

# STRONG APPROXIMATION IN RANDOM TOWERS OF GRAPHS.

YAIR GLASNER

**ABSTRACT.** The term *strong approximation* is used to describe phenomena where an arithmetic group as well as all of its Zariski dense subgroups have a *large image* in the congruence quotients. We exhibit analogues of such phenomena in a probabilistic, rather than arithmetic, setting.

Let  $T$  be the binary rooted tree,  $\text{Aut}(T)$  its automorphism group. To a given  $m$ -tuple  $\underline{a} = \{a_1, a_2, \dots, a_m\} \in \text{Aut}(T)^m$  we associate a tower of  $2m$ -regular Schreier graphs

$$\dots \rightarrow X_n \rightarrow X_{n-1} \rightarrow \dots \rightarrow X_0.$$

The vertices of  $X_n$  are the  $n^{\text{th}}$  level of the tree and two such  $x, y \in X_n$  are connected by an edge if  $y = x^{a_i}$  or if  $x = y^{a_i}$  for some  $i$ .

When  $\{a_i\} \subset \text{Aut}(T)$  are independent Haar-random elements we retrieve the standard model for iterated random 2-lifts studied in [BL04],[AL02],[ALM02]. If  $\underline{w} = \{w_1, w_2, \dots, w_l\} \subset F_m$  are words in the free group, the random substitutions  $\underline{w}(\underline{a}) := \{w_1(\underline{a}), \dots, w_l(\underline{a})\}$  give rise to new models for random towers of  $2l$ -regular graphs.

$$\dots \rightarrow Y_n \rightarrow Y_{n-1} \rightarrow \dots \rightarrow Y_0.$$

**Theorem A.** With the above notation, the following hold almost surely, whenever  $\Delta := \langle \underline{w} \rangle$  is a non-cyclic subgroup of  $F_m$ :

- The graphs  $Y_n$  have a bounded number of connected components,
- these connected components form a family of expander graphs.
- the closure  $\overline{\Delta}$  has positive Hausdorff dimension as a subgroup of the (metric) group  $\text{Aut}(T)$ .

Some of this is generalized to more general trees.

## 1. INTRODUCTION

**1.1. Strong approximation from arithmetic to probability.** The term *strong approximation* is used to describe phenomena where an arithmetic group as well as all of its Zariski dense subgroups have a *large image* in the congruence quotients of the group. For example

---

2000 *Mathematics Subject Classification.* Primary 20E08, 60J80; Secondary 20B27, 28A78, 11F06.

consider the group  $\Gamma = \Gamma_{\text{arith}} \stackrel{\text{def}}{=} \text{PSL}_2(\mathbf{Z})$ , with its congruence maps  $\Phi_N : \Gamma \rightarrow \text{PSL}_2(\mathbf{Z}/N\mathbf{Z})$ . The classical strong approximation theorem says that  $\Gamma$  maps onto all of its congruence quotients. In other words, if  $\Gamma = \langle \underline{a} \rangle$  then the Cayley graphs  $\mathbf{C}_N(\underline{a}) \stackrel{\text{def}}{=} \mathcal{C}(\text{PSL}_2(\mathbf{Z}/N\mathbf{Z}), \Phi_N(\underline{a}))$  are connected. A theorem of Lubotzky-Phillips-Sarnak shows that these graphs are connected in a very strong way: they form a collection of expander graphs.

**Definition.** (Expander graphs) A regular graph  $X$  is called an  $\alpha$ -*expander* if

$$h(X) \stackrel{\text{def}}{=} \inf_{A \subset X} \frac{e(A, \overline{A})}{\min\{|A|, |\overline{A}|\}} \geq \alpha,$$

A family of regular graphs is called an *expander family* if they are all  $\alpha$ -expanders for some  $\alpha > 0$ . The quantity  $h(X)$  defined above is known as the *Cheeger constant* of the graph.

A much more modern question studies the image  $\Phi_N(\Delta)$ , where  $\Delta = \langle \underline{w} \rangle = \langle w_1, w_2, \dots, w_l \rangle < \Gamma$  is Zariski dense (as a subgroup of  $\text{SL}_2(\mathbf{C})$ ). In the specific setting of  $\text{PSL}_2$  this only means that  $\Delta$  is non-cyclic. The strong approximation theorem for linear groups of Nori and Weisfeiler shows the existence of a bound  $M = M(\Delta)$  on the number of connected components for the Cayley graphs  $\mathcal{C}_N(\underline{w})$ . Expansion was established only very recently, and for a limited set of congruence maps. A recent theorem of Bourgain and Gamburd [BG08c] shows that the graphs  $\mathcal{C}_p(\underline{w})$ , where  $p$  ranges over the prime numbers, form a family of expanders whenever  $\Delta = \langle \underline{w} \rangle < \text{PSL}_2(\mathbf{Z})$  is Zariski dense. We elaborate on additional results of Bourgain and Gamburd in section 6.8

This paper exhibits analogues to all of the above strong approximation type theorems in the probabilistic setting described below.

**1.2. Random towers of graphs.** Let  $T = T_d$  be a  $d$ -ary rooted tree,  $\text{Aut}(T)$  its automorphism group. We replace the arithmetic group by the group  $\Gamma = \Gamma_{\text{prob}} \stackrel{\text{def}}{=} \langle a_1, a_2, \dots, a_m \rangle < \text{Aut}(T)$  generated by  $m$ -independent Haar-random elements; congruence maps are replaced by the natural homomorphisms  $\Psi_n : \text{Aut}(T) \rightarrow \text{Aut}_n(T)$  onto the automorphism groups of the truncated trees; finally instead of Zariski dense subgroups I consider all non-cyclic subgroups of  $\Gamma$ .

As defined in the abstract we denote by  $X_n = \mathcal{G}(\Gamma, \underline{a}, L(n))$  the Schreier graphs coming from the action of  $\Gamma$  on the  $n^{\text{th}}$  level of the tree -  $L(n)$ . For a finite set  $\underline{w} \subset F_m^l$  we denote by  $Y_n(\underline{w}) = \mathcal{G}(\Delta, \underline{w}, L(n))$  the Schreier graphs coming from the action of the subgroup  $\Delta =$

$\langle w_1(\underline{a}), \dots, w_l(\underline{a}) \rangle < \Gamma < \text{Aut}(T)$  which is generated by the random substitution in these words. We will often refer to these graphs also as  $Y_n(\Delta)$  or, when the group is understood, just  $Y_n$ . This abuse of notation indicates that we are interested in properties of these graphs that depend only on the group  $\Delta$  rather than on a specific choice of a finite generating set. The graphs  $Y_n$  form a *tower*; they come with covering maps

$$\dots \rightarrow Y_{n+1}(\Delta) \xrightarrow{p_{n,n+1}} Y_n(\Delta) \xrightarrow{p_{n-1,n}} Y_{n-1}(\Delta) \rightarrow \dots$$

Where the covering map  $p_{n,n+1} : Y_{n+1} \rightarrow Y_n$  takes a vertex  $x \in Y_{n+1} = L(n+1)$  to its unique ancestor in  $Y_n = L(n)$ .

An important variation on the standard model of random graph covering, allows for a parameter: a transitive permutation group  $H < \text{Sym}(d)$  of degree  $d$ . In our terminology this amounts to replacing the group  $\text{Aut}(T_d)$  with the infinite iterated wreath product  $W = W(H) \stackrel{\text{def}}{=} H \wr H \wr H \dots < \text{Aut}(T_d)$  and  $\text{Aut}_n(T)$  with the  $(n)$ -fold iterated wreath product

$$W_n = W_n(H) \stackrel{\text{def}}{=} \overbrace{H \wr H \wr \dots \wr H}^{n \text{ times}}.$$

The group  $W(H)$  admits a natural action on the tree  $T = T_d$  and this action restricted to the  $n^{\text{th}}$  level of the tree factors through the homomorphism  $\Psi_n : W(H) \rightarrow W_n(H)$ . We will adopt this more general notation, fixing throughout the paper a transitive permutation group  $H$  of degree  $d$ . Upon setting  $H = \text{Sym}(d)$  we retrieve the special case discussed so far  $W = W(\text{Sym}(d)) = \text{Aut}(T_d)$ . The groups  $W(H)$  will be treated in greater detail in Section 2.5 below.

**1.3. Bounded orbits and Expansion.** Assume that  $H < \text{Sym}(d)$  is a transitive permutation group,  $W = W(H) < \text{Aut}(T_d)$  the corresponding iterated wreath product,  $\Gamma = \langle \underline{a} \rangle$  a subgroup generated by  $m$  independent Haar-random elements of  $W$ . Finally define the Schreier graphs  $X_n = \mathcal{G}(\Gamma, \underline{a}, L(n))$  and  $Y_n(\Delta) = \mathcal{G}(\Delta, \underline{w}, L(n))$  for every finitely generated  $\Delta = \langle \underline{w} \rangle < \Gamma$  as above.

**Theorem.** *Almost surely, for every every non cyclic finitely generated subgroup  $\Delta < \Gamma$ , the associated tower of Schreier graphs satisfies the following properties:*

- (1) *The number of connected components of  $Y_n(\Delta)$  is bounded.*
- (2) *If  $d = 2$ , the collection of connected components of the graphs  $Y_n(\Delta)$  form a family of expander graphs.*

**1.4. Positive Hausdorff dimension.** By now the reader probably asked herself why, in the arithmetic setting, we work with Cayley graphs, while the probabilistic version, Theorem 1.3 above is stated for Schreier graphs. Indeed one could ask about connectedness and expansion of the Cayley graphs  $C_n(\underline{w}) = \mathcal{C}(W_n(H), \underline{w})$ . Since there are natural covering maps  $C_n(\underline{w}) \rightarrow Y_n(\underline{w})$  such statements would be stronger than the statements that we prove about the Schreier graphs. Unfortunately, in general, the Cayley graphs even fail to have a bounded number of connected components, even for the group  $\Gamma$  itself. We explain more on the reasons for this in Section 6.3. The following statement, stated in terms of Hausdorff dimension, is the next best thing that one could hope for.

**Theorem.** *Assume that  $H = \mathbf{Z}/p\mathbf{Z}$ . Then almost surely  $\dim_H(\overline{\Delta}) > 0$  for every non cyclic subgroup  $\Delta < \Gamma$ . Explicitly this means that for every such  $\Delta$  there exists a positive number  $h = h(\Delta) > 0$  such that*

$$\liminf \frac{|\Psi_n(\Delta)|}{|W_n(T)|^h} \geq 1.$$

Where  $W_n(T) = H \wr H \wr \dots \wr H$  is the  $n$ -fold iterated wreath product.

