

Sub-dominant cogrowth behaviour and the viability of deciding amenability numerically*

Murray Elder

University of Technology Sydney,
Ultimo, New South Wales 2007, Australia
murray.elder@uts.edu.au

Cameron Rogers

Launceston Church Grammar School,
Mowbray Heights, Tasmania 7248, Australia
cameron.m.rogers@gmail.com

June 27, 2017

Abstract

We critically analyse a recent numerical method due to the first author, Rechnitzer and van Rensburg, which attempts to detect amenability in a finitely generated group by numerically estimating its asymptotic cogrowth rate. We identify two potential sources of error. We then propose a modification of the method that enables it to easily compute surprisingly accurate estimates for initial terms of the cogrowth sequence.

1 Introduction

Researchers studying the amenability Thompson's group F will be familiar with a distrust of experimental methods applied to this problem. Part of this scepticism stems from the fact that (if it is amenable) F is known to have a very quickly growing *Følner function* [22]. However, experimental algorithms investigating amenability are rarely based on Følner's criteria directly, and to date no identification is made in the literature of a mechanism by which a quickly growing Følner function could interfere with a given experimental method.

In this paper we identify such a mechanism for a recent algorithm proposed by first author, A. Rechnitzer, and E. J. Janse van Rensburg [9], which was designed to experimentally detect amenability via the Grigorchuk-Cohen characterisation in terms of the cogrowth function. We will refer to this as the ERR algorithm in the sequel.

*Research supported by Australian Research Council grant FT110100178

We show that, in the ERR algorithm, estimates of the asymptotic cogrowth rate are compromised by sub-dominant behaviour in the reduced-cogrowth function.

However, even though sub-dominant behaviour in the cogrowth function may interfere with estimates of the asymptotic growth rate, the ERR algorithm can still be used to estimate other properties of the cogrowth function to high levels of accuracy. In particular we are able re-purpose the algorithm to quickly estimate initial values of the cogrowth function even for groups for which the determination of the asymptotic cogrowth rate is not possible (for example finitely generated groups with unsolvable word problem).

The present work started out as an independent verification by the second author of the experimental results in [9], as part of his PhD research at the University of Newcastle. More details can be found in [27].

The article is organised as follows. In Section 2 we give the necessary background on amenability, random walks and cogrowth. In Section 3 a function quantifying the sub-dominant properties of the reduced-cogrowth function is defined. In Section 4 the ERR algorithm is summarised, followed by an analysis of two types of pathological behaviour in Section 5. The first of these is easily handled, while the second is shown to depend on sub-dominant terms in the reduced-cogrowth function. In Section 6 the ERR method is modified to provide estimates of initial cogrowth values. Using this the first 2000 terms for the cogrowth function of Thompson's group F are estimated.

2 Preliminaries

The following characterisation of amenability is due to Grigorchuk [15] and Cohen [8]. A shorter proof of the equivalence of this criteria with amenability was provided by Szwarc [28].

Definition 2.1. Let G be a finitely generated non-free group with symmetric generating set S . Let c_n denote the number of freely reduced words of length n over S which are equal to the identity in G . Then G is amenable if and only if

$$\limsup_{n \rightarrow \infty} c_n^{1/n} = |S| - 1.$$

Equivalently, let d_n denote the number of words (reduced and unreduced) of length n over S which are equal to the identity. Then G is amenable if and only if

$$\limsup_{n \rightarrow \infty} d_n^{1/n} = |S|.$$

The function $n \mapsto c_n$ is called the *reduced-cogrowth function* for G with respect to S , and $n \mapsto d_n$ the *cogrowth function*.

Kesten's criteria for amenability is given in terms of the probability of a random walk on the group returning to its starting point.

Definition 2.2. Let G be a finitely generated group, and let μ be a symmetric measure on G . The random walk motivated by μ is a Markov chain on the group starting at the identity where the probability of moving from x to y is $\mu(x^{-1}y)$. Note the distribution after n steps is given by the n -fold convolution power of μ , which we denote as μ_n . That is, $\mu_n(g)$ is the probability that an n -step walk starting at e ends at g . By Kesten's criteria [18] a group is amenable if and only if

$$\limsup_{n \rightarrow \infty} (\mu_n(e))^{1/n} = 1.$$

Pittet and Saloff-Coste proved that the asymptotic decay rate of the probability of return function is independent of measure chosen, up to the usual equivalence [24]. For finitely generated groups we can choose the random walk motivated by the uniform probability measure on a finite generating set. This random walk is called a *simple random walk* and corresponds exactly with a random walk on the Cayley graph. For this measure the probability of return is given by

$$\mu_n(e) = \frac{d_n}{|S|^n}, \quad (1)$$

where the (reduced and non-reduced) cogrowth terms d_n are calculated with respect to the support of the measure. Thus the cogrowth function arises from a special case of return probabilities.

Følner's characterisation of amenability [13] can be phrased in several ways. Here we give the definition for finitely generated groups.

Definition 2.3. Let G be a group with finite generating set S . For each finite subset $F \subseteq G$, we denote by $|F|$ the number of elements in F . The *boundary* of a finite set F is defined to be

$$\partial F = \{g \in G : g \notin F, gs \in F \text{ for some } s \in S\}.$$

A finitely generated group G is amenable if and only if there exists a sequence of finite subsets F_n such that

$$\lim_{n \rightarrow \infty} \frac{|\partial F_n|}{|F_n|} = 0.$$

Vershik [29] defined the following function as a way to quantify how much of the Cayley graph must be considered before sets with a given isoperimetric profile can be found.

Definition 2.4. The Følner function of a group is

$$f(n) = \min \left\{ |F| : \frac{|\partial F|}{|F|} < \frac{1}{n} \right\}.$$

Significant literature exists on Følner functions. It is known that there exist finitely presented amenable groups with Følner functions growing faster than n^{n^n} ([20] Corollary 6.3) and finitely generated groups (iterated wreath product

of k copies of \mathbb{Z}) with Følner functions growing faster than $n^{n^{\dots}}$ of height k for arbitrary k [11].

Richard Thompson's group F is the group with presentation

$$\langle a, b \mid [ab^{-1}, a^{-1}ba], [ab^{-1}, a^{-2}ba^2] \rangle \quad (2)$$

where $[x, y] = xyx^{-1}y^{-1}$ denotes the commutator of two elements. See for example [6] for a more detailed introduction to this group. Whether or not F is amenable has attracted a large amount of interest, and has so far evaded many different attempts at a proof of both positive and negative answers. Experimental work previously done on this question includes [5, 1, 10, 17, 9].

Justin Moore [22] (2013) has shown that if F were amenable then its Følner function would increase faster than a tower of $n - 1$ twos,

$$2^{2^{2^{\dots}}}$$

This result has been proposed as an obstruction to all computational methods for approximating amenability; a computationally infeasibly large portion of the Cayley graph must be considered before sets with small boundaries can be found. However, in all but one of the experimental algorithms ([1]) listed above computing Følner sets was not the principle aim. In order to understand how a bad Følner function affects the performance of these methods, we need to understand the connection between convergence properties of the respective limits in the various characterisations of amenability.

3 Quantifying sub-dominant cogrowth behaviour

The Følner function quantifies the rate of convergence of the limit in Definition 2.3. We propose the following definitions to quantify the rate of convergence of the limits in Definition 2.1.

Definition 3.1. Let G be a finitely generated group with symmetric generating set S . Let c_n be the number of all reduced trivial words of length n and let $C = \limsup c_n^{1/n}$. Define

$$\mathcal{R}(n) = \min \left\{ k : \frac{c_{2k+2}}{c_{2k}} > C^2 - \frac{1}{n} \right\}$$

Definition 3.1 uses only even word lengths (and hence C^2 instead of C). This is necessary because group presentations with only even length relators have no odd length trivial words. For this paper we will only consider the function \mathcal{R} for amenable groups, in which case $C = |S| - 1$ except when the group is free (infinite cyclic).

A similar definition may be made for the cogrowth function.

Definition 3.2. For G a finitely generated group with symmetric generating set S we may define

$$\mathcal{R}'(n) = \min \left\{ k : \frac{d_{2k+2}}{d_{2k}} > D^2 - \frac{1}{n} \right\}$$

where d_n be the number of all (reduced and non-reduced) trivial words of length n and $D = \limsup c_n^{1/n}$.

Literature already exists studying the convergence properties of return probabilities, and we suspect that the function \mathcal{R}' is a reformulation of the L^2 -isoperimetric function [3].

Aside from the trivial group (where $\mathcal{R}(n) = 0$), it is usually easier to compute \mathcal{R}' (or its asymptotics) than it is to obtain \mathcal{R} . For this reason we first consider \mathcal{R}' functions for various groups, and then prove that for infinite, amenable, non-free groups \mathcal{R}' and \mathcal{R} have the same asymptotic behaviour.

