.

6.1 Convergence to the Markov measure

So far our work has been for maximal components of a semisimple graph structure. In this section we consider the whole graph structure, still in the semisimple case.

Let Γ be a semisimple graph structure for G with transition matrix M of spectral radius $\lambda > 1$. Let v_i be the vertices of the graph, and let v_0 be a vertex of large growth, which we take as the initial vertex. Then recall that $e_i^T M^n e_j$ is the number of paths of length n from v_i to v_j. Since M is semisimple, the limit

$$M_\infty := \lim_{n\to\infty} \frac{M^n}{\lambda^n}$$

exists. In particular, in keeping with notation at the beginning of § 4, we denote by e_i the ith vector of the standard basis, and define

$$\rho_i := \lim_{n\to\infty} \frac{e_i^T M^n 1}{\lambda^n} \quad \text{and} \quad u_i := \lim_{n\to\infty} \frac{e_0^T M^n e_i}{\lambda^n}.$$

By construction, $\rho = (\rho_i)$ satisfies $\rho = M_\infty 1$ and $M\rho = \lambda\rho$, while $u = (u_i)$ satisfies $u^T M = \lambda u^T$. Finally, $\sum_i u_i = \rho_0$ and $\sum_i u_i \rho_i = \rho_0$.

Note that vertices v_i for which $\rho_i > 0$ and $u_i > 0$ are precisely vertices of components of maximal growth. The large growth vertices are those with $\rho_i > 0$.

As before, we use a standard construction to define a Markov measure $\mathbb{P}$ on the space Ω of infinite paths starting at any vertex of Γ. First define the initial distribution of the Markov chain to start at vertex v_i with probability $\pi_i := u_i \rho_i / \rho_0$. Then assign to an edge from v_i to v_j the probability $\rho_j / \lambda \rho_i$ so that the transition probability from v_i to v_j is $p_{ij} := M_{ij} \rho_j / \lambda \rho_i$. Obviously, $\mathbb{P}$ is supported on paths that are entirely contained in components of maximal growth. We denote by $\mathbb{P}_n$ the distribution on the space of paths of length n induced by the Markov measure $\mathbb{P}$.

Remark 6.1. We remark that the induced measure on each maximal component C of Γ rescales to give the Markov measure on C previously considered. This follows immediately from the construction.

The following result relates the Markov measure on the semisimple graph structure to the counting measure. For its statement, let v_0 be any vertex of large growth. For each n, consider the path given by selecting uniformly a path γ starting at v_0 of length n, and take its subpath $\tilde{\gamma}$ from position $\lfloor \log n \rfloor$ to position $n - \lfloor \log n \rfloor$. To avoid writing the integer part every time, we set $\lg(n) := \lfloor \log n \rfloor$. Let $\tilde{\lambda}_n$ denote the distribution of $\tilde{\gamma}$.

LEMMA 6.2. *With notation as above, the total variation*

$$\|\mathbb{P}_{n-2\lg(n)} - \tilde{\lambda}_n\|_{TV} \to 0$$

as $n \to \infty$.

Proof. Denote $n' := n - 2\lg(n)$. Let γ be a path in the graph, starting at v_i and ending at v_j. Then by definition the proportion of paths of length n, starting at v_0, that have γ as 'middle subpath' of length n' is

$$\tilde{\lambda}_n(\gamma) = \frac{(e_0^T M^{\lg(n)} e_i)(e_j^T M^{\lg(n)} 1)}{e_0^T M^n 1}.$$

On the other hand,

$$\mathbb{P}_{n'}(\gamma) = \begin{cases} \dfrac{\pi_i \rho_j}{\rho_i \lambda^{n'}} & \text{if } v_i \text{ has large growth,} \\ 0 & \text{otherwise,} \end{cases}$$

which is nonzero if both v_i and v_j belong to a maximal component. In this case,

$$\frac{d\mathbb{P}_{n'}}{d\tilde{\lambda}_n}(\gamma) = \frac{\lambda^{\lg(n)}}{e_0^T M^{\lg(n)} e_i} \cdot \frac{\lambda^{\lg(n)}}{e_j^T M^{\lg(n)} 1} \cdot \frac{e_0^T M^n 1}{\lambda^n} \cdot \frac{\pi_i \rho_j}{\rho_i}$$

$$\longrightarrow \frac{1}{u_i} \cdot \frac{1}{\rho_j} \cdot \frac{\rho_0}{1} \cdot \frac{\pi_i \rho_j}{\rho_i} = 1$$

using that $\pi_i = u_i \rho_i / \rho_0$. Moreover, if $S_n^{i,j}$ denotes the set of paths of length n' from v_i to v_j, we have

$$\tilde{\lambda}_n(S_n^{i,j}) = \frac{(e_i^T M^{n'} e_j)(e_0^T M^{\lg(n)} e_i)(e_j^T M^{\lg(n)} 1)}{e_0^T M^n 1}$$

$$\leq \frac{(e_i^T M^{n'} 1)(e_0^T M^{\lg(n)} e_i)(e_j^T M^{\lg(n)} 1)}{e_0^T M^n 1}$$

$$\to \frac{\rho_i u_i \rho_j}{\rho_0},$$

hence such a probability tends to 0 unless both v_i and v_j belong to a maximal component.

2001

Finally, if we denote by $\mathcal{L}_n$ the set of paths of length n' which lie entirely in a maximal component, we have for any set A,

$$|\mathbb{P}_{n'}(A) - \tilde{\lambda}_n(A)| \leq \sum_{x \in A \cap \mathcal{L}_n} \left| \frac{\mathbb{P}_{n'}(x)}{\tilde{\lambda}_n(x)} \tilde{\lambda}_n(x) - \tilde{\lambda}_n(x) \right| + \tilde{\lambda}_n(A \setminus \mathcal{L}_n)$$

$$\leq \sup_{x \in \mathcal{L}_n} \left| \frac{\mathbb{P}_{n'}(x)}{\tilde{\lambda}_n(x)} - 1 \right| + \tilde{\lambda}_n(\mathcal{L}_n^c)$$

and both terms tend to 0 as $n \to \infty$, independently of A. $\qquad\square$

6.2 Central limit theorem for the counting measure in the semisimple case

We are now ready to prove the following result. For its statement, let S_n denote the set of length n paths beginning at the initial vertex v_0.

