

# Uniform rank gradient, cost and local-global convergence

Miklós Abért and László Márton Tóth\*

December 4, 2017

## Abstract

We analyze the rank gradient of finitely generated groups with respect to sequences of subgroups of finite index that do not necessarily form a chain, by connecting it to the cost of p.m.p. actions. We generalize several results that were only known for chains before. The connection is made by the notion of local-global convergence.

In particular, we show that for a finitely generated group  $\Gamma$  with fixed price  $c$ , every Farber sequence has rank gradient  $c-1$ . By adapting Lackenby's trichotomy theorem to this setting, we also show that in a finitely presented amenable group, every sequence of subgroups with index tending to infinity has vanishing rank gradient.

## 1 Introduction

For a finitely generated group  $\Gamma$  let  $d(\Gamma)$  denote the minimal number of generators (or rank) of  $\Gamma$ . For a subgroup  $H \leq \Gamma$  of finite index let

$$r(\Gamma, H) = (d(H) - 1) / |\Gamma : H|.$$

The *rank gradient* of  $\Gamma$  with respect to a sequence  $(\Gamma_n)$  of finite index subgroups is defined to be

$$\text{RG}(\Gamma, (\Gamma_n)) = \lim_{n \rightarrow \infty} r(\Gamma, \Gamma_n)$$

when this limit exists. This notion has been introduced by Lackenby [Lac] and further investigated in the literature, mainly for *chains* of subgroups.

---

\*The authors were supported by the Hungarian National Research, Development and Innovation Office, NKFIH grant K109684 and the ERC Consolidator Grant 648017.

Recall that a chain in  $\Gamma$  is a decreasing sequence  $\Gamma = \Gamma_0 > \Gamma_1 > \dots$  of subgroups of finite index in  $\Gamma$ . In this case, it is easy to see that  $r(\Gamma, \Gamma_n)$  is non-increasing and so the limit exists.

The main goal of this paper is to offer a general framework for understanding the rank gradient of an arbitrary sequence of subgroups in  $\Gamma$  using the cost of probability measure preserving (p.m.p.) actions of  $\Gamma$ . For chains this has been done by the first author and Nikolov in [AbN]. In [AGN] arbitrary sequences were analyzed for a special class of groups called right angled groups.

Let  $(G_n)$  be a sequence of finite graphs with an absolute degree bound. We define the *edge density* as

$$e(G_n) = \lim_{n \rightarrow \infty} \frac{|E(G_n)|}{|V(G_n)|}$$

when this limit exists and the *lower edge density*  $\underline{e}$  to be the  $\liminf$  of the same sequence. A *rewiring* of  $(G_n)$  is another sequence of graphs  $H_n$  on the same vertex set as of  $G_n$ , such that the bi-Lipshitz distortion of the maps  $\text{id}_{V(G_n)}$  stay bounded in  $n$ . The *combinatorial cost*  $\text{cc}(G_n)$  is defined as the infimum of the lower edge densities of possible rewirings of  $(G_n)$ . This notion has been introduced by Elek [El] as a discrete analogue of the notion of cost.

Our first result shows that actually combinatorial cost is more than an analogue and, when making an additional convergence assumption, it can be expressed as the cost of a limiting graphing.

Local-global convergence of graphs has been introduced by Bollobás and Riordan [BoR] under the name partition metric, while the limiting object and most of the known results were obtained by Hatami, Lovász and Szegedy [HLSz]. On the group theory side, the notion is related to the work of Kechris on weak containment, see [Kec] and [AbE1]. We postpone its definition to Section 2. For now, it suffices to know that every sequence has a convergent subsequence and that the limit is a graphing in the sense of Gaboriau [Gab].

**Theorem 1.** *Let  $G_n$  be a local-global convergent graph sequence. Then we have*

$$\text{cost}(\lim G_n) = \text{cc}(G_n).$$

*Moreover, one can choose rewirings such that the limit defining the edge density exists.*

Note that similar results to Theorem 1 and its consequences (until Corollary 5) have been obtained independently by A. Carderi, D. Gaboriau and M. de la Salle [CGS, 2017].

This result can be effectively used to give direct proofs of results on combinatorial cost using the already established theory of cost. For instance, Theorem 1 immediately implies the following theorem of Elek [El].

**Corollary 2.** *Let  $G_n$  be a graph sequence with girth tending to infinity such that  $e(G_n)$  exists. Then we have*

$$\text{cc}(G_n) = e(G_n).$$

Indeed, by the girth assumption, any subsequential local-global limit of  $G_n$  will be a so called treeing and by Gaboriau [Gab], the cost of a treeing equals its expected degree divided by two.

In Theorem 1, our graphs a priori have nothing to do with groups. When they do come from a sofic approximation of  $\Gamma$ , the limiting graphing gives rise to an essentially free probability measure preserving action of  $\Gamma$ , that is unique up to weak equivalence in the sense of Kechris [Kec]. This case of local-global convergence has been analyzed by the first author and Elek [AbE2].

Following Gaboriau, we say that  $\Gamma$  has fixed price  $c$ , if every essentially free probability measure preserving action of  $\Gamma$  has cost  $c$ . Applying Theorem 1 gives us the following new result.

**Theorem 3.** *Let  $\Gamma$  be a finitely generated group of fixed price  $c$ . Then*

$$\text{cc}(G_n) = c$$

*for any sofic approximation  $(G_n)$  of  $\Gamma$ .*

This gives an alternate proof of another result of Elek [El] that for an amenable group  $\Gamma$ , any sofic approximation of  $\Gamma$  has combinatorial cost 1. Indeed, by the Ornstein-Weiss theorem [OrW], amenable groups have fixed price 1.

A sequence of subgroups is *Farber*, if the quotient Schreier graphs  $\text{Sch}(\Gamma, \Gamma_n, S)$  form a sofic approximation of  $\Gamma$ . We can now connect the cost to the rank gradient as follows.

**Theorem 4.** *Let  $\Gamma$  be a finitely generated group of fixed price  $c$ . Then we have*

$$\text{RG}(\Gamma, (\Gamma_n)) = c - 1$$

*for any Farber sequence  $(\Gamma_n)$  in  $\Gamma$ .*

The same result is proved in [AbN] for Farber *chains*. Also, in [AGN] it is proved that any Farber sequence in a right angled group has rank gradient zero. Right angled groups have fixed price 1 by Gaboriau [Gab], so this now immediately follows from Theorem 4. Since amenable groups also have fixed price 1, we get the following.

**Corollary 5.** *Let  $\Gamma$  be a finitely generated amenable group. Then we have*

$$\mathrm{RG}(\Gamma, (\Gamma_n)) = 0$$

*for any Farber sequence  $(\Gamma_n)$  in  $\Gamma$ .*

When the sequence is not Farber, Corollary 5 is clearly not true, already for the standard lamplighter group (see [AJN]). However, one can show that it still holds for *finitely presented* amenable groups.

**Theorem 6.** *Let  $\Gamma$  be a finitely presented amenable group. Then we have*

$$\mathrm{RG}(\Gamma, (\Gamma_n)) = 0$$

*for any sequence  $(\Gamma_n)$  of distinct subgroups in  $\Gamma$ .*

Behind this is the following extension of [AJN] that generalized a theorem of Lackenby for normal chains [Lac].

We call a sequence of finite graphs  $G_n$  *dispersive* if for any subsequential local-global limit  $\mathcal{G}$  of  $G_n$ ,  $\mathcal{G}$  has no strongly ergodic component of positive measure. For the notion of strong ergodicity and a graph theoretic reformulation see Section 5.

**Theorem 7.** *Let  $\Gamma$  be a finitely presented group generated by a finite symmetric set  $S$ . Let  $(\Gamma_n)$  be an arbitrary sequence of subgroups of finite index in  $\Gamma$ . Then at least one of the following holds:*

- 1) *the sequence  $\mathrm{Sch}(\Gamma, \Gamma_n, S)$  is not dispersive;*
- 2)  *$\mathrm{RG}(\Gamma, (\Gamma_n)) = 0$ ;*
- 3) *there exists some  $n$  such that  $\Gamma_n$  decomposes as a non-trivial amalgamated product.*

We show that sequences in amenable groups are dispersive, and clearly they cannot decompose as a non-trivial amalgamated product. Thus Theorem 6 follows as a corollary of Theorem 7. Note that for a chain of subgroups, being dispersive is equivalent to saying that the limiting profinite action is not strongly ergodic. Hence, Theorem 7 implies [AJN, Theorem 3]. When the  $\Gamma_n$  are normal in  $\Gamma$ , being dispersive is equivalent to saying that  $(\Gamma_n)$  has no subsequence with Lubotzky's property  $(\tau)$ . So Theorem 7 also generalizes Lackenby's trichotomy theorem [Lac, Theorem 1.1.].