Example 3.3. For any finite group the rate of growth of d_n is the dominant eigenvalue of the adjacency matrix of the Cayley graph, and some simple analysis shows that $\mathcal{R}'(n)$ is at most logarithmic in n .

Define $f \lesssim g$ if there exist constants $a, b > 0$, such that for x large enough, $f(x) \leq ag(bx)$. Then $f \sim g$ (f and g are asymptotic) if $f \lesssim g$ and $g \lesssim f$.

Table 1 provides a sample of amenable groups for which the asymptotics of $\mathcal{R}'(n)$, the Følner function and probabilities of return are known [11, 25, 26].

The results for the asymptotics of $\mathcal{R}'(n)$ were derived directly from the known asymptotics for μ_n . A discussion of these methods appears in [27]. In practice however it proved quicker to guess the asymptotics and then refine using the following method.

Proposition 3.4. *The asymptotic results for $\mathcal{R}'(n)$ in Table 1 are correct.*

Proof. For a given group suppose $\mu_n(e) \sim g(n)$ where g is a continuous real valued function, as in Table 1. Then $d_n \sim |S|^n g(n)$.

Finding $\mathcal{R}'(n)$ requires solving the equation

$$\frac{d_{2k+2}}{d_{2k}} = |S|^2 - \frac{1}{n} \tag{3}$$

for $k = k(n)$. This is equivalent to solving

$$1 = n \left(|S|^2 - \frac{d_{2k+2}}{d_{2k}} \right)$$

for k .

Suppose $f(n)$ is a function where

$$L = \lim_{n \rightarrow \infty} n \left(|S|^2 - \frac{d_{2f(n)+2}}{d_{2f(n)}} \right) \tag{4}$$

| Example | $\mathcal{F}(n)$ | $\mu_n(e)$ | $\mathcal{R}'(n)$ |
|--|--|--|--------------------------|
| trivial | $\sim \text{constant}$ | $\sim \text{constant}$ | $\sim \text{constant}$ |
| \mathbb{Z}^k | $\sim n^k$ | $\sim n^{-k/2}$ | $\sim n$ |
| $BS(1, N)$ | $\sim e^n$ | $\sim e^{-n^{1/3}}$ | $\sim n^{3/2}$ |
| $\mathbb{Z} \wr \mathbb{Z}$ | n^n | $\sim e^{-n^{1/3}(\ln n)^{2/3}}$ | $\sim \ln(n)n^{3/2}$ |
| $\mathbb{Z} \wr \mathbb{Z} \wr \dots \wr \mathbb{Z}$ ($d+1$)-fold wreath product | $n^{n^{\dots^n}}$ (tower of $d+1$ n 's) | $\sim e^{-n^{\frac{d}{d+2}}(\ln n)^{\frac{2}{d+2}}}$ | $\sim \ln(n)n^{(d+2)/2}$ |

Table 1: Comparing asymptotics of the probabilities of return, the Følner function \mathcal{F} , and \mathcal{R}' for various groups.

exists and is non-zero.

If $L = 1$ then

$$\left(|S|^2 - \frac{d_{2f(n)+2}}{d_{2f(n)}} \right) \sim \frac{1}{n}$$

and so

$$\frac{d_{2f(n)+2}}{d_{2f(n)}} \sim |S|^2 - \frac{1}{n}.$$

Then $k(n) \sim f(n)$ satisfies Equation 3. Therefore $\mathcal{R}'(n)$ is asymptotic to $f(n)$.

If L exists and is non-zero then

$$\left(|S|^2 - \frac{d_{2f(n)+2}}{d_{2f(n)}} \right) \sim \frac{L}{n}.$$

Then

$$\left(|S|^2 - \frac{d_{2f(Ln)+2}}{d_{2f(Ln)}} \right) \sim \frac{L}{Ln} = \frac{1}{n}$$

and so $\mathcal{R}'(n) \sim f(Ln)$.

The derivations of candidates for $f(n)$ in each case in Table 1 is performed in [27]. The results in the table do not include the constant L since the probabilities of return used as input are only correct up to scaling. We leave the calculation of Equation 4 for the results from Table 1 as an exercise. \square

3.1 Converting from cogrowth to reduced-cogrowth

We now prove an equivalence between the sub-dominant behaviour of the cogrowth and reduced-cogrowth functions. This allows us to borrow the previously listed results for \mathcal{R}' when discussing \mathcal{R} and the ERR method. The dominant and sub-dominant cogrowth behaviour can be analysed from the generating functions for these sequences.

Definition 3.5. Let d_n denote the number of trivial words of length n in a finitely generated group. The *cogrowth series* is defined to be

$$D(z) = \sum_{n=0}^{\infty} d_n z^n.$$

Let c_n denote the number of reduced trivial words. Then

$$C(z) = \sum_{n=0}^{\infty} c_n z^n$$

is said to be the *reduced-cogrowth series*.

D and C are the generating functions for d_n and c_n respectively, and are related in the following way. Let $|S| = 2p$ be the size of a symmetric generating set. Then from [19, 30]

$$C(z) = \frac{1 - z^2}{1 + (2p - 1)z^2} D\left(\frac{z}{1 + (2p - 1)z^2}\right) \quad (5)$$

and

$$D(z) = \frac{1 - p + p\sqrt{1 - 4(2p - 1)z^2}}{1 - 4p^2z^2} C\left(\frac{1 - \sqrt{1 - 4(2p - 1)z^2}}{2(2p - 1)z}\right). \quad (6)$$

The dominant and sub-dominant growth properties of the cogrowth functions may be analysed by considering the singularities of these generating functions. For a detailed study of the relationship between singularities of generating functions and sub-dominant behaviours of coefficients see [12].

We now outline an example of how the composition of functions (as in Equations 5 and 6) effects the growth properties of the series coefficients.

Example 3.6. Consider

$$f(z) = \left(1 - \frac{z}{r}\right)^{-p}.$$

Then (for positive p) $f(z)$ has a singularity at $z = r$, and this defines the radius of convergence of $f(z)$ and the asymptotic growth rate of the series coefficients of the expansion of $f(z)$. It also determines the principle sub-dominant term contributing to the growth of the coefficients. In this example, the coefficients will grow like $n^{p-1}r^{-n}$.

We wish to investigate what happens to this growth behaviour when we compose the function f with a function g . Consider $f(g(z))$ for some function g for which $g(0) = 0$. The singularities of g are inherited by $f(g(z))$; if g is analytic everywhere then the only singularities of $f(g(z))$ will occur when $g(z) = r$. In this case, the new radius of convergence will be the minimum $|z|$ such that $g(z) = r$. Importantly, however, the principle sub-dominant growth term of the series coefficients will remain polynomial of degree $p - 1$.

A variation on this behaviour will occur if there is an r_0 for which $g(z)$ is analytic on the ball of radius r_0 , and $g(z) = r$ for some z in this region. Again, when this occurs, the new radius of convergence is obtained by solving $g(z) = r$ and the type of the principle sub-dominant term in the growth of the coefficients remains unchanged.

If there does not exist such an r_0 , the principle singularity of $g(z)$ will dominate the growth properties of the coefficients.

Proposition 3.7. *Let G be an infinite amenable group generated by p elements and their inverses. Then the principle sub-dominant terms contributing to the growth of d_n and c_n are asymptotically equivalent, except when the group is infinite cyclic.*

Proof. For an amenable group generated by p elements and their inverses the radius of convergence for $D(z)$ is exactly $1/2p$. This follows immediately from Definition 2.1.

Now from Equation 5, the reduced-cogrowth series is obtained by composing the cogrowth series with

$$p(z) = \frac{z}{1 + (2p - 1)z^2}$$

and then multiplying by

$$q(z) = \frac{1 - z^2}{1 + (2p - 1)z^2}.$$

Both of these functions are analytic inside the ball of radius $1/\sqrt{2p - 1}$.

Now

$$p\left(\frac{1}{2p - 1}\right) = \frac{1}{2p}, \tag{7}$$

the singularity of $D(z)$. Hence, $1/(2p - 1)$ is a singularity of $D(p(z))$, and hence of $C(z)$. Note that if the group is infinite cyclic, then $p = 1$ and $1/(2p - 1)$ and $1/\sqrt{2p - 1}$ are equal. In this scenario the radius of convergence of $p(z)$ is reached at the same moment that $p(z)$ reaches the radius of convergence of $D(z)$. This means that both p and q contribute to the principle singularity, and this explains why the reduced and non-reduced cogrowth functions for the infinite cyclic group exhibit such different behaviour.

If $p > 1$ then $1/(2p - 1)$ is inside the ball of radius $1/\sqrt{2p - 1}$ (ie, inside the region of convergence for p and q). Thus, the singularity of D is reached before z approaches the singularity of p and q .