THEOREM 6.3. *Let Γ be a semisimple, thick, coherent graph structure for a nonelementary group G of isometries of a δ-hyperbolic space (X, d), and let $o \in X$ be a base point. Then there exist $\ell \geq 0$, $\sigma \geq 0$ such that for any $a < b$ we have*

$$\lim_{n \to \infty} \frac{1}{\#S_n} \# \left\{ g \in S_n : \frac{d(o, go) - \ell n}{\sqrt{n}} \in [a, b] \right\} = \int_a^b d\mathcal{N}_\sigma(t).$$

In the following proof and later on, we will use the notation $N_\sigma(x) := \int_{-\infty}^x d\mathcal{N}_\sigma(t)$.

Remark 6.4. Note that if the graph structure Γ is semisimple and has a unique maximal component, then Theorem 6.3 holds even without assuming that the structure is coherent.

Proof. Let $C_1, \ldots, C_k$ be the maximal components, and let $\mu_n^{(i)}$ be the distribution on the space of paths of length n induced by the Markov measure associated to that component, as in §4. Theorem 4.1 shows a CLT for all such measures, and by Lemmas 5.6 and 5.7 all such measures have the same drift and variance, which we denote by ℓ and σ.

Now, since the starting probability (π_i) in the above construction is nonzero precisely on the set of vertices which belong to a maximal component, there exist weights $c_i \geq 0$ with $\sum_i c_i = 1$ such that

$$\mathbb{P}_n = \sum_{i=1}^k c_i \mu_n^{(i)}$$

for any n. Thus, for any $x \in \mathbb{R}$,

$$\mathbb{P}_n(g : \varphi(g) \leq x) = \sum_{i=1}^k c_i \mu_n^{(i)}(g : \varphi(g) \leq x) \to N_\sigma(x), \tag{8}$$

where we recall that

$$\varphi(g) = \frac{d(o, go) - \ell \|g\|}{\sqrt{\|g\|}}.$$

We now use that the counting measure can be approximated by the distribution on finite paths for the Markov chain. If g is a path of length n, we denote $g = g_0 g_1 g_2$ where g_0 is the prefix of length $\lg(n)$, g_1 is the middle part of length $n - 2\lg(n)$ and g_2 is the final part of length $\lg(n)$. By Remark 5.4, there exists n_0 such that

$$|\varphi(g) - \varphi(g_1)| \leq \epsilon \tag{9}$$

for any $n \geq n_0$ and g with $\|g\| = n$.

2002

Fix $x \in \mathbb{R}$ and $\epsilon > 0$. Then we have

$$
\begin{aligned}
\lambda_n(g : \varphi(g) \leq x) &= \lambda_n(g = g_0 g_1 g_2 : \varphi(g) \leq x) \\
&\leq \lambda_n(g = g_0 g_1 g_2 : \varphi(g_1) \leq x + \epsilon) && \text{by (9), for } n \text{ large} \\
&= \tilde{\lambda}_n(g_1 : \varphi(g_1) \leq x + \epsilon) && \text{by definition of } \tilde{\lambda}_n \\
&\leq \mathbb{P}_{n-2\lg(n)}(g_1 : \varphi(g_1) \leq x + \epsilon) + \epsilon && \text{by Lemma 6.2, for } n \text{ large.}
\end{aligned}
$$

Hence, by (8) we obtain

$$
\limsup_{n \to \infty} \lambda_n(g : \varphi(g) \leq x) \leq N_\sigma(x + \epsilon) + \epsilon
$$

and, by taking ϵ smaller and smaller and using the continuity of N_σ,

$$
\limsup_{n \to \infty} \lambda_n(g : \varphi(g) \leq x) \leq N_\sigma(x).
$$

The lower bound follows analogously. $\qquad\square$

Indeed, the same proof shows the following stronger statement. Let $\lambda_n^{(l)}$ denote the counting measure on the set of paths of length n starting at v_l.

COROLLARY 6.5. *Let Γ be a semisimple, thick, coherent graph structure for a nonelementary group G of isometries of a δ-hyperbolic space (X, d), and let $o \in X$ be a base point. Then there exist $\ell \geq 0$, $\sigma \geq 0$ such that for any vertex v_l of large growth for Γ and any $a < b$ we have*

$$
\lim_{n \to \infty} \lambda_n^{(l)}\left(g : \frac{d(o, go) - \ell \|g\|}{\sqrt{\|g\|}} \in [a, b] \right) = \int_a^b dN_\sigma(t).
$$

Proof. Let us fix a vertex v_l of large growth for M. Then we can define a Markov measure $\mathbb{P}^{(l)}$ on the space of infinite paths as follows. The transition probabilities will always be the same $p_{ij} = \rho_j / \lambda^p \rho_i$, while for each vertex v_l one finds a different set of starting probabilities $\pi_i^{(l)}$ given by

$$
\pi_i^{(l)} := \frac{u_i^{(l)} \rho_i}{\rho_l}, \quad \text{where} \quad u_i^{(l)} := \lim_{n \to \infty} \frac{e_l^T M^n e_i}{\lambda^n}.
$$

Just as before, there exist constants $c_i^{(l)} \geq 0$ such that $\sum_i c_i^{(l)} = 1$ and

$$
\mathbb{P}_n^{(l)} = \sum_{i=1}^k c_i^{(l)} \mu_n^{(i)}.
$$

The proof then proceeds exactly as for Theorem 6.3. $\qquad\square$

7. The central limit theorem for displacement and translation length

Now suppose that Γ is an almost semisimple graph structure for G with transition matrix M. Then M has some period $p \geq 1$ such that M^p is semisimple. We denote by Γ^p the corresponding p-step graph structure on G. That is, Γ^p is the graph with the same vertex set as Γ and an edge joining v_i to v_j for each directed path from v_i to v_j of length p, whose label is the word in G spelled by the corresponding path. The transition matrix for Γ^p is M^p, hence Γ is a semisimple graph structure for G.