The structure of the paper is as follows. In Section 2 we define the basic notions and state some lemmas that we need for our main result. In Section 3 we prove Theorems 1 and 3. We introduce the analogous notions and results for group actions in Section 4 and prove Theorem 4. We prove the results on finitely presented groups in Section 5. Finally, in Section 6 we list some open problems and suggest further directions of research.

## 2 Preliminaries

In this section we define the basic objects of investigation of the paper and state some known results.

### 2.1 Local-global convergence

Let  $U_r$  denote the set of connected, rooted graphs with radius at most  $r$  with all degrees bounded by some integer  $D$ . For any graph  $G$ , if we pick a vertex  $v \in V(G)$  and look at its  $r$ -neighborhood  $B_G(r, v)$  rooted at  $v$  we get an element of  $U_r$ . Picking  $v$  uniformly at random gives us a probability measure on  $U_r$  which we will denote  $P_{G,r}$ , and refer to as the  $r$ -neighborhood statistics of  $G$ .

For any finite set  $X$  let  $M(X)$  denote the set of probability measures on  $X$ . We say that a sequence of graphs  $(G_n)$  is *locally* (or Benjamini-Schramm) convergent, if for any  $r$  the sequence of probability measures  $P_{G_n,r} \in M(U_r)$  converge to a limit distribution as  $n \rightarrow \infty$ .

We will work with a more refined notion of convergence, and following notation from [HLSz] we introduce a colored version of the neighborhood statistics. Let  $K(k, G) = \{\varphi : V(G) \rightarrow \{1, \dots, k\}\}$  denote the set of  $k$ -colorings of the vertices of  $G$ . Let  $U_r^k$  denote the set of rooted, connected,  $k$ -colored graphs of radius at most  $r$ . For any coloring  $\varphi \in K(k, G)$  we can associate a colored neighborhood statistic  $P_{G,r}[\varphi] \in M(U_r^k)$  as before, by choosing a uniform random vertex  $v$ , and then considering its  $r$ -neighborhood  $B_G(r, v)$ , this time together with the coloring  $\varphi|_{B_G(r,v)}$ .

For  $\eta_1, \eta_2 \in M(U_r^k)$  let

$$d_{TV}(\eta_1, \eta_2) = \sup_{A \subseteq U_r^k} |\eta_1(A) - \eta_2(A)|.$$

Note that  $d_{TV}$  is the total variation distance. As we are operating in a finite dimensional space all the usual norms are equivalent.

Intuitively, a sequence of graphs  $(G_n)$  is local-global convergent if for any  $r, k \in \mathbb{N}$  and for  $i, j$  large enough the colored neighborhood distribution  $P_{G_i,r}[\varphi]$  for any  $k$ -coloring  $\varphi$  can be approximately modeled on  $G_j$ , that is we can find some coloring  $\psi$  such that  $P_{G_i,r}[\varphi]$  and  $P_{G_j,r}[\psi]$  are arbitrarily close.

For a finite graph  $G$  let  $Q_{G,r}^k$  denote the finite set of possible colored neighborhood statistics arising from a graph  $G$ :

$$Q_{G,r}^k = \{P_{G,r}[\varphi] \mid \varphi \in K(k, G)\} \subseteq M(U_r^k).$$

**Definition 8.** We say that a sequence of graphs  $(G_n)$  is local-global convergent if for every  $r, k \in \mathbb{N}$  the compact sets  $(Q_{G_n, r}^k)$  converge in the Hausdorff distance on  $(M(U_r^k), d_{TV})$ .

In [HLSz] the authors show that every sequence of bounded degree graphs has a locally-globally convergent subsequence, and that *graphings* can be considered as the limit objects of convergent sequences.

**Definition 9** ([HLSz] Definition 3.1). Let  $X$  be a Polish topological space and let  $\mu$  be a probability measure on the Borel sets in  $X$ . A graphing (with degree bound  $D$ ) is a graph  $\mathcal{G}$  on  $V(\mathcal{G}) = X$  with Borel edge set  $E(\mathcal{G}) \subset X \times X$  in which all degrees are at most  $D$  and

$$\int_A e(x, B) d\mu(x) = \int_B e(x, A) d\mu(x) \quad (1)$$

for all measurable sets  $A, B \subseteq X$ , where  $e(x, S)$  is the number of edges from  $x \in X$  to  $S \subseteq X$ .

Every finite graph  $G$  is a graphing with  $X = V(G)$  and  $\mu$  the uniform distribution on  $V(G)$ .

The colored neighborhood statistics  $P_{\mathcal{G}, r}[\varphi]$  can easily be defined for a graphing  $\mathcal{G}$ , provided that the coloring  $\varphi : X \rightarrow \{1, \dots, k\}$  is chosen to be Borel. We pick a random vertex  $x \in X$  according to  $\mu$ , and consider its colored  $r$ -neighborhood in  $\mathcal{G}$ .

As opposed to the finite case we now have to take the closure of all possible such statistics in order to obtain a compact set. Let

$$Q_{\mathcal{G}, r}^k = \overline{\{\mathcal{P}_{\mathcal{G}, r}[\varphi] \mid \varphi : V(\mathcal{G}) \rightarrow \{1, \dots, k\} \text{ Borel}\}}^{d_{TV}} \subseteq M(U_r^k).$$

The graphing  $\mathcal{G}$  is a local-global limit of the sequence  $(G_n)$  if  $Q_{G_n, r}^k \rightarrow Q_{\mathcal{G}, r}^k$  in the Hausdorff distance for all  $r$  and  $k$ .

We say that two graphings  $\mathcal{G}$  and  $\mathcal{H}$  are local-global equivalent, if the sets  $Q_{\mathcal{G}, r}^k$  and  $Q_{\mathcal{H}, r}^k$  are the same. Note that the limit is unique only up to local-global equivalence. Although we will only be dealing with sequences of finite graphs, observe that the above definition of convergence makes sense for sequences of graphings as well.

## 2.2 Cost

For a graphing  $\mathcal{G}$  and a vertex  $x \in X$  let  $[x]_{\mathcal{G}}$  denote the connected component of  $x$  in  $\mathcal{G}$ . For two graphings  $\mathcal{G}$  and  $\mathcal{H}$  on the same vertex set  $X$  we write

$\mathcal{G} \sim \mathcal{H}$  if they have the same connected components, that is  $[x]_{\mathcal{G}} = [x]_{\mathcal{H}}$   $\mu$ -almost surely.

Let  $\mathcal{R}_{\mathcal{G}} \subseteq X \times X$  denote the measurable equivalence relation generated by  $\mathcal{G}$ , where two points are in the same equivalence class if they are in the same connected component of  $\mathcal{G}$ . Clearly  $\mathcal{G} \sim \mathcal{H}$  if and only if  $\mathcal{R}_{\mathcal{G}} = \mathcal{R}_{\mathcal{H}}$  up to measure zero. Note that every component of  $\mathcal{G}$  is countable.

We will introduce a way of measuring edge sets of graphings. Let  $\tilde{\mu}$  be the measure on  $X \times X$  obtained the following way. For a measurable subset  $C \subseteq X \times X$  let

$$\tilde{\mu}(C) = \int_X \#\{y \mid (x, y) \in C, y \in [x]_{\mathcal{G}}\} d\mu(x).$$

In other words on each fiber  $\{x\} \times X$  we consider the counting measure concentrated on  $\{x\} \times [x]_{\mathcal{G}}$ , and integrate these with respect to  $\mu$  on the first coordinate.

This measure  $\tilde{\mu}$  is  $\sigma$ -finite, it is concentrated on  $\mathcal{R}_{\mathcal{G}}$ , and it is easy to see that in fact it only depends on the relation  $\mathcal{R}_{\mathcal{G}}$ . We can similarly define  $\tilde{\mu}'$  by taking the counting measures on the fibers  $[x]_{\mathcal{G}}$  and integrate with respect to  $\mu$  over the second coordinate. A standard argument shows that condition (1) in Definition 9 is equivalent to  $\tilde{\mu} = \tilde{\mu}'$ .

The *cost* of  $\mathcal{G}$  is defined to be

$$\text{cost}(\mathcal{G}) = \frac{1}{2} \inf \{ \tilde{\mu}(E(\mathcal{H})) \mid \mathcal{H} \sim \mathcal{G} \}.$$

The normalization factor  $\frac{1}{2}$  is included to account for counting every edge twice and to ensure coherence with [Gab]. Note that the  $\tilde{\mu}$  measure of the edge set of a graphing is half the expected degree of a  $\mu$ -random point. It is clear that if  $\mathcal{H} \sim \mathcal{G}$ , then their cost is the same, in fact the cost only depends on  $\mathcal{R}_{\mathcal{G}}$ .

The following lemma proved in [Gab] sheds some light on the bi-Lipschitz condition used in the definition of combinatorial cost. We will also use it in the proof of Theorem 1.

