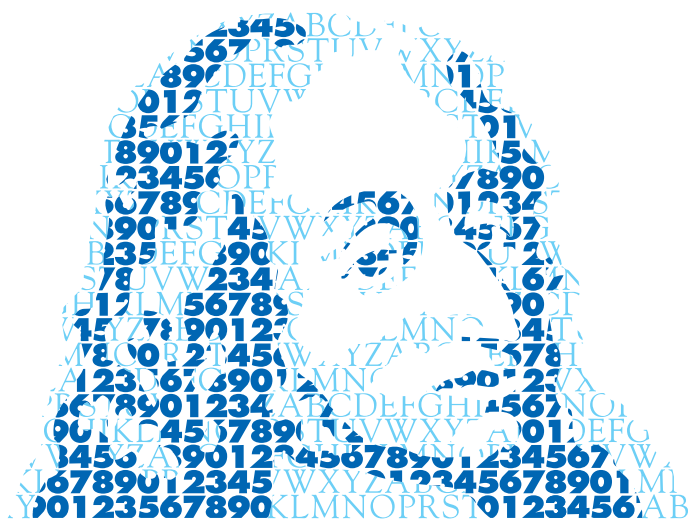


# ANNALES MATHÉMATIQUES



## BLAISE PASCAL

GOULNARA N. ARZHANTSEVA, VICTOR S. GUBA,  
MARTIN LUSTIG AND JEAN-PHILIPPE PRÉAUX

Testing Cayley graph densities

Volume 15, n° 2 (2008), p. 233-286.

[http://ambp.cedram.org/item?id=AMBP\\_2008\\_\\_15\\_2\\_233\\_0](http://ambp.cedram.org/item?id=AMBP_2008__15_2_233_0)

© Annales mathématiques Blaise Pascal, 2008, tous droits réservés.

L'accès aux articles de la revue « Annales mathématiques Blaise Pascal » (<http://ambp.cedram.org/>), implique l'accord avec les conditions générales d'utilisation (<http://ambp.cedram.org/legal/>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

*Publication éditée par le laboratoire de mathématiques  
de l'université Blaise-Pascal, UMR 6620 du CNRS  
Clermont-Ferrand — France*

cedram

Article mis en ligne dans le cadre du  
Centre de diffusion des revues académiques de mathématiques  
<http://www.cedram.org/>

# Testing Cayley graph densities

GOULNARA N. ARZHANTSEVA

VICTOR S. GUBA

MARTIN LUSTIG

JEAN-PHILIPPE PRÉAUX

## Abstract

We present a computer-assisted analysis of combinatorial properties of the Cayley graphs of certain finitely generated groups: given a group with a finite set of generators, we study the density of the corresponding Cayley graph, that is, the least upper bound for the average vertex degree (= number of adjacent edges) of any finite subgraph. It is known that an  $m$ -generated group is amenable if and only if the density of the corresponding Cayley graph equals to  $2m$ . We test amenable and non-amenable groups, and also groups for which amenability is unknown. In the latter class we focus on Richard Thompson's group  $F$ .

*Tester les densités de graphes de Cayley.*

## Résumé

Nous présentons une analyse assistée par ordinateur de propriétés combinatoires des graphes de Cayley de certains groupes de type fini : donné un groupe et un ensemble fini de générateurs, nous étudions la densité du graphe de Cayley correspondant, c'est à dire, la borne supérieure de la valence de sommet (= nombre d'arêtes adjacentes) moyenne de tous ses sous-graphes finis. Il est connu qu'un groupe ayant  $m$  générateurs est moyennable si et seulement si la densité du graphe de Cayley correspondant est  $2m$ . Nous testons des groupes moyennables et non-moyennables, ainsi que d'autres dont la moyennabilité est inconnue. Dans cette dernière classe nous nous intéressons au groupe  $F$  de Thompson.

## 1. Introduction

Let  $G$  be a group with a finite set of generators  $X$  of cardinality  $m$ . There is an associated *Cayley graph*  $\mathcal{C} = \mathcal{C}(G, X)$  (see §2), which has vertex

---

This work was supported by the Swiss National Science Foundation, No. PP002-68627.

*Keywords:* Amenability, Thompson's group  $F$ , computer-assisted analysis.

*Math. classification:* 20-04, 20F05.

set in bijection to  $G$ , and at every vertex there are precisely  $2m$  adjacent edges. The combinatorial properties of the Cayley graph reflect the algebraic structure of the group  $G$ . In this paper we investigate the *density* of Cayley graphs, introduced first in [5]. This is a numerical parameter  $\delta(\mathcal{C})$ , defined below, which takes values between 0 and twice the number of group generators. It strongly depends on the isoperimetric properties of the Cayley graph and hence on those of  $G$ , which are often expressed in terms of the graph isoperimetric constant  $\iota_*(\mathcal{C})$  (see §2 for the definition). It is known that  $\iota_*(\mathcal{C}) + \delta(\mathcal{C}) = 2m$ , see [5]. A group is *amenable* if and only if  $\iota_*(\mathcal{C}) = 0$ , or, equivalently,  $\delta(\mathcal{C}) = 2m$ .

In order to estimate the density of a Cayley graph, one can compute densities of certain of its finite subgraphs. We propose a simple algorithm to construct an optimized subgraph (*i.e.* with a greater density) from any given finite subgraph of the Cayley graph. We apply the algorithm to amenable groups, non-amenable groups, and to groups for which it is not known whether they are amenable. More specifically, we investigate the (amenable with polynomial growth) free abelian group  $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$ , the non-amenable Baumslag-Solitar group  $BS(2, 2)$ , the restricted wreath product  $\mathbb{Z} \wr \mathbb{Z}$  (amenable with exponential growth), and Richard Thompson's group  $F$  whose amenability is unknown (other testing examples can be found on our websites). We analyze empirical data obtained by a  $C^{++}$  implementation of our algorithm.

We quote here only one of the numerical results obtained from our algorithm, which we find particularly interesting:

**Sample result:** There is a subset of cardinality 10169678 in Thompson's group  $F$  that has density 2.89577 with respect to the classical generating system of cardinality  $m = 2$ .

*Acknowledgements.* Part of this work was conducted while the authors were visiting the Centre de Recerca Matemàtica in Barcelona. We would like to acknowledge the warm hospitality and support. Specifically we would like to thank José Burillo and Enric Ventura for their interest in our work, and for useful discussions.

## 2. Amenability and Følner families

Let  $G$  be a group generated by a finite set  $X$ . Let  $\mathcal{C} = \mathcal{C}(G, X)$  be the corresponding (right) Cayley graph. Recall that the set of vertices of  $\mathcal{C}$  is  $G$ ,

and that the set of oriented edges is  $G \times (X \dot{\cup} X^{-1})$ . For any edge  $e = (g, x)$  the initial vertex is  $g$ , and the terminal vertex is  $gx$ . The *inverse* of the edge  $e$ , considered here separately from  $e$ , is the edge  $e^{-1} = (gx, x^{-1})$ . The *label* of  $e = (g, x)$  is defined to be the generator  $x \in X^{\pm 1}$ . The group  $G$  acts canonically on  $\mathcal{C}$  from the left (by left multiplication of the vertices of  $\mathcal{C}$ ). Notice that with the above convention, if  $x$  and  $x^{-1}$  both belong to  $X$ , then altogether there are 4 edges (two for each orientation) with endpoints  $g$  and  $gx$ . This holds in particular if  $x \in X$  has order 2.

Throughout the paper we consider finite graphs  $A$  which are typically subgraphs of  $\mathcal{C}$ . We always require that with any edge  $e$  also the inverse edge  $e^{-1}$  belongs to  $A$ .

The *density* of a non-empty finite graph  $A$  is defined by:

$$\delta(A) = \frac{\sum_{v \in V(A)} \deg(v)}{\#V(A)}$$

where  $\deg(v)$  denotes the number of oriented edges with initial vertex  $v$ ,  $V(A)$  is the set of vertices of  $A$ , and  $\#V(A)$  is the cardinality of  $V(A)$ .

We define the *density* of the Cayley graph  $\mathcal{C} = \mathcal{C}(G, X)$  as supremum:

$$\delta(\mathcal{C}) = \sup_A \delta(A)$$

where  $A$  runs over all non-empty finite subgraphs of  $\mathcal{C}$ .

Similarly, for any subgraph  $A$  of  $\mathcal{C}$  one defines the isoperimetric constant:

$$\iota(A) = \frac{\#\partial A}{\#V(A)}$$

where  $\partial A$  denotes the set of vertices of  $A$  that have adjacent edges in both,  $A$  and  $\mathcal{C} - A$ .

**Lemma 2.1.** *For every finite subgraph  $A$  of  $\mathcal{C}$  one has:*

$$2m - (2m - 1)\iota(A) \leq \delta(A) \leq 2m - \iota(A).$$

*Proof.* Note that for any finite subset  $A \subset \mathcal{C}$  the complement of  $\partial A$  in  $A$  consists entirely of vertices of degree  $2m$ , hence:

$$\delta(A) = \frac{\sum_{v \in V(A)} \deg(v)}{\#V(A)} = \frac{\sum_{v \in \partial A} \deg(v)}{\#V(A)} + \frac{\sum_{v \in V(A) - \partial A} 2m}{\#V(A)}.$$

But a vertex in  $\partial A$  has at least degree 1, so that:

$$\frac{\#\partial A}{\#V(A)} + 2m \frac{\#V(A) - \#\partial A}{\#V(A)} \leq \delta(A),$$

and at most degree  $2m - 1$  in  $A$ , so that:

$$\delta(A) \leq (2m - 1) \frac{\#\partial A}{\#V(A)} + 2m \frac{\#V(A) - \#\partial A}{\#V(A)},$$

which shows that  $2m - (2m - 1)\iota(A) \leq \delta(A) \leq 2m - \iota(A)$ . □

A variation on the invariant  $\iota(A)$  is given by the invariant  $\iota_*(A)$  which is defined in precisely the same way except that the *inner boundary*  $\partial A$  is replaced by the *Cheeger boundary* of  $A$ , *i.e.* the number of edges in  $\mathcal{C}$  that have one endpoint in  $A$  and one endpoint in  $\mathcal{C} - A$ . The Cheeger boundary behaves a little better than the inner boundary; for example we derive directly from the definition the equation  $\delta(A) + \iota_*(A) = 2m$ . Similarly, one sees directly that:

$$\iota(A) \leq \iota_*(A) \leq 2m \iota(A).$$

The infimum of the values of  $\iota_*(A)$ , over all non-empty finite subsets  $A$  of  $\mathcal{C}$ , is called the *isoperimetric constant* of the graph  $\mathcal{C}$ , and is denoted by  $\iota_*(\mathcal{C})$ .

Isoperimetric properties of graphs play an important role in the study of amenable groups. There are many equivalent characterizations of amenability in the literature, see for example [3] and the references given there. We use the following one.

**Theorem 2.2.** *A finitely generated group  $G$  is amenable if and only if for some (or, equivalently, for any) finite generating set  $X$  the Cayley graph  $\mathcal{C} = \mathcal{C}(G, X)$  satisfies:*

$$\iota_*(\mathcal{C}) = 0.$$

A family of finite subsets  $A_n$  of  $\mathcal{C}$  is called a *Følner family* (or a *family of Følner sets*) if:

$$\lim_{n \rightarrow \infty} \iota_*(A_n) = 0.$$

In light of the above discussion this is equivalent to:

$$\lim_{n \rightarrow \infty} \delta(A_n) = 2m.$$

Hence, the group  $G$  is amenable if and only if there exists a family of Følner sets  $A_n \subset \mathcal{C}$ .

For certain classes of groups there are well known Følner families. For example, if  $G$  is of polynomial (or subexponential) growth, then one knows that with respect to any finite generating system  $X$  of  $G$  the set of *balls*  $B(n)$ , which consists of all points in  $\mathcal{C}$  of simplicial distance smaller or

equal to  $n$  from the neutral element  $1 \in G$ , is a Følner family [4, Proposition, Ch.VII.C.34]. Here we mean by simplicial distance the distance in the metric space obtained from  $\mathcal{C}$  if one gives to every edge the length 1. Examples for groups of polynomial growth are free abelian groups and certain Baumslag-Solitar groups (e.g.  $BS(1, -1)$ ).

However, Baumslag-Solitar groups (other than  $BS(1, 1)$  or  $BS(1, -1)$ ) are of exponential growth, but some of them (not all !) are still amenable. The same is true for the wreath product  $\mathbb{Z} \wr \mathbb{Z}$ , also considered below. In this case a Følner family exists in  $\mathcal{C}$ , but the balls  $B(n)$  will not constitute such a family: There is a uniform upper bound strictly smaller than  $2m$  to the density of every  $B(n)$ . Of course, this last statement is true also if  $G$  is non-amenable.

### 3. Group presentations and normal forms

In order to compute in a finitely generated group  $G$ , one needs a *normal form* for the elements of  $G$ : For example, in  $\mathbb{Z} \times \mathbb{Z} = \langle a, b \mid aba^{-1}b^{-1} = 1 \rangle$  the element  $(2, 1)$  can be written as  $a^2b$ ,  $aba$ ,  $baa$ , but also as  $aba^{-1}bab^{-1}a$  or  $a^{-69}ba^{71}$ . It is an essential restriction on the class of groups  $G$  considered here that we require the existence of a uniquely determined normal form for the elements of  $G$ , and that this normal form can be recursively calculated. Notice that the generating set of  $G$  used in the normal form may well differ from the system  $X$  which is used to build the Cayley graph; in some cases this discrepancy is a rather convenient from a computational point of view.

#### 3.1. Free abelian groups

The free abelian group of rank  $m$  is defined by the presentation:

$$\langle x_1, x_2, \dots, x_m \mid x_i x_j = x_j x_i \text{ for all } 1 \leq i < j \leq m \rangle.$$

A word in the canonical generators  $x_1, x_2, \dots, x_m$  and their inverses is in normal form if and only if it is of the form:

$$x_1^{p_1} x_2^{p_2} \cdots x_m^{p_m}$$

for some  $p_1, p_2, \dots, p_m \in \mathbb{Z}$ .

### 3.2. Baumslag-Solitar groups

Let  $p, q \geq 1$  be integers, and let  $BS(p, q)$  denote the Baumslag-Solitar group defined by the presentation:

$$\langle a, b \mid ab^p a^{-1} = b^q \rangle.$$

A word in  $a^{\pm 1}, b^{\pm 1}$  is in normal form whenever it is written in a reduced form (in the sense of HNN extensions, see [6]):

$$b^{p_0} a^{q_1} b^{p_1} \dots a^{q_n} b^{p_n}$$

with  $n \geq 0$  and  $p_0, p_1, \dots, p_n \in \mathbb{Z}$ ,  $q_1, q_2, \dots, q_n \in \mathbb{Z} \setminus \{0\}$ , and such that for  $i = 1, 2, \dots, n - 1$  one has  $p_i > 0$ , and  $p_n \geq 0$ . Furthermore, if  $q_i > 0$  then one has  $p_i < |p|$ , and if  $q_i < 0$  then  $p_i < |q|$ .

### 3.3. The wreath product $\mathbb{Z} \wr \mathbb{Z}$

We define this group by the non-finite presentation:

$$\langle a, \dots, x_{-1}, x_0, x_1, \dots \mid x_i^a = x_{i+1}, x_i x_j = x_j x_i \text{ for all } i, j \in \mathbb{Z} \rangle.$$

A word in the generators or their inverses is in normal form whenever it is of the form:

$$a^n x_{i_1}^{p_1} x_{i_2}^{p_2} \dots x_{i_n}^{p_n}$$

where  $n \in \mathbb{Z}$ ,  $i_1 < i_2 < \dots < i_n$ , and  $p_1, p_2, \dots, p_n \in \mathbb{Z} \setminus \{0\}$ .

In fact, this group can be generated by  $a$  and  $x_0$ . We will refer to this as the canonical set of generators.

### 3.4. Thompson's group $F$

Thompson's group  $F$  (cf. [2]) is the group of all piecewise-linear orientation preserving self-homeomorphisms of the unit interval such that (i) singular points are on dyadic numbers, and (ii) all slopes are integer powers of 2.

The group  $F$  admits the following infinite presentation:

$$\langle x_0, x_1, x_2, \dots \mid x_j x_i = x_i x_{j+1} \text{ if } i < j \rangle.$$

It turns out that it has a finite presentation on two generators  $x_0, x_1$ . We will use this canonical set of generators in our computations. The generators  $x_0$  and  $x_1$  are given by the following functions, see also Figure 1.