Since the previous results require this structure to be thick and coherent, we need the following lemma.

Lemma 7.1. *The following properties pass to the p-step graph structure.*

- *If v is a large growth vertex of Γ, then it is also a large growth vertex of Γ^p.*
- *If Γ is a thick structure, then Γ^p is also thick.*
- *If Γ is coherent, then so is Γ^p.*

Proof. The first statement holds because any path from v that ends in a component of maximal growth can be extended to a path whose length is a multiple of p by adding on a path in that component of length less than p.

Now suppose that Γ is thick. Let v be a vertex in a maximal component of Γ^p. Then v is also a vertex in a maximal component of Γ. Let $\Gamma_{v,p}$ be the semigroup of loops based at v of lengths multiple of p. Consider the semigroup homomorphism

$$f \colon \Gamma_v \to \mathbb{N} \to \mathbb{N}/p\mathbb{N}$$

given by taking the length and reducing it mod p. Clearly, the image of f is a subsemigroup of $\mathbb{N}/p\mathbb{N}$, which is a finite group, hence the image contains the inverse of each element and so is also a group. Let $\gamma_i, \ldots, \gamma_k \subseteq \Gamma_v$ be a set of representatives for each remainder class in the image of f. Now let $\gamma \in \Gamma_v$. Then $\|\gamma\|$ belongs to the image of f, hence there exists γ_i (the representative of the inverse modulo p), such that $\gamma\gamma_i$ has length a multiple of p, hence it belongs to $\Gamma_{v,p}$. Hence, by setting B' the set $\{\gamma_i^{-1} : 1 \leq i \leq k\}$, we have $\Gamma_v \subseteq \Gamma_{v,p}B'$ in the group. Since Γ is thick, there exists B'' such that $G = B''\Gamma_v B''$, hence also $G = B''\Gamma_{v,p}B'B''$, hence Γ^p is also thick.

Finally, note that any path g in Γ^p of length k can be naturally thought of as a path $g_\dagger$ in Γ of length pk such that, for all $i \geq 0$, $\overline{g_\dagger(pi)} = \overline{g}(i)$. From this, it follows that Γ^p is bounded whenever Γ is. Moreover, if $B \subset G$ and $\mathcal{B} \geq 0$ are as in the definition of coherence of Γ (Definition 2.2) and g, h are paths in Γ^p with $\overline{g} = b_1 \overline{h} b_2$ for $b_1, b_2 \in B$, then we also have that $\overline{g_\dagger} = b_1 \overline{h_\dagger} b_2$. Then coherence of Γ implies that

$$p \cdot |\|g\| - \|h\|| = |\|g_\dagger\| - \|h_\dagger\|| \leq \mathcal{B},$$

and so coherence of Γ^p follows. $\qquad\square$

Now let us consider the semisimple matrix M^p. Note that irreducible components of M^p may be proper subsets of irreducible components of M. Given a vertex v_i, let us denote by $\lambda_k^{(i)}$ the counting measure on paths starting at v_i of length k for Γ. Note that if $k = np$, then this also counts paths of length n in Γ^p starting at v_i.

Since a bounded, thick graph structure Γ is almost semisimple (Lemma 2.3), by applying Corollary 6.5 to Γ^p, we immediately obtain the following result.

Corollary 7.2. *Let Γ be a thick, coherent structure of period p for a nonelementary group G of isometries of a δ-hyperbolic space (X, d), and let $o \in X$ be a base point. Then there exist ℓ, σ such that the following assertion holds. For any vertex v_i of large growth for Γ and for any x, we have*

$$\lambda_{pn}^{(i)}\left(g : \frac{d(o, go) - \ell\|g\|}{\sqrt{\|g\|}} \leq x\right) \to \int_{-\infty}^{x} d\mathcal{N}_\sigma(t)$$

as $n \to \infty$.

We are now ready to prove the following theorem. Recall that S_n denotes the set of length n paths beginning at the initial vertex v_0.

Theorem 7.3. *Let Γ be a thick, coherent graph structure for a nonelementary group G of isometries of a δ-hyperbolic space (X, d), and let $o \in X$ be a base point. Then there exist $\ell \geq 0$,*

$\sigma \geq 0$ *such that for any $a < b$ we have*

$$\lim_{n \to \infty} \frac{1}{\#S_n} \# \left\{ g \in S_n : \frac{d(o, go) - \ell n}{\sqrt{n}} \in [a, b] \right\} = \int_a^b d\mathcal{N}_\sigma(t).$$

Proof. Let v_0 be the initial vertex, let S_n be the set of paths of length n based at v_0, and let λ_n be the uniform measure on S_n.

Let us fix $0 \leq r \leq p - 1$. Then we can write the counting measure on S_{pn+r}, starting at the initial vertex v_0, by first picking randomly a path g_0 of length r from v_0 with a certain probability μ, and then picking a random path starting at $v_i = t(g_0)$ with respect to the counting measure on the set of paths of length n starting at v_i.

To compute μ, let us consider a path g_0 of length r starting at v_0 and ending at v_i. Then, if v_i is of large growth for Γ^p,

$$\frac{\#\{\text{paths from } v_i \text{ of length } pn\}}{\#\{\text{paths from } v_0 \text{ of length } pn + r\}} = \frac{e_i M^{pn} 1}{e_0 M^{pn+r} 1} \to \frac{e_i M_\infty 1}{e_0 M^r M_\infty 1}.$$

Thus, we define

$$\mu(g_0) := \frac{e_i M_\infty 1}{e_0 M^r M_\infty 1}.$$

Note that $\mu(g_0) = 0$ if the end vertex of g_0 has small growth and, moreover,

$$\sum_{\|g_0\| = r} \mu(g_0) = \sum_i \mu(g_0) \#\{g_0 \in S_r : t(g_0) = v_i\} = \sum_i e_0 M^r e_i \frac{e_i M_\infty 1}{e_0 M^r M_\infty 1} = 1.$$

Let λ'_{pn+r} be the measure on S_{pn+r} given by first taking randomly a path g_0 of length r from v_0 with distribution μ and then taking uniformly a path of length pn starting from $t(g_1)$.