**1.5. Resolving dependence.** Iterated random graph covers (i.e. the case  $\Delta = \Gamma$ ) were studied by many people and in particular all of the above strong approximation type theorems were already known in this case. Abért and Virág [AV05] show that the number of connected components of  $X_n$  is bounded. In the specific case where  $d = p$  is prime,  $H = \mathbf{Z}/p\mathbf{Z}$  and  $m \geq 3$  they prove that the Hausdorff dimension of the group  $\overline{\Gamma}$  is almost surely one. In other words, when  $\Delta = \Gamma$ , they prove that the conclusion of Theorem 1.4 holds almost surely with  $h = 1 - \epsilon$ . For the binary tree, Bilu and Linial [BL04] prove that the connected components of  $X_n$  almost surely form a family of expander graphs. The novelty of the current paper is that we establish similar results not for the group  $\Gamma$  itself but rather for every non-cyclic subgroup  $\Delta < \Gamma$  thereby imitating the treatment of Zariski dense subgroups in the arithmetic case.

The main problem that we face is that the action of  $\Delta$  is not as nice as that of  $\Gamma$ . Group theoretically it was proved by Bhattacharjee in [Bha95] that  $\Gamma$ , and therefore also  $\Delta$ , are almost surely free groups. But the free generators of  $\Delta$  are no longer distributed according to Haar measure. Even worse, these generators need not be independent. To deal with this situation we develop a general method for “resolution of dependencies”. The main technical theorem 3.1 provides a general mechanism for taking random subgroups generated by non-uniform and

highly dependent elements such as  $\Delta$ , and reducing questions about them to questions about “nice” random subgroups that are generated by independent Haar-random generators such as  $\Gamma$ . Theorems 1.3 and 1.4 are then deduced from the main technical theorem, combined with known results of Abért-Virág and of Bilu-Linial that were mentioned above.

**1.6. Notation and arrangement of this paper.** The theorems, definitions are not numbered themselves but rather referred to by the number of the appropriate section. For example Theorem 1.3 is the (unique) theorem appearing in Section 1.3.

After introducing all the notation we state the main technical Theorem 3.1 in Section 3 we use this theorem to deduce Theorems 1.3(1) and 1.4. Section 4 is dedicated to the proof of the main technical theorem. In Section 5 we prove Theorem 1.3(2), namely the expansion of Schreier graphs. The last section 6 examines more closely the analogy with strong approximation for arithmetic groups.

**1.7. Acknowledgments.** I thank Shlomo Hoory and Bálint Virág who pointed out that the proof of Bilu and Linial actually yields the precise statement needed in this paper. I thank Miklós Abért for many stimulating talks on random group actions, and for explaining to me many of the details of the paper [AV05]. Much of this paper was written in Geneva, I am very thankful to the Swiss science foundation that enabled that visit and to the math department at Geneva for their warm hospitality. This material is based on research done at the institute for advanced studies and supported by U.S. National Science Foundation under agreement DMS-0111298. The work was also supported by ISF grant 888/07 and BSF grant 2006-222.

## 2. BACKGROUND AND NOTATION

**2.1. Graphs.** A graph  $X$  consists of two sets, a set of vertices also denoted by  $X$  and a set of edges denoted  $EX$ . Each edge  $e \in EX$  is incident to two (possibly identical) vertices  $(o(e), t(e)) \in X^2$ . Each edge  $e$  has an inverse edge  $\bar{e}$  such that  $\bar{\bar{e}} = e$ ,  $o(\bar{e}) = t(e)$ ,  $t(\bar{e}) = o(e)$ . A *geometric edge*  $[e]$  is a pair of opposite directed edges  $[e] \stackrel{\text{def}}{=} \{e, \bar{e}\}$ .

**2.2. Rooted trees.** Let  $T = T_d$  be the infinite rooted  $d$ -ary tree. The vertices of  $T$  correspond in a natural way to finite  $d$ -adic expansions. The empty expansion corresponds to the base vertex. A vertex  $v$  is a descendant of  $w$  if and only if the expansion of  $w$  is a prefix of the expansion corresponding to  $v$ . The subset  $L(n) \subset T$  of vertices corresponding to expansions of length  $n$  is called the  $n^{\text{th}}$  level of the

tree. The level of a given vertex  $v$ , denoted  $\text{Level}(v)$ , is just the length of its expansion. Given a vertex  $v$  we denote by  $T_v$  the subtree containing  $v$  and all of its descendants.  $T_v$  is isomorphic to  $T$ , in fact there is a natural isomorphism between these trees  $\tau_v : T_v \rightarrow T$  given by the truncation of the first  $\text{Level}(v)$  symbols of the expansion. It will be convenient to denote these functions from the right  $w \in T_v \mapsto w^{\tau_v} \in T$ .

**2.3. Group actions.** Let  $G$  be any group acting (from the right) on  $T$ , denote by  $G_v = \{g \in G \mid v^g = v\}$  the stabilizer of the vertex  $v$  and by  $v^G = \{v^g \mid g \in G\}$  the orbit,  $G^{[n]} = \cap_{v \in L(n)} G_v$  the pointwise stabilizer of the  $n^{\text{th}}$  level and  $G_n = G/G^{[n]}$ . The action of  $G$  on the first  $n$ -levels of the tree factors through the homomorphism  $\Psi_n : G \rightarrow G_n$ . In fact  $G^{[n]}$  is precisely the kernel of this action. In particular the map

$$\begin{aligned} \Psi_1 : G &\rightarrow G_1 < \text{Sym}(\{0, 1, \dots, d-1\}) \\ g &\mapsto \bar{g} \stackrel{\text{def}}{=} \Psi_1(g) \end{aligned}$$

captures the action of elements of the first level of the tree.

**2.4. The local cocycle.** The canonical identification of every subtree of the from  $T_v$  with  $T$  enables us to define the *local cocycle* of an automorphism  $g$  at a vertex  $v$

$$\begin{aligned} \beta : \text{Aut}(T) \times T &\rightarrow \text{Aut}(T) \\ (g, v) &\mapsto \beta(g, v) \stackrel{\text{def}}{=} \tau_v^{-1} g \tau_{v^g} \end{aligned}$$

This map satisfies the cocycle equality  $\beta(gh, v) = \beta(g, v)\beta(h, v^g)$  and in particular the map  $g \mapsto \beta(g, v)$  becomes a homomorphism when it is restricted to  $G_v$ . The image of this map  $\beta(G_v, v) = \{\beta(g, v) \mid g \in G_v\} < \text{Aut}(T)$  will play an important role in our discussion. The 1-*local cocycle*  $\text{Aut}(T) \times T \rightarrow \text{Sym}(d)$  is given by

$$\begin{aligned} \bar{\beta} : \text{Aut}(T) \times T &\rightarrow \text{Sym}(d) \\ (g, v) &\mapsto \overline{\beta(g, v)} = \Psi_1(\beta(g, v)). \end{aligned}$$

**2.5. The groups  $W(H)$ .** Throughout this paper we fix a transitive permutation group  $H < \text{Sym}(\{0, 1, \dots, d-1\})$  of degree  $d$ . Consider the group  $W = W(H) = \{g \in \text{Aut}(T) \mid \overline{\beta(g, v)} \in H \forall v \in T\}$ . It is easy to verify, using the cocycle condition, that  $W(H)$  is a group. In fact  $W(H)$  is naturally isomorphic to the infinite iterated wreath product of  $H$  acting on  $T$ . In particular  $W(\text{Sym}(d))$  is the full automorphism group of  $T$ . The group  $W(H)$  has the nice property that  $\beta(g, v) \in W(H) \forall g \in W(H), v \in T$ .

$W = W(H)$  is a compact (profinite) group and therefore admits a unique probability Haar measure. For any  $a \in W(H)$  we set  $v(a) =$

$\max_n \{\Psi_n(a) = 1\} \in \mathbf{N} \cup \{\infty\}$ . The group  $W(H)$  admits a natural invariant ultrametric which is compatible with the topology on the group

$$\delta(a, b) = d^{-v(a^{-1}b)}. \quad (1)$$

**2.6. The probability space.** Let  $F = F_m = \langle x_1, x_2, \dots, x_m \rangle$  be a free group on  $m$  generators. Our probability space will be  $\text{Hom}(F_m, W(H)) \cong W(H)^m$  endowed with the product Haar measure. In other words we consider random actions of  $F$  on  $T$  given by a choice of  $m$  independent Haar-random elements of  $W$ .

**2.7. Hausdorff dimension.** Viewing  $W(H)$  as a metric space, with the metric given by equation 1 it makes sense to talk about the Hausdorff dimension of a closed subset and in particular of a closed subgroup  $G < W(H)$ . It was shown by Barnea and Shalev [BS97] that the Hausdorff dimension of a closed subgroup  $G < W(H)$  is given by the formula

$$\dim_H G = \liminf \gamma_n(G). \quad (2)$$

Where  $\gamma_n(G)$ , the *density sequence*, is given by  $\gamma_n(G) = \log |G_n| / \log |W_n|$ . For convenience of notation we will use Equation (2) as the definition of Hausdorff dimension even for subgroups that are not closed. Note that this does not really extended the scope of our discussion because the quantity  $\dim_H G$  thus defined, depends only on the closure  $\overline{G}$ .

**2.8. Schreier graphs.** Let  $\Gamma \curvearrowright X$  be a (right) group action on a set and  $S \subset \Gamma$  a finite symmetric subset of  $\Gamma$ . The *Schreier graph*  $\mathcal{G}(\Gamma, S, X)$  is a directed graph with vertex set  $X$  and edge set  $X \times S$ . Where the graph structure is given by the following maps:

$$\overline{(x, s)} = (x^s, s^{-1}), \quad o((x, s)) = x, \quad t((x, s)) = x^s.$$

The projection on the second factor  $X \times S \rightarrow S$  is referred to as the natural edge labeling of the Schreier graph. If  $S \subset \Gamma$  is a finite subset that is not necessarily symmetric, we define  $\mathcal{G}(\Gamma, S, X) \stackrel{\text{def}}{=} \mathcal{G}(\Gamma, S \cup S^{-1}, X)$ . Note that in the definition of a Schreier graph we do not require that  $S$  be a generating set of  $\Gamma$ . The Schreier graph  $\mathcal{G}(\Gamma, S, X)$  will be connected if and only if  $\langle S \rangle < \Gamma$  is a transitive subgroup.

The *Cayley graph* of a group with respect to a finite set of generators, is defined as the Schreier graph coming from the right regular action of the group on itself  $\mathcal{C}(G, S) = \mathcal{G}(G, S, G)$ .

### 3. THE MAIN TECHNICAL RESULT AND ITS APPLICATIONS

Both Theorems 1.3 and 1.4 follow from known results due to Abért-Virág [AV05] and Bilu-Linial [BL04] concerning actions of groups generated by independent Haar-random elements combined with our main technical theorem stated in the following section:

#### 3.1. A theorem about resolution of independence.

**Theorem.** *Let  $\Delta < F$  be a non-cyclic subgroup and  $K \in \mathbf{N}$ . Then there exists a random variable  $N = N(K) \in \mathbf{N}$ . And for every  $\Delta$  orbit  $Z \in L(N)/\Delta$  there exist random variables as follows:*

- *A random vertex  $v = v(Z) \in Z$ ,*
- *$K$  random group elements  $\{\alpha_k = \alpha_k(Z) \in \Delta_v \mid 1 \leq k \leq K\}$ .*

*Such that the automorphisms  $\beta(\alpha_k(Z), v(Z)) = \beta(\alpha_k, v)$  are  $K$  independent Haar-random elements of  $W$ .*

In plain words: given  $K$  we can almost surely find a level  $N$  and orbit representatives  $v(Z) \in L(N)$  and subgroups  $A(Z) := \langle \alpha_k(Z) \mid 1 \leq k \leq K \Delta_{v(Z)} \rangle$ . In such a way that the action of the group  $A(Z)$  on  $T_{v(Z)}$  admits the distribution of a group generated by  $K$  independent Haar-random elements.