In this case the substitutions in Equation 5 change the location of the principle singularity, but do not change the type of the singularity, or the form of the principle sub-dominant term contributing to the growth of the series coefficients. \square

Corollary 3.8. *Suppose G is a finitely generated, infinite amenable group that is not the infinite cyclic group. Then \mathcal{R} is asymptotically equivalent to \mathcal{R}' .*

Remark 3.9. An alternate proof of the Grigorchuk/Cohen characterisation of amenability is easily constructed from an analysis of the singularities of $C(z)$ and $D(z)$. For example, Equation 7 proves the first result from Definition 2.1. This argument also picks up that the infinite cyclic group presents a special case. Though amenable, $\limsup_{n \rightarrow \infty} c_n \neq |S| - 1$. For this group we have $\mathcal{R}(n) \sim 0$ while $\mathcal{R}'(n) \sim n$.

3.2 Sub-dominant behaviour in the cogrowth of F

The groups $BS(1, N)$ limit to $\mathbb{Z} \wr \mathbb{Z}$ in the space of marked groups. This implies that the growth of the function \mathcal{R}' and hence \mathcal{R} for $BS(1, N)$ increases with N . This is consistent with Table 1, since these results do not include scaling constants. This leads to the following result.

Proposition 3.10. *If Thompson's group F is amenable, its \mathcal{R} function grows faster than the \mathcal{R} function for any $BS(1, N)$. In particular, it is asymptotically super-polynomial.*

Proof. By the convergence of $BS(1, N)$ to $\mathbb{Z} \wr \mathbb{Z}$ in the space of marked groups we have that, for any N , the function \mathcal{R}' for $BS(1, N)$ grows slower than the corresponding function for $\mathbb{Z} \wr \mathbb{Z}$. In [24] it is proved that, for finitely generated groups, the probability of return cannot asymptotically exceed the probability of return of any finitely generated subgroup. This implies that, for finitely generated amenable groups, the \mathcal{R}' function of the group must grow faster than the \mathcal{R}' function of any finitely generated subgroup. Since there is a subgroup of F isomorphic to $\mathbb{Z} \wr \mathbb{Z}$, $\mathcal{R}'(n)$ for F must grow faster than $\mathcal{R}'(n)$ for $\mathbb{Z} \wr \mathbb{Z}$ and hence $BS(1, N)$.

Since F contains every finite depth iterated wreath products of Z ([16] Corollary 20), the probability of return for F decays faster than

$$e^{-n^{\frac{d}{d+2}} (\ln n)^{\frac{2}{d+2}}}$$

for any d . Taking the limit as d approaches infinity of the corresponding values for \mathcal{R}' and then performing the conversion from \mathcal{R}' to \mathcal{R} gives the result. \square

Note that if F is non-amenable, then even though it still contains these subgroups, they do not affect the \mathcal{R}' function. In this scenario it is still true that the return probability for F decays faster than the iterated wreath product, because F would have exponentially decaying return probability. For non-amenable groups the return probability does not identify the principle sub-dominant term in d_n , and hence does not correlate directly with \mathcal{R}' .

4 The ERR algorithm

We start by summarising the original work by the first author, Rechnitzer and van Rensburg. Only the details directly pertinent to the present paper are discussed here, for a more detailed analysis of the random walk algorithm and a derivation of the stationary distribution we refer the reader to [9]. For the sake of brevity the random walk performed by the algorithm will be referred to as the *ERR walk*.

Recall that a group presentation, denoted $\langle S \mid R \rangle$, consists of a set S of formal symbols (generators) and a set R of words written in $S^{\pm 1}$ (relators) and corresponds to the quotient of the free group on S by the normal closure of the relators R . In our paper, as in [9], all groups will be finitely presented: both S and R will be finite. Furthermore, the implementation of the algorithm assumes both S and R to be symmetric, that is, $S = S^{-1}$ and $R = R^{-1}$. In addition, for convenience R is enlarged to be closed under cyclic permutation. Recall that c_n counts the number of reduced words in S of length n which represent the identity in the group (that is, belong to the normal closure of R in the free group).

4.1 The ERR walk

The ERR walk is not a random walk on the Cayley graph of a group, but instead a random walk on the set of trivial words for the group presentation. This makes the algorithm extremely easy to implement, since it does not require an easily computable normal form or even a solution to the word problem. The walk begins at the empty word, and constructs new trivial words from the current trivial word using one of two moves:

- (conjugation by $x \in S$). An element is chosen from S according to a predetermined probability distribution. The current word is conjugated by the chosen generator and then freely reduced to produce the new candidate word.
- (insertion of a relator). A relator is chosen from R according to a predetermined distribution and inserted into the current word at a position chosen uniformly at random. In order to maintain *detailed balance* (from which the stationary distribution is derived) it is necessary to allow only those insertions which can be immediately reversed by inserting the inverse of the relator at the same position. To this end the notion of *left insertion* is introduced; after relators are inserted free reduction is done on only the left hand side of the relator. If after this the word is not freely reduced the move is rejected.

Transition probabilities are defined which determine whether or not the trivial word created with these moves is accepted as the new state. These probabilities involve parameters $\alpha \in \mathbb{R}$ and $\beta \in (0, 1)$ which may be adjusted to control the distribution of the walk.

Let the current word be w and the candidate word be w' .

- If w' was obtained from w via a conjugation it is accepted as the new current state with probability

$$\min \left\{ 1, \left(\frac{|w'| + 1}{|w| + 1} \right)^{1+\alpha} \beta^{|w'| - |w|} \right\}.$$

- If w' was obtained from w via an insertion it is accepted as the new state with probability

$$\min \left\{ 1, \left(\frac{|w'| + 1}{|w| + 1} \right)^\alpha \beta^{|w'| - |w|} \right\}.$$

If w' is not accepted the new state remains as w .

These probabilities are chosen so that the distribution on the set of all trivial words given by

$$\pi(u) = \frac{(|u| + 1)^{1+\alpha} \beta^{|u|}}{Z},$$

(where Z is a normalizing constant) can be proved to be the unique stationary distribution of the Markov chain, and the limiting distribution of the random walk.

Proposition 4.1 ([9]). *As β approaches*

$$\beta_c = \frac{1}{\limsup_{n \rightarrow \infty} (c_n)^{1/n}}$$

the expected value of the word lengths visited approaches infinity.

This leads to the following method for estimating the value of β_c . For each presentation, walks are run with different values of β . Average word length is plotted against β . The results obtained for Thompson's group F are reproduced in Figure 1. The values for β at which the data points diverge gives an indication of β_c , and hence the amenability or otherwise of the group.

Random walks were run on presentations for a selection of amenable and non-amenable groups, including Baumslag-Solitar groups, some free product examples whose cogrowth series are known [21], the genus 2 hyperbolic surface group, a finitely presented group related to the basilica group, and Thompson's group F .

The data in Figure 1 appears to show fairly convincingly that the location of β_c is a long way from the value of $\frac{1}{3}$ expected were the group amenable.

It is noted in [9] that a long random walk may be split into shorter segments, and the variation in average word lengths of the segments gives an estimation of the error in the estimated expected word length.

Remark 4.2. In the original work reported in [9], the algorithm was coded in C++, words were stored as linked lists, the GNU Scientific Library was used to

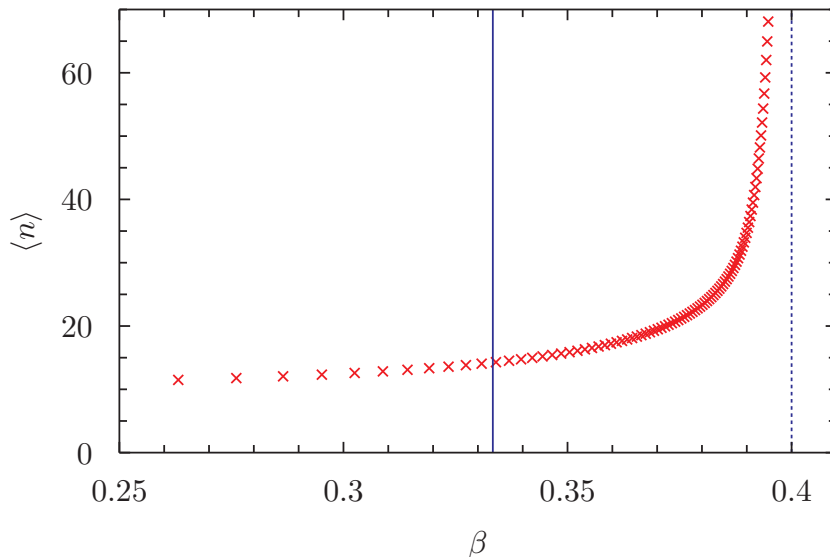


Figure 1: The results from [9] (Figure 10) of the ERR algorithm applied to the standard presentation of Thompson’s group F . Each data point plots the average word length of an ERR walk against the parameter β used.

generate pseudo-random numbers, and *parallel tempering* was used to speed up the convergence of the random walk. For independent verification the second author coded the algorithm in Python, kept words as strings, used the Python package *random*, and no tempering was used. Results obtained were consistent with those in [9]. The experimental analysis and modifications described in this paper use the Python version of the second author.