Now we show that the CLT holds for λ'_{pn+r}. Let ℓ, σ be given by Corollary 7.2, and let $\varphi(g) := (d(o, go) - \ell\|g\|)/\sqrt{\|g\|}$. By Corollary 7.2, for any vertex v_i of large growth, we have

$$\lambda_{pn}^{(i)}(g : \varphi(g) \leq x) \to N_\sigma(x).$$

Then if $g = g_0 g_1$, and $t(g_0)$ denotes the (index of the) end vertex of g_0,

$$\lambda'_{pn+r}(g : \varphi(g) \leq x) = \sum_{g_0 \in S_r} \mu(g_0) \lambda_{pn}^{(t(g_0))}(g_1 : \varphi(g_0 g_1) \leq x)$$

$$\longrightarrow \sum_{g_0} \mu(g_0) N_\sigma(x) = N_\sigma(x),$$

where we used that φ is uniformly bicontinuous as in the proof of Theorem 6.3.

Now we prove that

$$\|\lambda'_{pn+r} - \lambda_{pn+r}\|_{TV} \to 0$$

as $n \to \infty$. Indeed, if $\gamma = g_0 g_1$ is a path from v_0 of length $pn + r$ and g_0 is its prefix of length r ending at a vertex v_i of large growth, then

$$\frac{\lambda'_{pn+r}(\gamma)}{\lambda_{pn+r}(\gamma)} = \frac{\mu(g_0) \cdot (1/e_i M^{pn} 1)}{1/e_0 M^{pn+r} 1} \to 1.$$

On the other hand, if the end vertex of g_0 is of small growth, then $\lambda'_{pn+r}(g) = 0$, and also

$$\lambda_{pn+r}(g = g_0 g_1 : g_0 \text{ ends at a small growth vertex}) \to 0$$

2005

as $n \to \infty$. Now, let $A_x := \{g : \varphi(g) \le x\}$ and L_r be the set of paths starting at v_0 whose prefix of length r ends in a vertex of large growth. Then

$$\lambda_{pn+r}(g : \varphi(g) \le x) = \lambda_{pn+r}(g \in L_r : \varphi(g) \le x) + \lambda_{pn+r}(g \notin L_r : \varphi(g) \le x)$$
$$= \frac{\lambda_{pn+r}(A_x \cap L_r)}{\lambda'_{pn+r}(A_x \cap L_r)} \lambda'_{pn+r}(A_x \cap L_r) + \lambda_{pn+r}(A_x \setminus L_r)$$
$$\to 1 \cdot N_\sigma(x) + 0 = N_\sigma(x).$$

We have thus obtained a CLT for λ_{pn+r}, for any $0 \le r \le p-1$, always with the same ℓ, σ. Since there are only finitely many values r, the claim follows. $\square$

7.1 A central limit theorem for translation length
We now prove a more general version of our second main result, Theorem 1.1(ii).

THEOREM 7.4. *Let Γ be a thick, coherent graph structure for a nonelementary group G of isometries of a δ-hyperbolic space (X,d), let $o \in X$ be a base point, and let ℓ, σ be as in Theorem 7.3. Then for any $a < b$ we have*

$$\lim_{n\to\infty} \frac{1}{\#S_n} \#\left\{ g \in S_n : \frac{\tau(g) - \ell n}{\sqrt{n}} \in [a,b] \right\} = \int_a^b d\mathcal{N}_\sigma(t).$$

Proof. Let us recall that the translation length of an isometry g of a δ-hyperbolic space can be computed by (see, for example, [MT18, Proposition 5.8])

$$\tau(g) = d(o, go) - 2(go, g^{-1}o)_o + O(\delta) \tag{10}$$

where $O(\delta)$ is a constant which only depends on the hyperbolicity constant of X. Now, by choosing $f(n) = \epsilon\sqrt{n}$ in [GTT20, Proposition 5.8], for any ϵ we have

$$\lambda_n(g : (go, g^{-1}o)_o \le \epsilon\sqrt{n}) \to 1$$

as $n \to \infty$. The claim then follows by combining this statement and the statement of Theorem 7.3 into formula (10). $\square$

7.2 Zero variance
We finally complete our main theorem by characterizing the case where $\sigma = 0$. First, we give a general criterion.

PROPOSITION 7.5. *In the hypotheses of Theorem 7.3 we have $\sigma = 0$ if and only if there is $C \ge 0$ such that for all finite length paths g in Γ,*

$$|d(o, go) - \ell \|g\|| \le C.$$

We note that the proposition implies that $\sigma > 0$ whenever the action $G \curvearrowright X$ is nonproper.

Proof. Suppose that $\sigma = 0$ for the CLT for the counting measure. Then, by our previous discussion, we have $\sigma = 0$ also for the Markov chain on any maximal components. Then, by Theorem 4.2, we also have $\sigma = 0$ for the random walk on the loop semigroup driven by $\check{\mu}_v$. Hence, as in [BQ16, Proof of Theorem 4.7 (b)], for any n,

$$\frac{1}{n} \int (\eta_0(g, \xi))^2 d\check{\mu}_v^{*n}(g) d\nu_v(\xi) = 0$$

where η_0 is the centering of η. This implies

$$\eta_0(g, \xi) = 0$$

for any $g \in \Gamma_v^{-1}$ and ν_v-a.e. $\xi \in \overline{X}^h$. Thus, since $|\eta - \eta_0| \leq 2\|\psi\|_\infty$ is bounded, we have

$$|\beta_\xi(o, g^{-1}o) - \ell\|g\|| = |\eta(g, \xi)| \leq 2\|\psi\|_\infty,$$

hence by [Hor18, Corollary 2.3] there exists a constant C for which

$$|d(o, go) - \ell\|g\|| \leq C$$

for any g in the support of $\check{\mu}_v^{*n}(g)$.