**Remark.** We do not claim here any independence between groups associated with two different orbits  $A(Z), A(Z')$ . Nevertheless, a posteriori, one can actually prove this stronger statement. It is a good exercise to go over the proof of Theorem 3.1 below, assuming the validity of Theorem 1.3. A slight modifications of the argument yields  $KM(\Delta)$  elements  $\{\alpha_k(Z) \in \Delta_{v(Z)} \mid 1 \leq k \leq K, Z \in L(n)/\Delta\}$  that are jointly independent and each one admits the distribution of a Haar-random element in its corresponding tree. Here  $M(\Delta)$  denotes the bound on the number of orbits of  $\Delta$  given by Theorem 1.3(1). I do not know how to prove this stronger theorem without going over the argument twice.

**3.2. Bounded number of connected components.** We will use the following theorems due to Abért and Virág

**Theorem.** [AV05, Proposition 3.10] *Let  $H$  be any transitive permutation group, and  $m \geq 2$ . Then for almost every  $\phi \in \text{Hom}(F_m, W(H))$  the number of connected components of the graphs  $X_n$*

$$L(n)/\Psi_n(\phi(F_m))$$

*is bounded.*



*Proof of Theorem 1.3(1).* Since every non-cyclic subgroup of  $F$  contains a subgroup generated by two elements, it is enough to prove the theorem under the additional assumption that  $\Delta$  is generated by two elements. Furthermore, since there are only countably many subgroups of  $\Gamma$  that are generated by two elements, we are free to fix one such subgroup  $\Delta < F$  and prove the existence of a random variable  $M = M(\Delta) \in \mathbf{N}$ , bounding the number of orbits of  $\Delta$  independently of the level.

Now apply Theorem (3.1) with  $K = 2$ . By Theorem 3.2 there is an upper bound  $M(Z)$  on the number of orbits of  $A(Z) = \langle \alpha_k(Z) \mid 1 \leq k \leq K \rangle$  in its action on the tree  $T_{v(Z)}$  independent of the level. The number of orbits of  $\Delta$  will now be bounded by the random variable  $M(\Delta) \stackrel{\text{def}}{=} \sum_{Z \in L(N)/\Delta} M(Z)$ .  $\square$

**3.3. Positive Hausdorff dimension.** Theorem 1.4 follows directly from the following two theorems. The first one, due to Abért and Virág, asserts that a group generated by 3 independent Haar-random elements in  $W(\mathbf{Z}/p\mathbf{Z})$  has full Hausdorff dimension. The second one is a corollary of our main technical theorem.

**Theorem.** [AV05] *Let  $H = \mathbf{Z}/p\mathbf{Z}$  be a cyclic group of prime order. Then for almost every  $\phi \in \text{Hom}(F_3, W(H))$*

$$\dim_H(\phi(\Gamma)) = 1.$$

**Corollary.** *Assume that for a given transitive permutation group  $H < \text{Sym}(d)$  there exists a number  $K \in \mathbf{N}$  such that  $\dim_H(\phi(\overline{F_K})) > 0$  for almost every  $\phi \in \text{Hom}(F_K, W(H))$ . Fix an integer  $m \geq 2$ . Then for almost every  $\phi \in \text{Hom}(F_m, W(H))$  every non cyclic subgroup  $\Delta < F_m$*

$$\dim_H(\phi(\Delta)) \geq 0.$$

*Proof.* By the same argument used in the proof of Theorem 1.3(1) above, we are free to fix a subgroup  $\Delta < F$  generated by two elements and show that this specific subgroup almost surely has a positive Hausdorff dimension. Now apply Theorem (3.1) taking  $K$  to be the number given in the Assumption of Corollary 3.3. This assumption now implies that for any fixed connected component  $Z$  of  $X_N$  the group  $\beta(A, v) := \beta(A(Z), v(Z))$  almost surely acts with full Hausdorff dimension on  $T$ . In the calculation that follows we use the group associated

with one single orbit  $A = A(Z)$ .

$$\begin{aligned}
\dim_H(\Delta) &= \liminf_{n \rightarrow \infty} \frac{\log |\Delta_n|}{\log |W_n|} \\
&\geq \liminf_{n \rightarrow \infty} \frac{\log |A_n|}{\log |W_n|} \\
&\geq \liminf_{n \rightarrow \infty} \frac{\log |(\beta(A, v))_{n-N}|}{\log |W_n|} \\
&\geq \liminf_{n \rightarrow \infty} \frac{d^{n-N} - 1}{d^n - 1} \frac{\log |(\beta(A, v))_{n-N}|}{\log |W_{n-N}|} \\
&= \frac{\dim_H(\beta(A, v))}{d^N} > 0
\end{aligned}$$

In the last line we used our assumption that  $\dim_H(\beta(A, v))$  is almost surely positive. All the groups in this calculation are considered with their given action on the tree  $T$ . The group  $A$  acts as a subgroup of  $\Delta$  and in particular it fixes the vertex  $v$ . The action of the group  $\beta(A, v)$  on  $T$  is isomorphic to the action of  $A$  on  $T_v$ . In the line before last we used the explicit calculation  $\log |W_n| = \log |W_n(H)| = \log |H|^{\frac{d^n-1}{d-1}}$ .  $\square$

**3.4. Expansion.** The proof of Theorem 1.3(2), namely the expansion of the family of connected components of the Schreier graphs  $X_n$ , requires some of the techniques and terminology that is introduced in the proof of the main technical Theorem 3.1. We therefore defer it to Section (5).

#### 4. PROOF OF THE MAIN TECHNICAL THEOREM

**4.1. Some notation.** Let  $F = F_m$  be the group given in the statement of Theorem 3.1 - a free group with a free set of generators  $S = \{x_1, x_2, \dots, x_m\}$ . Since every non cyclic group contains a group generated by two elements it will be enough to prove Theorem (3.1) for a fixed subgroup  $\Delta = \langle w_1, w_2 \rangle < F$  generated by two elements  $T = \{w_1, w_2\} \subset F$ . We write  $w_i = w_i(x_1, \dots, x_s)$  as reduced words in the given generating set for  $F$ . Let  $l_i$  denote the length of the word  $w_i$  and set  $l = \max\{l_1, l_2\}$ . Given  $\phi \in \text{Hom}(F, W(H))$ , the associated action of  $F$  on the different level sets of  $T$  gives rise to Schreier graphs  $X_n = \mathcal{G}(F, S, L(n))$  and  $Y_n = \mathcal{G}(\Delta, T, L(n))$ . When  $\phi$  is taken to be random, we know a great deal about the properties of the graphs  $X_n$  and we would like to prove similar statements concerning the graphs  $Y_n$ . This notation will be fixed throughout the proof.

## 4.2. Schreier graphs and immersions.

**Definition.** An *immersion* of graphs is a map that is locally injective. A *covering* of graphs is a map that is

- locally a bijection,
- surjective and
- such that the cardinality of the inverse image of a point is independent of the point.

Note that the last two requirements are redundant if the image is connected.

For every  $n$  we obtain an immersion

$$\iota_n : Y_n \rightarrow X_n.$$

This map<sup>1</sup> is the identity on the vertices and it sends an edge  $e(v, w_i)$  in  $Y_n$  to the directed path  $e(v, w_i(x_1, \dots, x_s)) \subset X_n$ . This is the unique path of length  $l_i$  starting at  $v$  and traversing the path described by the letters of the word  $w_i(x_1, \dots, x_s)$ . For any pair of numbers  $n_1 < n_2$  we obtain covering maps:  $p_{n_2, n_1} : X_{n_2} \rightarrow X_{n_1}$  and  $p_{n_2, n_1} : Y_{n_2} \rightarrow Y_{n_1}$ . These coverings are the maps sending a vertex of  $X_{n_2}$  to its unique ancestor in level  $n_1$  of the tree. All these maps are compatible in the sense that the following diagram is commutative:

$$\begin{array}{ccc} Y_{n+1} & \xrightarrow{\iota_{N+1}} & X_{n+1} \\ \downarrow p_{n+1, n} & & \downarrow p_{n+1, n} \\ Y_n & \xrightarrow{\iota_N} & X_n \\ \downarrow p_{n, n-1} & & \downarrow p_{n, n-1} \\ Y_{n-1} & \xrightarrow{\iota_{N-1}} & X_{n-1} \end{array}$$

**4.3. The stable actions.** The following proposition, proved by Abért and Virág is very important in the sequel.

**Theorem.** [AV05, Proposition 4.1] *Any non trivial element  $f \in F$  fixes only finitely many vertices of  $T$ , almost surely.*

This theorem shows that the stable action of  $F$  and of  $\Delta$  on  $T$ , namely the action on high enough levels, has some nice properties. For example we will use following:

---

<sup>1</sup>This map is not really a graph morphism because it sends edges to longer paths, we could make this into a graph morphism by partitioning each edge labeled  $w_i$  into  $l_i$  segments and labeling them by the letters of the word  $w_i$ .