TESTING CAYLEY GRAPH DENSITIES

$$x_0(t) = \begin{cases} t/2 & 0 \leq t \leq 1/2 \\ t - 1/4 & 1/2 \leq t \leq 3/4 \\ 2t - 1 & 3/4 \leq t \leq 1 \end{cases} \quad x_1(t) = \begin{cases} t & 0 \leq t \leq 1/2 \\ t/2 + 1/4 & 1/2 \leq t \leq 3/4 \\ t - 1/8 & 3/4 \leq t \leq 7/8 \\ 2t - 1 & 7/8 \leq t \leq 1 \end{cases}$$

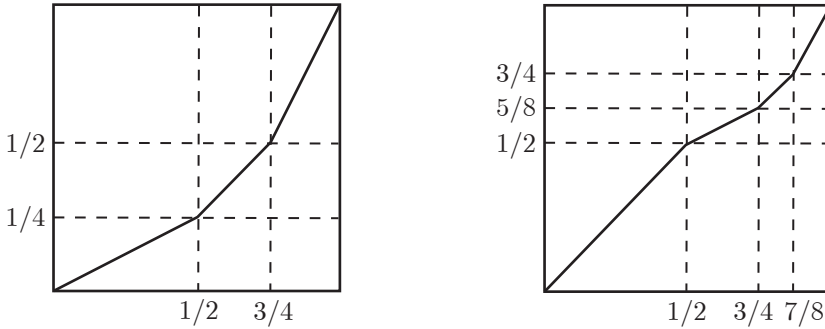


FIGURE 1. The canonical generators  $x_0$  and  $x_1$  of  $F$ .

We will consider two kinds of normal forms for the Thompson group  $F$ . The first one is given by words:

$$x_0^{p_0} x_1^{p_1} \cdots x_n^{p_n} x_n^{-q_n} \cdots x_1^{-q_1} x_0^{-q_0}$$

where  $n, p_0, p_1, \dots, p_n, q_0, q_1, \dots, q_n$  are non-negative integers such that:

- (i) exactly one of  $p_n$  or  $q_n$  is non-zero, and
- (ii) if  $p_k > 0$  and  $q_k > 0$  for some  $0 \leq k < n$ , then  $p_{k+1} > 0$  or  $q_{k+1} > 0$ .

The left half  $x_0^{p_0} x_1^{p_1} \cdots x_n^{p_n}$  is called the *positive part* of the word and the right half  $x_n^{-q_n} \cdots x_1^{-q_1} x_0^{-q_0}$  the *negative part*. A word is said to be *positive* (or *negative*) if its normal form only consists of its positive (or negative) part.

The second normal form is given by the so called *reduced forest diagrams*, see [1]. Recall that a binary forest is a finite sequence of binary trees, together with a pointer on one of the trees. The number of leaves in a binary forest is the sum of the numbers of leaves in its binary trees. A *forest diagram* is a pair of binary forests which have the same number of leaves. We speak of the *bottom forest* as well as of the *top forest*, see Figure 2.



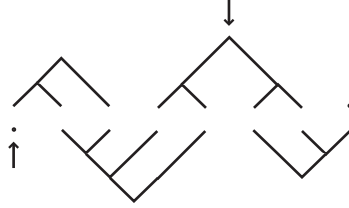


FIGURE 2. A forest diagram with 8 leaves

A *caret* in a tree is a pair of leaves with the same parent vertex. A forest diagram is reduced if it has no opposite pairs of (bottom and top) carets (cf. [1]). For example, the diagram of Figure 2 is reduced.

One associates to an arbitrary reduced forest diagram an element of  $F$  with normal form  $x_0^{p_0} x_1^{p_1} \cdots x_n^{p_n} x_n^{-q_n} \cdots x_1^{-q_1} x_0^{-q_0}$  as follows.

- Enumerate top and bottom leaves, as well as top and bottom trees, from the left to the right, starting at 1.
- The top (or bottom) forest gives the positive part (or negative part respectively) of the normal form.
- The exponent of  $x_i$ , for  $i > 0$ , equals to the maximal length of simple paths in the top forest starting at the  $i^{th}$  top leaf and following the top-to-right direction (the exponent is 0 whenever such a leaf does not exist).
- The exponent of  $x_i^{-1}$ , for  $i > 0$ , equals to the maximal length of simple paths in the bottom forest starting at the  $i^{th}$  bottom leaf and following the bottom-to-right direction.
- The exponent of  $x_0$  is  $n$  whenever the top pointer is on the  $(n + 1)^{st}$  tree.
- The exponent of  $x_0^{-1}$  is  $n$  whenever the bottom pointer is on the  $(n + 1)^{st}$  tree.

For example, the reduced forest diagram of Figure 2 gives the element:

$$x_0 x_1^2 x_4^2 x_6 x_7^{-1} x_6^{-1} x_2^{-3} .$$

Notice that adding a top and a bottom leaf on the right of a forest diagram does not change the corresponding element of  $F$ . Up to this trivial transformation, it turns out that each element of  $F$  can be represented by a unique reduced forest diagram (cf. [1]). This is the second normal form we are interested in.

## TESTING CAYLEY GRAPH DENSITIES

For example, the generators  $x_0, x_1$  and their inverses are represented by the reduced forest diagrams given in Figure 3.

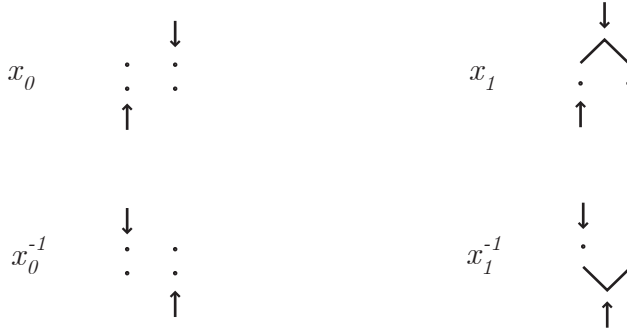


FIGURE 3. Forest diagrams for the generators  $x_0, x_1$  and their inverses

Below we use the following definitions: A binary forest is *trivial* if each of its subtrees consists of a single vertex only, and if the pointer is on the first of them. A forest diagram is *negative* (or *positive*) if its top (or bottom) tree is trivial; in such a case the normal form of the associated element of  $F$  is negative (or positive respectively). The *height* of a forest is the maximal height of one of its binary trees; it can take any value between 0 and the number of leaves minus 1 (or equivalently, the number of carets). For instance, the top forest and the bottom forest of the forest diagram in Figure 2 have height 2 and 3 respectively.

### 4. Special subsets in Thompson's group $F$

For the first three groups considered in this paper,  $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$ , Baumslag-Solitar group  $BS(2, 2)$ , and the wreath product  $\mathbb{Z} \wr \mathbb{Z}$ , all of our numerical experiments are performed on balls  $B(n)$  of radius  $n$  in the Cayley graph  $\mathcal{C}$ , centered around the vertex defined by the neutral element  $1 \in G$ . For the fourth group, Thompson's group  $F$ , we will work with balls, but also with other kinds of sets, which we specify now.

### 4.1. Left positive balls

Let  $n > 0$ . The *left positive ball of radius  $n$* , denoted by  $LP(n)$ , is defined to be the maximal subgraph in the Cayley graph  $\mathcal{C} = \mathcal{C}(F, \{x_0, x_1\})$  which contains only inverses of positive words:

$$x_0^{p_0} x_1^{p_1} \cdots x_n^{p_k} \quad \text{with} \quad p_0 + p_1 + \cdots + p_k \leq n$$

as vertices. Notice that for  $n \geq 6$  the left positive ball  $LP(n)$  cannot be a tree: indeed, our densification algorithm (see §5) deletes subtrees from some of the  $LP(n)$ , but it does not delete  $LP(6)$  (cf. §9.5). This shows in particular that  $LP(6)$  is not a tree. Furthermore, obviously  $LP(n)$  is always a subgraph of  $LP(n + 1)$ .

### 4.2. Negative forests

A *negative forest with  $n$  leaves*,  $n \in \mathbb{N}$ , denoted by  $NF(n)$ , is defined to be the maximal subgraph of the Cayley graph which contains only vertices that are given by group elements which are represented by a negative reduced forest diagram with at most  $n$  leaves. Obviously one has  $NF(n) \subset NF(n + 1)$ . The negative forest with 3 leaves,  $NF(3)$ , is given in Figure 4.

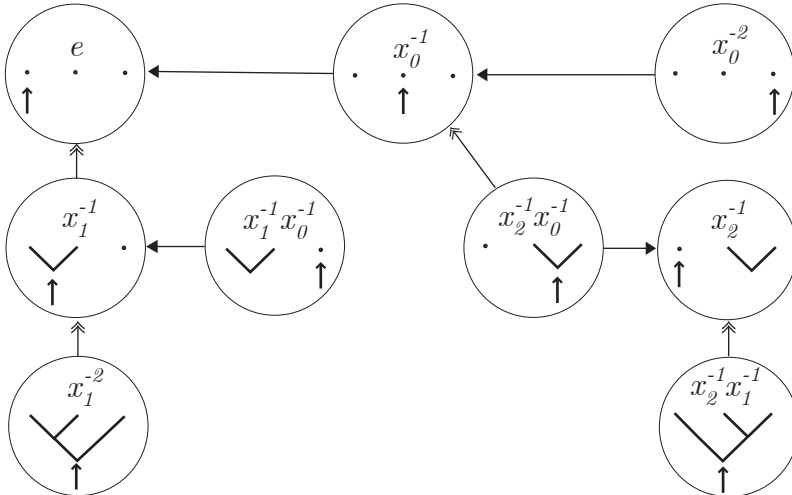


FIGURE 4. The negative forest  $NF(3)$

In the graphical representation of a negative forest we only draw the representative bottom forest. For example, in Figure 4 and Figure 5 below the generators  $x_0$  and  $x_1$  are given by simple and double arrows respectively.

Notice that even though  $NF(3)$  is a tree,  $NF(n)$  is not a tree for  $n \geq 5$  (see, for example, §9.6).

### 4.3. Belk-Brown sets

The *Belk-Brown set with  $n$  leaves and of height at most  $k$* , for  $n, k \in \mathbb{N}$  and  $k < n$ , denoted by  $BB(n, k)$ , is the maximal subgraph of the Cayley graph  $\mathcal{C}$  which contains as vertices only elements that have a negative reduced forest diagram with  $n$  leaves and height at most  $k$  (cf. [1]). For instance, Figure 5 represents the Belk-Brown set  $BB(4, 1)$ . Note that  $BB(4, 1)$  is not a tree (it contains the loop  $x_0^{-2}x_1^{-1}x_0^2x_1^{-1}x_0^{-1}x_1x_0x_1$ ).

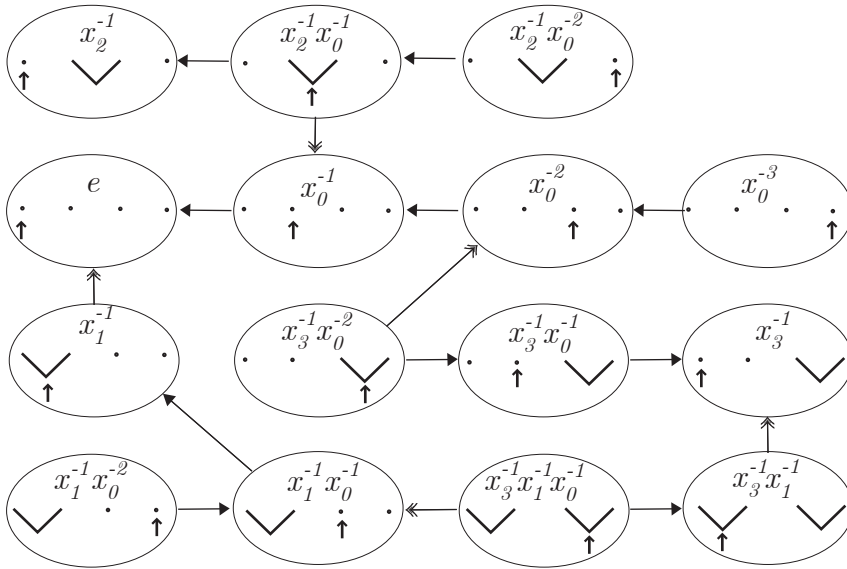


FIGURE 5. The Belk-Brown set  $BB(4, 1)$ .

Obviously one has  $BB(n, k) \subset BB(n+1, k)$  and  $BB(n, k) \subset BB(n, k+1)$ . In particular  $BB(n, k)$  is not a tree whenever  $n \geq 4$  and  $k \geq 1$ . Since

the height of a forest is at most the number of its leaves minus 1, we have  $BB(n, n - 1) = NF(n)$ . This shows that  $NF(n)$  is not a tree whenever  $n \geq 4$ , as already stated in the previous section.

## 5. The densification algorithm

In this section we describe the algorithm by which we can improve the density of a given finite graph, through passing over to a subgraph.

Given a finite subgraph  $A$  in the Cayley graph  $\mathcal{C}$  of a finitely generated group  $G$ , the algorithm applies a finite sequence of reductions (given in detail below), and returns a new “densified” subgraph  $\bar{A}$ . If the initial finite graph  $A$  is sufficiently dense, then the returned graph  $\bar{A}$  will have even higher density. Otherwise, for example if  $A$  is a cycle or a tree, it is possible that the algorithm will collapse  $A$  to a single vertex. As our applications concern all Cayley graphs built on 2-element generating systems, our algorithm is tuned to graphs with vertex degree uniformly bounded by 4. The necessary modifications for higher vertex degree, if needed, are fairly easy to device.

We first need to introduce some terminology. Let  $K$  be a finite graph. A *chain* is a maximal simple path in  $K$  where all of its vertices, except for the endpoints, have degree 2 in  $K$ . The length of a chain is the number of its vertices of degree 2. A *cycle* is a simple loop in  $K$  where all vertices have degree 2 in  $K$ . A *tripod* is a subgraph of  $K$  which consists of 3 chains which have precisely 1 vertex  $v$  in common, and  $v$  is an endpoint in each of the three chains. The length of a tripod is the sum of the lengths of its three chains. A *degenerated tripod* is a subgraph of  $K$  which consists of two chains  $c$  and  $c'$ , such that the two endpoints of  $c$  coincide with a vertex  $v$ , and precisely one of the endpoints of  $c'$  also coincides with  $v$ . The length of a degenerated tripod is the sum of the lengths of the two chains  $c$  and  $c'$ . We require that, in case of a tripod or a degenerated tripod, the vertex  $v$  has degree 3 in  $K$ ; *i.e.* there is no other edge adjacent to  $v$ .

We will now define four types of elementary reductions. They correspond to the removal of chains, cycles, and tripods (degenerated or not) from  $K$ , whenever their length is large enough. Any such transformation will in most cases increase the density, see Lemma 5.1.

For a finite graph  $K$  with density  $\delta(K)$  we define the following parameters: If  $\delta(K) \neq 2$ , let

$$N_c(K) = \max(0, \frac{2}{\delta(K) - 2}), N_t(K) = \max(0, \frac{4}{\delta(K) - 2} - 1)$$

and  $N_d(K) = \max(0, \frac{2}{\delta(K) - 2} - 1)$ . If  $\delta(K) = 2$ , we set  $N_c(K) = N_t(K) = N_d(K) = 0$ .

(R1): Remove any subtree of  $K$ .

(R2): Remove any cycle of  $K$ .

(R3): Remove all chains of length greater than  $N_c(K)$  from  $K$ .

(R4): Remove from  $K$  some tripod of length greater than  $N_t(K)$ , or some degenerated tripod of length greater than  $N_d(K)$ . Repeat this procedure as often as possible.

**Lemma 5.1.** *If  $K$  is a finite graph with density  $\delta(K) > 2$ , then any of the above elementary transformations (R1), (R2), (R3) or (R4) transforms  $K$  into a subgraph  $K'$  of strictly larger density of  $\delta(K)$ .*

*Proof.* It suffices to check that the number of edges removed in any of the elementary transformations is strictly smaller than  $\delta(K)$  times the number of vertices removed. This is trivially true for the transformations (R1) and (R2), since any tree has one more vertex than unoriented edges, and any cycle has equal number of vertices as unoriented edges. Since according to our conventions we have to count an edge and its inverse separately, this gives directly the desired inequality.

For the reduction (R3) we observe that any chain  $c$  of length  $n > N_c(K)$  has precisely  $2n + 2$  edges (counting again an edge and its inverse separately) and  $n + 2$  vertices, and while all of the edges are removed with  $c$ , only the  $n$  interior vertices of  $c$  are removed. Since  $\delta(K) > 2$  and  $N_c(K) = \max(0, \frac{2}{\delta(K) - 2})$ , one has  $n > \frac{2}{\delta(K) - 2}$  and hence  $\delta(K) > \frac{2}{n} + 2 = \frac{2n+2}{n}$ , which precisely what we need.

Any tripod of length  $n$  consists precisely of  $2n + 6$  edges and  $n + 4$  vertices, where  $n + 1$  of them (as well as all edges) will be removed. A degenerated tripod of length  $n$  consists precisely of  $2n + 4$  edges and  $n + 2$  vertices, of which  $n + 1$  will be removed. The further calculation for (R4) is very similar to the above one for (R3) and thus left to the reader.  $\square$

The algorithm proceeds as follows:

```

Given a finite graph  $K$ 
DO WHILE  the graph  $K$  is changing
  Apply successively reductions  $(R1), (R2), (R3), (R1), (R2),$ 
   $(R4)$ 
END
RETURN  the densified graph  $\overline{K}$ 

```

Below we will call each successive applications of the reductions  $(R1), (R2), (R3), (R1), (R2), (R4)$  a *round*.