Hence, by thickness we have that for any $g \in \Omega^*$ there exist $b_1, b_2 \in B$ and $h \in \Gamma_v^{-1}$ such that $\overline{h} = b_1 \overline{g} b_2$, thus by coherence and the triangle inequality

$$|d(o, go) - \ell\|g\|| \leq |d(o, ho) - \ell\|h\|| + \mathcal{B}_1 + \ell\mathcal{B},$$

thus there exists a constant C' such that

$$|d(o, go) - \ell\|g\|| \leq C'$$

for any $g \in \Omega^*$. This completes the proof. $\qquad \square$

We conclude with a corollary that applies when the graph structure is geodesic. For the action $G \curvearrowright X$, denote the translation length of h by $\tau_X(h)$. We use the notation $\tau_G(h)$ to denote the translation length of h with respect to the word metric d_G induced by the graph structure Γ:

$$\tau_G(h) = \lim_{n \to \infty} \frac{1}{n} d_G(1, h^n).$$

COROLLARY 7.6. *Suppose that Γ is a thick geodesic combing of G. If $\sigma = 0$ in the CLT, then for all $h \in G$,*

$$\tau_X(h) = \ell\tau_G(h),$$

where ℓ is the corresponding drift.

Proof. Recall that a geodesic combing is coherent by Lemma 2.4. Let g_n be a path in Γ representing h^n for $h \in G$. That is, $h^n = \overline{g}_n$. Since the structure is geodesic, $\|g_n\| = d_G(1, h^n)$. Applying Proposition 7.5, we get that

$$|d(o, h^n o) - \ell d_G(1, h^n)| = O(1).$$

The corollary follows after dividing by n and taking a limit. $\qquad \square$

8. Applications

The main theorem of § 1 now follows easily from the results in § 7.

Proof of Theorem 1.1. Since G has a thick geodesic combing with respect to S, the length $\|g\|$ of a path in the graph equals the word length with respect to S of its evaluation $\overline{g} \in G$, and the sphere of radius n in the Cayley graph of G is in bijection with the set of paths of length n in the graph. Moreover, the associated graph structure is coherent by Lemma 2.4. Then (i) follows immediately from Theorem 7.3, (ii) follows from 7.4 and (iii) from Corollary 7.6. $\qquad \square$

We now give proofs of the applications in the introduction. We first recall some examples of groups which admit thick geodesic combings; for further details, see also [GTT20].

LEMMA 8.1. *The following groups admit thick geodesic combings.*

(i) *A (word) hyperbolic group G admits a thick geodesic combing with respect to any generating set.*

2007

(ii) *If G is relatively hyperbolic with virtually abelian peripheral subgroups, then every finite generating set S' can be extended to a finite generating set S for G which admits a thick geodesic combing.*

(iii) *If G is a right-angled Artin group or right-angled Coxeter group that does not decompose as a product and S is the vertex generating set, then G admits a thick geodesic combing for S whose graph structure has only one maximal component, which is primitive.*

Proof. (i) By [Can84], a hyperbolic G has a geodesic combing with respect to any generating set. By [AL02, Theorem 3], this geodesic combing has the *growth quasitightness* property (see [GTT20, Definition 1.2], inspired by [AL02]). Since growth quasitightness implies thickness by [GTT20, Proposition 7.2], such a structure is thick. Alternatively, thickness can also be deduced from the proof of [GMM18, Lemma 4.6].

(ii) By [AC16, Corollary 1.9], the generating set S' of G can be enlarged to a generating set S, so that the pair (G, S) admits a geodesic combing. Yang [Yan19] proves that any relatively hyperbolic group has the growth quasitightness property with respect to any finite generating set, hence the proof is complete, observing as above that growth quasitightness implies thickness.

(iii) In [GTT20, Corollary 10.4], building on Hermiller and Meier [HM95], we proved that the language of lexicographically first geodesics in the vertex generators is parameterized by a thick graph structure. In fact, the graph structure we construct has only one maximal component, which is primitive. □

Proof of Theorem 1.2. Note that $\pi_1(M)$ is hyperbolic relative to its parabolic subgroups, which are virtually abelian since M has constant curvature. Hence, by Lemma 8.1(ii) the given generating set S' can be enlarged to a finite generating set S that is associated to a thick geodesic combing on $\pi_1(M)$. The theorem then follows from Theorem 1.1. Finally, $\sigma > 0$ by (iii) since the length spectrum is not arithmetic [GR86, Kim06]. □

Proof of Theorem 1.3. In the case where M has no rank 2 cusps, we have that $\pi_1(M)$ is hyperbolic. Indeed, by the tameness theorem [CG06, Ago04], M is the interior of a compact manifold $\overline{M}$, which by assumption does not have tori as boundary components. Then Thurston's hyperbolization theorem (see [Kap01]), $\overline{M}$ admits a convex cocompact hyperbolic structure on its interior. Hence, $\pi_1(M)$ is hyperbolic. The result now follows from Lemma 8.1 and Theorem 1.1.

For the moreover statement, the argument above gives that $\overline{M}$ admits a geometrically finite hyperbolic structure. Hence, $\pi_1(M)$ is hyperbolic relative to its rank 2 parabolic subgroups, which are virtually $\mathbb{Z} \times \mathbb{Z}$. The proof then proceeds as in Theorem 1.2 □

Proof of Theorem 1.4. First, since $\pi_1(M)$ is word hyperbolic, by Lemma 8.1(i) it has a thick geodesic combing with respect to any generating set.

Second, let $T = T_\Sigma$ be the dual tree associated to $\Sigma \subset M$. For details of this standard construction and the properties we need, see [Sha02, § 1.4]. Alternatively, T is the Bass–Serre tree associated to the splitting of $\pi_1(M)$ induced by Σ. Since Σ is not fiber-like, T is not the real line, and since the quotient $\mathcal{G}$ of the action $\pi_1(M) \curvearrowright T$ is compact (it is the underlying graph of the associated graph of groups), the action is nonelementary.