**Corollary.** *(Stable action) Given  $R \geq 1$ , there exists a random variable  $N = N(R)$  such that for every  $n \geq N(R)$  the following hold:*

- (1)  $\text{girth} X_n \geq R$ ,
- (2)  $\text{girth} Y_n \geq R$ ,
- (3) *The map  $\iota_n$  is injective on any edge of  $Y_n$ ,*
- (4)  *$\iota_n(\alpha)$  is contractible for any path  $\alpha$  of length  $\leq R$  in  $Y_n$ .*

*Proof.* All these properties are stable when passing to a cover, so it is enough to verify them for  $n = N$ . For (1) take  $N$  large enough that no element of  $F$  which is of word length at most  $R$  fixes any point in  $L(N)$ . For (2) do the same, making sure that no element of  $\Delta$  which is of word length less than  $R$  in the generators of  $\Delta$  fixes any point on  $L(N)$ . Recall that the generators of  $\Delta$  are written as reduced words  $w_i(x_1, \dots, x_s)$  in the generators of  $F$ . Assume that  $w_i$  is a word of length  $l_i$  and set  $l = \max\{l_1, l_2\}$ . If, using Property (1)  $\text{girth}(X_N)$  is chosen to be larger than  $l$ , then Property (3) follows. Finally if we take  $\text{girth}(X_N) \geq lR$  then property (4) holds. Indeed let  $\alpha = w_{i_1}w_{i_2} \dots w_{i_R}$  represent a path of length  $R$  in  $Y_N$ . The path  $\iota_S(\alpha)$  need not be reduced, so we cannot assume that  $\iota_S$  is injective on  $\alpha$ . However, the word  $w_{i_1}w_{i_2} \dots w_{i_R}$  is equivalent, in the free group  $F$ , to a unique reduced word  $w$ . Therefore  $\iota_N(\alpha)$  is homotopically equivalent to a path labeled by the word  $w$ . Since  $\text{length } w \leq \text{length } w_{i_1}w_{i_2} \dots w_{i_R} \leq lR$  the path  $w$  is just a line and therefore contractible. So  $\iota_N(\alpha)$  must be contractible as well.  $\square$

#### 4.4. Reformulation of the main technical theorem in combinatorial terms.

Keeping the notation defined in Section 4.1

**Theorem.** *Given  $K \in \mathbf{N}$ , there exists a random variable  $N = N(K) \in \mathbf{N}$  and for each connected component  $Z$  of  $Y_N$  there exist random variables as follows:*

- *A vertex  $v(Z) \in Z$ ,*
- *$K$  elements of the fundamental group*

$$\alpha_1(Z), \dots, \alpha_K(Z) \in \pi_1(Z, v(Z)),$$

- *$K$  random geometric edges*

$$\{e_1(Z), e_2(Z), \dots, e_K(Z)\} \subset EX_N,$$

*such that the directed path  $\iota_N(\alpha_j(Z))$  passes through the edge  $e_k(Z)$  exactly once if  $j = k$  and not at all if  $j \neq k$ .*

We will prove Theorem (4.4) in the next section, here we will assume Theorem (4.4) and show that Theorem (3.1) follows in a straight forward way.

Let  $N, v(Z), \{\alpha_k(Z)\}_{1 \leq k \leq K}, \{e_k(Z)\}_{1 \leq k \leq K}$  be the random variables provided by Theorem 4.4. Since there is no interaction between different connected components, we will fix the connected component  $Z \subset Y_N$  once and for all and drop it from the notation. Thus we will write  $\alpha_i = \alpha_i(Z), v = v(Z)$  etc'.

The group  $\Delta_v$  can be naturally identified with the fundamental group  $\pi_1(Z, v)$ , so we can identify  $\alpha_k$  as elements of  $\Delta_v$ . *Our goal will be to prove Theorem (3.1) by showing that the random group elements  $\{\xi(\alpha_k, v)\}_{1 \leq k \leq K}$  are independent and identically distributed according to Haar measure on the group  $W(H)$ .* In order to understand the statistical distribution of the elements  $\alpha_k$  and their action on the tree  $T_v$  we push forward to the graph  $X_N$ . We name the edges along this path  $\iota_N(\alpha_k) = a_{k,1}a_{k,2} \dots a_{k,M(k)}$ . Group theoretically, the labels across these edges give the expansion of the word  $\alpha_k$  in the generators of  $F$ . An example of such a path  $\iota_n(\alpha)$  is depicted in Figure (1). Using the

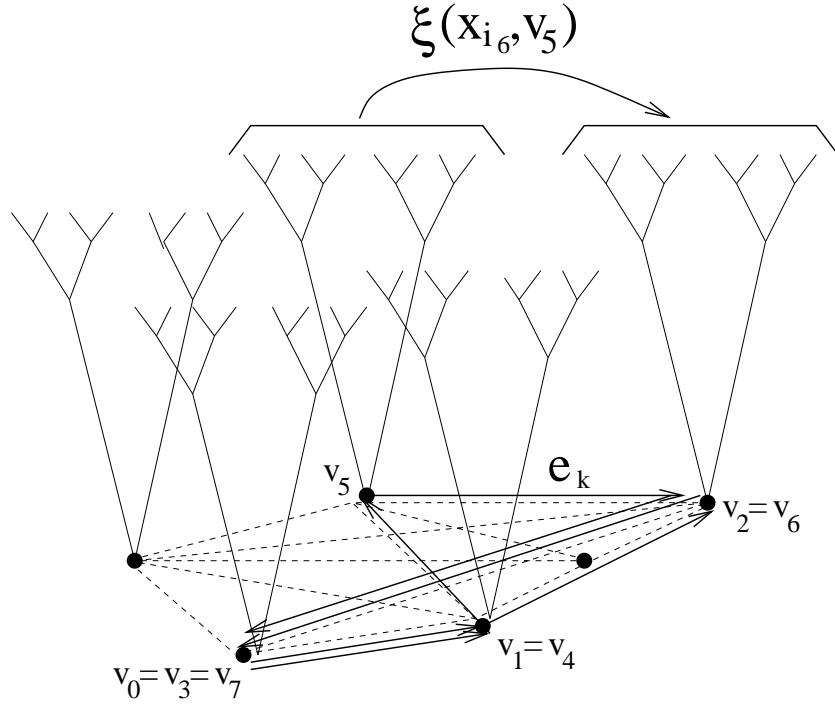


FIGURE 1. An example of a path of the form  $\iota(\alpha)$ , with a singled edge  $e$ .

cocycle equality we decompose the cocycle across the path

$$\xi(\alpha_k, v) = \xi(a_{k,1})\xi(a_{k,2}) \dots \xi(a_{k,M(K)}). \quad (3)$$

The advantage of this procedure is that now the statistical distribution of the factors on the right hand side is simple. Each factor is distributed according to Haar measure on  $W(H)$ . Factors that correspond to different geometric edges are jointly independent. Two factors that correspond to the same (resp. to opposite) directed edges, are identical (resp. inverse to each other).

In general the distribution of the expressions given by Equation (3) is complicated. Because of the special nature of Haar measure, the calculation becomes easy under the special conditions provided by Theorem (4.4). Recall that Theorem (4.4) singles out one edge, say  $e_k = a_{k,l(k)}$ , along the path  $\iota_N(\alpha_k)$ . This edge is unique and appears nowhere else. We collect terms in Equation 3 grouping together the factors that come before and after the special edge. In other words we can write

$$\xi(\alpha_k, v) = \beta_k, \gamma_k, \delta_k$$

where we set  $\beta_k = \xi(a_{k,1})\xi(a_{k,2})\dots\xi(a_{k,l(k)-1})$ ,  $\gamma_k = \xi(a_{k,l(k)})$ ,  $\delta_k = \xi(a_{k,l(k)+1})\dots\xi(a_{k,M(k)})$ . Note that the random elements  $\gamma_k \in W(H)$  are distributed according to Haar measure and completely independent of all the other random variables. Under these conditions Theorem 3.1 will follow by the following lemma.

**Lemma.** *Let  $G$  be a compact topological group with probability Haar measure  $P$ . Let  $\{\beta_k, \gamma_k, \delta_k \in G \mid 1 \leq k \leq K\}$  be random group elements, such that each one of the elements  $\gamma_k$  is distributed according to Haar measure and is independent of all the other random variables. Then the elements  $\eta_k = \beta_k \gamma_k \delta_k$  are distributed according to Haar measure and furthermore  $\{\eta_k \mid 1 \leq k \leq K\}$  are independent.*

*Proof.* The conclusion of the theorem is true, by the bi-invariance of Haar measure, if we condition on the value of the random variables  $\{\gamma_k, \delta_k\}$ . We conclude, using Fubini, by integrating over all these possible values.  $\square$

Thus we have reduced the proof of the theorem to the proof of the combinatorial reformulation 4.4.

**4.5. Proof of the combinatorial reformulation.** Given an edge  $e \in EX_n$  we denote by  $\zeta(e) = \{f \in EY_n \mid e \subset \iota_n(f)\}$ . More generally if  $E \subset EX_n$  is a collection of edges then  $\zeta(E) = \cup_{e \in E} \zeta(e)$ .

In order to satisfy Theorem (4.4) we need to come up with random variables

$$N \in \mathbf{N}, \quad v(Z) \in Z, \quad \{\alpha_k(Z) \in \pi_1(Z, v(Z))\}_{1 \leq k \leq K}, \quad \{e_k(Z) \in EX_N\}_{1 \leq k \leq K}$$

for every connected component  $Z \subset Y_N$ . Since we do not require any relationship between data associated with different connected components we will omit  $Z$  from the notation, keeping always in mind that all requirements should be satisfied for each connected component separately.

**4.6. Conditions imposed on the edges  $e_k$ .** We focus on the collection of edges  $E = E(Z) = \{e_k(Z)\}_{1 \leq k \leq K}$ . If we can arrange for the edges  $E$  to have the following good properties, then the rest of the construction will follow.

- (1)  $\iota_N$  is injective on edges of  $\zeta(E)$ .
- (2) If  $f_k \in \zeta(e_k)$  then  $E \cap \iota_N(f_k) = e_k$ ,
- (3)  $Z \setminus \zeta(E(Z)) \neq \emptyset$ ,
- (4)  $Z \setminus \zeta(E(Z))$  is connected.

Indeed assume that all these conditions are satisfied. Fix a representative  $f_k \in \zeta(e_k)$  for every  $1 \leq k \leq K$ . Using Condition (3) we can choose a vertex  $v(Z) \in Z \setminus \zeta(E(Z))$ . Since by Condition (4)  $Z \setminus \zeta(E)$  is connected, we can complete each edge  $f_k$  into a loop  $\alpha_k \in \pi_1(Z, v(Z))$  without using any other edge from  $\zeta(E)$ . Thus  $\iota_N(\alpha_k)$  can possibly intersect one of the edges from  $E$  only in the small segment  $\iota_N(f_k)$ . Finally by Condition (2) the edge  $e_j$  does not appear in  $\iota_N(\alpha_k)$  whenever  $k \neq j$ . And by condition (1)  $e_k$  appears in  $\iota_N(\alpha_k)$  exactly once.

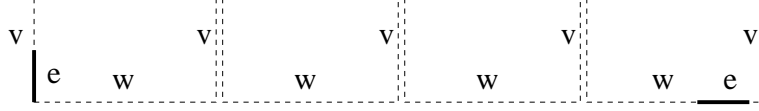
That the first three conditions above can be satisfied, almost surely, when  $N$  is large enough, follows directly from Corollary (4.3). Condition (4) is a very weak expansion property. We want to be able to erase a set of edges of the form  $\zeta(E)$  without disconnecting the graph, but we have to do that for every choice of generators  $\{w_1, w_2\} \subset F$ .

**4.7. Approximate random variables.** We start by constructing approximate random variables  $N', E' = \{e'_k\}, F' = \{f'_k\}, v'$ . These will satisfy the first three conditions and instead of Condition (4) they will satisfy the following auxiliary condition:

- (5) There are no trees among the connected components of the graph  $Z \setminus \zeta(E')$ .

Recall that each generator of  $\Delta$  is given as a word  $w_i(x_1, \dots, x_s)$  of length  $l_i$  in the generators of  $F$ . Set  $l = \max\{l_1, l_2\}$ . An edge of type  $w_i$  in  $Y_n$  is mapped by  $\iota$  to a path of length  $l_i$  in  $X_n$ . In particular there are at most  $D = 2K(l_1 + l_2)$  edges in  $\zeta(E')$  and this bound is independent of  $N'$ .