## 6. The algorithmic package

As mentioned in the introduction, our numerical results were obtained by means of computer calculations, executed by a program written in C++. In this section we give a brief description of the three parts I, II and III of our programmed algorithmic package, of its software routines and also of the assumptions and limitations involved. A fourth computational feature, concerning the linear interpolation of the numerical data obtained by parts I - III, and in particular calculating an *interpolated limit density* is performed using *Matlab* and is described in §8 below.

Part I of the algorithmic package consists of subprograms, one for each class of groups  $G$  considered here, that transform a given product of generators or their inverses into a word in normal form as introduced in §3.

Part II calculates, for a given parameter  $n$ , the finite graphs  $B(n)$ , or, in case where the group in question is Thompson's  $F$ , the finite graphs  $LP(n)$ ,  $NF(n)$ , or, for given  $k$  and  $n$ , the graph  $BB(n, k)$ , as defined in §4.

Part III calculates, for any finite graph  $A$  (= the graph computed in part II) a densified subgraph  $\overline{A}$  according to the algorithm presented in §5.

It is an important characteristic of our algorithmic package that it is organized in a strictly modular fashion where the different parts work independently from each other. This gives the possibility to easily improve specific parts without having to change the rest. For example, new classes of groups can be investigated by adding new subprograms to part I without changing parts II or III, new families of sets for the known groups can be investigated without changing I or III, and the densification algorithm could be embellished and reapplied to the groups and set families already programmed without ever changing I or II.

## 6.1. What the program can do

The program works in a console mode. A contextual menu allows the user to choose any of the actions. The actions, besides saving a copy of the outputs into a text file and offering some further options, consists mainly in:

- Choose one of the predefined groups that can be any free, finitely generated free abelian, Baumslag-Solitar, group,  $\mathbb{Z} \wr \mathbb{Z}$  or Thompson's group  $F$ ; all further computations will concern this group.
- Perform direct computations, like writing an element or a product of elements in normal form.
- Construct one (or a sequence) of the predefined finite graphs as explained in §4, and compute their density (or alternatively their isoperimetric constant).
- Apply to such a finite graph (or sequence of finite graphs) the algorithm of §5. The program provides some extra information, like the density at each step, and further details concerning the application of the elementary reductions (R1)–(R4).

## 6.2. What the program is made of

The program is written in standard  $C^{++}$  and can be compiled either on *Linux* or *Win32* platforms. It can be easily adapted to compilation on other platforms.

The program makes intensive use of the object-oriented abilities of  $C^{++}$ . Groups, graphs, vertices are all objects (or `class`); all the main algorithms correspond to general functions which takes data as input (like a group), and returns data. The functions which construct balls in the Cayley graph, and implement the algorithm of §5 are general and can be applied to any implemented group or finite graph.

Elements are given by strings of characters. This allows more choices when one encodes an abstract group element. Usually, strings look like words on given canonical generators and inverses. So they are really close to their mathematical meaning. However, for example in the case of Thompson's group  $F$ , they don't represent words on the canonical generators, but encode normal forms.



### 6.3. Limitations

The graphs are constructed in the physical memory (RAM) of the computer, and their size is almost proportional to the number of vertices. The main limitation of our computation is obviously the size of the computed graphs. This is closely related to the complexity of the group: In groups of polynomial growth, our computations of  $B(n)$  can easily be implemented for  $n$  going up to hundreds or thousands, while in groups of exponential growth  $n$  goes hardly up to 20. On the other hand, the program can handle free groups of rank at most 128, free abelian groups of rank at most 128, and words in wreath products with  $i_n$  at most 127. For the group  $F$  the program can handle normal forms  $x_0^{p_0} \dots x_n^{p_n} x_n^{-q_n} \dots x_0^{-q_0}$  with  $n$  up to 127.

## 7. Summary of experimental results

In this section we give an overview of our experimental results. For each of the classes of groups considered in §3 and for each type of the special subsets defined in §4 our presentation contains the following parts:

- (a) – known theoretical results for the group;
- (b) – best values of densities calculated by our program;
- (c) – analysis of the work of the densification algorithm;
- (d) – values of the interpolated limit density;
- (e) – comments.

In the next section we give a graphical interpretation of our experimental results, and in Section 9 we present the numerical data obtained.

### 7.1. Comparative analysis I: Amenable vs non-amenable

#### 7.1.1. Free abelian group $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$

(a) The group has a polynomial growth, and hence balls  $B(n)$  are known to be a family of Følner sets. The slow growth allows an easy implementation of balls  $B(n)$  for large  $n$  (for hundreds or for thousands). Also,

## TESTING CAYLEY GRAPH DENSITIES

theoretical values of the density of balls and of densified balls are very easy to calculate.

(b) The ball of radius 171 in  $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$  has density 5.94752 before and 5.94812 after the densification algorithm is applied.

(c) The densification algorithm does not change the initial density significantly: the increase of density is less than 1%.

(d) The interpolated values of the limit density coincide with the theoretical value : 6. Moreover, this is the case, both in small scale (for  $n = 1, \dots, 15$ ) and in large scale (for  $n = 1, \dots, 171$ ) calculations, which numerically confirms that balls constitute a Følner family.

### 7.1.2. Baumslag-Solitar group $BS(2, 2)$

(a) This group contains  $F_2 \times \mathbb{Z}$  as a subgroup (of index 2) and hence it is not amenable.

(b) The ball of radius 18 has density 2.58585. The densified ball of radius 18 has density 2.928.

(c) The densification algorithm induces a 14% increase in density and remove approximatively 40% of vertices.

(d) The interpolated values of the limit density of balls and of densified balls are 2.64 and 2.97, respectively.

(e) An interesting point is that the algorithm runs only through one round.

### 7.1.3. Wreath product $\mathbb{Z} \wr \mathbb{Z}$

(a) This is an amenable group of exponential growth. Balls  $B(n)$  are not a Følner family.

(b) The ball of radius 16 has density 2.32838. The densified ball of radius 16 has density 2.90938.

(c) The densification algorithm produces a 25% increase in density and removes approximatively  $2/3$  of vertices. Thus the algorithm is quite efficient in this case.

(d) The interpolated values of the limit densities of balls and of densified balls are 2.43 and 3 respectively.

## 7.2. Comparative analysis II: The Thompson group $F$

Amenability of  $F$  is unknown, but one knows that  $F$  grows exponentially (cf. [2]), so that balls will certainly not give a Følner family.

### 7.2.1. Balls $B(n)$ in $F$

(b) The ball of radius 15 has density 2.14905. The densified ball of radius 15 has density 2.7183.

(c) The densification algorithm induces a 25% increase in density. It removes more than 80% of vertices.

(d) The interpolated limit densities of balls and of the densified balls are 2.23 and 2.8 respectively.

(e) The results are quite similar to the above case of the wreath product  $\mathbb{Z} \wr \mathbb{Z}$ . An interesting point is that the densification algorithm performs at most three rounds for  $n < 15$ . However it suddenly takes 132 rounds to perform calculations for  $n = 15$  and the density increases a lot. This allows to believe that there may well exist a subset of much higher density than given by our interpolation.

### 7.2.2. Left positive balls $LP(n)$ in $F$

(b) The left positive ball  $LP(19)$  has density 2.15988. The densified left positive ball  $\overline{LP}(19)$  is of density 2.74349.

(c) The densification algorithm yields an increase in density of approximately 27%. This is one of the best values obtained. The densified left positive balls are particularly small: up to 90% of the vertices are removed by the densification algorithm.

(d) The interpolated limit densities of left positive balls and of densified left positive balls are 2.22 and 2.97 respectively.

(e) The densification algorithm appears to be most efficient in case of these particular graphs.

### 7.2.3. Negative forests $NF(n)$ for $F$

(b) The negative forest  $NF(14)$  has density 2.47619. The densification of this subgraph gives density 2.79448.

- (c) The densification algorithm is rather inefficient: it gives a 13% increase in density and removes less than 60% of the vertices.
- (d) The interpolation of the limit density gives values 2.67 and 3.03 for the negative forest and for the densified negative forest respectively.
- (e) The interpolated values of the limit density are exceptionally close to the calculated ones: the norm of the residues is approximatively  $10^{-5}$ .

#### 7.2.4. Belk-Brown sets $BB(n)$ for $F$

- (a) Since  $BB(n, n - 1) = NF(n)$  both implementations have the same behavior, even though distinct routines and functions are used.
- (b) The Belk-Brown set  $BB(17, 3)$  has density 2.82642. The densified set  $\overline{BB}(17, 2)$  is of density 2.89577. This is comparable to the case of  $\mathbb{Z} \wr \mathbb{Z}$ .
- (c) The densification algorithm increases density by less than 13% and removes approximatively 60% of vertices.
- (d) The interpolation of limit densities gives 3.18, both before and after the densification. Thus the densification algorithm seems to be inefficient in this case.
- (e) Our interpolations do not agree with theoretical values: it has been announced that the limit of density of Belk-Brown sets tends to 3.5, see [1].

Notice that the best value of density of  $BB(n, k(n))$  is obtained whenever  $k(n)$  increases, see Figure 27 and Figure 28. At the same time, the best value of densities of  $\overline{BB}(n, k(n))$  appear for  $k(n) = 3$  (or for slowly growing  $k(n)$ ).

For  $n$  fixed and  $k$  large enough all the  $\overline{BB}(n, k)$  have the same density and the same number of vertices. It is an interesting question whether these finite graphs are isomorphic as subgraphs of the Cayley graph.

## 8. Graphics and interpolation

We give a graphical interpretation of our experimental results. The main numerical results are given in §9. Each subsection below concerns the density of a family of finite graphs in a given group, and its behavior under application of the densification algorithm. We study successively:

$$(8.2) \text{ balls } B(n) \text{ in } \mathbb{Z} \times \mathbb{Z} \times \mathbb{Z},$$

$$(8.3) \text{ balls } B(n) \text{ in } BS(2, 2),$$

- (8.4) balls  $B(n)$  in  $\mathbb{Z} \wr \mathbb{Z}$ ,
- (8.5) balls  $B(n)$  in Thompson's group  $F$ ,
- (8.6) left positive balls  $LP(n)$  in  $F$ ,
- (8.7) negative forests  $NF(n)$  in  $F$ ,
- (8.8) Belk-Brown's sets  $BB(n)$  in  $F$ .

Except for Thompson's group  $F$ , amenability (or not) of these groups is well known (see §2). In order to estimate the limit (or limit superior) of the density of the families of subgraphs considered, we apply a first order approximation to the numerical data obtained from our experiments. This interpolation allows us, in a certain sense, to extrapolate this limit of densities by a value called "interpolated limit density" of this family of subgraphs. Of course, the reader has to be aware that for groups with exponential growth this does only estimate a lower bound to the density of the Cayley graph, compare the discussion at the end of §2.

### 8.1. Method of interpolation

We consider the densities  $\delta_n$  of a sequence of finite graphs  $S_n$ , for  $p+1 \leq n \leq q$ . We approximate the  $\delta_n$  by a real-valued function  $f(n)$ , specified below, defined on the domain  $\{p+1, p+2, \dots, q\}$ . We estimate the quality of the approximation first by constructing the vector in  $\mathbb{R}^{q-p}$  whose  $n^{\text{th}}$ -component is the residue  $\delta_n - f(n)$ , and later by considering its euclidian norm. We call this norm, the *norm of residues*. It is a non-negative number. Clearly, the smaller it is, the better the approximation will be, with the zero value for the norm of residues in the case of a perfect correspondence of  $f$  with the given values of  $\delta_n$ .

The approximating function  $f$  is set to be of type:

$$f(n) = \frac{an + b}{n + x}.$$

There are two main reasons: on one hand, these functions give the best experimental results. On the other hand, assume that the values  $\delta_n$  are well approximated by some rational function  $f(n) = \frac{P(n)}{Q(n)}$ . Then one easily sees (observing that  $\lim_{n \rightarrow \infty} \delta_n$  is neither zero nor infinite) that  $P$  and  $Q$  have

to be of the same degree, and thus  $f(n)$  may be just as well approximated by a function of the above type  $\frac{an+b}{n+x}$ .

The key points for the interpolation procedure now are the following: First, the interpolation reduces to the consideration of a parameter  $x$  and a linear interpolation  $\delta_n(n+x)$  in order to obtain  $a$  and  $b$ . Here the value of  $x$  is chosen such that  $\delta_n(n+x)$  is best distributed close to a line. That is, the corresponding norm of residues in a linear interpolation is the smallest one among all possibilities for the value of  $x$ . (In the search for the best  $x$ , using *Matlab*, we consider only large enough  $n$ , and the values for  $x$  are only considered up to  $10^{-1}$ .)

Observe that the limit of  $f(n)$  for  $n \rightarrow \infty$  is  $a$ , and hence this is the value, called *interpolated limit density*, which we use as parameter to estimate the limit of  $\delta_n$ . The latter is, after all, the information we are mainly interested in.

*Aside:* An interesting experimental result in the above described interpolation procedure is that the parameter  $a$  remains essentially unchanged whenever  $x$  is slightly modified. This stability with respect to perturbations seems interesting in light of the fact that  $a$  is related to the approximation of zero order: the line  $y = ax$  gives the asymptotic direction.

## 8.2. Free abelian group of rank 3

We proceed as above by comparing large scale and small scale interpolations with the true values. Numerical results are given in §9.1.

The “large scale” results for  $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$  appear in Figure 6. Let us restrict to  $n = 11, 21, \dots, 171$ . The densities of balls  $B(n)$ , rescaled via  $x + 0.6$ , are distributed along the line  $y = 6x - 5.4$ , see Figure 7. The norm of residues is approximately equal to 0.011. The interpolation gives:

$$\delta(B(n)) \approx \frac{6n - 5.4}{n + 0.6},$$

the interpolated limit density is equal to 6.

The norm of residues of the interpolation of the density of densified balls  $\overline{B}(n)$  is approximately 0.032, see Figure 8. One obtains the interpolation:

$$\delta(\overline{B}(n)) \approx \frac{6n + 4.3}{n + 2.2}.$$

The interpolated limit density is equal to 6.

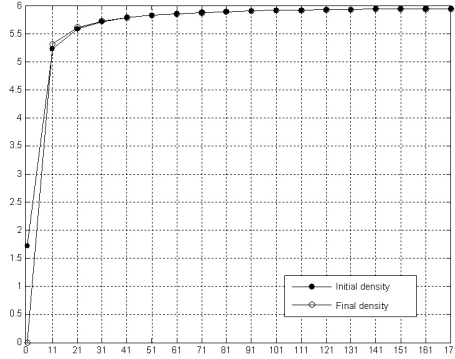


FIGURE 6. Large scale density of balls  $B(n)$  in  $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$

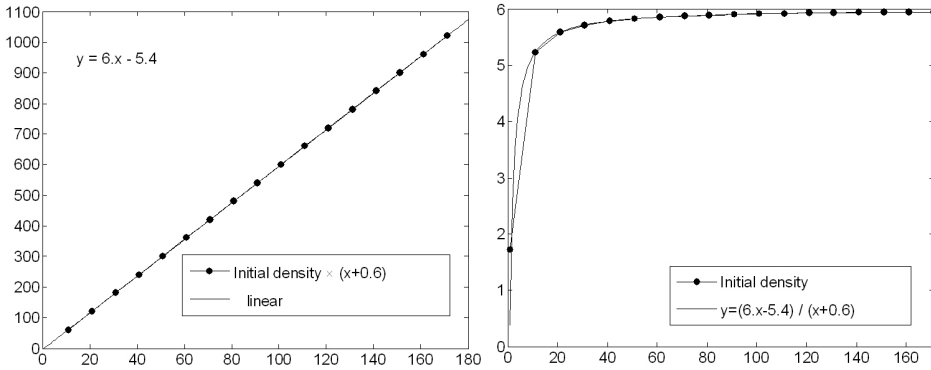


FIGURE 7. Interpolation of the large scale densities of balls  $B(n)$  in  $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$

The “small scale” results for  $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$ , for  $n = 1, 2, \dots, 15$ , are presented in Figure 9. We restrict ourselves to  $n = 3, 4, \dots, 15$ . Multiplying initial densities by  $x + 1$  gives the distribution along the line  $y = 6x - 3.6$ , Figure 10. The norm of residues of the density of balls  $B(n)$  is less than 0.05, Figure 10. The interpolation gives:

$$\delta(B(n)) \approx \frac{6n - 3.6}{n + 1}.$$

The interpolated limit density is equal to 6.