5 Investigating Pathological Behaviour

The theory underpinning the ERR walk is complete — the random walk is certain to converge to the stationary distribution. This does not preclude, however, convergence happening at a computationally undetectable rate. Since there are finitely presented groups with unsolvable word problem, there is no chance of deriving bounds on the rates of convergence of the walk in any generality. In the process of independently verifying the results in [9], however, we were able to identify two properties of group presentations which appear to slow the rate of convergence. The first of these is unconnected with the Følner function, and does not pose any problem to the implementation of the ERR algorithm to Thompson’s group F . It does, however, refute the claim in ([9] Section 3.7) that the method can be successfully applied to infinite presentations.

5.1 Walking on the wrong group

It is easy to see from the probabilistic selection criteria used by the ERR walk that moves which increase the word length by a large amount are rejected with high probability. This poses a problem for group presentations containing long relators since insertion moves that attempt to insert a long relator will be accepted much less often than moves which attempt to insert a shorter relation.

The following example makes this explicit.

Lemma 5.1. *All presentations of the form*

$$\langle a, b \mid abab^{-1}a^{-1}b^{-1}, a^n b^{-n-1} \rangle$$

describe the trivial group.

Proof. Since $a^n = b^{n+1}$ we have $a^n ba = b^{n+1}ba = bb^{n+1}a = ba^{n+1} = bab^{n+1}$. Since $aba = bab$ we have $a^i ba = a^{i-1}bab$ so $a^n ba = a^{n-1}bab = a^{n-2}bab^2 = \dots = bab^n$. Putting these results together gives $bab^n = bab^{n+1}$ and hence b is trivial. The result follows. \square

By increasing n we can make the second relator arbitrarily large without affecting the group represented by the presentation, or the group elements represented by the generators. This implies that ERR walks for each of these presentations should converge to the same stationary distribution.

Changing the presentation, however, does change the number of steps in the ERR walk needed to reach certain trivial words. ERR walks were performed on these presentations for $n = 1, 2, \dots, 19$. As well as recording the average word length of words visited, the number of *accepted* insertions of each relator was recorded.

Table 2 shows the sharp decline in the number of accepted insertions of the second relator as n increases. Indeed, for $n > 14$ there were no instances in which the longer relator was successfully inserted. Unsurprisingly, walks for large n did not converge to the same distribution as those where n was small, and for large n the data did not accurately predict the asymptotic growth rate of the cogrowth function. For these n the ERR walk was actually taking place on $\langle a, b \mid abab^{-1}a^{-1}b^{-1} \rangle$, a presentation for the 3-stand braid group, which is non-amenable.

Note that, given enough time, the longer relator would be successfully sampled, and that an infinite random walk is still guaranteed to converge to the theoretical distribution for the trivial group. Such convergence, however, may take a computationally infeasible amount of time.

Claim 5.2. *The presence of long relators in the input presentation slows the rate at which an ERR walk converges to the stationary distribution. Therefore, the ERR method cannot be reliably extended to accept infinite presentations.*

This result is not surprising. In [4] an infinitely presented amenable group is given for which any truncated presentation (removing all but a finite number of relators) is non-amenable. The ERR method could not expect to succeed on this

| n | number of steps | number of accepted insertions of small relator | number of accepted insertions for big relator |
|-----|-------------------|--|---|
| 1 | 2.0×10^8 | 2977228 | 7022772 |
| 2 | 3.6×10^8 | 4420185 | 5579815 |
| 3 | 6.1×10^8 | 6323376 | 3676624 |
| 4 | 9.0×10^8 | 8016495 | 1983505 |
| 5 | 1.2×10^9 | 9088706 | 911294 |
| 6 | 1.4×10^9 | 9621402 | 378598 |
| 7 | 1.5×10^9 | 9850251 | 149749 |
| 8 | 1.7×10^9 | 9943619 | 56381 |
| 9 | 1.8×10^9 | 9977803 | 22197 |
| 10 | 1.9×10^9 | 9991680 | 8320 |
| 11 | 2.1×10^9 | 9997122 | 2878 |
| 12 | 2.2×10^9 | 9998720 | 1280 |
| 13 | 2.2×10^9 | 9999585 | 415 |
| 14 | 2.3×10^9 | 9999938 | 62 |
| 15 | 2.4×10^9 | 10000000 | 0 |
| 16 | 2.6×10^9 | 10000000 | 0 |
| 17 | 2.7×10^9 | 10000000 | 0 |
| 18 | 2.8×10^9 | 10000000 | 0 |
| 19 | 2.9×10^9 | 10000000 | 0 |

Table 2: The ERR algorithm applied to the trivial group with presentation $\langle a, b \mid aba = bab, a^n = b^{n+1} \rangle$ for various n . As n increases, the longer relator is successfully inserted less frequently.

group even if long relators were sampled often; since the ERR walk can only be run for a finite time there can never be a representative sampling of an infinite set of relators, so ERR would incorrectly conclude this group is non-amenable.

The pathological presentations of the trivial group studied here form a sequence of presentations for amenable (trivial) groups which approach a non-amenable group in the space of marked groups. The failure of the ERR method to predict amenability for these groups suggests that one does not need particularly elaborate or large presentations to produce pathological behaviour.

However, we remark that this behaviour is easily monitored. In addition to counting the number of attempted moves of the walk, one should record the relative number of successful insertions of each relator. In the case of Thompson's group F the two relators have similar lengths, and in our experiments both were sampled with comparable frequency.

Further analysis of this phenomena appears [27].

5.2 Sub-dominant behaviour in cogrowth.

Recall that the solvable Baumslag-Solitar groups $BS(1, N) = \langle a, t \mid tat^{-1}a^{-N} \rangle$ are the only two generator, single-relator, amenable groups [7]; for each of these groups $\beta_c = 1/3$. In [9] walks were run on $BS(1, 1) = \mathbb{Z}^2$, $BS(1, 2)$ and $BS(1, 3)$ and for these groups the random walk behaved as predicted with divergence occurring at the moment when β exceeded β_c . It may be surprising then to see the output of some ERR walks for $BS(1, 7)$ shown in Figure 2.

It is clear that, for this group, the divergence for $\beta > \beta_c$ predicted by the theory is not occurring. This is further seen in Figure 3, which shows the progression over time of one of the random walks used to generate Figure 2. The results in Figure 3 show the word lengths visited for ten ERR walks (superimposed) performed on $BS(1, 7)$, with $\alpha = 3$ and $\beta = 0.34$. Since the group has only a single relator, which was successfully inserted into the word 10000 times, it is not an error of the type identified in Subsection 5.1. The ERR method relies on the divergence of the average word length to identify β_c , so application of the method in this case will not accurately identify the amenability of $BS(1, 7)$.

Divergence of the ERR walk (when $\beta > \beta_c$) relies on the abundance of long trivial words. For most presentations, at all points in an ERR walk there are always more moves which lengthen the word than shorten it, but the probabilistic selection criteria ensures balance. More specifically, the parameter β imposes a probabilistic barrier which increases exponentially with attempted increase in word length. When $\beta > \beta_c$ this exponential cap is insufficient, and the word length diverges.

Recall that for a given word length n the function $\mathcal{R}(n)$ quantifies how many reduced-trivial words there are of length similar to n . The results in Table 1 imply that, for many groups, large word lengths must be reached before the asymptotic growth rate is reflected by a local abundance of longer trivial words. We have noted in Section 3.2 that the convergence properties of $BS(1, N)$ in the space of marked groups requires $\mathcal{R}(n)$ to grow more quickly as N increases.

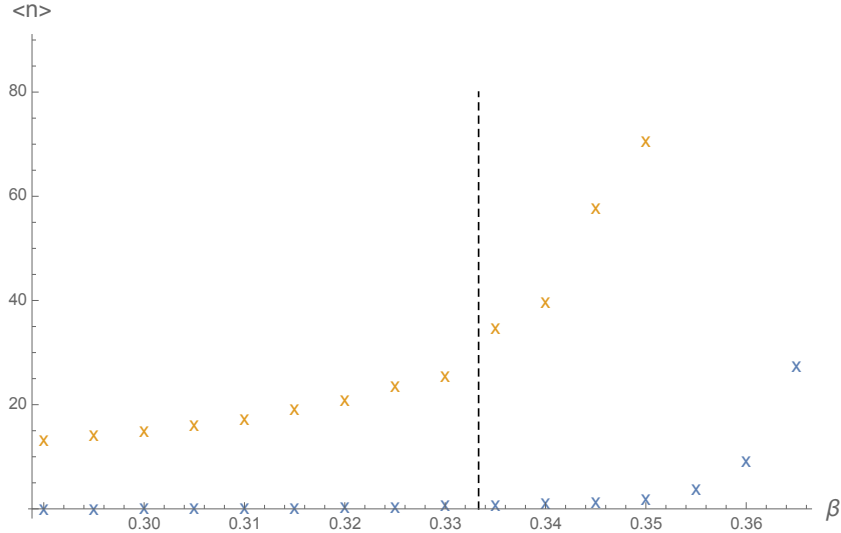


Figure 2: A graph (as in [9]), of average word length in ERR walks for $\langle a, t \mid t^{-1}ata^{-7} \rangle$ plotted against the parameter β . The orange points (top curve) come from walks where $\alpha = 3$, and the blue points (bottom curve) come from walks where $\alpha = 0$. The vertical line at $1/3$ marks the expected asymptote.