Finally, the intersection number $i(\gamma, \Sigma)$ equals the translation length of γ with respect to the action $\pi_1(M) \curvearrowright T$. To see this, note that the translation length of γ for this action is equal to the number of edges $\#_e\gamma$ crossed by the shortest representative of γ in $\mathcal{G}$. If we embed $\mathcal{G}$ in M dual to Σ, this shows that $i(\gamma, \Sigma) \leq \#_e\gamma$. For the opposite inequality, recall that there is a retraction $r \colon M \to \mathcal{G}$ mapping each component of Σ to the midpoint of some edge. Thus by taking a representative of γ intersecting Σ minimally, considering its image under the retraction,

and homotoping it off edges that it does not fully cross, we obtain that $\#_e \gamma \leq i(\gamma, \Sigma)$. Hence, $i(\gamma, \Sigma) = \ell(\gamma)$ for the action on T.

We now obtain the CLT by applying Theorem 1.1 to this action. If $\sigma = 0$, then Theorem 1.1(iii) implies that the action $\pi_1(M) \curvearrowright T$ is proper and hence vertex stabilizers are finite. However, this is impossible since only virtually free groups admit cocompact actions on trees with finite vertex stabilizers (see, for example, [SW79, Theorem 7.3]). □

For the following application, let us assume G is a hyperbolic group, let ∂G be its Gromov boundary, and let d be a metric on G. We define the *growth rate* of the metric d as

$$v := \limsup_{n \to \infty} \frac{1}{n} \log \#\{g \in G : d(1, g) \leq n\},$$

and for each $s > v$ let us consider the measure on $G \cup \partial G$:

$$\nu_s := \frac{\sum_{g \in G} e^{-sd(1,g)} \delta_g}{\sum_{g \in G} e^{-sd(1,g)}}.$$

Then any limit point of (ν_s) as $s \to v$ is supported on ∂G and is called a *Patterson–Sullivan measure*. By Coornaert [Coo93], any two limit measures are absolutely continuous with respect to each other, with bounded Radon–Nikodym derivative, so the Patterson–Sullivan *measure class* is well defined.

The following rigidity statement about Patterson–Sullivan measures is essentially due to Furman [Fur02], at least in the torsion-free case. For the reader's convenience, we assemble a complete proof from results in the literature. Let $G \curvearrowright X$ be a properly discontinuous and cocompact action on a geodesic Gromov hyperbolic metric space (X, d). Let $o \in X$, and ∂X be the Gromov boundary. Given $\zeta \in \partial X$, we define the Busemann function on the Gromov boundary by setting, for $x, y \in X$, $\mathfrak{b}_\zeta(x, y) := \liminf_{z \to \zeta} \beta_z(x, y)$, where z are points in X. A measure ν on ∂X is called *quasiconformal* with respect to d if there are constants $\alpha, C > 0$ such that

$$C^{-1} \exp(\alpha \mathfrak{b}_\zeta(go, o)) \leq \frac{dg\nu}{d\nu}(\zeta) \leq C \exp(\alpha \mathfrak{b}_\zeta(go, o))$$

for any $g \in G$, ν-a.e. $\zeta \in \partial X$. The constant α is called the *quasiconformal dimension* of ν. The Patterson–Sullivan measure on the boundary of a hyperbolic group is quasiconformal with respect to the associated word metric [Coo93].

PROPOSITION 8.2. *Let $G \curvearrowright X$ be a properly discontinuous and cocompact action on a metric space X. Let d_1 and d_2 be G-equivariant, geodesic, Gromov hyperbolic metrics on X. Let ν_1 and ν_2 be measures on ∂X, quasiconformal with respect to d_1 and d_2, respectively. Then ν_1 and ν_2 are in the same measure class if and only if there are constants $a, b > 0$ with $|ad_1(x, y) - d_2(x, y)| \leq b$ for any $x, y \in X$.*

Proof. Suppose $|ad_1 - d_2| \leq b$. Then, letting $\mathfrak{b}^i$ be the Busemann functions associated to d_i, we have $|a\mathfrak{b}^1_\zeta(\cdot, \cdot) - \mathfrak{b}^2_\zeta(\cdot, \cdot)| \leq 2b$ for any $\zeta \in \partial X$. Therefore any d_1-quasiconformal measure is d_2-quasiconformal and vice versa. By Coornaert [Coo93, Theorem 7.7], any two d_1-quasiconformal measures are absolutely continuous with bounded Radon–Nikodym derivative. Therefore ν_1 and ν_2 are in the same measure class.

To prove the converse, we assume that ν_1, ν_2 are in the same measure class. First, we prove that their Radon–Nikodym derivative is bounded, as follows. As claimed in [Fur02] for the torsion-free case and proved by Bader and Furman [BF17, Theorem 1.4] in general (see also [CDST18, Theorem 4.1] for a generalization beyond cocompact actions), the action of G on $\partial X \times \partial X$ is ergodic with respect to $\nu_i \times \nu_i$ for $i = 1, 2$. By [CDST18, Lemma 2.6], there

2009

exists for any $i = 1, 2$ a rescaling m_i of $d\nu_i \times d\nu_i$, known as *Bowen–Margulis measure*, which is G-invariant. Hence, by e.g. [GT20, Lemma 5.2], ν_1 and ν_2 have bounded Radon–Nikodym derivative.

Now, by the chain rule, the ratio between $dg\nu_1/d\nu_1$ and $dg\nu_2/d\nu_2$ is bounded independently of $g \in G$. Quasiconformality implies that there are $\alpha_1, \alpha_2, D > 0$ such that

$$|\alpha_1 \mathfrak{b}^1_\zeta(go, o) - \alpha_2 \mathfrak{b}^2_\zeta(go, o)| \leq D$$

for ν_i-a.e. $\zeta \in \partial X$ and every $g \in G$. Now recall that Busemann functions are coarsely continuous (see, for example, [CDST18, Proposition 2.1]), that is, there exists D_1 such that

$$\limsup_{\xi \to \zeta} \mathfrak{b}_\xi(x, y) - \liminf_{\xi \to \zeta} \mathfrak{b}_\xi(x, y) \leq D_1$$

for any $\zeta \in \partial X$, any $x, y \in X$. Since the ν_i have full support on the infinite ∂X [Coo93, Corollary 5.2], this implies that there exists D_2 such that

$$|\alpha_1 \mathfrak{b}^1_\zeta(go, o) - \alpha_2 \mathfrak{b}^2_\zeta(go, o)| \leq D_2$$

for every $\zeta \in \partial X$ and every $g \in G$. Since $d_i(o, go) = \sup_{\zeta \in \partial X} \mathfrak{b}^i_\zeta(go, o)$, we obtain $|\alpha_1 d_1(o, go) - \alpha_2 d_2(o, go)| \leq D_2$. Since the action of G is cocompact and the metrics d_i are G-equivariant, this implies $|\alpha_1 d_1 - \alpha_2 d_2| \leq D_3$ for some constant D_3, as required. $\quad\square$

Proof of Theorem 1.5. Since G is word hyperbolic, it has a thick geodesic combing by Lemma 8.1(i). The first statement then follows immediately from Theorem 1.1, by considering the action of G on the Cayley graph of G'.