## TESTING CAYLEY GRAPH DENSITIES

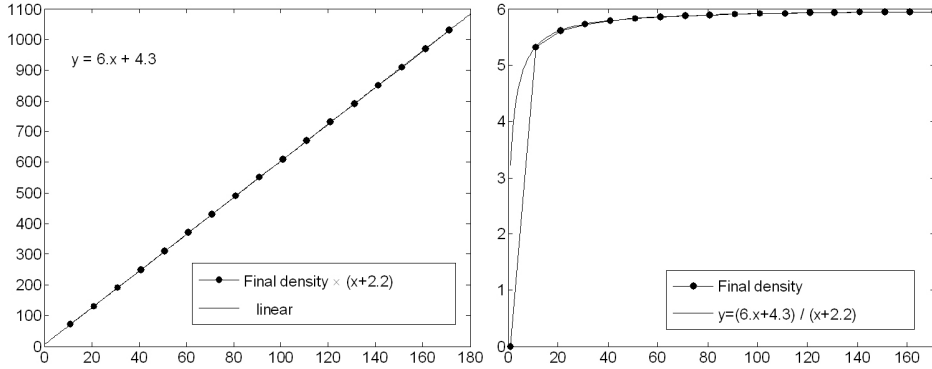


FIGURE 8. Interpolation of large scale density of densified balls  $\bar{B}(n)$  in  $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$

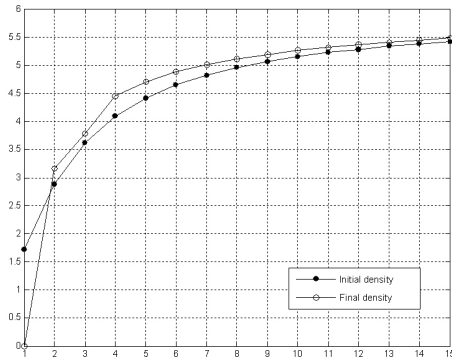


FIGURE 9. Small scale density of balls  $B(n)$  in  $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$

For  $n = 4, 5, \dots, 15$ , multiplying the density of densified balls  $\bar{B}(n)$  by  $x + 1.1$  gives distribution along the line  $y = 6x - 1.1$ , Figure 11. The norm of residues is close to 0.061. The interpolation gives:

$$\delta(\bar{B}(n)) \approx \frac{6n - 1.1}{n + 1.1},$$

and the interpolated limit density is equal to 6.

A ball  $B(n)$  has  $4n^3 + 2n$  edges and  $(4n^3 + 6n^2 + 8n + 3)/3$  vertices,

$$\delta(B(n)) = \frac{24n^3 + 12n}{4n^3 + 6n^2 + 8n + 3} \underset{+\infty}{\sim} \frac{6n}{n + 3/2}.$$



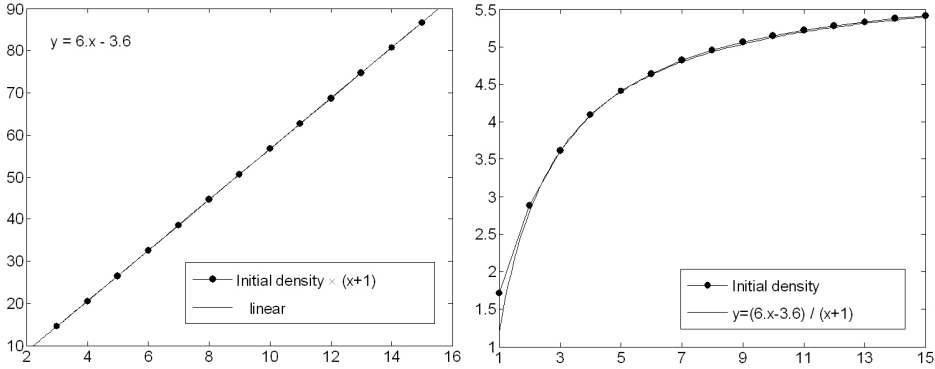


FIGURE 10. Interpolation of the small scale density of balls  $B(n)$  in  $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$

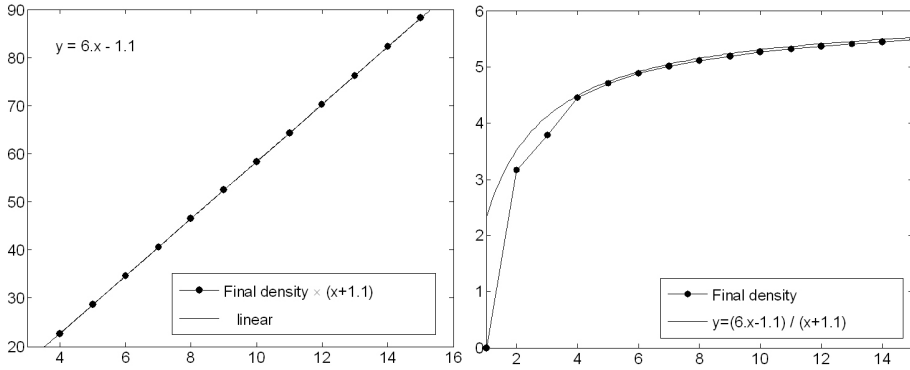


FIGURE 11. Interpolation of the small scale density of densified balls  $\overline{B}(n)$  in  $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$

The density of densified balls  $\overline{B}(n)$  can be computed for  $n \geq 4$ . The algorithm removes 12 vertices of valency 1 and  $12(n - 1)$  edges. Thus it removes  $12n$  vertices and  $24n - 12$  edges. We deduce that:

$$\delta(\overline{B}(n)) = \frac{24n^3 - 132n + 72}{4n^3 + 6n^2 - 28n + 3} \underset{+\infty}{\sim} \frac{6n}{n + 3/2}.$$

Our (small and large scale) interpolations are not exactly the same, but they are not so far off either. The key point is that in all cases one recovers the theoretical limit density 6: The given subsets are a Følner family.

### 8.3. Baumslag-Solitar group $BS(2, 2)$

Graphics are given in Figure 12. Numerical results are found in §9.2. For

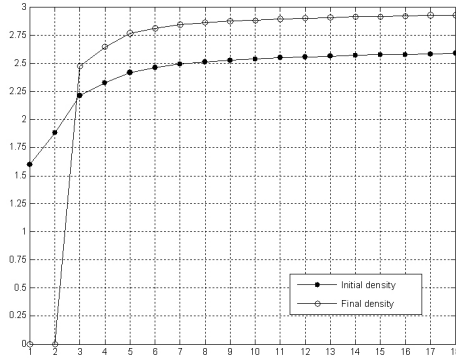


FIGURE 12. Density of balls  $B(n)$  of  $BS(2, 2)$

$n = 3, 4, \dots, 18$ , the density  $\delta(B(n))$ , multiplied by  $(x - 0.9)$ , is distributed close to the line  $y = 2.64x - 3.3$ , Figure 13. The norm of residues is approximatively 0.066. The interpolation gives:

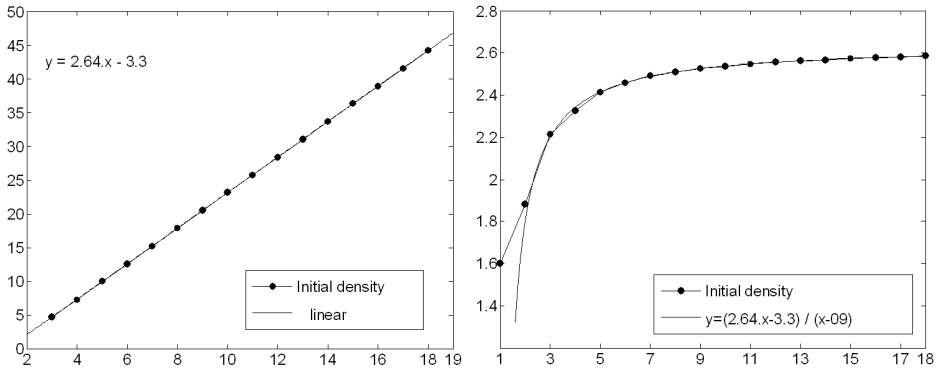


FIGURE 13. Interpolation of densities of balls  $B(n)$  in  $BS(2, 2)$

$$\delta(B(n)) \approx \frac{2.64n - 3.3}{n - 0.9}$$

and the interpolated limit density is equal to 2.64.

For  $n = 4, 5, \dots, 18$ , we multiply the density of the densified balls  $\overline{B}(n)$  by  $x - 1.9$ . This gives values that can be interpolated by the line  $y = 2.97x - 6.29$ , see Figure 14. The norm of residues is approximately 0.084. Our approximation gives:

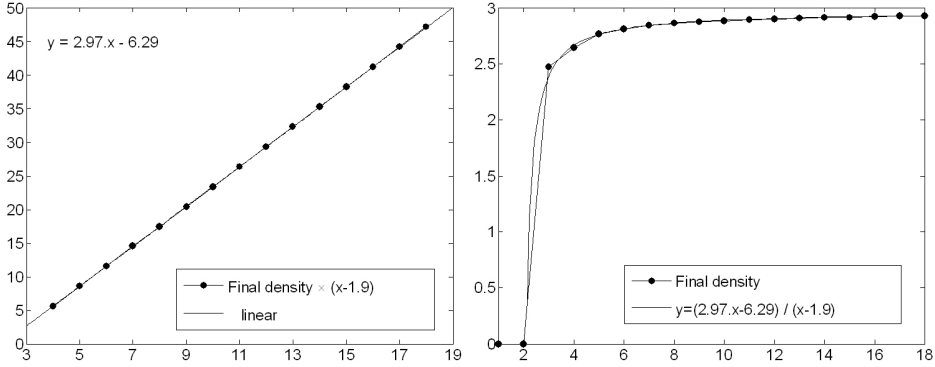


FIGURE 14. Interpolation of densities of densified balls  $\overline{B}(n)$  in  $BS(2, 2)$

$$\delta(\overline{B}(n)) \approx \frac{2.97n - 6.29}{n - 1.9}.$$

The interpolated limit density is equal to 2.97.

#### 8.4. Wreath product $\mathbb{Z} \wr \mathbb{Z}$

Graphics are given in Figure 15. Numerical results are given in §9.3. The initial density, multiplied by  $x + 0.9$ , is distributed close to the line  $y = 2.43x + 0.55$ , see Figure 16. The norm of residues is approximately equal to 0.25.

The interpolation of densities is given by:

$$\delta(B(n)) \approx \frac{2.43n + 0.55}{n + 0.9}.$$

The interpolated limit density is equal to 2.43.

## TESTING CAYLEY GRAPH DENSITIES

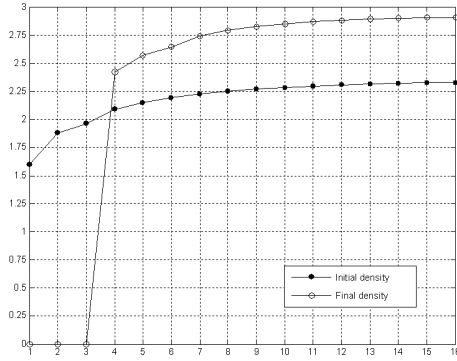


FIGURE 15. Density of balls  $B(n)$  in  $\mathbb{Z} \wr \mathbb{Z}$

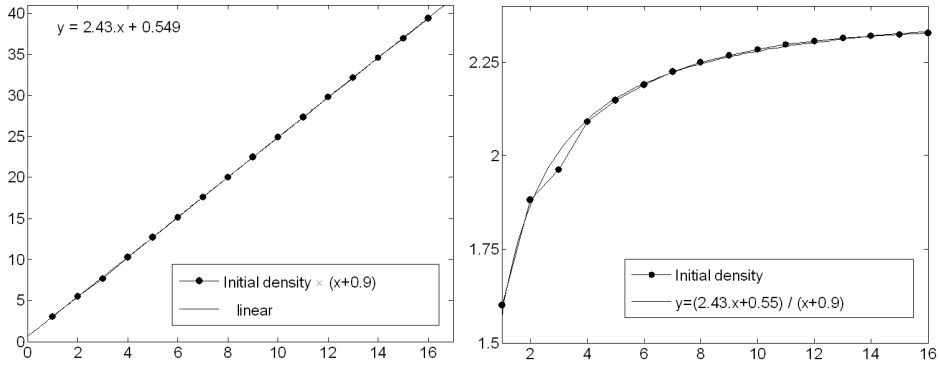


FIGURE 16. Interpolation of the densities of balls  $B(n)$  in  $\mathbb{Z} \wr \mathbb{Z}$

We interpolate the densities  $\delta(\overline{B}(n))$  by using our numerical results for  $n = 4, 5, \dots, 16$ . The norm of residues of our interpolation is approximately 0.23, see Figure 17. The density of densified balls  $\overline{B}(n)$  is interpolated by:

$$\delta(\overline{B}(n)) \approx \frac{3n - 7}{n - 1.9}.$$

The interpolated limit density is equal to 3.

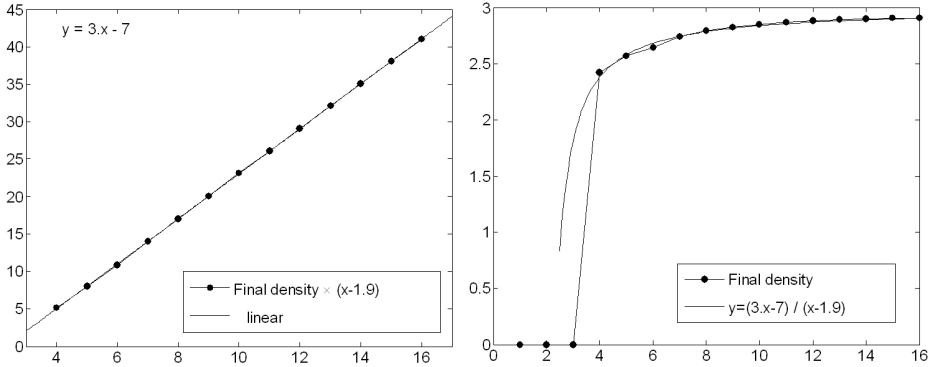


FIGURE 17. Interpolation of densities of densified balls  $\overline{B}(n)$  in  $\mathbb{Z} \wr \mathbb{Z}$

### 8.5. Balls $B(n)$ in Thompson’s group $F$

We now consider Thompson’s group  $F$ , for which amenability is unknown. We first investigate the density of balls  $B(n)$  in  $F$ . Our results are given in Figure 18. Numerical results are given in §9.4.

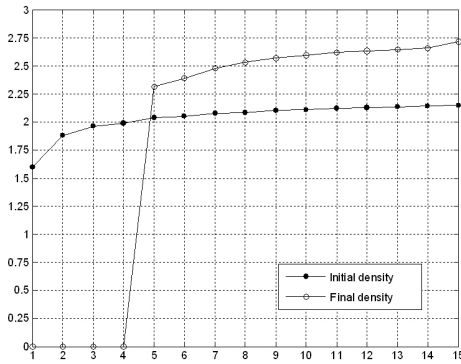


FIGURE 18. Density of balls  $B(n)$  in Thompson’s group  $F$

In order to interpolate the densities  $\delta(B(n))$  we only consider  $n = 3, 4, \dots, 15$ . Multiplication of the density of  $B(n)$  by  $n + 2.3$  gives values close to the line  $y = 2.23x + 3.65$ , see Figure 19. The norm of residues is approximatively 0.123.

## TESTING CAYLEY GRAPH DENSITIES

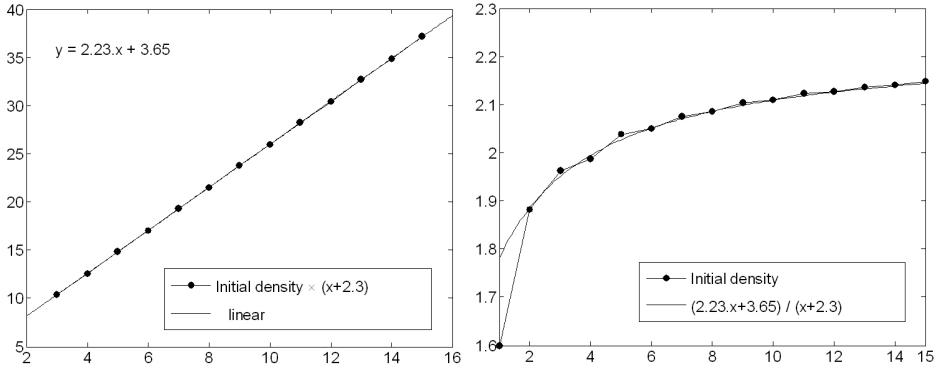


FIGURE 19. Interpolation of densities of balls  $B(n)$  in  $F$

The interpolation gives:

$$\delta(B(n)) \approx \frac{2.23n + 3.65}{n + 2.3}.$$

The interpolated limit density is equal to 2.2.

Now we consider the behavior of densities of balls after the densification algorithm is applied. For  $n = 3, 4, \dots, 14$ , the density of the  $\overline{B}(n)$ , multiplied by  $n - 1.7$ , is distributed close to line  $y = 2.8x - 6.4$ , see Figure 20. The norm of residues is 0.11395.