**Lemma 10** (Gaboriau). *Let  $\mathcal{G}$  be a graphing. For every  $\varepsilon > 0$  there exists some integer  $L$  and some  $\mathcal{H} \sim \mathcal{G}$  such that  $\tilde{\mu}(E(\mathcal{H})) < \text{cost}(\mathcal{G}) + \varepsilon$  and  $\mathcal{G}$  and  $\mathcal{H}$  are  $L$ -bi-Lipschitz equivalent, that is the graph metrics they define on the connected components are within a factor of  $L$  from each other.*

## 2.3 Combinatorial cost

The combinatorial analogue of cost for sequences of graphs is due to Elek [El]. Let  $(G_n)$  be a sequence of graphs with  $|V(G_n)| \rightarrow \infty$ , and degree bounded

by  $D$ . The sequence  $(H_n)$  is a *rewiring* of  $(G_n)$  – which we will denote  $(H_n) \sim (G_n)$  – if they have the same vertex set, and the distances defined by the graphs are uniformly bi-Lipschitz equivalent, that is  $V(G_n) = V(H_n)$  and there exists some natural number  $L$  such that for all  $n \in \mathbb{N}$

$$\frac{1}{L}d_{H_n}(x, y) \leq d_{G_n}(x, y) \leq Ld_{H_n}(x, y) \text{ for all } x, y \in V(G_n).$$

The *lower edge density* of a graph sequence is defined as follows.

$$\underline{e}((H_n)) = \liminf_{n \rightarrow \infty} \frac{|E(H_n)|}{|V(H_n)|}.$$

**Definition 11.** *The combinatorial cost of a sequence  $(G_n)$  is the infimum of the lower edge densities of its rewirings:*

$$\text{cc}((G_n)) = \inf \{ \underline{e}((H_n)) \mid (H_n) \sim (G_n) \}.$$

### 3 The cost of a local-global limit

In this section we will prove Theorems 1 and 3.

#### 3.1 Proof of the main result

We aim to show that the combinatorial cost of a locally-globally convergent graph sequence is equal to the cost of its limit. The idea of the proof is that if there is a cheap rewiring of the sequence  $(G_n)$ , then we can encode it into a coloring which then can be modeled with small error on the limit graphing  $\mathcal{G}$ . Using this coloring on the limit we can reconstruct a cheap graphing that (after some small modification) spans the same connected components as  $\mathcal{G}$ . In order to make this reconstruction process possible we will need to break the possible local symmetries of the graphs.

**Proof of Theorem 1.** Fix  $\varepsilon > 0$  and suppose that  $(H_n)$  is an  $L$ -rewiring of  $G_n$  such that  $\underline{e}(H_n) < \text{cc}(G_n) + \varepsilon$ . Set  $r = L^2 + 1$ ,  $R = 2r$ .

As the degrees of the  $G_n$  are bounded by  $D$ , there is a constant  $k$  such that each  $G_n$  can be vertex colored by  $k$  colors so that no two vertices within distance  $2R$  have the same color. Fix such a coloring  $\eta_n : V(G_n) \rightarrow \{1, \dots, k\}$  for each  $G_n$ . The role of these  $\eta_n$  is merely to break all possible symmetries of the  $R$ -neighborhoods.

For each vertex  $v \in V(G_n)$  define its *type* to be the following data. Let  $(\alpha_v, \eta_v)$  denote the colored  $R$ -neighborhood of  $v$  in  $G_n$ , that is  $B_{G_n}(R, v)$



rooted at  $v$ , together with  $\eta_n|_{B_{G_n}(R,v)}$ . Let  $F_v$  denote the set of edges of  $H_n$  that connect two vertices from  $B_{G_n}(R,v)$ . The type of  $v$  is the triple  $(\alpha_v, \eta_v, F_v)$ . We think of this as a rooted, vertex colored graph with some extra distinguished edges ( $F_v$ ) indicated. Note that  $\eta_v$  assigns distinct colors to the vertices of  $\alpha_v$ .

Let  $T$  denote the set of all possible types. Note that  $|T|$  is finite, as  $k$  and  $R$  are fixed. Now assigning each vertex its type can be considered as a coloring of  $V(G_n)$  by  $|T|$  colors. Let  $\varphi_n$  denote this coloring:

$$\varphi_n : V(G_n) \rightarrow T, \quad \varphi_n(v) = (\alpha_v, \eta_v, F_v) \text{ for all } v \in V(G_n).$$

Observe that  $\eta_n$  is a function of  $\varphi_n$ : for all  $v \in V(G_n)$ ,  $\eta_n(v)$  equals the color of the root of  $\varphi_n$ .

For any vertex  $v \in V(G_n)$  the edge  $(v, u) \in E(G_n)$  connecting  $v$  to its neighbor  $u$  can be traversed using at most  $L$  edges of  $H_n$ . Choose a shortest path between  $v$  and  $u$  in  $H_n$ . Because the length of the steps on the edges of  $H_n$  are bounded by  $L$  in terms of the graph distance in  $G_n$  the  $r = L^2 + 1$ -neighborhood of  $v$  in  $G_n$  already contains this shortest path. This holds for all neighbors  $u$ . This fact will be reflected in the type of  $v$ , namely for every neighbor of the root of  $\alpha_v$  there will be a path of length at most  $L$  using edges from  $F_v$  connecting the root to the neighbor. We will refer to this property by saying that  $F_v$  *witnesses  $L$ -bi-Lipschitz equivalence at the root*.

Since the  $G_n$  converge locally-globally to  $\mathcal{G}$  if we choose  $n$  large enough, we can find a Borel coloring  $\varphi : X \rightarrow T$  such that

$$d_{TV}(P_{G_n,r}[\varphi_n], P_{\mathcal{G},r}[\varphi]) < \delta,$$

that is we can model the local statistics of  $\varphi_n$  on  $\mathcal{G}$  with at most  $\delta$  error. Choose  $\delta$  such that  $\delta(D + \frac{1}{2}D^L) < \varepsilon$ .

The type of  $x$  gives a suggestion on how to construct a cheap graphing around  $x$ , which is  $F_x$ , the collection of distinguished edges. The idea is to consider the  $r$ -neighborhood of  $x$  in  $\mathcal{G}$ , and choose the edges of some graphing  $\mathcal{H}$  locally according to  $F_x$ . This  $\mathcal{H}$  would have the same connected components as  $\mathcal{G}$  because  $F_x$  witnesses bi-Lipschitz equivalence at the root, and would be cheap because the expected  $\mathcal{H}$ -degree of a point is close to the expected  $F_v$ -degree of the root in  $P_{G_n,r}[\varphi_n]$ . The problem is that the  $r$ -neighborhood of  $x$  in  $\mathcal{G}$  is a priori not the same as what the type of  $x$  suggests it is.

However, we will show that the above idea works for most of the points, and after a slight modification the resulting graphing will be a cheap generating graphing of the relation  $\mathcal{R}_{\mathcal{G}}$ .

First we construct a Borel coloring  $\eta : X \rightarrow \{1, \dots, k\}$  from  $\varphi$  imitating the way the  $\eta_n$  could be recovered from the  $\varphi_n$ . Let  $x \in X$ , and let  $\varphi(x) = (\alpha_x, \eta_x, F_x)$  be the type assigned to  $x$  by  $\varphi$ . The type suggests a color for the root, namely  $\eta_x(o)$  where  $o$  is the root of  $\alpha_x$ . So we set  $\eta(x) = \eta_x(o)$ . Observe that  $\eta_x$  is a coloring of the rooted graph  $\alpha_x$ , and a priori neither  $\alpha_x$  nor  $\eta_x$  has anything to do with the structure of  $\mathcal{G}$ . The value  $\eta(x)$  on the other hand is a concrete color from  $\{1, \dots, k\}$  that is assigned to the point  $x \in X$ .

It will turn out that that for most points  $x \in X$ , their  $\eta$ -colored neighborhood in  $\mathcal{G}$  is the same as the colored neighborhood  $(\alpha_x, \eta_x)$  suggested by thier type  $\varphi(x)$ , and  $\eta$  breaks the possible local symmetries of  $\mathcal{G}$  by being injective on the neighborhood. We define  $Y_1$  as the set of points where this does not hold up to distance  $r$ :

$$Y_1 = \left\{ x \in X \mid (B_{\mathcal{G}}(r, x), \eta|_{B_{\mathcal{G}}(r, x)}) \not\cong (B_{\alpha_x}(r, o), \eta_x|_{B_{\alpha_x}(r, o)}) \right\} \\ \bigcup \left\{ x \in X \mid \eta|_{B_{\mathcal{G}}(r, x)} \text{ is not injective} \right\}.$$

For any  $x$  outside  $Y_1$  we can identify  $B_{\alpha_x}(r, o)$  with  $B_{\mathcal{G}}(r, x)$  using their colorings. For any  $v \in V(\alpha_x)$  there exists a unique  $y \in B_{\mathcal{G}}(r, x)$  with  $\eta(y) = \eta_x(v)$ . Such a  $y$  exists because of the isomorphism, and the injectivity of the colorings implies uniqueness. Later on we will denote this unique  $y$  by  $y_{x,v}$ . The identification works the other way around as well, for every  $y$  in  $B_{\mathcal{G}}(r, x)$  we can find a unique  $v \in V(\alpha_x)$  such that  $\eta_x(v) = \eta(y)$ . Let us denote this unique  $v$  by  $v_{x,y}$ .