For the moreover statement, if $\sigma = 0$, Theorem 1.1(iii) implies that

$$|\|\phi(g)\|_{S'} - \ell\|g\|_S|$$

is bounded independently of $g \in G$, hence ϕ has finite kernel.

Now consider the factorization $G \xrightarrow{\pi} \overline{G} := (G/\ker\phi) \xrightarrow{\overline{\phi}} G'$, and define $\overline{S} := \pi(S)$. Then the Cayley graph of $\overline{G}$ carries the two metrics

$$d_1(g, h) := \|h^{-1}g\|_{\overline{S}} \quad d_2(g, h) := \|\overline{\phi}(h^{-1}g)\|_{S'},$$

and they satisfy

$$|d_1(g, h) - \ell d_2(g, h)| \leq C \tag{11}$$

for any $g, h \in \overline{G}$. In particular, $\phi\colon G \to G'$ is a quasi-isometric embedding and so it admits a continuous extension $\partial\phi\colon \partial G \to \partial G'$ to an embedding.

Now, by Proposition 8.2, inequality (11) holds if and only if the Patterson–Sullivan measure classes for d_1, d_2 on $\partial\overline{G}$ are the same. Finally, if ϕ has finite kernel, there exists $C > 0$ for which

$$|d_1(\pi(g), \pi(h)) - d_S(g, h)| \leq C$$

for any $g, h \in G$. Hence, the Patterson–Sullivan measure class for (G, d_S) on ∂G pushes forward to the Patterson–Sullivan measure class for $(\phi(G), d_{S'})$ if and only if $\sigma = 0$. $\quad\square$

Proof of Theorem 1.6. By Lemma 8.1(iii), a right-angled Artin or Coxeter group has a graph structure with respect to the vertex generating set, which is semisimple with only one maximal component. Hence, the CLT follows from Theorem 6.3 (see Remark 6.4). To complete the proof, we note that $\#_v(g)$ is equal to the displacement of g with respect to the action of G on the Bass–Serre tree for the hyperplane associated to v. The details are similar to those of Theorem 1.4. $\quad\square$

References

Ago04   I. Agol, *Tameness of hyperbolic 3-manifolds*, Preprint (2004), arXiv:math/0405568.

AC16   Y. Antolin and L. Ciobanu, *Finite generating sets of relatively hyperbolic groups and applications to geodesic languages*, Trans. Amer. Math. Soc. **368** (2016), 7965–8010.

AL02   G. N. Arzhantseva and I. G. Lysenok, *Growth tightness for word hyperbolic groups*, Math. Z. **241** (2002), 597–611.

BF17   U. Bader and A. Furman, *Some ergodic properties of metrics on hyperbolic groups*, Preprint (2017), arXiv:1707.02020.

BQ16   Y. Benoist and J.-F. Quint, *Central limit theorem on hyperbolic groups*, Izv. Math. **80** (2016), 3–23.

BH11   M. Björklund and T. Hartnick, *Biharmonic functions on groups and limit theorems for quasimorphisms along random walks*, Geom. Topol. **15** (2011), 123–143.

Bow75   R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Lecture Notes in Mathematics, vol. 470 (Springer, 1975).

CF10   D. Calegari and K. Fujiwara, *Combable functions, quasimorphisms, and the central limit theorem*, Ergodic Theory Dynam. Systems **30** (2010), 1343–1369.

CG06   D. Calegari and D. Gabai, *Shrinkwrapping and the taming of hyperbolic 3-manifolds*, J. Amer. Math. Soc. **19** (2006), 385–446.

Can84   J. W. Cannon, *The combinatorial structure of cocompact discrete hyperbolic groups*, Geom. Dedicata **16** (1984), 123–148.

Can21   S. Cantrell, *Statistical limit laws for hyperbolic groups*, Trans. Amer. Math. Soc. **374** (2021), 2687–2732.

CL12   M. Chas and S. P. Lalley, *Self-intersections in combinatorial topology: statistical structure*, Invent. Math. **188** (2012), 429–463.

CLM13   M. Chas, K. Li and B. Maskit, *Experiments suggesting that the distribution of the hyperbolic length of closed geodesics sampling by word length is Gaussian*, Exp. Math. **22** (2013), 367–371.

Coo93   M. Coornaert, *Mesures de Patterson-Sullivan sur le bord d'un espace hyperbolique au sens de Gromov*, Pacific J. Math. **159** (1993), 241–270.

CDST18   R. Coulon, R. Dougall, B. Schapira and S. Tapie, *Twisted Patterson-Sullivan measures and applications to amenability and coverings*, Mem. Amer. Math. Soc., to appear. Preprint (2018), arXiv:1809.10881.

EPC⁺92   D. Epstein, M. S. Paterson, J. W. Cannon, D. F. Holt, S. V. Levy and W. P. Thurston, *Word processing in groups* (AK Peters, 1992).

Fur02   A. Furman, *Coarse-geometric perspective on negatively curved manifolds and groups*, in *Rigidity in dynamics and geometry (Cambridge, 2000)* (Springer, Berlin, 2002), 149–166.

Fur63   H. Furstenberg, *Noncommuting random products*, Trans. Amer. Math. Soc. **108** (1963), 377–428.