We will take  $N'$  large enough for all the stability conditions from Corollary (4.3) to hold, specifying the parameter  $R$  as we go. Condition

FIGURE 2. The path  $\iota_{N'}(\alpha)$ .

(1) is satisfied Using Corollary (4.3)(3). By 4.3(2) we can make the connected components of  $Y_{N'}$  large enough to choose the edges  $e'_k$  in such a way that  $d_{X_{N'}}(e'_{k_1}, e'_{k_2}) > 2lD$  for every  $k_1 \neq k_2$ . This is more than enough to ensure that Condition (2) holds and that there are enough vertices in order to satisfy Condition (3).

Finally assume that the auxiliary Condition (5) does not hold. In other words that there is a connected component of  $Z' \subset Z \setminus \zeta(E)$  that is isomorphic to a tree. But a tree that can be cut out of a 4-regular graph by erasing  $D$  edges has at most  $D - 2$  edges. Consider a path  $\alpha \subset Z$  that is labeled by  $w_1^A$ , with the property that the first and last edges are in  $\zeta(E)$ , but all the rest of the edges are contained in  $Z'$ . A non empty path like this always exists, upon possibly replacing  $w_1$  by  $w_2$ . Since  $Z'$  is a tree with at most  $D - 2$  edges we know that  $A \leq D$ . The image of this pass  $\iota_N(\alpha)$  in  $X_{N'}$  intersects  $E$  twice. In fact  $\iota_{N'}(\alpha)$  has to pass through the same edge  $e \in E$  twice, because we have chosen the edges in  $E$  in such a way that  $d_{X_{N'}}(e'_{k_1}, e'_{k_2}) \geq 2lD$ . This is demonstrated in Figure (2), where the two highlighted edges should be identified, these represent the edge  $e$ . It is not true that the path  $\iota_{N'}(\alpha)$  is a path without backtracking in the graph  $X_{N'}$ . But, if we write  $w_1 = v w v^{-1}$  where  $w$  is non-trivial cyclically reduced word, then there is a homotopically equivalent path labeled  $v w^d v^{-1}$  which is a path without backtracking. By minimality of  $\alpha$  this new path still passes through the edge  $e$  twice and therefore it is not homotopically trivial. Now we reach a contradiction if we require  $R$  to be large enough so that  $\text{girth}(X_{N'}) \geq R \geq 2lD$ .

By Theorem [AV05, Proposition 3.10], which is our Theorem 1.3(1) for the case  $\Delta = \Gamma$ , there is an upper bound on the number of orbits of the group  $\Gamma$  itself, independent of the level. In other words there exists a random variable  $M$ , such that the number of connected components of  $X_m$  is independent of  $m$  for  $m \geq M$ . We will impose the additional condition  $N' \geq M$ .

**4.8. Passing to a cover.** In order to satisfy Condition (4) we will choose some bigger  $N = N(K) \geq N'$ , thereby passing to a cover  $p = p_{N,N'} : X_N \rightarrow X_{N'}$ . By our auxiliary condition, for every connected



component  $Z$  of  $Y_{N'}$ , every connected component  $Z'$  of  $Z \setminus \zeta(E(Z))$  is not a tree. Let  $\alpha(Z')$  denote a non-trivial circle inside every such  $Z'$ . The importance of these circles is that, by choosing  $N$  big enough and using Corollary 4.3(2), we can require every connected component of  $p^{-1}(Z')$  to be as large as we want. In fact if we make  $\text{girth}(Y_N) \geq R$  then every lifting of every circle  $\alpha(Z')$  is at least of length  $R$ . Specifically, we choose  $R = \max\{|Z| \mid Z \text{ is a connected component of } Y_{N'}\}$  so *for every  $Z'$  every connected component of  $p^{-1}(Z')$  is at least as large as  $Z$* . Figure (3) depicts a situation where one of the components of  $Z \setminus \zeta(E)$  is contractible and therefore cannot grow when we pass to a cover.

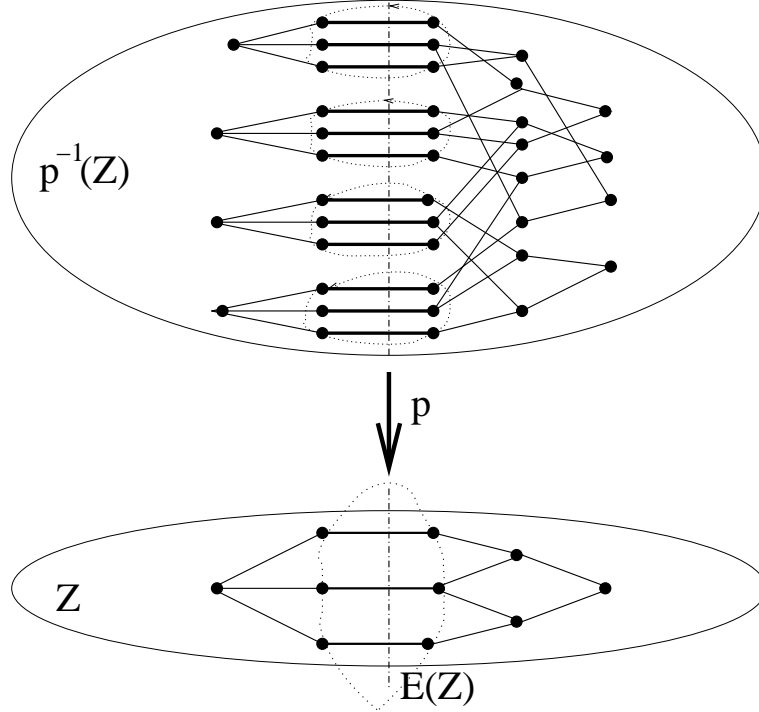


FIGURE 3. A component of  $Z \setminus \zeta(E)$  that does not grow when passing to covers.

**4.9. Connected complements.** Let  $W$  be any connected component of  $Y_N$ . Note that  $Z = p(W)$  is a connected component of  $Y_{N'}$ . In fact the map  $p$  restricted to  $W$  becomes a covering map of connected graphs  $p : W \rightarrow Z$ , say of degree  $d = d(W)$ . We choose the edges  $F(W) = \{f_k(W) \mid 1 \leq k \leq K\}$  as liftings of the edges  $f'_k(Z)$ , namely

in such a way that

$$p(f_k(W)) = f'_k(Z). \quad (4)$$

The edges  $E(W) = \{e_k(W) \mid 1 \leq k \leq K\}$  are then determined by the requirement that  $e_k(W)$  is the first edge in the path  $\iota_N(f_k(W))$ . The condition  $p(e_k(W)) = e'_k(Z)$  as well as Conditions (1), (2) and (3) from the statement of the theorem are automatically satisfied by our choices.

Assume by way of contradiction that Condition (4) cannot be satisfied by such a construction. In other words for every choice of edges  $F(W)$  satisfying the Equation (4) the graph  $W \setminus \zeta(F(W))$  is disconnected. There are  $d$  different liftings of the edge  $f'_k(Z)$ , we give them numbers as follows:  $p^{-1}(f'_k(Z)) = \{f_k(W, i) \mid 1 \leq i \leq d\}$ . And we focus our attention on  $d$  specific choices  $F(W, i) = \{f_k(W, i) \mid 1 \leq k \leq K\}$ . Since  $\zeta(E(W, i)) \cap \zeta(E(W, i')) = \emptyset \forall i \neq i'$  and since by our contradiction assumption  $W \setminus \zeta(E(W, i))$  is disconnected for every  $i$  we conclude that:

$$W \setminus \zeta(E(W, 1) \cup E(W, 2) \cup E(W, 3) \cup \dots \cup E(W, d)),$$

has at least  $d + 1$  connected components. But each one of these connected components covers a connected component of  $Z \setminus \zeta(E'(Z))$ , so by our choice of  $N$  each one of these connected components is of size at least  $|Z|$ . So  $W$  has at least  $(d + 1)|Z|$ -vertices, contradicting the fact that  $p : W \rightarrow Z$  is a covering of degree  $d$ .

**4.10. Measurability.** One has to verify the measurability of the random variables that we have constructed <sup>2</sup>. In fact we can even make all these random variables continuous, on a subset of measure one in  $\Omega$ . The first thing to note, is that Conditions (1-4) are open conditions. If conditions (1-4) are satisfied for some  $\phi \in \Omega$ , then they will be satisfied, with the same values of  $N, F, E, v$ , for any other  $\phi' \in \Omega$  as long as

$$\Psi_N \circ \phi = \Psi_N \circ \phi'. \quad (5)$$

Recall that  $\Psi_N : W(H) \rightarrow W_N(H)$  is the homomorphism that remembers only the action on the first  $N$  levels of  $T$ . Thus Equation (5) above just means that the actions  $\phi(F)$  and  $\phi'(F)$  coincide on the first  $N$  levels of the tree. Note that it is meaningful to say that the values of  $F, E, v$  are the same, because Equation (5) implies, in particular, that the graphs  $X_N, Y_N$  are the same for  $\phi$  and  $\phi'$ .

---

<sup>2</sup>For notation purposes, we forget now all the attributes of the random variables and think of them only as functions on  $\Omega$ . Thus for example  $F = F(\phi)$  is understood as a function associating to  $\phi \in \Omega$  a collection of edges  $\{f_k(Z)(\phi)\}$  in the graph  $Y_{N(\phi)}$  which is itself dependent on  $\phi$ .

We can now enumerate the possible values of the 4-tuple  $(N, E, F, v)$ , making sure that lower values of  $N$  appear first in the list. The value  $(N(\phi), E(\phi), F(\phi), v(\phi))$  is then chosen to be the first one in the list, that still satisfies conditions (1 – 4) of the theorem. All the discussion above shows that this is well defined for all but a null subset of  $\Omega$ . Moreover this construction ensures that if Equation (5) holds then  $(N(\phi), E(\phi), F(\phi), v(\phi)) = (N(\phi'), E(\phi'), F(\phi'), v(\phi'))$ , thereby proving continuity of the random variables on a co-null set. This completes the proof of Theorem 3.1.

## 5. EXPANDERS

The goal of this section is to prove Theorem 1.3(2); namely to establish the expansion property, for the connected components of the Schreier graphs  $Y_n$  - coming from the action of a subgroup of a random group on the levels of the tree. In accordance with our general strategy; we will use a known result of Bilu and Linial concerning expansion of random towers of graphs  $X_n = \mathcal{G}(F_K, S, L(n))$  and then combine it with our Theorem 3.1 to obtain a proof of Theorem 1.3(2).