We use this identification to reconstruct our rewiring on  $\mathcal{G}$ . Define the edges of a graphing  $\mathcal{H}_0$  around  $x \in X \setminus Y_1$  as follows: for every edge  $(o, v) \in F_x$  that connects the root  $o$  of  $\alpha_x$  with some other point  $v \in V(\alpha_x)$  we include the edge  $(x, y_{x,v})$  in  $\mathcal{H}_0$ .

Recall that we aim to show that  $\mathcal{H}_0$  (with some small later adjustments) spans the same connected components as  $\mathcal{G}$ .

**Definition 12** (Perfect points). *Call a point  $x \in X$  perfect, if the following conditions hold:*

- 1)  $F_x$  witnesses  $L$ -bilipchitz equivalence at the root;
- 2) all the edges in  $F_x$  that  $\varphi$  suggests in  $B_{\alpha_x}(r, o)$  are indeed chosen to be in the edge set of  $\mathcal{H}_0$ .

For a type  $\varphi(x)$ , which is a rooted graph of radius  $R = 2r$  with some additional decorations we write  $\varphi(x)|_r$  for the graph where we simply forget

everything outside radius  $r$ . Similarly when  $v \in B_{\alpha_x}(r, o)$  write  $\varphi(x)|_{r,v}$  for the rooted, decorated graph we get by considering  $v$  as the root, and then forgetting everything outside radius  $r$  from  $v$ .

**Definition 13** (Problematic points). *Let  $Y_2$  be the set of points where one of the following holds.*

- i)  $\varphi$  does not witness the bi-Lipschitz connectivity at the root;*
- ii)  $\varphi$  fails to capture the local  $\eta$ -colored structure (up to distance  $R$ );*
- iii)  $\eta$  is not injective up to radius  $R$ ;*
- iv) there is some  $y$  close to  $x$  where  $\varphi(y)|_r$  differs from  $\varphi(x)|_{r,y_x}$ .*

*We call these points problematic.*

$$\begin{aligned}
Y_2 = & \left\{ x \in X \mid F_x \text{ does not witness bi-Lipschitz equivalence at the root} \right\} \\
& \bigcup \left\{ x \in X \mid (B_{\mathcal{G},R}(x), \eta|_{B_{\mathcal{G},R}(x)}) \not\cong (\alpha_x, \eta_x) \right\} \\
& \bigcup \left\{ x \in X \mid \eta|_{B_{\mathcal{G},R}(x)} \text{ is not injective} \right\} \\
& \bigcup \left\{ x \in X \setminus Y_1 \mid \exists v \in B_{\alpha_x}(r, o) \text{ s.t. } \varphi(x)|_{r,v} \not\cong \varphi(v_x)|_r \right\}.
\end{aligned}$$

The next lemma shows that because no such incoherencies happen in  $G_n$  the measure of the problematic points will be small. Also note that  $Y_1 \subset Y_2$ , as in  $Y_2$  we include all points where the local structure is not captured up to distance  $R$  instead of  $r$ .

**Lemma 14.** *The points in  $(X \setminus Y_2)$  are perfect and  $\mu(Y_2) < \delta$ .*

**Proof.** The  $P_{G_n,r}[\varphi_n]$  and  $P_{\mathcal{G},r}[\varphi]$  are probability distributions on the set  $U_R^T$  of rooted,  $T$ -colored graphs of radius at most  $R$ , where  $T$  is the set of all possible types.

Let  $(\beta, o_\beta, \psi)$  denote such a graph with root  $o_\beta$  and coloring  $\psi : V(\beta) \rightarrow T$ . For a vertex  $u \in V(\beta)$  its type  $\psi(u) \in T$  is the rooted  $k$ -colored graph  $(\alpha_v, \eta_v)$  with the additional distinguished edges  $F_v$ . There is some root  $o_{\alpha_v}$  of  $\alpha_v$ , and this way we define the coloring  $\eta_\psi : V(\beta) \rightarrow \{1, \dots, k\}$  by  $\eta_\psi(v) = \eta_v(o_{\alpha_v})$ . This  $\eta_\psi$  is defined from  $\psi$  the same way as  $\eta$  (on  $X$ ) is defined from  $\varphi$ . Now let  $(\beta, o_\beta, \psi)$  be random with distribution  $P_{G_n,r}[\varphi_n]$ , then

$$\mathbb{P}_{P_{G_n,r}[\varphi_n]} [(\alpha_{o_\beta}, \eta_{o_\beta}) \cong (\beta, \eta_\psi)] = 1.$$

The above equality just restates that the type encodes the local colored structure (specifically the color of the root), but this formulation shows that this property will be inherited with small error when the total variation distance is small.

This isomorphism again enables us to identify  $\alpha_{o_\beta}$  with  $\beta$ . For any  $v \in V(\alpha_{o_\beta})$  write  $y_{o_\beta, v}$  for the unique  $y \in V(\beta)$  for which  $\eta_{o_\beta}(v) = \eta_\psi(y)$ .

We restate that  $(H_n)$  is a rewiring by saying that the distinguished edges witness the  $L$ -bi-Lipschitz equivalence at the root:

$$\mathbb{P}_{P_{G_n, R}[\varphi_n]} [F_{o_\beta} \text{ witnesses } L\text{-bi-Lipschitz equivalence at the root}] = 1.$$

We also restate the fact that the type of  $v$ , which is all the information up to distance  $R = 2r$ , includes all the information in the  $r$ -neighborhood of some other point  $u$ , provided that  $u$  is within distance  $r$  from  $v$ .

$$\mathbb{P}_{P_{G_n, R}[\varphi_n]} \left[ \psi(o_\beta)|_{r, v} \cong \psi(y_{o_\beta, v})|_r \text{ for all } v \in V(B_{\alpha_{o_\beta}}(r, o_{\alpha_{o_\beta}})) \right] = 1.$$

Finally we restate that  $\eta$  distinguishes all points in the  $R$ -neighborhoods.

$$\mathbb{P}_{P_{G_n, R}[\varphi_n]} [\eta_\psi \text{ is injective}] = 1.$$

We see that the four events together hold with probability 1 with respect to  $P_{G_n, R}[\varphi_n]$ . Since  $P_{\mathcal{G}, r}[\varphi]$  is close to  $P_{G_n, R}[\varphi_n]$  we get that the same holds for  $\varphi$  and  $\mathcal{G}$  with probability at least  $1 - \delta$ , which implies  $\mu(Y_2) < \delta$ .

If  $x \in X \setminus Y_2$  and  $y \in B_{\mathcal{G}}(r, x)$  then  $y \in X \setminus Y_1$ , which means all the distinguished edges starting from  $y$  suggested by  $\varphi(y)$  are indeed in  $\mathcal{H}_0$ , and by the definition of  $Y_2$  we know that these are exactly the ones that  $\varphi(x)$  would suggest. It is also clear that  $F_x$  has to witness generation, as otherwise  $x$  would be in  $Y_2$ . This implies that  $x$  is perfect. ■

Adding all the edges leaving the points in  $Y_2$  we get  $\mathcal{H}$ :

$$\mathcal{H} = \mathcal{H}_0 \cup \{(x, y) \in E(\mathcal{G}) \mid x \in Y_2, y \in X\}.$$

**Lemma 15.**  $\mathcal{H}$  has the same connected components as  $\mathcal{G}$ , and  $\tilde{\mu}(E(\mathcal{H})) \leq \frac{|E(H_n)|}{|V(H_n)|} + \varepsilon$ .

**Proof.**  $\mathcal{H}$  will have the same connected components as  $\mathcal{G}$ , because for any edge  $(x, y) \in E(\mathcal{G})$  where  $x$  is perfect the connection is witnessed by the  $r$ -neighborhood of  $x$ . If  $x$  is not perfect, then  $x \in Y_2$ , so  $(x, y) \in E(\mathcal{H})$  by definition.