We now show that the growth rate of $\mathcal{R}(n)$ is sufficient to cause the pathological behaviour noted above.

To this end we postulate a hypothetical cogrowth function for which we can explicitly identify and control $\mathcal{R}(n)$.

Example 5.3. Suppose that for some group on two generators and $q > 0$, $p \in (0, 1)$, the reduced-cogrowth is known to be exactly

$$c_n = 3^{n-qn^p}.$$

Then $\limsup_{n \rightarrow \infty} c_n^{1/n} = 3$ and so the group is amenable. It may easily be verified by the methods outlined in Proposition 3.4 that

$$\mathcal{R}(n) = (9 \log(3)qp2^pn)^{\frac{1}{1-p}}.$$

Note that as p approaches 1, the exponent $\frac{1}{1-p}$ approaches infinity. This increases both the degree of the polynomial in n , and the coefficient $(9 \log(3)qp2^pn)^{\frac{1}{1-p}}$.

Even though we do not know a group presentation with precisely this cogrowth function, by varying p and q this hypothetical example models the groups listed in Table 1.

Figure 4 shows the effect of increasing the parameter p on the ERR walk distribution. Note that this figure is not the output of any computer simulation, rather it models the distributions for an ERR walk on an amenable group with

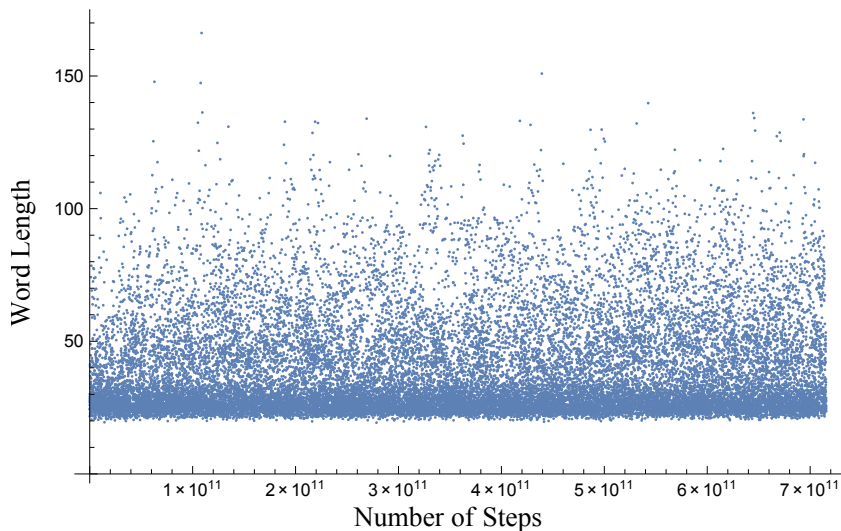


Figure 3: The distribution of ERR walks on $BS(1, 7)$ with $\alpha = 3$ and $\beta = 0.34$. This is a plot of word length against number of steps taken. The data represents ten ERR walks overlaid on top of each other. As can be seen, none of the walks diverged. Each dot represents the average word length over 10000 accepted relator insertions. There is no divergence at this β value, even though the group is amenable.

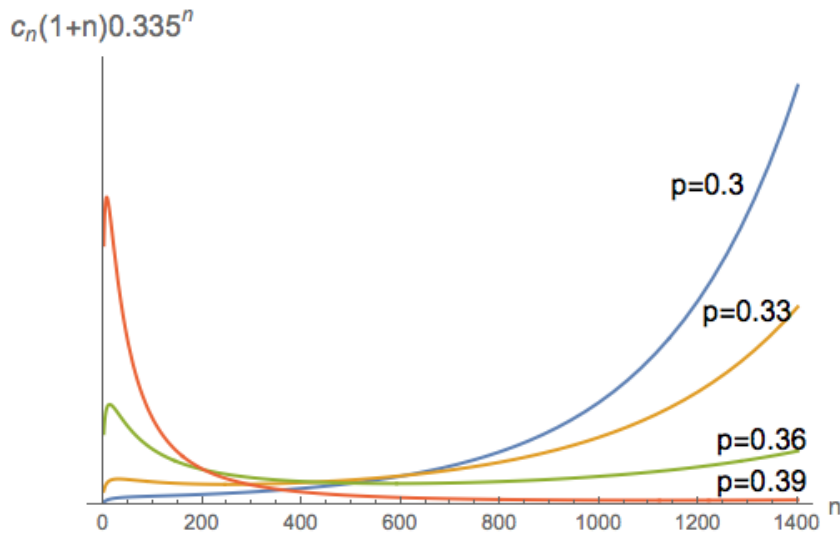


Figure 4: Graphs of $c_n(n + 1)0.335^n$ for $c_n = 3^{n-n^p}$.

the hypothetical cogrowth function, for $\alpha = 0, \beta = 0.335$ and $q = 1$. Recall that for $\beta < \beta_c$ the theoretical distribution of word lengths visited by the ERR walk is

$$\Pr(n) = \frac{c_n(n+1)^{\alpha+1}\beta^n}{Z}$$

where Z is a normalizing constant. For $\beta > \beta_c$ the distribution cannot be normalised. In this case the function $c_n(n+1)^{\alpha+1}\beta^n$ still contains information about the behaviour of the walk. If the walk reaches a word of length x then the relative heights of $c_n(n+1)^{\alpha+1}\beta^n$ either side of x describe the relative probabilities of increasing or decreasing the word length in the next move.

From Figure 4 we see that, for $p = 0.3$, the slope of $c_n(n+1)^{\alpha+1}\beta^n$ is always positive, so at all word lengths probabilities are uniformly in favour of increasing the word length. However, as p increases (and the growth rate for $\mathcal{R}(n)$ increases) a ‘hump’ appears at short word lengths. A walk for such a group would tend to get stuck in the ‘hump’. Indeed, for $p = 0.39$ the distribution looks much less like a walk diverging towards infinite word lengths and much more like the distributions for $BS(1, 7)$ used to produce Figure 2, where the average word length in the ERR walk remained finite.

The distributions in Figure 4 exhibit a mechanism which can explain anomalous behaviour previously observed. When $\mathcal{R}(n)$ increases quickly the ERR walk may adhere to the behaviour predicted by the theory and simultaneously give anomalous results about the asymptotics of the cogrowth function. In this sense if [9] contains incorrect answers it is because the original ERR algorithm as it was initially proposed asks the wrong question. The ERR walk does not measure asymptotic properties of the cogrowth function; it provides information about the cogrowth function only for word lengths visited by the walk. This observation forms the basis of Section 6.

Note that increasing the parameter α pushes the algorithm towards longer word lengths. Thus, any pathological behaviour caused by the growth of $\mathcal{R}(n)$ could theoretically be overcome by increasing α . If $\mathcal{R}(n)$ is known, then it may be used to calculate how large words have to get before divergence occurs. A method to do this is outlined by the following example.

Suppose that ERR walks are run on a two generator group with $\beta = 0.34$ (as in Figure 3). If we eliminate the α term of the stationary distribution (which, being polynomial, becomes insignificant for long word lengths) the divergence properties are controlled by the contest between 0.34^n and c_n . That is, divergence will occur when $c_{2n+2}/c_{2n} > 1/0.34^2 = 3 - 1/17$; the word length at which divergence will occur is $\mathcal{R}(17)$. If this value is known α may be increased until the walk visits words of this length. This process, however, requires specific information about $\mathcal{R}(n)$ including all scaling constants. It is hard to imagine a group for which the sub-dominant cogrowth behaviour was known to this level of precision, but dominant cogrowth behaviour (and hence the amenability question for the group) was still unknown.

5.3 Reliability of the ERR results for Thompson's group F

In Proposition 3.10 we saw that the \mathcal{R} function for F grows faster than that of any iterated wreath product of \mathbb{Z} 's, and certainly faster than that of any $BS(1, N)$ group. Since the ERR method fails to predict the amenability of these groups for N as low as 7, and this behaviour is consistent with the pathological behaviour caused by \mathcal{R} , we conclude that the data encoded in Figure 1 does not imply the non-amenability of F , and so the conclusion of the paper [9] that F appears to be non-amenable based on this data is unreliable.