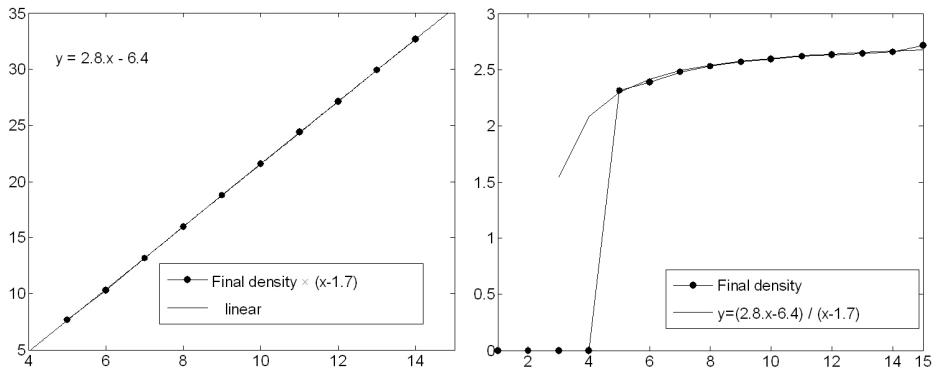


FIGURE 20. Interpolation of the densities of densified balls  $\overline{B}(n)$  in  $F$

The estimation is given by:

$$\frac{2.8n - 6.4}{n - 1.7}.$$

The interpolated limit density is equal to 2.8.

These results are clearly comparable with those of the amenable group  $\mathbb{Z} \wr \mathbb{Z}$ .

### 8.6. Left-positive balls $LP(n)$ in $F$

Numerical results are given in §9.5, graphical data are given in Figure 21.

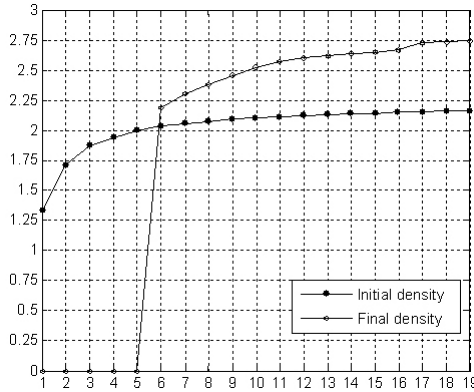


FIGURE 21. Density of left-positive balls  $LP(n)$  in  $F$

The density of the  $LP(n)$ , multiplied by  $n+0.343$ , is distributed close to the line  $y = 2.22x - 0.427$ , see Figure 22. The calculated norm of residues is equal to 0.038316.

This gives the approximation:

$$\delta(LP(n)) \approx \frac{2.22n - 0.427}{n + 0.343}.$$

The density of the densified sets  $\overline{LP}(n)$ , multiplied by  $n - 0.7$ , is distributed close to the line  $y = 2.97x - 6.25$ , see Figure 23. The norm of residues equals to 0.65841. This provides the approximation:

$$\delta(\overline{LP}(n)) \approx \frac{2.97n - 6.25}{x - 0.7}.$$

## TESTING CAYLEY GRAPH DENSITIES

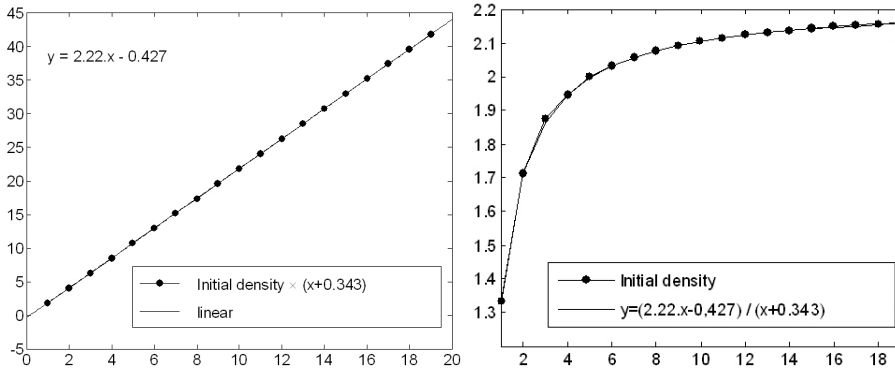


FIGURE 22. Interpolation of the densities of the sets  $LP(n)$  in  $F$

The interpolated limit density is equal to 3.

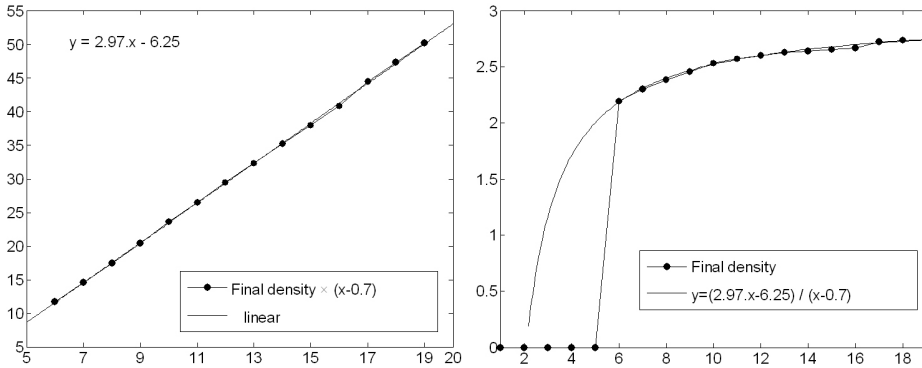


FIGURE 23. Interpolation of the densities of the densified sets  $\overline{LP}(n)$  in  $F$

### 8.7. Negative forests $NF(n)$ in $F$

Numerical results are given in §9.6. A graphical interpretation is given in Figure 24.

The best approximation of the density of  $NF(n)$  is obtained by multiplying it by  $n$ , see Figure 25. The norm of residues is exceptionally low. It



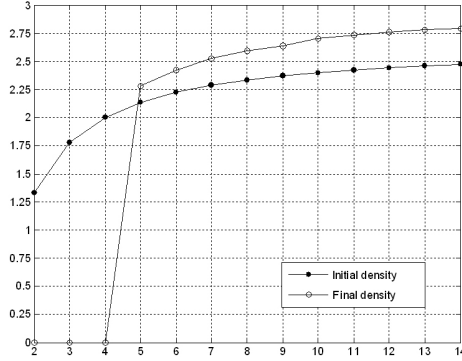


FIGURE 24. Density of negative forests  $NF(n)$  in  $F$

is equal to  $6.4769 \cdot 10^{-5}$ . Notice that both our data and the results of the interpolation have a margin of error of  $10^{-5}$ . The density is distributed close to the line  $2.6667x - 2.6667$ . The densities of  $NF(n)$  are particularly well approximated for  $n = 2, 3, \dots, 14$  by:

$$\delta(NF(n)) \approx \left(2 + \frac{2}{3}\right) \frac{n-1}{n}.$$

For  $n = 5, 6, \dots, 14$ , the density of the densified negative forests  $\overline{NF}(n)$ ,

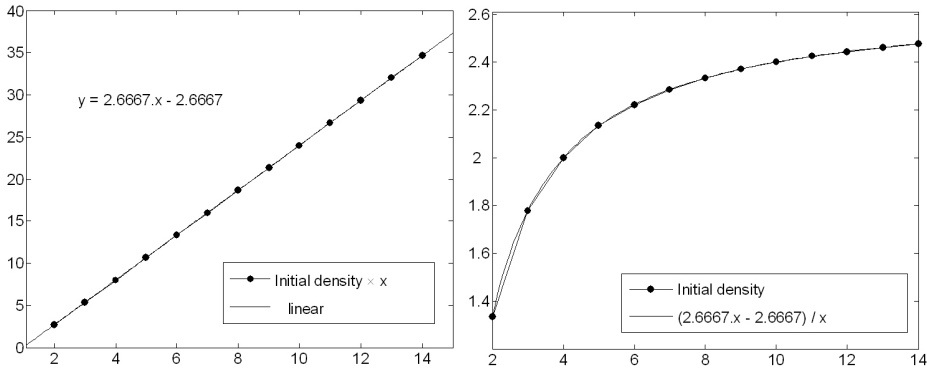


FIGURE 25. Interpolation of densities of  $NF(n)$  in  $F$

multiplied by  $n - 1.1$ , is distributed along line  $y = 3.03x - 6.28$ , Figure 26. The approximation is given by:

## TESTING CAYLEY GRAPH DENSITIES

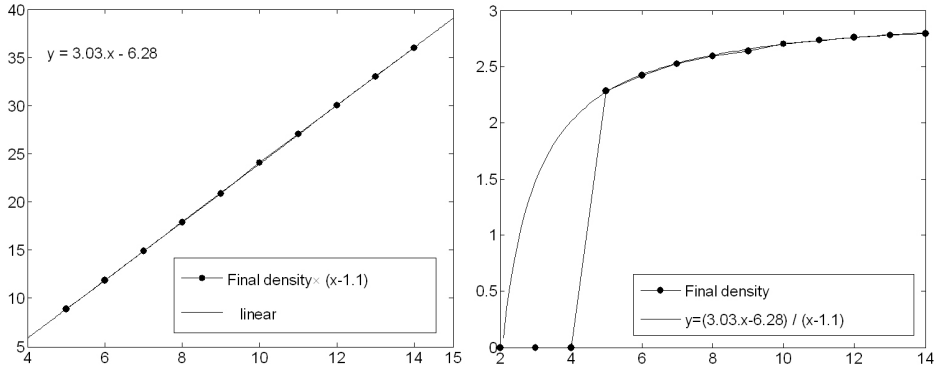


FIGURE 26. Interpolation of densities of  $\overline{NF}(n)$  in  $F$

$$\delta(\overline{NF}(n)) = \frac{3.03n - 6.28}{n - 1.1}$$

and the interpolated limit density is equal to 3.03.

### 8.8. Belk-Brown sets $BB(n, k)$ in $F$

Numerical results are given in §9.7. The densities of the  $BB(n, k)$  are given in Figure 27. The densities of the densified Belk-Brown sets  $\overline{BB}(n, k)$  are given in Figure 28.

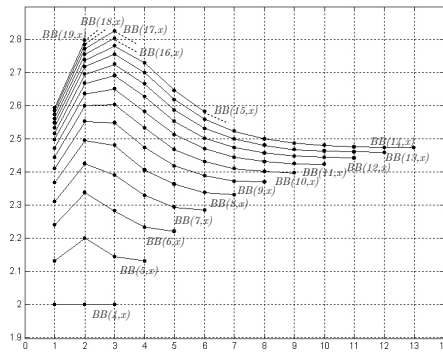


FIGURE 27. Densities of Belk-Brown sets  $BB(n, k)$  in  $F$

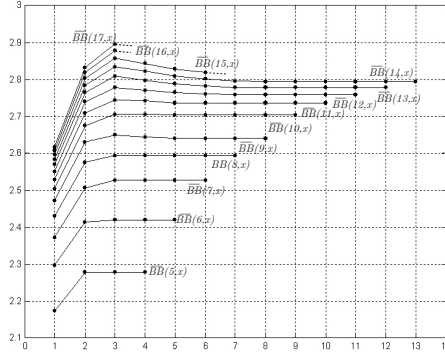


FIGURE 28. Final densities of Belk-Brown sets  $BB(n, k)$  in  $F$

For  $k$  large enough all the densified  $\overline{BB}(n, k)$  have the same density and the same number of vertices. Since  $BB(n, k) \subset BB(n, k + 1)$ , one should expect that the sets obtained from our densification algorithm are the same. The best density we obtain is the density of  $\overline{BB}(17, 3)$ , which is equal to 2.89577.

We construct new sequences as follows. For each  $n$  we consider the maximum of densities of  $BB(n, k)$  as well as the maximum of densities of densified Belk-Brown sets  $\overline{BB}(n, k)$ . The sequences are given in Figure 29.

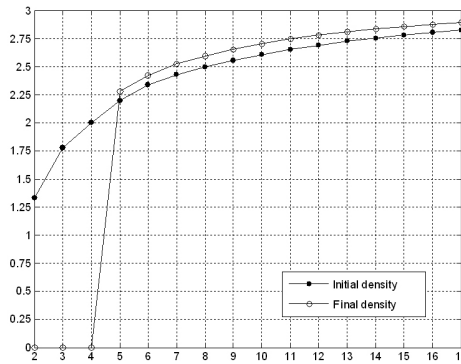


FIGURE 29. Best densities of Belk-Brown sets

For  $n = 3, 4, \dots, 17$  the best initial densities, multiplied with  $n + 1.7$ , is distributed close to  $y = 3.18x - 1.19$ , see Figure 30. The norm of residues is approximatively 0.25. This gives as approximation:

## TESTING CAYLEY GRAPH DENSITIES

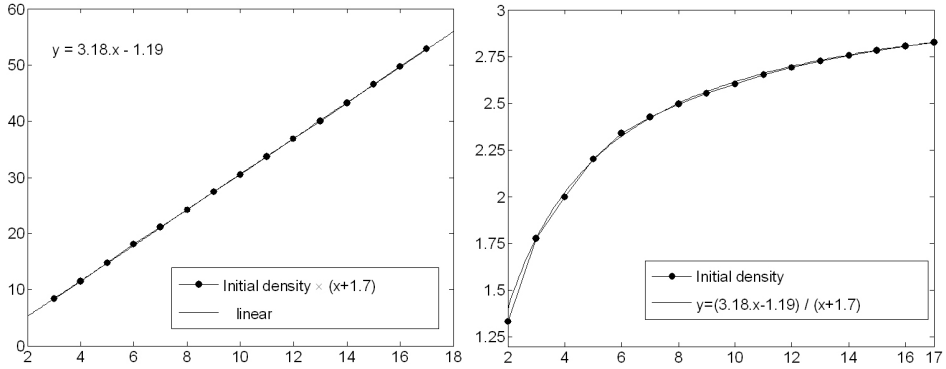


FIGURE 30. Interpolation of the best densities of the  $BB(n, k)$

$$\delta(n) = \frac{3.18x - 1.19}{x + 1.7}.$$

The interpolated limit density is equal to 3.18.

For  $n = 5, \dots, 17$ , the best densities of the densified Belk-Brown sets  $\overline{BB}(n, k)$ , multiplied by  $n + 0.69$ , are close to the line  $y = 3.18x - 2.92$ , see Figure 31. The norm of residues is less than 0.096. This gives as ap-

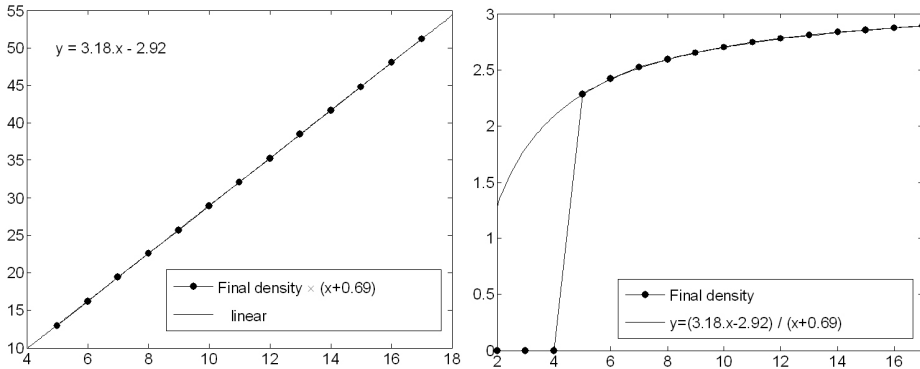


FIGURE 31. Interpolation of best densities of the densified Belk-Brown sets  $\overline{BB}(n, k)$

proximation:

$$\delta(n) = \frac{3.18x - 2.92}{x + 0.69}.$$

Thus, the interpolated limit density for the Belk-Brown sets  $BB(n, k)$  before densification and the one for the densified Belk-Brown sets  $\overline{BB}(n, k)$  agree: they are both approximately equal to 3.18. Observe that the best theoretical limit density of Belk-Brown sets is known to be 3.5 (cf. [1]), more than the computed value. Best densities we have obtained are in case  $k = 3$ . However, a careful analysis of Figure 27 suggests that for large values of  $k$  and  $n$  best densities might be obtained for greater values of  $k$ . The same behavior appears once the densification algorithm is applied.

## 9. Numerical data

In this section we will give some of the numerical data, obtained from our algorithmic package, in the form of tables. Each table corresponds to a fixed group and to a fixed family of subgraphs of the Cayley graph, as presented and discussed in the previous sections. For example the table given in §9.2 corresponds to the Baumslag-Solitar group  $BS(2, 2)$  and to the family of balls  $B(n)$  in its Cayley graph. Each table is organised as follows.