We now aim to show that  $\mathcal{H}$  is indeed a cheap generator for the equivalence relation.

$$\begin{aligned}\tilde{\mu}(E(\mathcal{H})) &\leq \tilde{\mu}(E(\mathcal{H}_0)) + \tilde{\mu}(\{(x, y) \in E(\mathcal{G}) \mid x \in Y_2, y \in X\}) \leq \\ &\leq \tilde{\mu}(E(\mathcal{H}_0)) + \delta D.\end{aligned}$$

For every point  $x \in X$  let  $\deg_{F_x}(o)$  denote the number of edges in  $F_x$  leaving the root of  $\alpha_x$ . It also makes sense to talk about the expectation of this  $F$ -degree with respect to colored neighborhood statistics, as the  $F$ -degree of the root can be determined from its type.

$$\begin{aligned}\tilde{\mu}(E(\mathcal{H}_0)) &= \frac{1}{2} \int_X \deg_{\mathcal{H}_0}(x) \, d\mu \leq \frac{1}{2} \int_X \deg_{F_x}(o) \, d\mu = \frac{1}{2} \mathbb{E}_{P_{\mathcal{G}, r}[\varphi]} \left[ \deg_F(o) \right] \leq \\ &\leq \frac{1}{2} \left( \mathbb{E}_{P_{G_n, r}[\varphi_n]} \left[ \deg_F(o) \right] + \delta D^L \right) = \frac{|E(H_n)|}{|V(H_n)|} + \frac{1}{2} \delta D^L.\end{aligned}$$

Here we used the fact that there can be no more than  $D^L$  edges leaving the root in  $F_v$  (because of the bi-Lipschitz condition), and that the two distributions are close in total variation. Putting all this together and using that we chose  $\delta$  to ensure that  $\delta(D + \frac{1}{2}D^L) < \varepsilon$  we get

$$\tilde{\mu}(E(\mathcal{H})) \leq \frac{|E(H_n)|}{|V(H_n)|} + \varepsilon.$$

■

By the choice of  $(H_n)$  we can assume that  $\frac{|E(H_n)|}{|V(H_n)|} \leq \text{cc}(G_n) + 2\varepsilon$ , which implies  $\tilde{\mu}(E(\mathcal{H})) < \text{cc}(G_n) + 3\varepsilon$ . This shows the inequality  $\text{cost}(\mathcal{G}) \leq \text{cc}(G_n)$ .

The other inequality is proved exactly the same way. The condition that  $(H_n)$  is a rewiring was only used to ensure that the bi-Lipschitz constant  $L$  does not depend on  $n$ , only on  $\varepsilon$ . To prove that  $\text{cost}(\mathcal{G}) \geq \text{cc}(G_n)$  we start by picking a cheap  $L$ -bi-Lipschitz generator for the single graphing  $\mathcal{G}$  using Lemma 10, and by local-global convergence we know that for  $n$  large enough we can copy it to  $G_n$  with small error.

As for any large enough  $n$  and  $m$  the graphs  $G_n$  and  $G_m$  are arbitrarily close in the local-global topology we can do the same copying argument between the two. We fix the constant  $L$  first, and then choose  $n$  and  $m$  accordingly. This shows that (for all  $L$ ) the rewirings  $(H_n)$  can indeed be chosen such that the limits defining the edge densities exist. This finishes the proof of Theorem 1. ■

### 3.2 Sofic approximations

Using Theorem 1 we will show that sofic approximations of a group with fixed price  $c$  have combinatorial cost  $c$  as well.

**Proof of Theorem 3.** The sequence  $G_n$  of  $S$ -edge-labeled graphs converges to  $\text{Cay}(\Gamma, S)$  in the Benjamini-Schramm sense, so any subsequential local-global limit will be a graphing of an essentially free action of  $\Gamma$ , which by the fixed price assumption implies that it has cost  $c$ .

First pick a locally-globally convergent subsequence  $G_{n_k}$  with limit  $\mathcal{G}_1$ .

$$\text{cc}(G_n) \leq \text{cc}(G_{n_k}) = \text{cost}(\mathcal{G}_1) = c.$$

Now assume that  $\text{cc}(G_n) < c$ . We pick an  $L$  large enough such that there is some  $L$ -bi-Lipschitz rewiring  $(H_n)$  with  $\underline{e}(H_n) < c$ . As  $\underline{e}$  is defined by a liminf we can choose a subsequence  $n_l$  such that

$$\lim \frac{|E(H_{n_l})|}{|V(H_{n_l})|} < c.$$

Now by passing to a further subsequence we can assume that the  $(G_{n_l})$  converge locally-globally to some  $\mathcal{G}_2$ . The  $H_{n_l}$  witness that  $\text{cc}(G_{n_l}) < c$ , while local-global convergence implies  $\text{cc}(G_{n_l}) = \text{cost}(\mathcal{G}_2) = c$  by Theorem 1. This is clearly a contradiction, hence  $\text{cc}(G_n) = c$ . ■

## 4 Group actions

The same notions and results exist in the world of measure preserving group actions, where convergence with respect to the weak containment topology takes the place of local-global convergence. The analogous definitions and statements will be introduced in this section.

### 4.1 Groupoid cost

Let  $\Gamma$  be a finitely generated group, generated by the finite symmetric set  $S = S^{-1}$ . Let  $(X, \mu)$  be either a standard Borel probability space or  $X$  a finite set with  $\mu$  the uniform measure on  $X$ . A probability p.m.p. action  $f$  of  $\Gamma$  is a homomorphism from  $\Gamma$  to the group of measure preserving transformations of  $(X, \mu)$ . The image of some  $\gamma \in \Gamma$  under this homomorphism will be denoted by  $f_\gamma$ .

Any such p.m.p. action gives rise to a groupoid denoted  $M_f$ : endow  $\Gamma$  with the discrete topology and counting measure, consider  $M_f = X \times \Gamma$  with

the product Borel structure and product measure  $\tilde{\mu}$ . We also define a partial product on  $X \times \Gamma$ :  $(x_1, \gamma_1) \cdot (x_2, \gamma_2) = (x_1, \gamma_1 \gamma_2)$  whenever  $x_2 = f_{\gamma_1}(x_1)$ . The inverse is defined by  $(x, \gamma)^{-1} = (f_{\gamma}(x), \gamma^{-1})$ , and so  $X \times \Gamma$  becomes a groupoid with respect to this partial product. We think of the element  $(x, \gamma)$  as an arrow pointing from  $x$  to  $f_{\gamma}(x)$ , with the arrow labeled by  $\gamma$ .

The notion of a generating subset of the groupoid is just as one would expect it: a subset generates, if all elements of  $M_f$  can be written as a product of elements and their inverses chosen from the subset.

For  $A, B \subseteq M_f$  we will write

$$A \cdot B = \{a \cdot b \mid a \in A, b \in B, \text{ and } a \cdot b \text{ is defined}\}.$$

Also let  $E = X \times \{e\}$ , where  $e$  is the identity element of  $\Gamma$ . Using our notation  $A$  generates  $M_f$  if and only if

$$M_f = \bigcup_{n=1}^{\infty} (A \cup A^{-1} \cup E)^n.$$

The groupoid cost of  $f$  is

$$\text{gcost}(f) = \inf\{\tilde{\mu}(A) \mid A \text{ generates } M_f\}.$$

The generators  $S$  of the group give rise to a specific generating subset of  $M_f$ , namely  $X_S = X \times S$ . We will say that a generating subset  $A$  is  $L$ -bi-Lipschitz, if all the elements of  $X_S$  can be generated by using at most  $L$  arrows from  $A$  and vice versa. More precisely we require that

$$X_S \subseteq (A \cup A^{-1} \cup E)^L, \text{ and } A \subseteq (X_S \cup X_S^{-1} \cup E)^L.$$

Note that while the actual value of the bi-Lipschitz constant  $L$  may depend on the choice of  $S$ , the property of  $A$  being a bi-Lipschitz generating subset (with some bi-Lipschitz constant) does not.

In this setting Lemma 10 was stated by Abért and Nikolov [AbN]. It says that by paying an arbitrarily small amount, we can choose the generating subset to be bi-Lipschitz. That is, for any  $\varepsilon > 0$  there exists some integer  $L$  and an  $L$ -bi-Lipschitz generating subset  $A \subseteq M_f$  such that  $\tilde{\mu}(A) < \text{gcost}(f) + \varepsilon$ .

## 4.2 The weak containment topology

The notion of weak containment of actions was introduced by Kechris [Kec]. The topology described below on the weak equivalence classes was defined by

Abért and Elek in [AbE2] and then studied further by Carderi in [Car]. They showed that the topology of local-global convergence is a compact topology on the weak equivalence classes of actions.

For an action  $f$  of the group  $\Gamma$  and a point  $x \in X$  the  $\Gamma$ -orbit of  $x$  admits a Schreier graph structure: for two points  $y, z \in \Gamma x$  in the orbit draw an oriented edge from  $y$  to  $z$  labeled by some  $s \in S$  if  $f_s(y) = z$ . Denote this graph by  $\text{Sch}(\Gamma, f, x)$ .

The only difference compared to the local-global convergence of graph sequences and graphings is that in this case we consider the neighborhoods in the Schreier graphs together with the edge labeling by the generators  $S$ .

To an action  $f$  we again associate a set  $Q_{f,r}^k$  that is the closure of all local statistics arising from Borel  $k$ -colorings with respect to the total variation distance. We say that an action  $f$  *weakly contains* another action  $g$  (denoted  $f \succeq g$ ) if  $Q_{g,r}^k \subseteq Q_{f,r}^k$  for all  $r$  and  $k$ . This means that all colorings of  $g$  can be modeled on  $f$  with arbitrarily small error. The actions are *weakly equivalent* if they both weakly contain the other, that is  $Q_{g,r}^k = Q_{f,r}^k$ .