**5.1. The paper of Bilu and Linial.** Bilu and Linial [BL04] treat only the case of 2-lifts; in our terminology this corresponds to the case  $d = 2$  and  $W = W(\mathbf{Z}/2\mathbf{Z})$ , the full automorphism group of the binary tree. It is not difficult to generalize their results to the general setting of the group  $W(H)$ . But since this takes us too far from the methods of the current paper we will not peruse this generalization here. The following theorem is not explicitly stated by Bilu and Linial but it does follow directly from their proof:

**Theorem.** (*Bilu-Linial*) *Let  $T$  be the binary tree,  $\phi \in \Omega$  a random action of  $F = F_m = \langle S \rangle$  on  $T$  and set  $\eta = 1/(4m^6)$ . Then with probability of at least  $(1 - 2\eta)^{1/(1-\eta)}$ , the Schreier graphs  $X_n = \mathcal{G}(F, S, L(n))$  admit a uniformly bounded spectrum:*

$$\text{Spec} X_n \subset \{\pm 2m\} \cup \left[ -C\sqrt{2m \log^3(2m)}, C\sqrt{2m \log^3(2m)} \right] \quad \forall n.$$

*In particular, for large enough values of  $m$ , the graphs  $X_n$  will be a family of expanders with high probability.*

*Proof.* We will introduce some adjustments to Bilu and Linial's notation in order to be consistent with the rest of our paper. In particular we use the following substitutions  $n \mapsto 2^n, d \mapsto r = 2m$ . Set  $\gamma(r) = 10\sqrt{r \log_2 r}$ .

**Definition.** The graph  $X_n$ , is called  $(\beta, t)$ -sparse if for every  $u, v \in \{0, 1\}^{X_n}$  with  $|\text{Supp}(u) \cup \text{Supp}(v)| \leq t$ ,

$$u^t A_n v \leq \beta \|u\| \|v\|.$$

In the above definition we used  $A_n$  for the adjacency matrix of  $X_n$ . Denote by  $A_{n,s}$  the signed adjacency matrix corresponding to the covering map  $X_{n+1} \rightarrow X_n$ . The matrix  $A_{n,s}$  is obtained from  $A_n$ , by adding signs to the matrix entries, where the sign of each edge is determined according to the value of the 1-local cocycle across this edge. It is easy to see that the graph  $X_{n+1}$  inherits all the eigenvalues of  $X_n$ , in fact the corresponding eigenfunctions are obtained by pulling back via the covering map  $p_{n,n+1} : X_{n+1} \rightarrow X_n$ . It is shown in [BL04] that the eigenvalues of  $A_{n,s}$  are exactly the “new eigenvalues”; these that are not inherited from  $X_n$  in this fashion. Let  $E(n+1)$  denote the event that both of the following hold:

- (1)  $\forall u, v \in \{-1, 0, 1\}^{2^n} : |u^t A_{n,s} v| \leq \gamma(r) \|u\| \|v\|$
- (2)  $X_{n+1}$  is  $(\gamma(r), n+1)$ -sparse.

The proof of [BL04, Lemma 3.4(2)] establishes that, conditioned on the assumption that  $X_n$  is  $(\gamma(r), n)$ -sparse the event  $E(n+1)$  holds with probability of at least

$$1 - 2 \left( 2^{n(4-6 \log_2 r)} \right) = 1 - 2\eta^n.$$

Writing out the union bound at the end of the proof gives an explicit lower bound for the probability of (2)

$$P\{(2)\} \geq 1 - 2^{(n+1)[\log_2 r(1-20 \log_2 e)+2]} \geq 1 - \eta^n.$$

In particular if  $X_{n+1}$  is  $(\gamma(r), n+1)$ -sparse then 1 cannot be violated for vectors of  $u, v$  such that  $|\text{Supp} u \cup \text{Supp} v| \leq n$ . The calculation in the beginning of the proof shows that the probability that condition (1) is satisfied by all vectors  $u, v$  such that  $|\text{Supp} u \cup \text{Supp} v| > n$  is at least

$$1 - \eta^n.$$

Now the estimate in (1) follows by applying the union bound.

The distribution of the graph  $X_{n+1}$ , conditioned on the graph  $X_n$ , is the distribution of a random 2-cover of the graph  $X_n$ . Thus the probability that the conditions  $E(n)$  hold for all  $n$  is at least as large as the infinite product:

$$\prod_{n=1}^{\infty} (1 - 2\eta^n) \geq (1 - 2\eta)^{\frac{1}{1-\eta}}.$$

Where the last inequality follows from the standard estimate on convergence of infinite products. This completes the proof of the proposition.  $\square$

**5.2. Expansion is independent of a generating set.** As mentioned earlier we are interested in properties of Schreier graphs that depend only on the action, and not on the specific choice of a generating set. While the value of the Cheeger constant depends on the choice of generating set, the qualitative property of expansion does not.

**Lemma.** *Let  $\Delta = \langle W \rangle$  be a finitely generated group and  $\Sigma = \langle V \rangle < \Delta$  a finitely generated subgroup. Then there exists a constant  $C$  with the following property. Whenever  $\Delta$  acts on a finite set  $L$ , with corresponding Schreier graphs  $Y = \mathcal{G}(\Delta, W, L)$ ,  $Z = \mathcal{G}(\Sigma, V, L)$ , Then*

$$h(Z) \leq Ch(Y).$$

Here  $h(\cdot)$  stands for the Cheeger constant of the graph as in Definition (1.1). Recall that a family of graphs regular graphs is an expander family if their Cheeger constants are bounded below by a positive constant.

*Proof.* Set  $m = \max\{|v| : v \in V\}$  and , where  $|\cdot|$  stands for word length with respect to the generating set  $W$ . As in the proof of the main theorem we have an immersion of graphs  $\iota : Z \rightarrow Y$ . This immersion takes an edge of  $Z$  to a path in  $X$  whose length is at most  $m$ . For an edge  $e \in EY$  let  $\zeta(e) = \{f \in EZ : e \subset \iota f\}$  and for  $E \subset EY$  we let  $\zeta(E) = \cup_{e \in E} \zeta(e)$ . Clearly for any set  $E \subset EY$  we have  $|\zeta(E)| \leq 2m|V||E|$ .

For a subset  $A \subset X$ , if  $f \in e_Z(A, \overline{A})$  is an edge connecting  $A$  to its complement then at least one of the edges in the path  $\iota(f)$  must also pass this boundary. Therefore  $e_Z(A, \overline{A}) \subset \zeta(e_Y(A, \overline{A}))$ . We assume without loss of generality that  $|A| \geq |\overline{A}|$  and write

$$\begin{aligned} h(Z) &= \min_{A \subset X; |A| \leq |\overline{A}|} \frac{|e_Z(A, \overline{A})|}{|A|} \\ &\leq \min_{A \subset X; |A| \leq |\overline{A}|} \frac{2m|V||e_Y(A, \overline{A})|}{|A|} = 2m|V|h(X) \end{aligned}$$

This concludes the proof, after setting  $C = 2|V| \max\{|v| : v \in V\}$   $\square$

In particular we obtain the following

**Corollary.** *Let a group  $\Delta$  act on a tree (or on any other collection of finite sets  $\{L(n) : n \in \mathbf{N}\}$ ) the expansion of the corresponding sequence of Schreier graphs is independent on the choice of the set*

of generators. Even though the exact value of the Cheeger constants changes with the generators.

*Proof.* This follows by applying the above lemma with  $\Delta = \Sigma$ .  $\square$

**5.3. Changing generators for  $\Delta$ .** By Theorem 1.3(1) there is, almost surely, a bound  $M$ , on the number of connected components of  $Y_n$ . By Theorem (5.1) there is a number  $K \in \mathbf{N}$  such that, with probability  $1 - \epsilon/M$ , a random action of  $F_K$  on  $T$  gives rise to a family of expanding Schreier graphs. Now apply Theorem (3.1) with this value of  $K$ . We obtain random variables as follows: (i) a level  $N$ , (ii) orbit representatives  $v_1, v_2, \dots, v_M \in L(N)$ , (iii) for each  $1 \leq j \leq M$  group elements  $\alpha_1(v_j), \dots, \alpha_K(v_j) \in \Delta_{v_j}$  such that the automorphisms  $\{\beta(\alpha_k(v_j), v_j)\}_{1 \leq k \leq K}$  are  $K$  independent Haar-random elements of  $W(H)$ .

We would like to extend the definition of the elements  $\alpha_i(v)$  also to vertices that do not happen to be orbit representatives. This is done by arbitrarily choosing, for every vertex  $v \in L(N)$ , a group element  $\delta_v \in \Delta$  such that  $\delta_v(v) = v_j$  for some  $j \in \{1, 2, \dots, M\}$ . We then define

$$\alpha_k(v) \stackrel{\text{def}}{=} \delta_v \alpha_k(v_j) \delta_v^{-1}.$$

Now for every  $v \in L(N)$  the group  $A_v \stackrel{\text{def}}{=} \langle \beta(\alpha_k(v), v) \mid 1 \leq k \leq K \rangle$  is generated by  $K$  independent Haar-random elements of  $W(H)$ .

The elements  $\alpha_k(v)$  are by no means independent when we vary  $v$ . No independence is guaranteed when  $v, w$  are vertices from different  $\Delta$  orbits. Moreover If  $v, w$  are two vertices in the same  $\Delta$  orbit then  $A_w$  and  $A_v$  are very much dependent, they are conjugate subgroups of  $W(H)$ ; but exactly because of this if one of them gives rise to a family of expanders so does the other one. By our choice of  $K$ , for every  $\Delta$  orbit there is a probability of at least  $1 - \epsilon/M$  that the tower of Schreier graphs coming from the action of  $A_{v_j}$  on the tree is a family of expanders. By the union bound, with probability of at least  $1 - \epsilon$ , this happens for every orbit representative and therefore for every vertex  $v \in L(n)$ . Let  $\hat{h}$  be the infimum of all the Cheeger constants on all of these graphs

$$\hat{h} \stackrel{\text{def}}{=} \inf_{v \in L(N)} \inf_n \{h(\mathcal{G}(A_v, \{\beta(\alpha_k(v), v)\}_{1 \leq k \leq K}, L(n)))\}.$$

By all we have just said

$$\mathcal{P}\{\hat{h} > 0\} > 1 - \epsilon.$$

We conclude by showing that whenever  $\hat{h} > 0$  the connected components of the graphs  $Y_n = \mathcal{G}(\Delta, W, L(n))$  form a family of expanders.

By Corollary 5.2 we are free to change the set of generators. We will in fact enlarge the original generating set, setting

$$V = W \cup \{\alpha_i(v) \mid 1 \leq i \leq K, v \in L(N)\}$$

we will prove that the family of connected components of the graphs  $Z_n \stackrel{\text{def}}{=} \mathcal{G}(\Delta, V, L(n))$  form a family of expander graphs.

**5.4. Using the expansion above every vertex.** Let us assume by way of contradiction, that one can find a level  $n$  and a connected component  $Z \subset Z_n$  with  $h(Z) < \eta$  for arbitrarily small values of  $\eta$ . Let  $p : Z \rightarrow X_N$  be the canonical covering map and  $A \subset L(n)$  a set realizing the Cheeger constant of  $Z$ . In other words we assume that  $|A| \leq |\bar{A}|$  and still  $e(A, \bar{A}) = h(Z)|A| \leq \eta|A|$ .