6 Appropriation of the ERR algorithm

The original implementation of the ERR walk uses only the mean length of words visited in an attempt to estimate asymptotic behaviour of the cogrowth function. In this section we show that, using the full distribution of word lengths visited, it is possible to estimate specific values of the cogrowth function.

When doing a long ERR walk, the probability of arriving at a word of length n can be estimated by multiplying the number of words of that length by the asymptotic probability that the walk ends at a word of this length, $\pi(n)$. That is,

$$\Pr(n) \approx c_n \pi(n) = c_n \frac{(n+1)^\alpha \beta^n}{Z}.$$

The proportion of the time that the walks spends at words of length n , however, gives us another estimate of $\Pr(n)$. If we let W_n be the number of times the walk visits a word of length n then we have that

$$\Pr(n) \approx \frac{W_n}{Y},$$

where Y is equal to the length of the walk. From this we obtain

$$\frac{W_n}{Y} \approx c_n \frac{(n+1)^\alpha \beta^n}{Z}.$$

For two different values, n and m , we obtain the result

$$\frac{W_m}{W_n} \approx \frac{c_m (m+1)^\alpha \beta^m}{c_n (n+1)^\alpha \beta^n},$$

Thus,

$$c_m \approx c_n \frac{W_m}{W_n} \left(\frac{n+1}{m+1} \right)^\alpha \beta^{n-m}. \quad (8)$$

Equation 8 provides a method of estimating the value of c_m using some known or previously estimated value of c_n and the distribution of word lengths visited from an ERR walk. Let's try a quick implementation of this for Thompson's group F , where the first 48 cogrowth terms of which are known [17].

| n | W_n | n | W_n | n | W_n |
|-----|-------------|-----|------------|-----|----------|
| 0 | 32547326274 | 22 | 3360907182 | 36 | 65983874 |
| 10 | 56273373521 | 24 | 1905434239 | 38 | 37210588 |
| 12 | 31613690578 | 26 | 1121735814 | 40 | 20642387 |
| 14 | 26477475739 | 28 | 638093341 | 42 | 11332618 |
| 16 | 13576713156 | 30 | 367320461 | 44 | 6243538 |
| 18 | 9684082360 | 32 | 208025510 | 46 | 3421761 |
| 20 | 5444250723 | 34 | 118432982 | 48 | 1863477 |

Table 3: Data collected from an ERR walk of length $Y = 1.8 \times 10^{11}$ with $\alpha = 3$ and $\beta = 0.3$ on the standard presentation for Thompson’s group F .

We ran an ERR walk of length exceeding 10^{12} steps on the standard presentation (Equation 2) for $\alpha = 3$ and $\beta = 0.3$. The frequency of word length visited is shown in Table 3.

We used Equation 8 and the data in Table 3 to estimate c_{10} from c_0 , and then this estimate was used to estimate c_{12} . (Note that the shortest non-empty trivial words are of length 10. Since the relators in the standard presentation of F are even in length there are no odd length relators.) Using the data and the previous estimate for c_{n-2} , estimates were made of the first 48 terms, and these compared to the correct value in Table 4.

This implementation of Equation 8 may be refined in several ways. Firstly, in many groups we have exact initial values of c_n for more than the trivial result $c_0 = 1$. In this case these initial values can be used to estimate subsequent terms. In this paper we are primarily concerned with testing the efficacy of this method for determining cogrowth, and so do not make use of such data.

Secondly, in the above implementation the only cogrowth value used to estimate c_n was c_{n-2} . Instead, estimates for c_n may be made from c_k for any $k < n$. These estimates may then be averaged to form an estimate for c_n . Note, however, that if only one ERR walk is used, and each of the c_k is itself estimated from previous values of the same distribution there may be issues with interdependence.

This leads naturally to the following refinement — to obtain several independent estimates for a given cogrowth value several ERR walks can be run with different values for the parameters α and β .

6.1 The ERR-R algorithm.

The ERR-R algorithm accepts as input a group presentation and the cogrowth value with $c_0 = 1$. As above, recursive application of Equation 8 is used to produce estimates for longer word lengths. However, in each step previous estimates for a range of c_n are used to produce new estimates. A detailed analysis of the error incurred with each application of Equation 8 is performed in Section 6.5. All error bounds which appear in subsequent graphs are constructed using these techniques.

| n | exact | estimate | percentage error |
|-----|-------------------------|-------------------------|------------------|
| 10 | 20 | 19.9988 | .006 |
| 12 | 64 | 63.9928 | .01 |
| 14 | 336 | 335.969 | .01 |
| 16 | 1160 | 1160.23 | .02 |
| 18 | 5896 | 5893.13 | .05 |
| 20 | 24652 | 24667.2 | .06 |
| 22 | 117628 | 117588 | .03 |
| 24 | 531136 | 530650 | .09 |
| 26 | 2559552 | 2551340 | .3 |
| 28 | 12142320 | 12116600 | .2 |
| 30 | 59416808 | 59353400 | .1 |
| 32 | 290915560 | 290848000 | .02 |
| 34 | 1449601452 | 1453990000 | .3 |
| 36 | 7269071976 | 7206930000 | .8 |
| 38 | 36877764000 | 36583500000 | .8 |
| 40 | 1.8848×10^{11} | 1.8461×10^{11} | 2 |
| 42 | 9.7200×10^{11} | 9.3078×10^{11} | 4 |
| 44 | 5.0490×10^{12} | 4.7504×10^{12} | 6 |
| 46 | 2.6423×10^{13} | 2.4308×10^{13} | 8 |
| 48 | 1.3920×10^{14} | 1.245×10^{14} | 10 |

Table 4: Estimate of the first 48 terms of the cogrowth function for Thompson’s group F , constructed from an ERR walk of $Y = 1.8 \times 10^{11}$ steps with $\alpha = 3$ and $\beta = 0.3$. Exact values from [17].

Unsurprisingly, the error analysis in Section 6.5 predicts that the largest errors are incurred when data is used from the tails of random walk distributions. Ideally then, a separate walk should be run for each c_n , with parameters α and β chosen so that the sampled word lengths occupy the peaks of the distribution. If many estimates are to be made this is computationally infeasible. Instead we performed ERR walks using a range of α and β values, which can be chosen so that all word lengths of interest are visited often.

When estimating c_m , one estimate was made from each walk distribution and from each c_n , $m - 100 < n < m$. To avoid using the tails of distributions only data points which were greater than 10% of the maximum height were used.

Using Equation 9 each estimate was assigned a weight equal to the inverse of the estimated error. The final value for c_m was taken as the weighted average of the estimates, and the error in c_m was taken to be the weighted average of the individual error estimates.

Random walk data was obtained as before using the Python code of the second author as described in Remark 4.2.

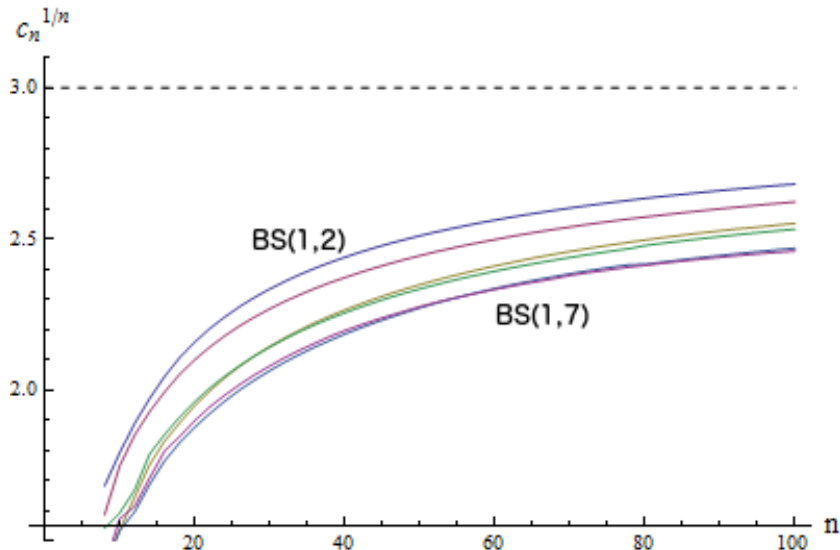


Figure 5: Estimates for $c_n^{1/n}$ for the groups $BS(1, N)$, $N = 2, \dots, 7$. As N increases the curves takes longer to approach the asymptote.

6.2 Application to the examples in Section 5

Applying the ERR-R algorithm can be used to analyse in more detail the pathological behaviours analysed above. Unsurprisingly, for the presentations of the trivial group given in 5.1 which ignore the long relator, the ERR-R estimates for cogrowth values align closely with the three strand braid group. For $BS(1, N)$ we can use estimates of initial cogrowth to analyse how \mathcal{R} increases with N . This is shown, for example in Figure 5 which exhibits the behaviour predicted by the convergence to $\mathbb{Z} \wr \mathbb{Z}$ in the space of marked groups. Further analysis of these presentations appears in [27].