Convergence with respect to the weak containment topology is defined by the convergence of  $Q_{f_n,r}^k$  for all  $r, k$  as compact sets with respect to the Hausdorff distance. The intuitive meaning of this convergence is the same as the one for local-global convergence. Abért and Elek showed that the topology induced by this convergence notion is compact, in particular every convergent sequence has a limit [AbE2].

Kechris showed that if  $f$  and  $g$  are free p.m.p. actions and  $f \succeq g$ , then  $\text{cost}(\mathcal{R}_f) \leq \text{cost}(\mathcal{R}_g)$  [Kec, Corollary 10.14]. Here  $\mathcal{R}_f$  denotes the orbit equivalence relation generated by the action  $f$ . Abért and Weiss extended this beyond free actions in [AbW]: for any actions with  $f \succeq g$  the groupoid cost satisfies  $\text{gcost}(f) \leq \text{gcost}(g)$ . This implies that the groupoid cost is well defined on weak equivalence classes, and studying the continuity properties of the groupoid cost with respect to the weak containment topology makes sense.

### 4.3 The groupoid cost of weak containment limits

Following the proof of Theorem 1 we get a result for group actions.

**Proposition 16.** *Suppose that the sequence  $f_1, f_2, \dots$  of p.m.p. actions is convergent in the weak containment topology to the p.m.p. action  $f$ . Then*

$$\limsup_{n \rightarrow \infty} \text{gcost}(f_n) \leq \text{gcost}(f). \quad (2)$$



This is a semicontinuity result for the groupoid cost with respect to the weak containment topology. The proof follows exactly the same steps as in Theorem 1: we choose a cheap bi-Lipschitz generating set for the groupoid  $M_f$ , record all local information into a coloring of  $X$ , model this coloring on the  $f_n$  when  $n$  is large enough with some small error and build a cheap generating set for  $M_{f_n}$  by decoding the coloring.

The whole process is actually slightly easier in this setting, because there is no need to break the local symmetries of the graphs as the Schreier edge labeling already takes care of that. As we are not imposing a uniform bound on the complexity of generation by talking about "combinatorial groupoid cost", we only get an inequality.

However, for the inequality we only need that colorings of  $f$  can be modeled with small error on the  $f_n$ , and we don't have to require it the other way around. That is, if the sequence "asymptotically weakly contains"  $f$ , then we have (2). This can be thought of as an asymptotic version of the monotonicity results by Kechris [Kec], and Abért-Weiss [AbW].

**Remark (The ultraproduct technique).** These results, together with Theorem 1 for graphings of free p.m.p. actions can be obtained by using the ultraproduct techniques introduced in [AbE2] and [Car], Carderi's result on ultraproduct actions being weakly equivalent to some standard action and the monotonicity results of Kechris and Abért-Weiss.

If one modifies (the somewhat arbitrary) choice of lower edge density in the definition of the combinatorial cost to *edge density along an ultrafilter*  $\omega$  by taking an ultralimit instead of a liminf, then this modified combinatorial cost of the sequence will equal the cost of the ultraproduct graphing.

#### 4.4 Rank gradient in groups with fixed price

We need one further tool to prove Theorem 4. The following lemma is stated in [AGN, Lemma 21].

**Lemma 17.** *Let  $\Gamma$  be a countable group, and  $H$  a subgroup of finite index in  $\Gamma$ . Let  $f$  be the right coset action of  $\Gamma$  on  $\Gamma/H$ . Then we have*

$$r(\Gamma, H) = \frac{\text{rank}(H) - 1}{|\Gamma : H|} = \text{gcost}(f) - 1.$$

**Proof of Theorem 4.** First we show that  $\liminf r(\Gamma, \Gamma_n) \geq c - 1$ . We select a subsequence  $\Gamma_{n_k}$  such that  $r(\Gamma, \Gamma_{n_k})$  converges to the liminf. Taking the diagonal product of the corresponding group actions  $f_{n_k}$  we get an action  $f$  of  $\Gamma$  that factors onto each  $f_{n_k}$ , which implies  $\text{gcost}(f) \leq \text{gcost}(f_{n_k})$  for

every  $n_k$ . The measure of the set of fixed points of a group element can only increase for factros, which implies that  $f$  is essentially free because of the Farber condition. Using Lemma 17 and  $\text{gcost}(f) = c$  we get

$$\lim_{k \rightarrow \infty} r(\Gamma, \Gamma_{n_k}) = \lim_{k \rightarrow \infty} (\text{gcost}(f_{n_k}) - 1) \geq \text{gcost}(f) - 1 = c - 1.$$

Similarly we can choose a subsequence such that

$$\limsup r(\Gamma, \Gamma_n) = \lim r(\Gamma, \Gamma_{n_l}).$$

By passing to a further subsequence we can also assume that the actions  $f_{n_l}$  converge in the weak containment topology to some action  $\hat{f}$ . This limit is essentially free by the Farber condition, so  $\text{gcost}(\hat{f}) = c$ . Using Proposition 16 we get

$$\lim_{l \rightarrow \infty} r(\Gamma, \Gamma_{n_l}) = \lim_{l \rightarrow \infty} (\text{gcost}(f_{n_l}) - 1) \leq \text{gcost}(\hat{f}) - 1 = c - 1.$$

■

**Remark** (alternative proof). The second part of the proof can be obtained without Proposition 16 – which we only sketched – by using a result from [AGN].

After choosing a subsequence such that  $\limsup r(\Gamma, \Gamma_n) = \lim r(\Gamma, \Gamma_{n_l})$ , [AGN, Theorem 8] states that  $\lim r(\Gamma, \Gamma_{n_l}) \leq \text{cc}(\text{Sch}(\Gamma, \Gamma_{n_l}, S)) - 1$ . Using Theorem 3 we get

$$\lim_{l \rightarrow \infty} r(\Gamma, \Gamma_{n_l}) \leq \text{cc}(\text{Sch}(\Gamma, \Gamma_{n_l}, S)) - 1 = c - 1.$$

## 5 The trichotomy theorem

In this section we introduce strong ergodicity, and prove the results on finitely presented groups.

### 5.1 Strong ergodicity

Let  $f$  be a p.m.p. action of the countable group  $\Gamma$  on a standard Borel space  $(X, \mu)$ . A sequence  $A_n$  of measurable subsets is called *almost invariant*, if

$$\lim_{n \rightarrow \infty} \mu(f_\gamma A_n \triangle A_n) = 0, \text{ for all } \gamma \in \Gamma.$$

The action  $f$  is *strongly ergodic*, if for any almost invariant sequence  $A_n$  we have  $\lim_{n \rightarrow \infty} \mu(A_n)(1 - \mu(A_n)) = 0$ .

We will make use of the following result of Abért and Weiss [AbW, Theorem 3].

**Theorem 18.** *Let  $f$  be an ergodic p.m.p. action of a countable group  $\Gamma$  on a standard Borel space  $(X, \mu)$ . If  $f$  is not strongly ergodic, then  $f$  is weakly equivalent to  $f \times I$ , which is the diagonal action on  $(X, \mu) \times [0, 1]$  with  $\Gamma$  acting trivially on the second coordinate.*

## 5.2 Dispersive actions

Let  $\text{Sch}(\Gamma, \Gamma_n, S)$  be a sequence of Schreier graphs, and let  $f_n$  denote the corresponding finite actions. We call the sequence  $\text{Sch}(\Gamma, \Gamma_n, S)$  *dispersive* if for any subsequential weak containment limit  $f$  of  $f_n$ ,  $f$  has no strongly ergodic, ergodic component of positive measure.

**Lemma 19.** *Let  $\Gamma$  be a group generated by the finite set  $S$  and  $\Gamma_n$  a sequence of subgroups such that the corresponding Schreier graphs  $\text{Sch}(\Gamma, \Gamma_n, S)$  form a dispersive sequence. Then for every  $\varepsilon > 0$  and  $k \in \mathbb{N}$  we can find some  $n$  such that the vertex set  $V$  of  $\text{Sch}(\Gamma, \Gamma_n, S)$  can be partitioned into  $k$  sets  $A_1, \dots, A_k$  such that*

1.  $\frac{1}{k} - \varepsilon \leq \frac{|A_i|}{|V|} \leq \frac{1}{k} + \varepsilon$  for all  $A_i$ ,
2.  $\sum |SA_i \setminus A_i| < \varepsilon |V|$ .

**Proof.** Pick a subsequence  $\Gamma_{n_i}$  such that the  $\text{Sch}(\Gamma, \Gamma_{n_i}, S)$  converge in the weak containment topology to some  $\Gamma$  action  $f$  on a standard Borel space  $(X, \mu)$ . As the sequence is dispersive we know that  $f$  has no strongly ergodic, ergodic components of positive measure.