GTT18   I. Gekhtman, S. J. Taylor and G. Tiozzo, *Counting loxodromics for hyperbolic actions*, J. Topol. **11** (2018), 379–419.

GTT19   I. Gekhtman, S. J. Taylor and G. Tiozzo, *A central limit theorem for closed geodesics: proof of the Chas-Li-Maskit conjecture*, Adv. Math. **358** (2019), 106852.

GTT20   I. Gekhtman, S. J. Taylor and G. Tiozzo, *Counting problems in graph products and relatively hyperbolic groups*, Israel J. Math. **237** (2020), 311–371.

GT20   I. Gekhtman and G. Tiozzo, *Entropy and drift for Gibbs measures on geometrically finite manifolds*, Trans. Amer. Math. Soc. **373** (2020), 2949–2980.

GMM18   S. Gouëzel, F. Mathéus and F. Maucourant, *Entropy and drift in word hyperbolic groups*, Invent. Math. **211** (2018), 1201–1255.

GR86   Y. Guivarc'h and A. Raugi, *Products of random matrices: convergence theorem*, Contemp. Math. **50** (1986), 31–53.

HM95   S. Hermiller and J. Meier, *Algorithms and geometry for graph products of groups*, J. Algebra **171** (1995), 230–257.

HRR17   D. F. Holt, S. Rees and C. E. Röver, *Groups, languages and automata*, London Mathematical Society Student Texts, vol. 88 (Cambridge University Press, Cambridge, 2017).

Hor18   C. Horbez, *Central limit theorems for mapping class groups and $Out(F_N)$*, Geom. Topol. **22** (2018), 105–156.

HS09   M. Horsham and R. Sharp, *Lengths, quasi-morphisms and statistics for free groups*, Contemp. Math. **14** (2009), 219–237.

Kac47   M. Kac, *On the notion of recurrence in discrete stochastic processes*, Bull. Amer. Math. Soc. **53** (1947), 1002–1010.

Kap01   M. Kapovich, *Hyperbolic manifolds and discrete groups*, vol. 183 (Springer, 2001).

Kim06   I. Kim, *Length spectrum in rank-1 symmetric spaces is not arithmetic*, Proc. Amer. Math. Soc. **134** (2006), 3691–3696.

Lal87   S. P. Lalley, *Distribution of periodic orbits of symbolic and Axiom A flows*, Adv. Appl. Math. **8** (1987), 154–193.

MT18   J. Maher and G. Tiozzo, *Random walks on weakly hyperbolic groups*, J. Reine Angew. Math. **742** (2018), 187–239.

MS20   P. Mathieu and A. Sisto, *Deviation inequalities for random walks*, Duke Math. J. **169** (2020), 961–1036.

MT04   I. Melbourne and A. Török, *Statistical limit theorems for suspension flows*, Israel J. Math. **144** (2004), 191–209.

Mor88   T. Morita, *Deterministic version lemmas in ergodic theory of random dynamical systems*, Hiroshima Math. J. **18** (1988), 15–29.

Mos97   L. Mosher, *Central quotients of biautomatic group*, Comment. Math. Helv. **72** (1997), 16–29.

Par64   W. Parry, *Intrinsic Markov chains*, Trans. Amer. Math. Soc. **112** (1964), 55–66.

PP90   W. Parry and M. Pollicott, *Zeta functions and the periodic orbit structure of hyperbolic dynamics*, Astérisque **187–188** (1990).

PS98   M. Pollicott and R. Sharp, *Comparison theorems and orbit counting in hyperbolic geometry*, Trans. Amer. Math. Soc. **350** (1998), 473–499.

PS01   M. Pollicott and R. Sharp, *Poincaré series and comparison theorems for variable negative curvature*, in *Topology, ergodic theory, real algebraic geometry*, American Mathematical Society Translations: Series 2, vol. 202 (American Mathematical Society, Providence, RI, 2001), 229–240.

PS06   M. Pollicott and R. Sharp, *Correlations for pairs of closed geodesics*, Invent. Math. **163** (2006), 1–24.

Rat73   M. Ratner, *The central limit theorem for geodesic flows on n-dimensional manifolds of negative curvature*, Israel J. Math. **16** (1973), 181–197.

Riv10   I. Rivin, *Growth in free groups (and other stories)—twelve years later*, Illinois J. Math. **54** (2010), 327–370.

Rue04   D. Ruelle, *Thermodynamic formalism* (Cambridge University Press, 2004).

Sar20   O. Sarig, *Lecture notes on ergodic theory*, Preprint (2020), available online at http://www.weizmann.ac.il/math/sarigo/sites/math.sarigo/files/uploads/ergodicnotes.pdf.

SW79   P. Scott and T. Wall, *Topological methods in group theory*, in *Homological group theory*, London Mathematical Society Lecture Note series, vol. 36 (Cambridge University Press, 1979), 137–203.

Sha02   P. Shalen, *Representations of 3-manifold groups*, in *Handbook of geometric topology*, eds R. Daverman and R. Sher (North-Holland, Amsterdam, 2002), 955–1044.

Sin60   Y. G. Sinai, *The central limit theorem for geodesic flows on manifolds of constant negative curvature*, Soviet Math. Dokl. **1** (1960), 938–987.

Sin72   Y. G. Sinai, *Gibbs measures in ergodic theory*, Uspekhi Mat. Nauk **27** (1972), 21–64.

Wal00   P. Walters, *An introduction to ergodic theory*, Graduate Texts in Mathematics, vol. 79 (Springer, 2000).

Yan19   W. Y. Yang, *Statistically convex-cocompact actions of groups with contracting elements*, Int. Math. Res. Not. IMRN **2019** (2019), 7259–7323.

Ilya Gekhtman   ilyagekh@gmail.com

Faculty of Mathematics, Technion-Israel Institute of Technology, Haifa 3200003, Israel

Samuel J. Taylor   samuel.taylor@temple.edu

Department of Mathematics, Temple University, 1805 North Broad Street, Philadelphia, PA 19122, USA

Giulio Tiozzo   tiozzo@math.toronto.edu

Department of Mathematics, University of Toronto, 40 St George St, Toronto, ON M5S 2E4, Canada