A line in the table corresponds to a fixed choice of parameters for the family of subgraphs, thus specifying a particular subgraph. For example, the 5-th line of the table of §9.2 corresponds to the ball  $B(4)$ , for the group  $BS(2, 2)$ . The first column, labelled “Density”, states the name of the subgraph considered. The second column, labelled “Before”, states the density of this graph before applying the densification algorithm. The third column, labelled “#”, states the number of rounds performed, *i.e.* the number of times the algorithm runs once through the sequence of steps specified at the end of §5 with at least one edge or vertex deleted (in which case it tries to repeat the maneuver, rising the value of # by one). The fourth column, labelled “After”, states the density of the subgraph obtained as final result of our densification program. The fifth column, called “Increase”, explicits the amount of density gained by the densification procedure, as well (in paranthesis) the percentage this increase means with respect to the density before applying the program. Finally, the last column, labelled “Deleted vertices”, states the number of vertices deleted in the densification procedure from the originally given subgraph. In paranthesis it states the percentage this amounts to, with respect to the number of all vertices in the originally given graph.

TESTING CAYLEY GRAPH DENSITIES

**Remark:** If the densification algorithm is applied to a non-empty tree, then the number # of rounds the algorithm repeats the densification procedure will be equal to 1, and the final density must be equal to 0. This, however, is not specific for the case of non-empty trees: it will also happen, for example, if the given subgraph is a cycle, or any other graph obtained from gluing trees to a cycle.

9.1. The group  $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$

Large scale balls $B(n)$ in $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$					
Density	Before	#	After	Increase	Deleted vertices
B(1)	1.71429	1	0	-1.71429(-100.00%)	7/7(-100.00%)
B(11)	5.22325	1	5.3201	+0.0968509(+1.85%)	132/2047(-6.45%)
B(21)	5.5823	1	5.61473	+0.0324311(+0.58%)	252/13287(-1.90%)
B(31)	5.71457	1	5.73058	+0.0160031(+0.28%)	372/41727(-0.89%)
B(41)	5.78326	1	5.79276	+0.0095005(+0.16%)	492/95367(-0.52%)
B(51)	5.82531	1	5.83159	+0.00628328(+0.11%)	612/182207(-0.34%)
B(61)	5.8537	1	5.85816	+0.00446129(+0.08%)	732/310247(-0.24%)
B(71)	5.87415	1	5.87748	+0.00333071(+0.06%)	852/487487(-0.17%)
B(81)	5.88959	1	5.89217	+0.00258064(+0.04%)	972/721927(-0.13%)
B(91)	5.90165	1	5.90371	+0.00205851(+0.03%)	1092/1021567(-0.11%)
B(101)	5.91134	1	5.91302	+0.0016799(+0.03%)	1212/1394407(-0.09%)
B(111)	5.91929	1	5.92069	+0.00139713(+0.02%)	1332/1848447(-0.07%)
B(121)	5.92593	1	5.92711	+0.00118017(+0.02%)	1452/2391687(-0.06%)
B(131)	5.93156	1	5.93257	+0.00100994(+0.02%)	1572/3032127(-0.05%)
B(141)	5.9364	1	5.93727	+0.000874043(+0.01%)	1692/3777767(-0.04%)
B(151)	5.9406	1	5.94136	+0.000763893(+0.01%)	1812/4636607(-0.04%)
B(161)	5.94427	1	5.94495	+0.000673294(+0.01%)	1932/5616647(-0.03%)
B(171)	5.94752	1	5.94812	+0.000597954(+0.01%)	2052/6725887(-0.03%)

Small scale balls $B(n)$ in $\mathbb{Z} \times \mathbb{Z} \times \mathbb{Z}$					
Density	Before	#	After	Increase	Deleted vertices
B(1)	1.71429	1	0	-1.71429(-100.00%)	7/7(-100.00%)
B(2)	2.88	1	3.15789	+0.277895(+9.65%)	6/25(-24.00%)
B(3)	3.61905	1	3.78947	+0.170426(+4.71%)	6/63(-9.52%)
B(4)	4.09302	1	4.44444	+0.351421(+8.59%)	48/129(-37.21%)
B(5)	4.41558	1	4.70175	+0.28617(+6.48%)	60/231(-25.97%)
B(6)	4.64721	1	4.87869	+0.231473(+4.98%)	72/377(-19.10%)
B(7)	4.82087	1	5.01018	+0.189314(+3.93%)	84/575(-14.61%)
B(8)	4.95558	1	5.11262	+0.157037(+3.17%)	96/833(-11.52%)
B(9)	5.06299	1	5.19505	+0.132067(+2.61%)	108/1159(-9.32%)
B(10)	5.15054	1	5.26301	+0.112467(+2.18%)	120/1561(-7.69%)
B(11)	5.22325	1	5.3201	+0.0968509(+1.85%)	132/2047(-6.45%)
B(12)	5.28457	1	5.3688	+0.0842314(+1.59%)	144/2625(-5.49%)
B(13)	5.33697	1	5.41087	+0.0739012(+1.38%)	156/3303(-4.72%)
B(14)	5.38225	1	5.44759	+0.0653448(+1.21%)	168/4089(-4.11%)
B(15)	5.42176	1	5.47994	+0.0581827(+1.07%)	180/4991(-3.61%)

## 9.2. The Baumslag-Solitar group $BS(2, 2)$

Balls $B(n)$ in $BS(2, 2)$					
Density	Before	#	After	Increase	Deleted vertices
B(1)	1.6	1	0	-1.6(-100.00%)	5/5(-100.00%)
B(2)	1.88235	1	0	-1.88235(-100.00%)	17/17(-100.00%)
B(3)	2.21277	1	2.47619	+0.263425(+11.90%)	26/47(-55.32%)
B(4)	2.32479	1	2.64407	+0.319281(+13.73%)	58/117(-49.57%)
B(5)	2.41455	1	2.76423	+0.349682(+14.48%)	152/275(-55.27%)
B(6)	2.4576	1	2.81388	+0.35628(+14.50%)	308/625(-49.28%)
B(7)	2.49029	1	2.84224	+0.351948(+14.13%)	624/1391(-44.86%)
B(8)	2.51032	1	2.85974	+0.349421(+13.92%)	1292/3053(-42.32%)
B(9)	2.526	1	2.87281	+0.346812(+13.73%)	2696/6635(-40.63%)
B(10)	2.53755	1	2.88305	+0.3455(+13.62%)	5668/14313(-39.60%)
B(11)	2.54712	1	2.89163	+0.344501(+13.53%)	11936/30695(-38.89%)
B(12)	2.55501	1	2.89894	+0.343929(+13.46%)	25148/65509(-38.39%)

TESTING CAYLEY GRAPH DENSITIES

B(13)	2.56183	1	2.90531	+0.343485(+13.41%)	52920/139235(-38.01%)
B(14)	2.56776	1	2.91092	+0.343161(+13.36%)	111188/294881(-37.71%)
B(15)	2.57303	1	2.91591	+0.342875(+13.33%)	233168/622559(-37.45%)
B(16)	2.57774	1	2.92037	+0.342623(+13.29%)	488044/1310685(-37.24%)
B(17)	2.58199	1	2.92438	+0.342383(+13.26%)	1019624/2752475(-37.04%)
B(18)	2.58585	1	2.928	+0.342154(+13.23%)	2126468/5767129(-36.87%)

9.3. The wreath product  $\mathbb{Z} \wr \mathbb{Z}$

Balls $B(n)$ in $\mathbb{Z} \wr \mathbb{Z}$					
Density	Before	#	After	Increase	Deleted vertices
B(1)	1.6	1	0	-1.6(-100.00%)	5/5(-100.00%)
B(2)	1.88235	1	0	-1.88235(-100.00%)	17/17(-100.00%)
B(3)	1.96226	1	0	-1.96226(-100.00%)	53/53(-100.00%)
B(4)	2.0915	1	2.4242	+0.332739(+15.91%)	120/153(-78.43%)
B(5)	2.14727	1	2.56881	+0.421539(+19.63%)	312/421(-74.11%)
B(6)	2.19022	1	2.64264	+0.45242(+20.66%)	792/1125(-70.40%)
B(7)	2.2254	2	2.74425	+0.518849(+23.31%)	2198/2937(-74.84%)
B(8)	2.25023	2	2.7895	+0.539267(+23.96%)	5442/7537(-72.20%)
B(9)	2.26973	2	2.82382	+0.554091(+24.41%)	13502/19093(-70.72%)
B(10)	2.285	2	2.84949	+0.564495(+24.70%)	33390/47881(-69.74%)
B(11)	2.29693	2	2.86788	+0.570949(+24.86%)	82190/119133(-68.99%)
B(12)	2.30638	2	2.88159	+0.575211(+24.94%)	201546/294585(-68.42%)
B(13)	2.31387	2	2.89178	+0.577903(+24.98%)	492598/724869(-67.96%)
B(14)	2.31984	2	2.89937	+0.579533(+24.98%)	1200726/1776717(-67.58%)
B(15)	2.32459	2	2.90508	+0.580484(+24.97%)	2920614/4341425(-67.27%)
B(16)	2.32838	2	2.90938	+0.580996(+24.95%)	7092194/10582177(-67.02%)



9.4. Balls in Thompson’s group  $F$

Balls $B(n)$ in $F$					
Density	Before	#	After	Increase	Deleted vertices
B(1)	1.6	1	0	-1.6(-100.00%)	5/5(-100.00%)
B(2)	1.88235	1	0	-1.88235(-100.00%)	17/17(-100.00%)
B(3)	1.96226	1	0	-1.96226(-100.00%)	53/53(-100.00%)
B(4)	1.98758	1	0	-1.98758(-100.00%)	161/161(-100.00%)
B(5)	2.03789	1	2.31579	+0.277895(+13.64%)	418/475(-88.00%)
B(6)	2.05069	1	2.39106	+0.340374(+16.60%)	1202/1381(-87.04%)
B(7)	2.07632	2	2.48073	+0.404413(+19.48%)	3412/3957(-86.23%)
B(8)	2.08597	2	2.53365	+0.44768(+21.46%)	9959/11237(-88.63%)
B(9)	2.10377	2	2.57062	+0.46685(+22.19%)	26994/31589(-85.45%)
B(10)	2.11048	2	2.59635	+0.485875(+23.02%)	75036/88253(-85.02%)
B(11)	2.12304	3	2.61761	+0.494578(+23.30%)	203765/244823(-83.23%)
B(12)	2.12823	3	2.63324	+0.505008(+23.73%)	558984/676061(-82.68%)
B(13)	2.13765	3	2.64741	+0.50976(+23.85%)	1512760/1857029(-81.46%)
B(14)	2.14177	3	2.65825	+0.51648(+24.11%)	4120532/5082969(-81.07%)
B(15)	2.14905	132	2.7183	+0.569249(+26.49%)	12420620/13856005(-89.64%)

9.5. Left-positive balls in Thompson’s group  $F$

Left-positive balls $LP(n)$ in $F$					
Density	Before	#	After	Increase	Deleted vertices
LP(1)	1.33333	1	0	-1.33333(-100.00%)	3/3(-100.00%)
LP(2)	1.71429	1	0	-1.71429(-100.00%)	7/7(-100.00%)
LP(3)	1.875	1	0	-1.875(-100.00%)	16/16(-100.00%)
LP(4)	1.94444	1	0	-1.94444(-100.00%)	36/36(-100.00%)
LP(5)	2	1	0	-2(-100.00%)	81/81(-100.00%)
LP(6)	2.03297	1	2.19355	+0.160581(+7.90%)	151/182(-82.97%)
LP(7)	2.05868	1	2.30189	+0.243207(+11.81%)	356/409(-87.04%)
LP(8)	2.07835	1	2.38571	+0.307368(+14.79%)	779/919(-84.77%)
LP(9)	2.09395	3	2.45662	+0.362674(+17.32%)	1846/2065(-89.39%)
LP(10)	2.10647	12	2.52926	+0.422796(+20.07%)	4247/4640(-91.53%)

TESTING CAYLEY GRAPH DENSITIES

LP(11)	2.11682	21	2.57116	+0.454332(+21.46%)	9379/10426(-89.96%)
LP(12)	2.1255	38	2.60121	+0.475717(+22.38%)	20789/23427(-88.74%)
LP(13)	2.13283	71	2.62354	+0.490717(+23.01%)	46116/52640(-87.61%)
LP(14)	2.13908	136	2.6407	+0.501627(+23.45%)	102464/118281(-86.63%)
LP(15)	2.14445	265	2.65425	+0.509793(+23.77%)	227988/265775(-85.78%)
LP(16)	2.14911	522	2.66511	+0.515999(+24.01%)	507883/597191(-85.05%)
LP(17)	2.15318	914	2.72507	+0.571886(+26.56%)	1237673/1341876(-92.23%)
LP(18)	2.15674	793	2.73535	+0.578606(+26.83%)	2761427/3015168(-91.58%)
LP(19)	2.15988	1029	2.74349	+0.583608(+27.02%)	6164348/6775021(-90.99%)

9.6. Negative forests in Thompson’s group  $F$

Note that the following results, for negative forest  $NF(n)$ , correspond precisely to the results for Belk-Brown sets  $B(n, n - 1)$ , see below:

Negative forests $NF(n)$ in $F$					
Density	Before	#	After	Increase	Deleted vertices
NF(2)	1.33333	1	0	-1.33333(-100.00%)	3/3(-100.00%)
NF(3)	1.77778	1	0	-1.77778(-100.00%)	9/9(-100.00%)
NF(4)	2	1	0	-2(-100.00%)	28/28(-100.00%)
NF(5)	2.13333	1	2.27907	+0.145736(+6.83%)	47/90(-52.22%)
NF(6)	2.22222	2	2.42105	+0.19883(+8.95%)	183/297(-61.62%)
NF(7)	2.28571	3	2.52717	+0.24146(+10.56%)	633/1001(-63.24%)
NF(8)	2.33333	2	2.59580	+0.262469(+11.25%)	1955/3432(-56.96%)
NF(9)	2.37037	2	2.64099	+0.270623(+11.42%)	6299/11934(-52.78%)
NF(10)	2.4	6	2.70468	+0.304678(+12.69%)	28117/41990(-66.96%)
NF(11)	2.42424	5	2.73583	+0.311589(+12.85%)	94931/149226(-63.62%)
NF(12)	2.44444	5	2.75949	+0.315048(+12.89%)	326375/534888(-61.02%)
NF(13)	2.46154	5	2.77912	+0.317579(+12.90%)	1142627/1931540(-59.16%)
NF(14)	2.47619	5	2.79448	+0.31829(+12.85%)	4031727/7020405(-57.43%)

9.7. Belk-Brown sets in Thompson's group  $F$

Belk-Brown sets $BB(n, k)$ in $F$					
Density	Before	#	After	Increase	Deleted vertices
BB(2,1)	1.33333	1	0	-1.33333(-100.00%)	3/3(-100.00%)
BB(3,1)	1.71429	1	0	-1.71429(-100.00%)	7/7(-100.00%)
BB(3,2)	1.77778	1	0	-1.77778(-100.00%)	9/9(-100.00%)
BB(4,1)	2	1	0	-2(-100.00%)	15/15(-100.00%)
BB(4,2)	2	1	0	-2(-100.00%)	24/24(-100.00%)
BB(4,3)	2	1	0	-2(-100.00%)	28/28(-100.00%)
BB(5,1)	2.13333	1	2.17391	+0.0405796(+1.90%)	7/30(-23.33%)
BB(5,2)	2.2	1	2.27907	+0.0790696(+3.59%)	17/60(-28.33%)
BB(5,3)	2.14634	1	2.27907	+0.132728(+6.18%)	39/82(-47.56%)
BB(5,4)	2.13333	1	2.27907	+0.145736(+6.83%)	47/90(-52.22%)
BB(6,1)	2.24138	1	2.29787	+0.056493(+2.52%)	11/58(-18.97%)
BB(6,2)	2.34014	1	2.41509	+0.0749583(+3.20%)	41/147(-27.89%)
BB(6,3)	2.28326	2	2.42105	+0.137791(+6.03%)	119/233(-51.07%)
BB(6,4)	2.23488	2	2.42105	+0.186177(+8.33%)	167/281(-59.43%)
BB(6,5)	2.22222	2	2.42105	+0.19883(+8.95%)	183/297(-61.62%)
BB(7,1)	2.31193	1	2.37363	+0.0616999(+2.67%)	18/109(-16.51%)
BB(7,2)	2.42735	1	2.50714	+0.0797923(+3.29%)	71/351(-20.23%)
BB(7,3)	2.39258	2	2.52717	+0.134593(+5.63%)	279/647(-43.12%)
BB(7,4)	2.33064	3	2.52717	+0.196538(+8.43%)	497/865(-57.46%)
BB(7,5)	2.29515	3	2.52717	+0.232024(+10.11%)	601/969(-62.02%)
BB(7,6)	2.28571	3	2.52717	+0.24146(+10.56%)	633/1001(-63.24%)
BB(8,1)	2.36816	1	2.43023	+0.0620732(+2.62%)	29/201(-14.43%)
BB(8,2)	2.49695	1	2.57676	+0.0798025(+3.20%)	124/821(-15.10%)
BB(8,3)	2.48144	2	2.59581	+0.114371(+4.61%)	489/1778(-27.50%)
BB(8,4)	2.40847	2	2.5958	+0.18733(+7.78%)	1167/2644(-44.14%)
BB(8,5)	2.36387	2	2.5958	+0.231935(+9.81%)	1667/3144(-53.02%)
BB(8,6)	2.33967	2	2.5958	+0.256135(+10.95%)	1891/3368(-56.15%)
BB(8,7)	2.33333	2	2.5958	+0.262469(+11.25%)	1955/3432(-56.96%)
BB(9,1)	2.41096	1	2.4717	+0.060739(+2.52%)	47/365(-12.88%)
BB(9,2)	2.55485	1	2.63139	+0.0765369(+3.00%)	233/1896(-12.29%)