We claim that  $X$  can be partitioned into  $k$  disjoint Borel sets  $B_1, \dots, B_k$  of approximately equal measure that are almost invariant, namely

$$\sum_{s \in S} \sum_{1 \leq i \leq k} \mu(sA_i \setminus A_i) < \varepsilon/2.$$

Assume first that  $f$  is ergodic. Then it is not strongly ergodic, and we can use Theorem 18. The base space of  $f \times I$  can be easily partitioned into  $k$  invariant subsets, specifically  $B'_i = (X, \mu) \times [(i-1)/k, i/k]$ . Weak equivalence guarantees that this partition can be modeled with arbitrarily small error (e.g.  $\varepsilon/2$ ) for  $f$  on the set  $X$ , and thus we have the desired  $B_i$ .

When  $f$  is not ergodic the we can divide  $X$  into two invariant sets  $X_1$  and  $X_2$  of positive measure. If  $f$  is ergodic on one of the  $X_i$ , then it is not strongly ergodic on that part and hence we can use the above argument to partition that  $X_i$ . If  $f$  is not ergodic on  $X_i$ , then we can again divide it into invariant subsets of positive measure.

Since all positive measure ergodic components are not strongly ergodic, we can continue this procedure and find a partition into some tiny (measure  $\varepsilon/100$ ) invariant sets and some non-strongly-ergodic components, which we can partition into almost invariant sets. Putting these blocks together into  $k$  sets we can get an almost invariant partition into pieces with measure in  $[1/k - \varepsilon/2, 1/k + \varepsilon/2]$ .

Now since  $f$  is a limit, for some  $n_l$  large enough we can model the partition  $B_1, \dots, B_k$  with error  $\varepsilon/2$  on  $\text{Sch}(\Gamma, \Gamma_{n_l}, S)$ , and get the desired  $A_i$ . ■

### 5.3 Finitely presented groups

We now briefly discuss how we will present finite index subgroups of finitely presented groups. Notation and general framework follows [AJN].

Let  $\Gamma = \langle S \mid R \rangle$  be a finitely presented group, and  $H \subseteq \Gamma$  a finite index subgroup. Let  $\mathcal{T}$  denote a spanning tree of the Schreier graph  $\text{Sch}(\Gamma, H, S)$ . We select a transversal  $T$  for the subgroup  $H$  as follows: for each coset  $\gamma H$  we consider the unique path in  $\text{Sch}(\Gamma, H, S)$  from the root  $H$  to  $\gamma H$ , and select the corresponding  $S$ -word to be in  $T$ . This  $T$  is called the left *Schreier transversal* corresponding to  $\mathcal{T}$  with respect to  $S$ . For a group element  $\gamma \in \Gamma$  let  $\tilde{\gamma}$  denote the unique element in  $T$  such that  $\gamma H = \tilde{\gamma} H$ .

For every edge  $e = (\gamma H, s\gamma H)$  of  $\text{Sch}(\Gamma, H, S)$  we put  $T(e) = (\tilde{s}\tilde{\gamma})^{-1}s\tilde{\gamma}$ . It is known that the  $\{T(e)\}$  belong to and generate  $H$ . Note that if  $e \in E(\mathcal{T})$ , then  $T(e) = 1$ .

For a relation  $r = s_l \dots s_1 \in R$  and group element  $t \in T$  let  $r_t = t^{-1}rt$ . This  $r_t$  is an element of  $H$ , and can be considered as a word in the  $T(e)$ :  $r_t = T(e_l) \dots T(e_1)$ , where  $e_i = (s_{i-1} \dots s_1 t H, s_i \dots s_1 t H)$ . We are going to use the fact that these relations give a presentation of  $H$ :

$$H = \langle \{T(e)\}_{e \in E(\text{Sch}(\Gamma, H, S)) \setminus E(\mathcal{T})} \mid \{r_t\}_{r \in R, t \in T} \rangle.$$

Suppose a group  $H$  has subgroups  $H_i \subseteq H$  ( $1 \leq i \leq k$ ) which all contain a fixed subgroup  $L$  and  $H \cong *_L H_i$ . We say that this decomposition is non-trivial, if  $L$  has index at least 3 in at least two of the subgroups.

**Proof of Theorem 7.** Assume that  $\text{Sch}(\Gamma, \Gamma_n, S)$  is dispersive and  $\text{RG}(\Gamma, (\Gamma_n)) > 0$ . We will show that some  $\Gamma_n$  decomposes as a non-trivial

amalgamated product. We can pass to a subsequence and assume that  $\frac{d(\Gamma_n)-1}{[\Gamma:\Gamma_n]} > c > 0$  for all  $n$ . Choose an integer  $k$  such that

$$\left(\frac{3}{2}|S| + 1\right) \frac{1}{k} \leq c/2.$$

Let  $M$  be the sum of the lengths of the relations in  $R$ . As the sequence is dispersive, using Lemma 19 we can choose some  $n$  such that the vertex set  $V(\text{Sch}(\Gamma, \Gamma_n, S))$  can be split into the disjoint union of  $k$  sets  $A_1, \dots, A_k$  such that

1.  $\frac{[\Gamma:\Gamma_n]}{k} - \frac{[\Gamma:\Gamma_n]}{2k} < |A_j| < \frac{[\Gamma:\Gamma_n]}{k} + \frac{[\Gamma:\Gamma_n]}{2k}$  for all  $j \in \{1, \dots, k\}$  and
2.  $|\partial(A_1, \dots, A_k)| < \frac{1}{k(1+M^2)}[\Gamma:\Gamma_n]$ , where

$$\partial(A_1, \dots, A_k) = \{e \in E(\text{Sch}(\Gamma, \Gamma_n, S)) \mid e = (x, y), x \in A_j, y \in A_l, j \neq l\}.$$

As we have  $[\Gamma:\Gamma_n] \rightarrow \infty$  we can also make sure that we choose  $n$  large enough so that  $\frac{k-1}{[\Gamma:\Gamma_n]} \leq c/2$ . Put  $H = \Gamma_n$  and follow the above construction for a presentation of  $H$ . Note that  $|V(\text{Sch}(\Gamma, H, S))| = [\Gamma:H]$ .

Define  $Y$  to be the collection of generators  $T(e)$  that share a relation with an inbetween edge. More precisely let  $T(e) \in Y$  if either  $e \in \partial(A_1, \dots, A_k)$  or there exists a relation  $r_t = T(e_1)^{\pm 1} \dots T(e_l)^{\pm 1}$  for which some  $e_j \in \partial(A_1, \dots, A_k)$  and some  $e_m = e$ . Let  $X_i$  be the set of generators  $T(e)$  for which both end-points of  $e$  are in  $A_i$ .

Define the subgroups  $L = \langle Y \rangle$  and  $H_i = \langle Y \cup X_i \rangle$ . Clearly  $L \leq H_i$  for all  $i$ .

**Lemma 20.**  *$H$  decomposes as the amalgamated product of the  $H_j$  over  $L$ ,  $H \cong *_L H_i$ .*

We postpone the proof of the lemma, and show that this decomposition is non-trivial. Suppose that  $L$  has index at most 3 in  $H_1, H_2, \dots, H_{k-1}$ . Then each of these  $H_i$  ( $1 \leq i \leq k-1$ ) are generated by  $L$  and at most 1 other element. Thus  $H$  is generated by  $H_k$  (which includes  $L$ ) and at most  $k-1$  other elements,  $d(H) \leq d(H_k) + k - 1$ .

It is easy to bound the cardinality of  $|X_k|$ :

$$|X_k| \leq |S||A_k| \leq |S| \left( \frac{[\Gamma:H]}{k} + \frac{[\Gamma:H]}{2k} \right).$$

Let  $r = s_l \dots s_1$  be a relation of  $\Gamma$  of length  $l$ . Note that there are at most  $l|\partial(A_1, \dots, A_k)|$  different lifts  $r_{\tilde{g}} = T(e_1)^{\pm 1} \dots T(e_l)^{\pm 1}$  of  $r$  for which some  $e_j \in \partial(A_1, \dots, A_k)$ . Also for each such relation of  $H$  we have at most  $l$

generators  $T(e)$  of  $H$  that are getting into  $Y$ . Thus, if  $\{l_j\}$  is the set of lengths of the relations of  $\Gamma$  (so  $M = \sum l_j$ ), then we can bound the cardinality of  $Y$ .

$$|Y| \leq |\partial(A_1, \dots, A_k)|(1 + \sum l_j^2) \leq |\partial(A_1, \dots, A_k)|(1 + M^2) \leq \frac{[\Gamma : H]}{k}.$$

Putting our bounds together we get

$$d(H) \leq d(H_k) + k - 1 \leq |X_k| + |Y| + k - 1 \leq \left(\frac{3}{2}|S| + 1\right) \frac{[\Gamma : H]}{k} + k - 1.$$