Given any vertex  $v \in p(Z)$  we denote by  $A(v) \stackrel{\text{def}}{=} p^{-1}(v) \cap A$  and  $\bar{A}(v) = p^{-1}(v) \cap \bar{A}$ . The condition  $\hat{h} > 0$  implies that for any given vertex  $v \in p(Z)$  either  $A(v)$  or  $\bar{A}(v)$  is very big. The expanding graph  $\mathcal{G}(A_v, \{\beta(\alpha_k(v), v)\}, L(n - N))$  is embedded in  $p^{-1}(v) \subset Z_n$ , using the expansion of this graph and setting  $\eta \leq \hat{h}/5|Y_N|$  we obtain the following estimate:

$$\begin{aligned} \min \left\{ \frac{|A(v)|}{|p^{-1}(v)|}, \frac{|\bar{A}(v)|}{|p^{-1}(v)|} \right\} &\leq \frac{|e(A(v), \bar{A}(v))|}{\hat{h}|p^{-1}(v)|} \\ &\leq \frac{|e(A, \bar{A})|}{\hat{h}|p^{-1}(v)|} \leq \frac{\eta|A|}{\hat{h}|p^{-1}(v)|} \\ &\leq \frac{\eta|Z|}{2\hat{h}|p^{-1}(v)|} = \eta \frac{|p(Z)|}{2\hat{h}} \\ &\leq \eta \frac{|Y_N|}{2\hat{h}} \leq \frac{1}{10}. \end{aligned}$$

**5.5. Passing to the graph below.** Set  $A^* = \{v \in p(Z) \mid \frac{|A(v)|}{|p^{-1}(v)|} \geq 9/10\} \subset Z_N$  and  $\bar{A}^* = \{v \in p(Z) \mid \frac{|\bar{A}(v)|}{|p^{-1}(v)|} \geq 9/10\} \subset Z_N$ . We have established in the previous paragraph that  $p(Z) = A^* \amalg \bar{A}^*$  is a partition of the graph  $p(Z)$ , mimicking the partition  $Z = A \amalg \bar{A}$  upstairs. Recall that  $p^{-1}(v) = d^{n-N}$ , the degree of the covering map, is independent of the vertex  $v$ . Let  $e \in e_{p(Z)}(A^*, \bar{A}^*)$  be an edge connecting  $v \in A^*$  and  $w \in \bar{A}^*$ . At least 90% of the  $d^{n-N}$  vertices covering  $v$  (resp.  $w$ ) are in  $A^*$  (resp.  $\bar{A}^*$ ). Therefore at least 80% of the  $d^{n-N}$  edges covering  $e$  are

in  $e_Z(A, \overline{A})$ , so that  $|e_Z(A, \overline{A})| \geq 4d^{n-N}|e_{p(Z)}(A^*, \overline{A}^*)|/5$ , so

$$\begin{aligned} \eta \geq h(Z) &= \frac{|e_Z(A, \overline{A})|}{|A|} \geq \frac{|e_Z(A, \overline{A})|}{|Z|} \\ &\geq \frac{4d^{n-N}|e_{p(Z)}(A^*, \overline{A}^*)|/5}{d^{n-N}|p(Z)|} \geq \frac{4}{5|p(Z)|} \end{aligned}$$

contradicting our assumption that  $\eta$  can be chosen to be arbitrarily small and completing the proof.  $\square$

## 6. THE ANALOGY WITH ARITHMETIC GROUPS - AN EXPLICIT EXAMPLE.

In this section we focus on a very specific arithmetic example of strong approximation and compare some of the known arithmetic results surveyed in the introduction to the probabilistic results that we obtained in the current paper. We chose an arithmetic example that comes particularly close to our probabilistic setting because it too involves an action on a tree. In fact it was with this example in mind that I started to look for probabilistic analogues for strong approximation theorems. This section cannot be considered as a survey of strong approximation, since the example I treat here is very specific. For good such surveys we refer the reader to [LS03, Window: strong approximation for linear groups], [Nik09] [PR94, chapter 7].

**6.1. A specific arithmetic example.** Consider the group  $\Gamma_{\text{arith}} = \text{PSL}_2(\mathbf{Z})$ , fix a prime  $p$ . The standard embedding  $\mathbf{Z} \hookrightarrow \mathbf{Z}_p$  into the ring of  $p$ -adic integers gives rise to a group embedding  $\Gamma_{\text{arith}} \hookrightarrow \text{PSL}_2(\mathbf{Z}_p) \hookrightarrow \text{PSL}_2(\mathbf{Q}_p)$ . It is well known that the group  $\text{PSL}_2(\mathbf{Q}_p)$  admits a natural action on a tree  $T$ , its Bruhat-Tits tree, and that the group  $\text{PSL}_2(\mathbf{Z}_p)$  fixes a point  $O \in T$ . Therefore the above embedding renders an action of  $\Gamma_{\text{arith}}$  on the rooted tree  $(T, O)$ . Note that this is not exactly the  $p$ -ary rooted tree because the first level has  $p+1$  vertices. This fact is very important in other settings because it means that  $\text{Aut}(T)$  is not compact but it will not play an important role in our current discussion. In particular we will work only with the compact group  $A = \text{Aut}(T, O)$  of all automorphisms fixing the basepoint  $O$ .

The congruence homomorphism  $\Psi_n : \Gamma_{\text{arith}} \rightarrow \Gamma_{\text{arith}, n}$  coming from the restriction of the action on the tree to the first  $n$ -levels, can be identified with the arithmetically defined congruence map

$$\begin{aligned} \Psi_n : \text{PSL}_2(\mathbf{Z}) &\rightarrow \text{PSL}_2(\mathbf{Z}/p^n\mathbf{Z}) \\ \begin{pmatrix} a & b \\ c & d \end{pmatrix} &\mapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \pmod{p^n} \end{aligned}$$



In fact it is possible to identify  $L(n)$ , the  $n^{\text{th}}$  level of the tree, with the projective line  $\mathbf{P}^1(\mathbf{Z}/p^n\mathbf{Z})$  in such a way that the action of  $\Gamma_{\text{arith},n} \curvearrowright L(n)$  is identified with the natural action of  $\text{PSL}_2(\mathbf{Z}/p^n\mathbf{Z}) \curvearrowright \mathbf{P}^1(\mathbf{Z}/p^n\mathbf{Z})$ . Or, what is the same thing, with the action  $\text{PSL}_2(\mathbf{Z}/p^n\mathbf{Z}) \curvearrowright \frac{\text{PSL}_2(\mathbf{Z}/p^n\mathbf{Z})}{B_n}$  on the cosets of the subgroup

$$B_n \stackrel{\text{def}}{=} \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \right\} < \text{PSL}_2(\mathbf{Z}/p^n\mathbf{Z}).$$

A word of caution is due: since  $\mathbf{Z}/p^n\mathbf{Z}$  is a ring and not a field one *cannot* identify the projective line with  $\mathbf{Z}/p^n\mathbf{Z} \cup \{\infty\}$ .

If we want the arithmetic example to be even more similar to the probabilistic example, we can replace  $\Gamma_{\text{arith}}$  with a finite index free subgroup of  $\Gamma_{\text{arith}}$ . This will change nothing in the discussion that follows.

**6.2. A specific probabilistic example.** Consider the same tree  $(T, O)$ , the  $(p+1)$  regular tree rooted at some base-point  $O$ . Let  $\Gamma_{\text{prob}}$  be a subgroup generated by  $m$  Haar random independent elements of  $A = \text{Aut}(T, O)$ . We consider the group  $\Gamma_{\text{prob}}$  with its non cyclic subgroups  $\Delta < \Gamma_{\text{prob}}$  and the congruence reductions

$$\psi_n : \Gamma_{\text{prob}} < A \rightarrow \text{Aut}_n(T).$$

coming from the restriction of the action to the  $n^{\text{th}}$  level of the tree. all the theorems that we proved hold in this setting.

**6.3. strong approximation - Cayley or Schreier graphs?** The classical strong approximation theorem says, in our restricted setting, that the congruence maps  $\psi_n : \Gamma_{\text{arith}} \rightarrow \text{PSL}_2(\mathbf{Z}/p^n\mathbf{Z})$  are surjective. Equivalently  $\Gamma_{\text{arith}} = \text{PSL}_2(\mathbf{Z})$  is dense in  $\text{PSL}_2(\mathbf{Z}_p)$ . Already this elementary observation, is false in the probabilistic setting. The group  $A = \text{Aut}(T, O)$  is not finitely generated as a topological group. Therefore there is no hope of  $\Gamma_{\text{prob}}$  ever mapping onto all of the groups  $\text{Aut}_n(T)$ . In order to see this note that the map

$$\begin{aligned} \text{Aut}(T, O) &\rightarrow \prod_{n \in \mathbf{N}} \mathbf{Z}/2\mathbf{Z} \\ \sigma &\mapsto \{\text{sgn}(\sigma|_{L(n)})\}_{n \in \mathbf{N}} \end{aligned}$$

sending each automorphism to the list of signatures of the permutations that it induces on the various levels of the tree, is surjective.

In order to say something meaningful we must weaken our requirements at this early stage of the discussion. There are three natural ways to do this, two of which are addressed in the current paper.

- Require  $\Gamma_{\text{prob}} < \text{Aut}(T, O)$  to be large in a weaker sense. We establish this in theorem 1.4, showing that the Hausdorff dimension of  $\Gamma_{\text{prob}}$  is positive.
- Instead of discussing the image of  $\Gamma_{\text{prob}}$  in  $\text{Aut}_n(T)$  and the corresponding Cayley graph  $\mathbf{C}_{\text{prob},n} := \mathcal{C}(\Gamma_{\text{prob}}, S, \text{Aut}_n(T))$  we can talk about the Schreier graphs  $X_{\text{prob},n} := \mathcal{G}(\Gamma_{\text{prob}}, S, L(n))$  coming from the action on the  $n^{\text{th}}$  level of the tree. In a similar way<sup>3</sup> one defines  $\mathbf{C}_{\text{arith},n} := \mathcal{C}(\Gamma_{\text{arith}}, S, \text{PSL}_2(\mathbf{Z}/p^n\mathbf{Z}))$  and  $X_{\text{arith},n} := \mathcal{G}(\Gamma_{\text{arith}}, S, L(n))$ . In both arithmetic and probabilistic settings we have a covering map  $\mathbf{C}_n \rightarrow X_n$  from the Cayley to the Schreier graph. It follows that the latter is connected (resp. a good expander) if the former is. This is because the Cayley graph inherits all the eigenvalues of the Schreier graph.

Let us note though that in the arithmetic case the difference between the two graphs is polynomial  $|\mathbf{C}_{\text{arith},n}| \cong |X_{\text{arith},n}|^3$ ; while in the probabilistic setting, the difference is exponential  $|\mathbf{C}_{\text{prob},n}| \cong (p!)^{|X_{\text{prob},n}|/(p-1)}$ . This gives another indication that the Schreier graphs are the right setting for the generalization in the probabilistic setting.