TESTING CAYLEY GRAPH DENSITIES

BB(9,3)	2.5493	2	2.65126	+0.101964(+4.00%)	977/4828(-20.24%)
BB(9,4)	2.47602	2	2.64423	+0.168205(+6.79%)	2749/8008(-34.33%)
BB(9,5)	2.4212	2	2.64099	+0.219797(+9.08%)	4555/10190(-44.70%)
BB(9,6)	2.39025	2	2.64099	+0.250741(+10.49%)	5691/11326(-50.25%)
BB(9,7)	2.37439	2	2.64099	+0.266608(+11.23%)	6171/11806(-52.27%)
BB(9,8)	2.37037	2	2.64099	+0.270623(+11.42%)	6299/11934(-52.78%)
BB(10,1)	2.4458	1	2.50432	+0.0585163(+2.39%)	76/655(-11.60%)
BB(10,2)	2.60032	3	2.67492	+0.0745924(+2.87%)	775/4331(-17.89%)
BB(10,3)	2.60551	4	2.70623	+0.100716(+3.87%)	3422/12994(-26.34%)
BB(10,4)	2.5348	5	2.70607	+0.171267(+6.76%)	10956/24136(-45.39%)
BB(10,5)	2.47021	6	2.70483	+0.234619(+9.50%)	19145/32998(-58.02%)
BB(10,6)	2.43355	6	2.70468	+0.271126(+11.14%)	24277/38150(-63.64%)
BB(10,7)	2.41258	6	2.70468	+0.292101(+12.11%)	26837/40710(-65.92%)
BB(10,8)	2.40245	6	2.70468	+0.302224(+12.58%)	27861/41734(-66.76%)
BB(10,9)	2.4	6	2.70468	+0.304678(+12.69%)	28117/41990(-66.96%)
BB(11,1)	2.47423	1	2.53026	+0.0560327(+2.26%)	123/1164(-10.57%)
BB(11,2)	2.63755	3	2.71078	+0.0732303(+2.78%)	1443/9800(-14.72%)
BB(11,3)	2.65202	4	2.74514	+0.0931263(+3.51%)	6978/34680(-20.12%)
BB(11,4)	2.58441	4	2.74205	+0.157632(+6.10%)	26855/72394(-37.10%)
BB(11,5)	2.51403	5	2.73747	+0.223436(+8.89%)	53749/106600(-50.42%)
BB(11,6)	2.47079	5	2.7358	+0.26501(+10.73%)	74443/128762(-57.81%)
BB(11,7)	2.44596	5	2.73583	+0.289869(+11.85%)	86515/140810(-61.44%)
BB(11,8)	2.43202	5	2.73583	+0.303807(+12.49%)	92243/146538(-62.95%)
BB(11,9)	2.4257	5	2.73583	+0.310129(+12.79%)	94419/148714(-63.49%)
BB(11,10)	2.42424	5	2.73583	+0.311589(+12.85%)	94931/149226(-63.62%)
BB(12,1)	2.49805	1	2.55154	+0.0534873(+2.14%)	199/2052(-9.70%)
BB(12,2)	2.66921	3	2.7401	+0.0708911(+2.66%)	2862/22008(-13.00%)
BB(12,3)	2.69209	4	2.779	+0.0869138(+3.23%)	14982/91965(-16.29%)
BB(12,4)	2.62814	4	2.77148	+0.143337(+5.45%)	66315/216154(-30.68%)
BB(12,5)	2.5533	5	2.7649	+0.211606(+8.29%)	153498/343946(-44.63%)
BB(12,6)	2.50352	5	2.76087	+0.257348(+10.28%)	229951/435268(-52.83%)
BB(12,7)	2.47504	5	2.76001	+0.284971(+11.51%)	280855/488584(-57.48%)
BB(12,8)	2.45827	5	2.76036	+0.302098(+12.29%)	309351/516520(-59.89%)

G. N. ARZHANTSEVA ET AL.

BB(12,9)	2.44917	5	2.75949	+0.310318(+12.67%)	320743/529256(-60.60%)
BB(12,10)	2.4453	5	2.75949	+0.314195(+12.85%)	325351/533864(-60.94%)
BB(12,11)	2.44444	5	2.75949	+0.315048(+12.89%)	326375/534888(-61.02%)
BB(13,1)	2.51823	1	2.56924	+0.0510149(+2.03%)	322/3593(-8.96%)
BB(13,2)	2.69615	2	2.76459	+0.0684388(+2.54%)	5845/49110(-11.90%)
BB(13,3)	2.7264	4	2.80896	+0.0825596(+3.03%)	35904/242478(-14.81%)
BB(13,4)	2.66664	4	2.79765	+0.131004(+4.91%)	166062/643068(-25.82%)
BB(13,5)	2.58797	5	2.78796	+0.199994(+7.73%)	442137/1108550(-39.88%)
BB(13,6)	2.53309	5	2.78229	+0.249196(+9.84%)	717635/1472390(-48.74%)
BB(13,7)	2.50062	5	2.77942	+0.2788(+11.15%)	914731/1700220(-53.80%)
BB(13,8)	2.48142	5	2.7789	+0.297486(+11.99%)	1037283/1827316(-56.77%)
BB(13,9)	2.4702	5	2.77912	+0.308918(+12.51%)	1102691/1891604(-58.29%)
BB(13,10)	2.46437	5	2.77912	+0.314748(+12.77%)	1130851/1919764(-58.91%)
BB(13,11)	2.46203	5	2.77912	+0.317089(+12.88%)	1140579/1929492(-59.11%)
BB(13,12)	2.46154	5	2.77912	+0.317579(+12.90%)	1142627/1931540(-59.16%)
BB(14,1)	2.53557	1	2.58423	+0.0486629(+1.92%)	521/6255(-8.33%)
BB(14,2)	2.71929	2	2.78529	+0.0659952(+2.43%)	12052/108982(-11.06%)
BB(14,3)	2.75635	4	2.83483	+0.0784769(+2.85%)	80539/636264(-12.66%)
BB(14,4)	2.70065	6	2.82311	+0.122467(+4.53%)	467063/1906645(-24.50%)
BB(14,5)	2.61918	8	2.80987	+0.190696(+7.28%)	1368089/3569029(-38.33%)
BB(14,6)	2.55993	10	2.80225	+0.242317(+9.47%)	2361586/4984631(-47.38%)
BB(14,7)	2.52346	5	2.79651	+0.273047(+10.82%)	3012087/5931157(-50.78%)
BB(14,8)	2.50193	5	2.79464	+0.292711(+11.70%)	3504047/6486437(-54.02%)
BB(14,9)	2.48895	5	2.79436	+0.305406(+12.27%)	3796015/6786933(-55.93%)
BB(14,10)	2.48154	5	2.79448	+0.312939(+12.61%)	3945199/6933877(-56.90%)
BB(14,11)	2.47786	5	2.79448	+0.316618(+12.78%)	4007151/6995829(-57.28%)
BB(14,12)	2.47647	5	2.79448	+0.318012(+12.84%)	4027631/7016309(-57.40%)
BB(14,13)	2.47619	5	2.79448	+0.31829(+12.85%)	4031727/7020405(-57.43%)
BB(15,1)	2.55062	1	2.59708	+0.0464547(+1.82%)	843/10835(-7.78%)
BB(15,2)	2.73947	1	2.80304	+0.0635667(+2.32%)	24893/240693(-10.34%)
BB(15,3)	2.78255	4	2.85759	+0.0750453(+2.70%)	186351/1662399(-11.21%)
BB(15,4)	2.73117	6	2.84478	+0.113609(+4.16%)	1193477/5636091(-21.18%)
BB(15,5)	2.64757	7	2.82836	+0.180787(+6.83%)	3994150/11478205(-34.80%)

## TESTING CAYLEY GRAPH DENSITIES

BB(15,6)	2.58416	8	2.81871	+0.234552(+9.08%)	7483649/16887924(-44.31%)
BB(15,7)	2.5442				20730535
BB(16,1)	2.56381	1	2.60821	+0.0443943(+1.73%)	1364/18687(-7.30%)
BB(16,2)	2.75722	1	2.81842	+0.0612032(+2.22%)	51425/529373(-9.71%)
BB(16,3)	2.80576	4	2.87778	+0.072022(+2.57%)	440405/4327228(-10.18%)
BB(17,1)	2.57547	1	2.61795	+0.0424783(+1.65%)	2207/32106(-6.87%)
BB(17,2)	2.77292	1	2.83186	+0.0589356(+2.13%)	106246/1160005(-9.16%)
BB(17,3)	2.82642	4	2.89577	+0.069357 (+2.45%)	1056462/11226140(-9.41%)
BB(18,1)	2.58584	1	2.62654	+0.0406988(+1.57%)	3571/54974(-6.50%)
BB(18,2)	2.78693	1	2.8437	+0.0567749(+2.04%)	219506/2533584(-8.66%)
BB(19,1)	2.59513	1	2.63418	+0.039046 (+1.50%)	5778/93845(-6.16%)
BB(19,2)	2.7995	1	2.85423	+0.054727 (+1.95%)	453495/5517456(-8.22%)
BB(20,1)	2.6035	1	2.64101	+0.037510 (+1.44%)	9349/159765(-5.85%)
BB(20,2)	2.81085	1	2.86364	+0.052791 (+1.88%)	936918/11983889(-7.82%)
BB(21,1)	2.61108	1	2.64716	+0.036081 (+1.38%)	15127/271321 (-5.58%)
BB(22,1)	2.61797	1	2.65272	+0.034750 (+1.33%)	24476/459743(-5.32%)
BB(23,1)	2.62427	1	2.65778	+0.033508(+1.28%)	39603/777432(-5.09%)
BB(24,1)	2.63005	1	2.6624	+0.032347(+1.23%)	64079/1312200(-4.88%)
BB(25,1)	2.63537	1	2.66663	+0.0312603(+1.19%)	103682/2211025(-4.69%)
BB(26,1)	2.64028	2	2.67104	+0.0307548(+1.16%)	585073/3719643(-15.73%)
BB(27,1)	2.64483	2	2.67506	+0.0302331(+1.14%)	946668/6248479(-15.15%)
BB(28,1)	2.64906	2	2.67876	+0.0296998(+1.12%)	1531741/10482351(-14.61%)

We now consider Belk-Brown sets  $B(n, k)$  with best density, for fixed  $n$  and any  $k$ , before (or after, in the subsequent table) applying the densification algorithm.

<b>Best density Belk-Brown sets (before densification)</b>					
Density	Before	#	After	Increase	Deleted vertices
BB(2,1)	1.33333	1	0	-1.33333(-100.00%)	3/3(-100.00%)
BB(3,2)	1.77778	1	0	-1.77778(-100.00%)	9/9(-100.00%)
BB(4,2)	2	1	0	-2(-100.00%)	24/24(-100.00%)
BB(5,2)	2.2	1	2.27907	+0.0790696(+3.59%)	17/60(-28.33%)
BB(6,2)	2.34014	1	2.41509	+0.0749583(+3.20%)	41/147(-27.89%)
BB(7,2)	2.42735	1	2.50714	+0.0797923(+3.29%)	71/351(-20.23%)
BB(8,2)	2.49695	1	2.57676	+0.0798025(+3.20%)	124/821(-15.10%)
BB(9,2)	2.55485	1	2.63139	+0.0765369(+3.00%)	233/1896(-12.29%)
BB(10,3)	2.60551	4	2.70623	+0.100716(+3.87%)	3422/12994(-26.34%)
BB(11,3)	2.65202	4	2.74514	+0.0931263(+3.51%)	6978/34680(-20.12%)
BB(12,3)	2.69209	4	2.779	+0.0869138(+3.23%)	14982/91965(-16.29%)
BB(13,3)	2.7264	4	2.80896	+0.0825596(+3.03%)	35904/242478(-14.81%)
BB(14,3)	2.75635	4	2.83483	+0.0784769(+2.85%)	80539/636264(-12.66%)
BB(15,3)	2.78255	4	2.85759	+0.0750453(+2.70%)	186351/1662399(-11.21%)
BB(16,3)	2.80576	4	2.87778	+0.072022(+2.57%)	440405/4327228(-10.18%)
BB(17,3)	2.82642	4	2.89577	+0.069357 (+2.45%)	1056462/11226140(-9.41%)

<b>Best density Belk-Brown sets (after densification)</b>					
Density	Before	#	After	Increase	Deleted vertices
BB(2,1)	1.33333	1	0	-1.33333(-100.00%)	3/3(-100.00%)
BB(3,2)	1.77778	1	0	-1.77778(-100.00%)	9/9(-100.00%)
BB(4,2)	2	1	0	-2(-100.00%)	24/24(-100.00%)
BB(5,2)	2.2	1	2.27907	+0.0790696(+3.59%)	17/60(-28.33%)
BB(6,3)	2.28326	2	2.42105	+0.137791(+6.03%)	119/233(-51.07%)
BB(7,3)	2.39258	2	2.52717	+0.134593(+5.63%)	279/647(-43.12%)
BB(8,3)	2.48144	2	2.59581	+0.114371(+4.61%)	489/1778(-27.50%)
BB(9,3)	2.5493	2	2.65126	+0.101964(+4.00%)	977/4828(-20.24%)
BB(10,3)	2.60551	4	2.70623	+0.100716(+3.87%)	3422/12994(-26.34%)

## TESTING CAYLEY GRAPH DENSITIES

BB(11,3)	2.65202	4	2.74514	+0.0931263(+3.51%)	6978/34680(-20.12%)
BB(12,3)	2.69209	4	2.779	+0.0869138(+3.23%)	14982/91965(-16.29%)
BB(13,3)	2.7264	4	2.80896	+0.0825596(+3.03%)	35904/242478(-14.81%)
BB(14,3)	2.75635	4	2.83483	+0.0784769(+2.85%)	80539/636264(-12.66%)
BB(15,3)	2.78255	4	2.85759	+0.0750453(+2.70%)	186351/1662399(-11.21%)
BB(16,3)	2.80576	4	2.87778	+0.072022(+2.57%)	440405/4327228(-10.18%)
BB(17,3)	2.82642	4	2.89577	+0.069357 (+2.45%)	1056462/11226140(-9.41%)

## 10. Outlook

Throughout this section we use essentially the same notation as introduced in the beginning of the paper, and as used in the description of our densification algorithm (§5):  $A$  always denotes a (not necessarily connected) finite graph, which we think of as “virtually” embedded as subgraph into an infinite ambient graph  $\mathcal{C}$ . The latter is usually the Cayley graph  $\mathcal{C} = \mathcal{C}(G, X)$  of a finitely generated group  $G$  with respect to a generating system of finite cardinality  $m \in \mathbb{N}$ , on which  $G$  acts on the left. In any case we always assume that  $\mathcal{C}$  has a uniform bound  $2m$  for the degree of any of its vertices. Note that the ambient graph is “virtual” in that, contrary to the subgraph  $A$ , it exists only as theoretical construct, and hence any finite piece  $A^*$  of it (typically with  $A \subset A^*$ ) has to be algorithmically constructed before it can be used in the algorithm.

For the purposes of this section it is easier to work with non-oriented edges. Thus every edge in this section corresponds to a pair of inversely oriented edges with same endpoints, in the notation of the earlier sections. Below we denote by  $V(A)$  the set of vertices of  $A$  and by  $E(A)$  the set of edges. By  $v(A)$  and  $e(A)$  we denote the cardinality of  $V(A)$  and  $E(A)$  respectively. We call  $e(A)$  the *volume* of  $A$ . The Euler characteristic of  $A$  is given by  $\chi(A) = v(A) - e(A)$ . For any vertex  $x \in V(A)$  the degree  $\deg_A(x)$  is the number of edge segments adjacent to  $x$ , which is consistent with the use of  $\deg(A)$  in the earlier sections in view of the above transition from oriented to non-oriented edges. A vertex  $x \in V(A)$  is called *branch point* if it has degree  $\deg(x) \geq 3$ .