6.3 Application to surface group

The fundamental group of a surface of genus 2 has presentation $\langle a, b, c, d \mid [a, b][c, d] \rangle$. The cogrowth of this group has received a lot of attention, and good upper and lower bounds are known for the asymptotic rate of growth [14, 23].

ERR walks were run on this surface group with $\alpha = 3, 30, 300$ and $\beta = 0.281, 0.286, 0.291, \dots, 0.351$. Estimates were made for c_n as well as the error Δc_n . The resultant upper and lower bounds for $c_n^{1/n}$ are shown in Figure 6.

6.4 Application to Thompson's group F

We now apply the more sophisticated implementation of the method to F . We again compare the first 48 values with exact values obtained by Haagerup

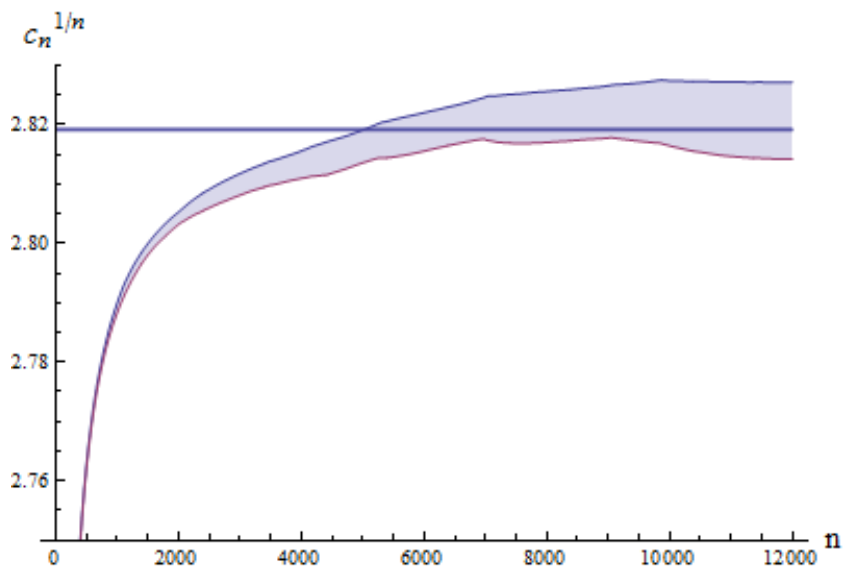


Figure 6: Upper and lower bounds for the n -th root of the cogrowth function for the fundamental group of a surface of genus 2 as calculated from ERR walks. The horizontal lines (indistinguishable at this scale) identify the known upper and lower bounds. Note that after 12000 recursive applications of Equation 8 the error in the n -th root is still only approximately 0.01.

| n | exact | estimate | error (%) | predicted error (%) |
|-----|-------------------------|--------------------------|-----------|---------------------|
| 10 | 20 | 19.9996 | 0.002 | .03 |
| 12 | 64 | 63.9981 | 0.003 | 0.06 |
| 14 | 336 | 335.999 | 0.0002 | 0.07 |
| 16 | 1160 | 1159.96 | 0.003 | 0.1 |
| 18 | 5896 | 5895.98 | 0.0003 | 0.1 |
| 20 | 24652 | 24653.1 | 0.005 | 0.1 |
| 22 | 117628 | 117625 | 0.003 | 0.2 |
| 24 | 531136 | 531098 | 0.007 | 0.2 |
| 26 | 2559552 | 2558950 | 0.02 | 0.2 |
| 28 | 12142320 | 12138200 | 0.03 | 0.3 |
| 30 | 59416808 | 59408300 | 0.01 | 0.3 |
| 32 | 290915560 | 290861000 | 0.02 | 0.3 |
| 34 | 1449601452 | 1449260000 | 0.02 | 0.3 |
| 36 | 7269071976 | 7268550000 | 0.007 | 0.4 |
| 38 | 36877764000 | 36876700000 | 0.003 | 0.5 |
| 40 | 1.8848×10^{11} | 1.88491×10^{11} | 0.003 | 0.5 |
| 42 | 9.7200×10^{11} | 9.7205×10^{11} | 0.005 | 0.5 |
| 44 | 5.0490×10^{12} | 5.05097×10^{12} | 0.04 | 0.6 |
| 46 | 2.6423×10^{13} | 2.64353×10^{13} | 0.05 | 0.6 |
| 48 | 1.3920×10^{14} | 1.39246×10^{14} | 0.03 | 0.7 |

Table 5: Estimate of the first 48 terms of the cogrowth function for Thompson’s group F , constructed from 60 ERR walks. Exact values from [17].

et al.. ERR walks were run on F with $\alpha = 3, 13, 23, 33, 53, 63$ and $\beta = 0.28, 0.29, \dots, 0.37$. Table 5 shows comparisons between estimates for $c_n^{1/n}$ and the actual values, for $n \leq 48$, as well as the estimates for the error obtained from the experimental data.

Remark 6.1. Table 5 shows a marked increase in the degree of accuracy of the estimates over those of Table 4. This suggests the method of using multiple distributions and weighted averages is effective. Note that there are approximately 10^{12} trivial words of length 48 so the walks could not possibly have visited each one. The sample of words visited by the walk seem to reflect the space as a whole reasonably accurately.

Our method allows us to go much further than this though. Figure 7 shows our estimates for upper and lower bounds of $c_n^{1/n}$ for $n \leq 2000$.

6.5 Error analysis

Here we identify a method by which error in cogrowth estimates may be estimated. We stress that this is a statistical measurement of error, rather than theoretical.

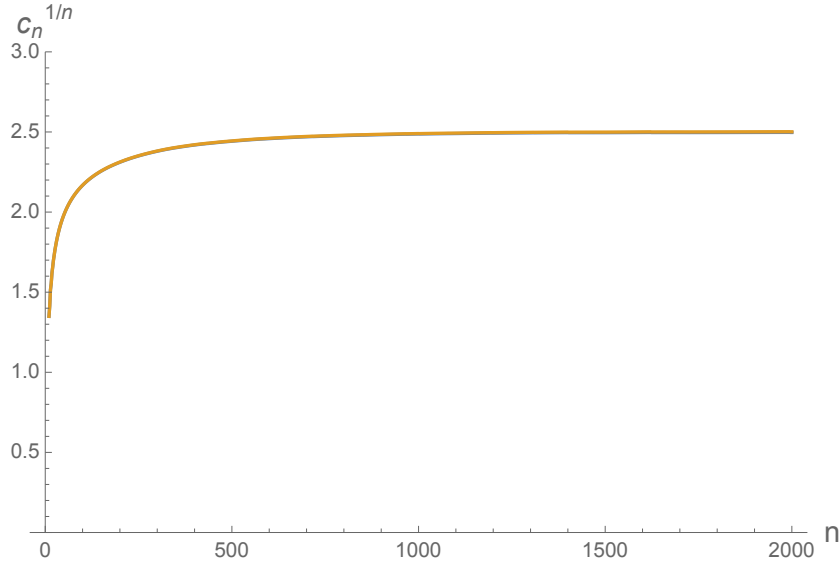


Figure 7: Estimates of $c_n^{1/n}$ for Thompsons group F for $n \leq 2000$, using the ERR-R method. The figure includes upper and lower bounds, but at this scale the estimated error is too small for the bounds to be distinguished.

Recall Equation 8. Suppose that c_n is known up to $\pm\Delta c_n$, and that the error in the measurements W_m and W_n are $\pm\Delta W_m$ and $\pm\Delta W_n$ respectively. Then, from elementary calculus, the error in c_m is given by

$$\begin{aligned}
\Delta c_m &\approx \frac{W_m}{W_n} \left(\frac{n+1}{m+1} \right)^\alpha \beta^{n-m} \Delta c_n \\
&\quad + \frac{c_n}{W_n} \left(\frac{n+1}{m+1} \right)^\alpha \beta^{n-m} \Delta W_m \\
&\quad + c_n \frac{W_m}{W_n^2} \left(\frac{n+1}{m+1} \right)^\alpha \beta^{n-m} \Delta W_n \\
&= c_n \frac{W_m}{W_n} \left(\frac{n+1}{m+1} \right)^\alpha \beta^{n-m} \left(\frac{\Delta c_n}{c_n} + \frac{\Delta W_m}{W_m} + \frac{\Delta W_n}{W_n} \right) \\
&\approx c_m \left(\frac{\Delta c_n}{c_n} + \frac{\Delta W_m}{W_m} + \frac{\Delta W_n}{W_n} \right). \tag{9}
\end{aligned}$$

Hence the proportional error in the estimate of c_m is approximately equal to the sum of the proportional errors in c_n , W_m and W_n . It is clear from this that if Equation 8 is used recursively (building new estimates based on previously estimated cogrowth values) the proportional error in c_n is certain to increase. Note, the factor controlling the rate of growth in the proportional error of estimates is the proportional error in ΔW_n . If this is constant as n increases the proportional error in c_n will grow linearly with n .