This however gives an upper bound on the rank quotient at  $H$ :

$$\frac{d(H) - 1}{[\Gamma : H]} \leq \left(\frac{3}{2}|S| + 1\right) \frac{1}{k} + \frac{k - 1}{[\Gamma : H]} \leq c/2 + c/2 = c.$$

This contradicts our assumption that each such quotient is more than  $c$ , hence the decomposition is non-trivial. ■

**Proof of Lemma 20.** The argument follows the one in [AJN, Section 3].

Consider the following sets of relations. Let  $R_i$  be the set of all the  $r_t = T(e_1)^{\pm 1} \dots T(e_l)^{\pm 1}$  where either all  $e_j$  have both endpoints in  $A_i$  or some  $e_j$  is in  $\partial(A_1, \dots, A_k)$ . Now  $R_i \cup R_j$  is the same set  $\bar{R}$  for all pairs  $(i, j)$ , that is the relations having an inbetween edge. Define the groups  $T_i$  by the presentations  $\langle X_i \cup Y \mid R_i \rangle$ , let  $\bar{T} = \langle Y \mid \bar{R} \rangle$ . We have a homomorphism  $\phi_j : \bar{T} \rightarrow T_i$  by the inclusion of  $Y$  into  $X_i \cup Y$ . From the presentations we see that  $H \cong *_\bar{T} T_i$ .

Each  $T_i$  surjects onto  $H_i$  (and  $\bar{T}$  surjects onto  $L$ ) by mapping the abstract generators to their counterparts in  $H$ . By the universal property of the amalgamated product one can see  $H \cong *_L H_i$ . ■

**Proposition 21.** *For a countable amenable group  $\Gamma$  all sequences  $(\Gamma_n)$  of distinct finite index subgroups are dispersive.*

**Proof.** It is a result of Schmidt [Sch, Theorem 2.4] that amenable groups admit no strongly ergodic actions. If any subsequential limit would have an ergodic component of positive measure that is strongly ergodic, then restricting the action to that component would contradict [Sch]. ■

As amenable groups cannot decompose as non-trivial amalgamated products this proves Theorem 6.

## 6 Open problems

Note that an even stronger form of Theorem 4 would simply express the rank gradient of a local-global convergent sequence as the cost of the limiting graphing of the sequence. As of this moment, we do not know how to prove or disprove this. The obstacle is that combinatorial cost handles sequences of generating sets with a bounded complexity with respect to some standard generating set. A priori, it could happen that actual small generating sets over the sequence need a very aggressive growth of complexity, and the local-global metric is too weak to connect such generating sets over the sequence. This is the same obstacle that makes the proof of Theorem 7 somewhat tricky.

The following problem connects two well-known unsolved problems, one in ergodic theory, the other in 3-manifold theory.

**Problem 22.** *Let  $\Gamma$  be a finitely generated group. Does there exist  $c$  such that for any Farber sequence  $(\Gamma_n)$  in  $\Gamma$ , we have  $\text{RG}(\Gamma, (\Gamma_n)) = c$ ?*

Equivalently, one can ask whether  $\text{RG}(\Gamma, (\Gamma_n))$  exists for any Farber sequence  $(\Gamma_n)$  in  $\Gamma$ . Indeed, an advantage of Farber sequences over Farber chains is that they are closed to merging.

By Theorem 4, a negative answer to Problem 22 would immediately give a negative answer to the Fixed Price problem of Gaboriau [Gab], that asks whether for an arbitrary countable group  $\Gamma$ , all essentially free p.m.p. actions of  $\Gamma$  have the same cost. A positive answer, on the other hand, would specifically show that in a finitely generated group, any two normal chains with trivial intersection have the same rank gradient, which by [AbN] would then solve the strong Rank vs Heegaard genus problem on hyperbolic 3-manifolds.

One possible approach to Problem 22 is through graph theory as follows. Let  $G$  be a finite, connected graph with maximal degree  $D$ . For  $L \geq 1$  let

$$c_L(G) = \min_H \frac{|E(H)|}{|V(G)|}$$

where  $H$  runs through all rewirings of  $G$  with bi-Lipschitz constant at most  $L$ . It is easy to see that

$$1 - o(1) \leq c_L(G) \leq D.$$

The following problem is related to the Fixed Price problem of Gaboriau.

**Problem 23.** *Let  $(G_n)$  be a Benjamini-Schramm convergent sequence of graphs of bounded degree. Does  $c_L(G_n)$  converge for every  $L \geq 1$ ?*

The connection is one sided.

**Proposition 24.** *An affirmative solution of Problem 23 implies an affirmative solution of Problem 22.*

Note, however, that this problem seems to be a real strengthening. Indeed, it could happen that for two large graphs  $G_1$  and  $G_2$  that are very close in the Benjamini-Schramm topology, one can find a cheap rewiring of  $G_1$  with a bi-Lipschitz constant  $L_1$  but only do the same to  $G_2$  with a much bigger constant.

**Problem 25.** *Let  $\Gamma$  be a finitely presented group generated by a finite symmetric set  $S$ . Let  $(\Gamma_n)$  be a sequence of subgroups of finite index in  $\Gamma$  and let  $G_n = \text{Sch}(\Gamma, \Gamma_n, S)$ . Assume that no  $\Gamma_n$  decomposes as a non-trivial amalgamated product and that the sequence  $(G_n)$  is dispersive. Is it true that  $\text{cc}(G_n) = 1$ ?*

In Theorem 7 we show that the rank gradient of  $(\Gamma_n)$  must vanish.

## References

- [AbE1] M. ABÉRT AND G. ELEK, Dynamical properties of profinite actions, *Ergodic Theory and Dynamical Systems* 32 (2012), issue 6, 1805–1835
- [AbE2] M. ABÉRT AND G. ELEK, The Space of Actions, Partition Metric and Combinatorial Rigidity, preprint
- [AbN] M. ABÉRT AND N. NIKOLOV, Rank gradient, cost of groups and the rank vs Heegaard genus conjecture, *J. Eur. Math. Soc.* 14 (2012), 1657–1677.
- [AJN] M. ABÉRT, N. NIKOLOV AND A. JAIKIN-ZAPIRAIN, The rank gradient from a combinatorial viewpoint, *Groups Geom. Dyn.* 5 (2011), no. 2, 213–230.
- [AGN] M. ABÉRT, T. GELANDER AND N. NIKOLOV, Rank, combinatorial cost and homology torsion growth in higher rank lattices, *Duke Math. J.* 166 (2017), no. 15, 2925–2964.
- [AbW] M. ABÉRT AND B. WEISS, Bernoulli actions are weakly contained in any free action, *Ergodic theory and dynamical systems* 33 (2013), issue 2, 323–333.



- [BoR] B. BOLLOBÁS AND O. RIORDAN, Sparse graphs: metrics and random models. *Random Structures and Algorithms* 39 (2010), no. 1, 1–38.
- [Car] A. CARDERI, Ultraproducts, weak equivalence and sofic entropy, preprint
- [CGS] A. CARDERI, D. GABORIAU AND M. DE LA SALLE, Non-standard limits of graphs and their orbit equivalence invariants
- [El] G. ELEK, The combinatorial cost, *Enseign. Math.* (2) 53 (2007), no. 3-4, 225–235.
- [Gab] D. GABORIAU, Coût des relations d’équivalence et des groupes. (French) [Cost of equivalence relations and of groups] *Invent. Math.* 139 (2000), no. 1, 41–98.
- [HLSz] H. HATAMI, L. LOVÁSZ AND B. SZEGEDY, Limits of locally-globally convergent graph sequences, *Geom. Funct. Anal.* 24 (2014), no. 1, 269–296.
- [Kec] A. KECHRIS, Global aspects of ergodic group actions. *Mathematical Surveys and Monographs*, 160. American Mathematical Society, 2010.
- [Lac] M. LACKENBY, Expanders, rank and graphs of groups, *Israel J. Math.* 146 (2005), 357–370.
- [OrW] D. S. ORNSTEIN AND B. WEISS, Ergodic theory of amenable group actions. I. The Rohlin lemma, *Bull. Amer. Math. Soc. (N.S.)* 2 (1980), no. 1, 161–164.
- [Sch] K. SCHMIDT, Amenability, Kazhdan’s property T, strong ergodicity and invariant means for ergodic group-actions, *Ergodic Theory Dynam. Systems* 1 (1981), no. 2, 223–236.

MIKLÓS ABÉRT, MTA ALFRÉD RÉNYI INSTITUTE OF MATHEMATICS,  
BUDAPEST, HUNGARY

*E-mail:* `abert.miklos@renyi.mta.hu`

LÁSZLÓ MÁRTON TÓTH, CENTRAL EUROPEAN UNIVERSITY AND MTA  
ALFRÉD RÉNYI INSTITUTE OF MATHEMATICS, BUDAPEST, HUNGARY

*E-mail:* `ltoth@renyi.mta.hu`