For any subset  $X \subset A$  we denote by  $cl(X)$  the smallest subgraph of  $A$  that contains  $X$ . For any subgraph  $K$  of  $A$  we define the  $A$ -*boundary*  $\partial_A K$  of  $K$  to be the 0-dimensional subgraph of  $K$  which consists of all vertices



that bound simultaneously an edge from  $A$  and an edge from  $cl(A - K)$ :

$$\partial_A K = A \cap cl(A - K).$$

Also,  $int_A K = K - \partial_A K$  denotes the  $A$ -interior of  $K$ , which is in general not a subgraph. A subgraph  $K$  of  $A$  is called *full*, if it contains all edges of  $A$  that have both endpoints in  $K$ . The density of  $A$  is given by:

$$\delta(A) = \sum_{x \in V(A)} \frac{\deg_A(x)}{v(A)} = 2 \frac{e(A)}{v(A)}.$$

We extend this notion in the obvious way to "graphs with some vertices missing" like the above set  $int_A(K)$ , for which one has  $\delta(int_A(K)) = 2 \frac{e(K)}{v(K) - v(\partial_A K)}$ .

The comments and improvements proposed below concern the following three aspects: (A) the algorithmic determination of subgraphs of  $A$  with higher density, (B) the deterministic construction of larger graphs  $A^* \subset C$  which contain  $A$  and have higher density, and (C) the non-deterministic construction of such  $A^*$ .

## A. Subgraphs with higher density

We observe that the improvements on the density by passing over to a subgraph  $\bar{A}$  of  $A$ , as performed by the subroutines (R1) - (R4) of our algorithm presented in §5, are all based on the following principle: the computer checks for the existence of subgraphs  $K$  of  $A$  of a certain (fairly simple) type, and, if it finds any of them, it replaces  $A$  by  $\bar{A} = A - int_A K$ . The type of subgraphs  $K$  in question assures that the density increases strictly in this process. This is ensured by the topology of  $K$ , which needs to be of low density itself, and with small  $A$ -boundary. More precisely, one has:

*Remark 10.1.* For any subgraph  $K \subset A$  the complementary subgraph  $\bar{A} = A - int_A K$  satisfies  $\delta(\bar{A}) > \delta(A)$  if and only if one has  $\delta(A) > \delta(int_A K)$ . Such a subgraph  $K$  is called *density increasing*.

Below we propose four further methods how to effectively find a density increasing subgraph  $K$ , in any given finite graph  $A$ .

(1) A first improvement of the algorithm used in our work can simply be obtained by embellishing the list of density increasing subgraphs  $K$ ,

which are integrated as fixed part of the algorithm without ever changing in the process. This is done by adding to the present list (*i.e.* trees, cycles, long chains, long tripods and long degenerated tripods, see §5) further subgraphs with low densities and small boundary. For example, any connected subgraph of  $A$  which is of class  $\mathcal{K}(k, l, n)$ , defined as set of all graphs  $K$  with  $v(K) = k$ ,  $|\partial_A K| \leq l$  and  $\chi(K) \geq n$ , is density increasing if  $\frac{k-n}{k-l} < \delta(A)$ , by Remark 10.1.

(2) We devise a new subroutine, where the computer searches for the set  $K_0$  of all vertices  $x \in V(A)$  with  $\deg_A(x) < \delta(A)$  (or  $\deg_A(x) < \delta(A) - C$  for some constant  $C > 0$ ), and assembles them into “clusters”, *i.e.* it builds iteratively full subgraphs  $K_i$  which have a high percentage of low-valence vertices. The subgraphs  $K_i$  are defined iteratively out of the connected components  $K_{i-1}^j$  of  $K_{i-1}$  by adding vertices and edges from their neighborhood in order to create larger connected components, with the goal to decrease the total  $A$ -boundary of the union of the  $K_{i-1}^j$ .

At any given state  $K_i$  the computer checks the cardinality of  $\partial_A K_i$ , and stops the subroutine if this check shows that  $K_i$  is density decreasing.

(3) A promising method to find interesting candidates for density decreasing subgraphs  $K \subset A$  seems to be the following: We consider a symmetric random walk on  $A$  where the starting measure on each vertex  $x$  is given by  $\mu_0(x) = 2m - \deg_A(x)$ . We then let the random walk proceed for some integer time  $t$ , thus distributing the measure to give a value of  $\mu_t(x)$  on any  $x \in V(A)$  via the formula  $\mu_t(x) = \sum \frac{\mu_{t-1}(y)}{\deg_A(y)}$ , where the sum is taken over all vertices  $y$  adjacent to  $x$  in  $A$ . For any  $h \geq 0$  we define the vertex sets  $V(h, t) = \{x \in V(A) \mid \mu_t(x) \geq h\}$ , and the subgraphs  $K(h, t)$  as the full subgraphs of  $A$  with vertex set  $V(h, t)$ . For any integer time  $t \geq 0$ , if we let  $h$  decrease monotonously from  $\max\{\mu_t(x) \mid x \in V(A)\}$  to 0, the family  $K(h, t)$  defines a (finite) increasing nested sequence of subgraphs of  $A$ , which we propose as candidates for density decreasing subgraphs.

A variation of this approach would be to iterate the random walk until  $t$  is large enough so that the measure  $\mu_t(x)$  approximates a stable equilibrium  $\mu_\infty(x)$ , for all vertices  $x \in V(A)$ . But this seems less interesting, as there is only one such limit distribution, and that is precisely given by  $\frac{1}{2m}$  times the density function.

Another, perhaps more promising variation comes from adding exterior measure sources or measure sinks, for example sinks for the high-density

vertices, or sources for the low-density vertices, to force an equilibrium state to assemble the measure in the neighborhood of certain subgraphs considered as possible candidates for density decreasing subgraphs. (Recall that low-density does not imply density increasing, as one also needs that the  $A$ -boundary of  $K$  is small.)

(4) More generally, improving the density of  $A$  by erasing interiors of subgraphs can also be viewed as improving the quality of  $A$  as an expander: We look for a “small” set  $Z$  of vertices (corresponding to  $\partial_A K$  in the above approaches) which cuts  $A$  into subgraphs  $K$  and  $\bar{A}$  (with union  $A$  and intersection  $Z$ ) that have rather different densities  $\delta(\bar{A}) > \delta(K)$ . If the difference of these densities is large with respect to the cardinality of  $Z$ , for example if  $\delta(\bar{A}) - \delta(K) > \frac{2m\#Z}{v(K) - \#Z}$ , then  $\delta(\bar{A})$  will be strictly bigger than  $\delta(A)$ .

There are also some interesting theoretical questions surrounding the algorithmic attempts to improve the density by erasing subgraphs:

(i) Is there an algorithm to find the (possibly non-uniquely determined) subgraph  $A_{\max}$  of highest density among all subgraphs of  $A$  ?

As  $A$  is finite, the answer is of course “yes”, but trying out all subgraphs is unfortunately not feasible in practise. Hence we rephrase the question as:

(ii) What is the minimal complexity of any algorithm that derives  $A_{\max}$  from a given finite graph  $A$ . In particular, is there a polynomial-time algorithm ?

(iii) Is there an algorithm for finding  $A_{\max}$  that uses only finitely many types of steps to pass from one intermediate subgraph  $A_i$  to  $A_{i+1}$  ? Here a “step” consists of modifying a subgraph of  $A_i$  of a given graph type into a new graph of given type.

(iv) Is there always a sequence of nested subgraphs  $A_i$  of increasing density and uniformly bounded volume difference  $e(A_i) - e(A_{i+1})$  connecting  $A$  to  $A_{\max}$  ? What is the minimal value for the volume difference bound needed to answer this question in the positive, in terms of the universal vertex degree bound  $2m$  ?

An important fact the reader should note is the observation that certain “wrong” initial improvements on  $A$  (by erasing the interior of some density increasing subgraph  $K$ ) can prevent the algorithm used in this paper,

as well as any of the above proposed improvements (1) - (4), from ever finding any of the really desired subgraphs  $\bar{A} \subset A$  with density  $\delta(\bar{A})$  close to  $\delta(A_{\max})$ . Indeed, it is not hard to find examples of graphs (for example built on two disjoint graphs connected by adding a long chain) which answer the following question in the negative:

Is any subgraph  $\bar{A}$  of  $A$  with  $\delta(\bar{A}) \geq \delta(A)$ , such that  $\bar{A}$  does not contain a subgraph of strictly larger density than  $\delta(\bar{A})$ , necessarily equal to  $A_{\max}$  ?

**B. Deterministic methods do increase  $A$  to a larger graph  $A^*$  with higher density**

We first notice that for any finite subgraph  $A$  of the Cayley graph  $\mathcal{C}$  of  $G$ , and for any  $g \in G$  with sufficiently large translation length in  $\mathcal{C}$ , the subgraph  $gA$  of  $\mathcal{C}$  is a disjoint isomorphic copy of  $A$ , and hence their union has the same density as  $A$ . On the other hand, if one finds an element  $g \in G$  such that  $A$  and  $gA$  intersect in a single vertex, then the density of the union  $A \cup gA$  is strictly larger than that of  $A$ . Of course, as the special case of a subtrees  $A$  shows, there are rather strict limits to this method in its crude form, but nevertheless it gives the right idea why the following is promising.

Since  $A$  is finite, the subset  $G_A \subset G$  defined by:

$$G_A = \{g \in G \mid A \cap gA \text{ non-empty}\}$$

is also finite, so that at least in principle one can calculate, for all subsets  $B \subset G_A$ , the density of the union:

$$A^B = \cup\{gA \mid g \in B\}.$$

It seems quite realistic that among the  $A^B$  one finds new graphs with substantially higher density than  $A$ , and that an iteration of this procedure leads to a very promising family of density test graphs for  $G$ .

On the other hand, the calculation of the unions  $A^B$  is tedious and requires much computing time. Hence the following suggestion may prove to be helpful:

Denote  $gA$  by  $A'$ , and let  $K = A \cap A'$  be the intersection subgraph. Then  $\delta(A \cup A')$  is calculated by the formula:

$$\delta(A \cup A') = \frac{4e(A) - 2e(K)}{2v(A) - v(K)}$$

and hence  $\delta(A \cup A') > \delta(A)$  if and only if  $\delta(A) > \delta(K)$ .

We may thus start out with a large  $g = g_q = x_q \dots x_1 \in G$ , so that  $gA$  is disjoint from  $A$  and then pass successively to  $g_{q-1}A$ , to  $g_{q-2}A$ , etc, for  $g_k = x_k \dots x_1$ , until  $A$  and  $g_kA$  meet. As small graphs have (a fortiori) small density, the first non-empty intersection graphs  $K = A \cap g_iA$  seem to be interesting candidates for the above procedure.

A very different deterministic approach to construct families  $A_i$  of increasing volume  $e(A_i)$  and increasing density  $\delta(A_i)$  consists of systematic "local" improvements implemented as follows:

A first computer program compiles a complete list  $\mathcal{L} = \mathcal{L}_n$ , for some integer  $n \geq 1$ , of all pairs of subgraphs  $K_i \subset L_i$  contained in the ball  $B_n(1)$  in  $\mathcal{C}$  of radius  $n$  around the trivial element  $1 \in G$ , which satisfy  $\delta(L_i) > \delta(K_i)$ . A second program then verifies, for any  $x \in A_i$ , whether  $B_n(x) \cap A_i = xK_i$ , and if so, replaces the subgraph  $xK_i$  of  $A_i$  by  $xL_i$  to obtain the new graph  $A_{i+1}$ .

Of course, if one can increase the index  $n$  of the list  $\mathcal{L}_n$  in the first computer program, the procedure performed by the second program will lead to better values. In principle one can also imagine an interactive procedure, where all pairs of graphs  $A_i \subset A_{i+1}$  produced by the second program are automatically added to the list of test pairs  $K_i \subset L_i$  from  $\mathcal{L}$ . The problem with this theoretically most promising approach is of course the huge amount of memory needed to store the list  $\mathcal{L}$ .

### C. Non-deterministic methods for enlarging the test graph $A$

A different concept for finding high-density subgraphs of  $\mathcal{C}$  comes from the observation that a random walk in a graph has the tendency to accumulate large amounts of measure (= "heat") in parts of the graph which are "heat preserving". There are several methods how to mimick random walks in more or less efficient ways on a computer:

(1) For any non-negative function  $\mu_t : \mathcal{C} \rightarrow \mathbb{R}$  which is equal to 0 outside a finite set  $A \subset \mathcal{C}$  we define  $\mu_{t+1} : \mathcal{C} \rightarrow \mathbb{R}$  via  $\mu_{t+1}(x) = \sum \frac{\mu_t(y)}{\deg_{\mathcal{C}}(y)}$ , where the sum is taken over all vertices  $y$  adjacent to  $x$  in  $\mathcal{C}$ . We then define, for fixed  $t$  and increasing  $h$ , a decreasing family of finite level sets  $A_{t,h} = \{x \in \mathcal{C} \mid \mu_t(x) > h\}$  with empty intersection, which hence can be computed for any value of  $t$  and any value of  $h$ . Of course, the computation is very time consuming.

(2) An approximation of the previous method is the following: at any time  $t$ , one only distributes the weight  $\mu_t(y)$  among all of its neighbors  $x_i$  (including possibly  $y$  itself) if  $\mu_t(y)$  is maximal or close to the maximum value of  $\mu_t(z)$  among all  $z \in A_t = A_{t,0}$ .

The idea here is that vertices with small measure will have to be ignored anyway, as their totality grows too much like balls and will hence have low density, in general.

(3) We can exploit the fact that our graph  $\mathcal{C}$  in question is not just any graph, but actually the Cayley graph of a group  $G$ , by denoting a (finite support) measure on  $\mathcal{C}$  as element in the group ring  $\mathbb{R}G$ . It is easy to see that convolution of (finite support) measures is nothing else than simply multiplying the corresponding elements in  $\mathbb{R}G$ . In particular, the classical nearest neighbor symmetric random walk on  $\mathcal{C}(G, X)$  is directly given by the powers  $\mu^t$  for  $t \rightarrow \infty$ , where  $\mu = \frac{1}{2\#X} \sum_{X \cup X^{-1}} x \in \mathbb{R}G$ .

(4) An interesting variation of the previous three approaches seems to be the following “discretization”: One decides ahead of time on a finite integer scale, (say, from 0 to  $N$ ), and rescale the heat function  $\mu_t$  at any time so that its maximal value on  $\mathcal{C}$  equals  $N$ . Furthermore, for every vertex  $x$  the value  $\mu_t(x)$  is decreased to  $\mu_t^*(x) = [\mu_t(x)]$ , *i.e.* to the largest integer smaller or equal to  $\mu_t(x)$ . This reduces on one hand the computational effort, and at the same time it cuts off the undesired very-low-heat vertices added by the pure random walk as described above in (1). Of course, if one choses the scale too coarsely by picking  $N$  too small, we may get nowhere, by cutting off at every “rescaling second half-step” precisely what has been gained right before by the “neighbor-heat-distribution” in the first half-step, throughout any step of our discretized random walk procedure.

## References

- [1] J. M. Belk and K. S. Brown, *Forest diagrams for elements of Thompson’s group  $F$* , Internat. J. Algebra Comput. **15** (5-6) (2005), 815–850.
- [2] J. W. Cannon, W. J. Floyd, and W. R. Parry, *Introductory notes on Richard Thompson’s groups*, L’Enseignement Mathématique **42** (2) (1996), 215–256.
- [3] T. Ceccherini-Silberstein, R. Grigorchuk, and P. de la Harpe, *Amenability and paradoxal decompositions for pseudogroups and for*

*discrete metric spaces*, Proc. Steklov Inst. Math. **224** (1) (1999), 57–97.

- [4] P. de la Harpe, *Topics in geometric group theory*, University of Chicago Press, Chicago, IL, 2000.
- [5] V. S. Guba, *On the properties of the Cayley graph of Richard Thompson's group  $F$* , Internat. J. Algebra Comput. **14** (5-6) (2004), 677–702.
- [6] R.S. Lyndon and P.E. Schupp, *Combinatorial group theory*, Springer-Verlag, Berlin, 2001, Reprint of the 1977 edition.

GOULNARA N. ARZHANTSEVA  
Section de Mathématiques  
Université de Genève  
CP 64, 1211 Genève 4  
SWITZERLAND  
Goulnara.Arjantseva@math.unige.ch

VICTOR S. GUBA  
Department of Mathematics  
Vologda State University  
6 S. Orlov St., Vologda 160600  
RUSSIA  
guba@uni-vologda.ac.ru

MARTIN LUSTIG  
LATP, UMR CNRS 6632  
Mathématiques  
Université d'Aix-Marseille III  
Avenue Escadrille Normandie-Niemen  
13397 Marseille cédex 20  
FRANCE  
martin.lustig@univ.u-3mrs.fr

JEAN-PHILIPPE PRÉAUX  
Centre de recherche de l'Armée de l'air  
Ecole de l'air  
13661 Salon-air  
FRANCE  
LATP, UMR CNRS 6632  
Université de Provence  
39 rue Joliot-Curie  
13453 Marseille cédex 13  
FRANCE  
preaux@cmi.univ-mrs.fr