- Finally not all of the groups  $W(H)$  fail to be finitely generated. For example it is shown in [BM00] that the group  $W(\text{Alt}(d))$  is finitely generated. Where  $\text{Alt}(d)$  is the alternating group. It is an interesting question whether a random finitely generated subgroup there is dense with positive probability. If so, one could ask to what extent the corresponding Cayley graphs exhibits any of the other *strong approximation phenomena* treated in this paper

**6.4. Path expansion - the combinatorial interpretation of passing to subgroups.** In two ways we seek to strengthen the statement that the Schreier graphs  $\{X_{\text{arith},n}\}$  are connected; or at least have a bounded number of connected components, independent of  $n$ . First requiring that the connected components of these graphs form a family of expanders; second requiring that the Schreier graphs coming from every non cyclic subgroup still have a bounded number of components. Ultimately, in Theorem 1.3, we end up combining these two properties, establishing expansion of the connected components for every non-cyclic subgroup.

---

<sup>3</sup>Note that this definition of the arithmetic Cayley graphs differs from the one we used in the introduction where we used  $\mathbf{C}_N$  as  $\mathcal{C}(\text{PSL}_2(\mathbf{Z}/N\mathbf{Z}))$ . The current notation is much more convenient here though as it agrees with the probabilistic notation.

**Definition.** A family of Schreier graphs  $\mathcal{G}(\Gamma, S, L(n))$  are called a family of *path-expanders* if for every non-cyclic subgroup  $\Delta < \Gamma$  there is a bound  $M = M(\Delta)$  on the number of connected components of  $\mathcal{G}(\Delta, \cdot, L(n))$ .

Like expansion, path expansion has an appealing combinatorial interpretation. Both notions measure “degree of connectivity” by abusing the graphs and verifying to what extent they remain connected. In expander graphs, this is done by erasing edges. In the case of path expansion this is done by allowing to pass only along edge paths that admit certain labeling - the labeling corresponding to a generating set of any non cyclic subgroup  $\Delta$ ; and requiring that the number of connected components remain bounded on all graphs. Of course the number of connected components does depends on  $\Delta$ . Indeed given any family of Schreier graphs  $X_n = \mathcal{G}(\Gamma, S, L(n))$  one can choose a finite index subgroup  $\Delta < \Gamma$  that will act trivially on the finite set  $L(n)$ . With respect to such a  $\Delta$  the graph  $\mathcal{G}(\Delta, \cdot, L(n))$  will just be a union of isolated vertices.

Here is the hierarchy of strong approximation properties that we treat in this paper:

**Combined**

The connected components of  $\mathcal{G}(\Delta, \cdot, L(n))$  form a family of expander graphs, for every non-cyclic  $\Delta < \Gamma$ .

**Path expansion**

Bounded number of connected components for  $\mathcal{G}(\Delta, \cdot, L(n))$ , for every non-cyclic  $\Delta < \Gamma$ .

**Expansion**

The connected components of  $\mathcal{G}(\Gamma, S, L(n))$  form a family of expanders.

**Connectedness**

Bounded number of connected components for  $\mathcal{G}(\Gamma, S, L(n))$ .

**6.5. Connectedness.** In the arithmetic setting the strong approximation theorem says that the maps  $\psi : \mathrm{PSL}_2(\mathbf{Z}) \rightarrow \mathrm{PSL}_2(\mathbf{Z}/N\mathbf{Z})$  are onto, and in particular all the relevant Schreier graphs are connected. In the probabilistic setting the fact that the Schreier graphs, almost surely have boundedly many connected components is due to Abért and virág [AV05, Proposition 3.10] here stated as Theorem 3.2.

**6.6. Expansion.** In the arithmetic setting, the expansion of Cayley and Schreier graphs of  $\mathrm{PSL}_2(\mathbf{Z}/N\mathbf{Z})$ , with a generating set coming from a generating set of  $\mathrm{PSL}_2(\mathbf{Z})$  is a theorem due to Lubotzky Phillips and Sarnak. Their proof of this fact is not purely graph or even group theoretic, as it passes through hyperbolic geometry and uses Selberg's  $\frac{3}{16}$  theorem.

In the probabilistic case the corresponding theorem is established only for binary trees and stated here as Theorem 5.1. I do not think that it should be difficult to generalize this result to more general trees.

**6.7. Path expansion.** In the arithmetic setting the path expansion of Cayley and Schreier graphs of congruence quotients is known as *the strong approximation theorem for linear groups*. Stated here for the specific case of  $\mathrm{SL}_n$ .

**Theorem.** (*Strong approximation theorem for linear groups*). *For every Zariski dense subgroup  $\Delta < \mathrm{SL}_n(\mathbf{Z})$ , there exists a number  $M = M(\Delta)$  such that*

$$[\psi(\mathrm{SL}_n(\mathbf{Z})) : \psi(\Delta)] < M$$

*for every congruence map  $\psi : \mathrm{SL}_n(\mathbf{Z}) \rightarrow \mathrm{SL}_n(\mathbf{Z}/N\mathbf{Z})$ .*

This theorem was first proved by Weisfeiler [Wei84] using the classification of finite simple groups. Later a classification free proof was given by Nori [Nor87] using methods of algebraic geometry. Other treatments of the theorem can be found in [LS03], [HP95], [Pin00]. The following follows directly by applying the theorem in the case  $n = 2$ , and recalling that the only subgroups of  $\mathrm{PSL}_2(\mathbf{Z})$  that fail to be Zariski dense are cyclic groups:

**Corollary.** *The family of Schreier and Cayley graphs of the groups  $\mathrm{PSL}_2(\mathbf{Z}/N\mathbf{Z})$ , with respect to a generating set coming from a generating set of  $\mathrm{SL}_2(\mathbf{Z})$ , forms a family of path expanders.*

In the probabilistic setting the analogous theorem is our Theorem 1.3(1).

**6.8. Combination.** Establishing expansion of Cayley and Schreier graphs of  $\mathrm{PSL}_2(\mathbf{Z}/N\mathbf{Z})$  with respect to generators coming from Zariski dense subgroups of  $\mathrm{SL}_2(\mathbf{Z})$  is just very recently established by Bourgain and Gamburd in [BG08a], [BG08c], [BG08b]; generalizing previous works of Gamburd [Gam02] and Shalom [Sha99], [Sha97]. The following theorem for example

**Theorem.** (*Bourgain-Gamburd* [BG08a]) *Cayley graphs of  $\mathrm{PSL}_d(\mathbf{Z}/p^n\mathbf{Z})$  are expanders with respect to the projection of any fixed elements in  $\mathrm{SL}_n(\mathbf{Z})$  generating a Zariski-dense subgroup*

Note that this is the best result known to date. The same result is not known for other algebraic groups, and even for different types of congruences. In [BG08c] a similar result is proved for the group  $\mathrm{SL}_2(\mathbf{F}_p)$  where  $p$  ranges over the prime numbers.

In the probabilistic setting the analogous result is our Theorem 1.3(2).

## REFERENCES

- [AL02] Alon Amit and Nathan Linial. Random graph coverings. I. General theory and graph connectivity. *Combinatorica*, 22(1):1–18, 2002.
- [ALM02] Alon Amit, Nathan Linial, and Jiří Matoušek. Random lifts of graphs: independence and chromatic number. *Random Structures Algorithms*, 20(1):1–22, 2002.
- [AV05] Miklós Abért and Bálint Virág. Dimension and randomness in groups acting on rooted trees. *J. Amer. Math. Soc.*, 18(1):157–192 (electronic), 2005.
- [BG08a] Jean Bourgain and Alex Gamburd. Expansion and random walks in  $\mathrm{SL}_d(\mathbf{Z}/p^n\mathbf{Z})$ . I. *J. Eur. Math. Soc. (JEMS)*, 10(4):987–1011, 2008.
- [BG08b] Jean Bourgain and Alex Gamburd. Random walks and expansion in  $\mathrm{SL}_d(\mathbf{Z}/p^n\mathbf{Z})$ . *C. R. Math. Acad. Sci. Paris*, 346(11-12):619–623, 2008.
- [BG08c] Jean Bourgain and Alex Gamburd. Uniform expansion bounds for Cayley graphs of  $\mathrm{SL}_2(\mathbf{F}_p)$ . *Ann. of Math. (2)*, 167(2):625–642, 2008.
- [Bha95] Meenaxi Bhattacharjee. The ubiquity of free subgroups in certain inverse limits of groups. *J. Algebra*, 172(1):134–146, 1995.
- [BL04] Yonatan Bilu and Nati Linial. Constructing expander graphs by 2-lifts and discrepancy vs. spectral gap. arXiv:math.CO/0312022, April 2004.
- [BM00] Mark Burger and Shahar Mozes. Groups acting on trees: From local to global structure. *Inst. Hautes Études Sci. Publ. Math.*, (92):113–150, 2000.
- [BS97] Yiftach Barnea and Aner Sahlev. Hausdorff dimension, pro- $p$  groups, and Kac-Moody algebras. *Trans. Amer. Math. Soc.*, 349(12):5073–5091, 1997.
- [Gam02] Alex Gamburd. On the spectral gap for infinite index “congruence” subgroups of  $\mathrm{SL}_2(\mathbf{Z})$ . *Israel J. Math.*, 127:157–200, 2002.
- [HP95] Ehud Hrushovski and Anand Pillay. Definable subgroups of algebraic groups over finite fields. *J. Reine Angew. Math.*, 462:69–91, 1995.
- [LS03] Alexander Lubotzky and Dan Segal. *Subgroup growth*, volume 212 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 2003.
- [Nik09] Nikolay Nikolov. Strong approximation methods in group theory, an lms/epsrc short course lecture notes. arXiv:0803.4165, 2009.
- [Nor87] Madhav V. Nori. On subgroups of  $\mathrm{gl}_n(\mathbf{F}_p)$ . *Invent. Math.*, 88(2):257–275, 1987.
- [Pin00] Richard Pink. Strong approximation for Zariski dense subgroups over arbitrary global fields. *Comment. Math. Helv.*, 75(4):608–643, 2000.

- [PR94] Vladimir Platonov and Andrei Rapinchuk. *Algebraic groups and number theory*, volume 139 of *Pure and Applied Mathematics*. Academic Press Inc., Boston, MA, 1994. Translated from the 1991 Russian original by Rachel Rowen.
- [Sha97] Yehuda Shalom. Expanding graphs and invariant means. *Combinatorica*, 17(4):555–575, 1997.
- [Sha99] Yehuda Shalom. Expander graphs and amenable quotients. In *Emerging applications of number theory (Minneapolis, MN, 1996)*, volume 109 of *IMA Vol. Math. Appl.*, pages 571–581. Springer, New York, 1999.
- [Wei84] Boris Weisfeiler. Strong approximation for Zariski-dense subgroups of semisimple algebraic groups. *Ann. of Math. (2)*, 120(2):271–315, 1984.

*E-mail address:* `yair@ias.edu`

SCHOOL OF MATHEMATICS, INSTITUTE FOR ADVANCED STUDY, EINSTEIN DRIVE SIMONYI HALL, PRINCETON, NJ 08540.