To calculate useful error margins for c_n it is necessary to quantify ΔW_n . Here we employ the same method used in [9]; walks are split into M segments and the number of times the walk visits words of length n is recorded for each segment. Let $x_{i,n}$ denote the number of times the walk visited words of length n in the i th segment. Then W_n is taken to be the average of $x_{i,n}$ for $i = 1, \dots, M$ and the error in W_n is calculated from the statistical variance of these values,

$$\Delta W_n = \sqrt{\frac{\text{Var}\{x_{i,n}\}_{1 \leq i \leq M}}{M-1}}. \quad (10)$$

Example 6.2. Equations 9 and 10 were used to produce the estimates of the error in the estimates contained in Table 5. Note that the estimated error is much larger than the actual error.

6.6 Error in the n -th root of c_n

We have noted that recursive uses of Equation 8 will result in an increasing proportional error in c_n . However, it is the n -th root of c_n which reflects the amenability of a group. Let $\gamma_n = c_n^{1/n}$ and $\Delta\gamma_n$ denote the error of the estimate for γ_n . Once again from elementary calculus we obtain that for a given n

$$\begin{aligned} \Delta\gamma_n &\approx \frac{1}{n} c_n^{\frac{1}{n}-1} \Delta c_n \\ &= \frac{1}{n} c_n^{\frac{1}{n}} \frac{\Delta c_n}{c_n} \\ &= \gamma_n \frac{1}{n} \frac{\Delta c_n}{c_n} \end{aligned}$$

and so $\frac{\Delta\gamma_n}{\gamma_n} \approx \frac{1}{n} \frac{\Delta c_n}{c_n}$. (11)

Thus, if $\frac{\Delta c_n}{c_n}$ increases at most linearly, $\frac{\Delta\gamma_n}{\gamma_n}$ can be expected to remain constant.

The values for c_n grow exponentially, so a linearly increasing proportional error in c_n corresponds with a massive increase in the absolute error in c_n . In contrast, γ_n approaches a constant, so the proportional error depends linearly on the absolute error. Thus it is not surprising that our experimental results show that even when the error in cogrowth estimates grows large, the error in the n -th root grows very slowly.

7 Conclusion

Several ideas emerge from this study. Firstly, researchers performing experimental mathematics to determine the amenability of a group need to take care that their algorithm is not susceptible to interference from sub-dominant behaviours. For the reduced-cogrowth function the sub-dominant behaviour is

identified by \mathcal{R} . Amenability is an asymptotic property, and the interference of sub-dominant behaviours on experimental algorithms can be subtle and nuanced. In particular, we have shown that, if Thompson’s group F is amenable, its function \mathcal{R} grows faster than any polynomial. This implies that the prediction of non-amenability of F in [9] is unreliable.

We have also shown that, despite potential inaccuracies in estimates of asymptotics, the ERR-R method can produce accurate results for initial cogrowth values. These are interesting in their own right. Indeed, if Thompson’s group is not amenable, then its \mathcal{R} function need not be super-polynomial and results from experimental methods might well inform the construction of conjectures regarding cogrowth.

In this context the original benefits of the ERR algorithm still stand: it requires no group theoretic computational software, no solution to the word problem, and remains a computationally inexpensive way to quickly gain insight into the cogrowth function of a finitely presented group.

Acknowledgements

The authors wish to thank Andrew Reznitser, Andrew Elvey-Price, and the anonymous reviewer for helpful feedback on this work.

References

- [1] G. N. Arzhantseva, V. S. Guba, M. Lustig, and J. Préaux. Testing Cayley graph densities. *Ann. Math. Blaise Pascal*, 15(2):233–286, 2008.
- [2] J. M. Belk and K. S. Brown. Forest diagrams for elements of Thompson’s group F . *Internat. J. Algebra Comput.*, 15(5-6):815–850, 2005.
- [3] A. Bendikov, C. Pittet, and R. Sauer. Spectral distribution and L^2 -isoperimetric profile of Laplace operators on groups. *Math. Ann.*, 354(1):43–72, 2012.
- [4] M. G. Benli, R. I. Grigorchuk, and P. de la Harpe. Amenable groups without finitely presented amenable covers. *Bull. Math. Sci.*, 3(1):73–131, 2013.
- [5] J. Burillo, S. Cleary, and B. Wiest. Computational explorations in Thompson’s group F . In *Geometric group theory*, Trends Math., pages 21–35. Birkhäuser, Basel, 2007.
- [6] J. W. Cannon, W. J. Floyd, and W. R. Parry. Introductory notes on Richard Thompson’s groups. *Enseign. Math. (2)*, 42(3-4):215–256, 1996.
- [7] T. G. Ceccherini-Silberstein and R. I. Grigorchuk. Amenability and growth of one-relator groups. *Enseign. Math. (2)*, 43(3-4):337–354, 1997.

- [8] J. M. Cohen. Cogrowth and amenability of discrete groups. *J. Funct. Anal.*, 48(3):301–309, 1982.
- [9] M. Elder, A. Reznitzner, and E. J. Janse van Rensburg. Random sampling of trivial words in finitely presented groups. *Exp. Math.*, 24(4):391–409, 2015.
- [10] M. Elder, A. Reznitzner, and T. Wong. On the cogrowth of Thompson’s group F . *Groups Complex. Cryptol.*, 4(2):301–320, 2012.
- [11] A. Erschler. On isoperimetric profiles of finitely generated groups. *Geom. Dedicata*, 100:157–171, 2003.
- [12] P. Flajolet and R. Sedgewick. *Analytic Combinatorics*. Cambridge University Press, 2009.
- [13] E. Følner. On groups with full Banach mean value. *Math. Scand.*, 3:243–254, 1955.
- [14] S. Gouëzel. A numerical lower bound for the spectral radius of random walks on surface groups. *Combin. Probab. Comput.*, 24(6):838–856, 2015.
- [15] R. I. Grigorchuk. Symmetrical random walks on discrete groups. In *Multi-component random systems*, volume 6 of *Adv. Probab. Related Topics*, pages 285–325. Dekker, New York, 1980.
- [16] V. S. Guba and M. V. Sapir. On subgroups of the R. Thompson group F and other diagram groups. *Mat. Sb.*, 190(8):3–60, 1999.
- [17] S. Haagerup, U. Haagerup, and M. Ramirez-Solano. A computational approach to the Thompson group F . *Internat. J. Algebra Comput.*, 25(3):381–432, 2015.
- [18] H. Kesten. Symmetric random walks on groups. *Trans. Amer. Math. Soc.*, 92:336–354, 1959.
- [19] D. Kouksov. On rationality of the cogrowth series. *Proc. Amer. Math. Soc.*, 126(10):2845–2847, 1998.
- [20] P. H. Kropholler and A. Martino. Graph-wreath products and finiteness conditions. *J. Pure Appl. Algebra*, 220(1):422–434, 2016.
- [21] D. Kuksov. Cogrowth series of free products of finite and free groups. *Glasg. Math. J.*, 41(1):19–31, 1999.
- [22] J. T. Moore. Fast growth in the Følner function for Thompson’s group F . *Groups Geom. Dyn.*, 7(3):633–651, 2013.
- [23] T. Nagnibeda. An upper bound for the spectral radius of a random walk on surface groups. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 240(Teor. Predst. Din. Sist. Komb. i Algoritm. Metody. 2):154–165, 293–294, 1997.

- [24] C. Pittet and L. Saloff-Coste. On the stability of the behavior of random walks on groups. *J. Geom. Anal.*, 10(4):713–737, 2000.
- [25] C. Pittet and L. Saloff-Coste. On random walks on wreath products. *Ann. Probab.*, 30(2):948–977, 2002.
- [26] C. Pittet and L. Saloff-Coste. Random walks on finite rank solvable groups. *J. Eur. Math. Soc. (JEMS)*, 5(4):313–342, 2003.
- [27] C. Rogers. *On the interconnectedness, via random walks, of cogrowth rates and the Følner function*. PhD thesis, The University of Newcastle, 2016.
- [28] R. Szwarc. A short proof of the Grigorchuk-Cohen cogrowth theorem. *Proc. Amer. Math. Soc.*, 106(3):663–665, 1989.
- [29] A. M. Vershik. Amenability and approximation of infinite groups. *Selecta Math. Soviet.*, 2(4):311–330, 1982. Selected translations.
- [30] W. Woess. Cogrowth of groups and simple random walks. *Arch. Math. (Basel)*, 41(4):363–370, 1983.