Random presentations and random subgroups: a survey

Frédérique Bassino, Cyril Nicaud, Pascal Weil

Random presentations and random subgroups*

Frédérique Bassino, bassino@lipn.univ-paris13.fr
Univ. Paris 13, Sorbonne Paris Cité, LIPN, CNRS UMR7030, F-93430 Villetaneuse, France

Cyril Nicaud, nicaud@univ-mlv.fr
Université Paris-Est, LIGM (UMR 8049), CNRS, ENPC, ESIEE Paris, UPEM,
F-77454, Marne-la-Vallée, France

Pascal Weil, pascal.weil@labri.fr
Univ. Bordeaux, LaBRI, CNRS UMR 5800, F-33400 Talence, France†

Contents

*Part of this work was done while the third author was Ada Peluso Visiting Professor, Mathematics Department, Hunter College, CUNY. He also received support from the ANR (project DeLTA, ANR-16-CE40-0007).

†LaBRI, Univ. Bordeaux, 351 cours de la Libération, 33400 Talence, France.

1 Introduction

In infinite group theory, it is a classical and natural question to ask what most groups look like, what a random group looks like. The question can and must be made more precise: it is actually a question about random finitely presented groups, and in most of the literature, in fact a question about random finite group presentations on a fixed set of generators. The specific questions may be whether a random finite group presentation satisfies a small cancellation property, whether the group it presents is hyperbolic, residually finite, etc.

Early on, Gromov gave an answer to this question: almost all groups are hyperbolic (see [28], and [49, 13, 45] for precise statements and complete proofs).

When a group G is fixed (*e.g.*, the free group $F(A)$ over a given finite set A of generators, a hyperbolic group, a braid group, the modular group), one may also ask what a random finitely generated subgroup looks like: is it free? is it malnormal? does it have finite index? in the case where $G = F(A)$, is the subgroup Whitehead minimal?

These questions have been abundantly studied in the literature. This chapter is a partial survey and as such, it contains no new results, but it offers a synthetic view of a part of this very active field of research. We also include a small number of proofs, in full or only sketched. We refer the reader to the survey by Ollivier [46] for more details on some of the topics discussed here, and to the survey by Dixon [16] for a discussion of probabilistic methods in finite group theory.

A specific aspect of the present survey is that we discuss both random presentations and random subgroups, unlike Ollivier [46].

Random presentations were considered first in the literature, and we will start with them as well (Section 2). We then proceed to a discussion of results on random subgroups (Section 3). Finally, Section 4 discusses recent results on non-uniform distributions.

This is an updated and extended version of a talk given at the conference GAGTA-8 in Newcastle, NSW, in July 2014.

1.1 Discrete representations

The very notion of randomness relies on a notion of probability, and in many cases, on a notion of counting discrete representations of finitely presented groups, or finitely generated subgroups, of a certain size: how many subgroups of $F(A)$ are there, with a tuple of $f(n)$ generators of length at most n for a given function f? how many whose Stallings graph (see Section 3.1) has at most n vertices? how many isomorphism classes of 1-relator groups are there, whose relator has length at most n? etc.

So we must first discuss the discrete representations we will use to describe subgroups and presentations.

Let A be a finite non-empty set and let $F(A)$ be the *free group on* A. The symmetrized alphabet $\tilde{A}$ is the union of A and a copy of A, $\bar{A} = \{\bar{a} \mid a \in A\}$, disjoint from A. We denote by $\tilde{A}^*$ the set of all words on the alphabet $\tilde{A}$. The operation $x \mapsto \bar{x}$ is extended to $\tilde{A}^*$ by letting $\bar{\bar{a}} = a$ and $\overline{ua} = \bar{a}\bar{u}$ for all $a \in \tilde{A}$ and $u \in \tilde{A}^*$. Recall that a word is *reduced* if it does not contain a factor of the form $a\bar{a}$ ($a \in \tilde{A}$). The (free group) *reduction* of a word $u \in \tilde{A}^*$ is the word $\rho(u)$ obtained from u by iteratively deleting factors of the form $a\bar{a}$ ($a \in \tilde{A}$). We can then think of $F(A)$ as the set of reduced words on $\tilde{A}$: the product in $F(A)$ is given by $u \cdot v = \rho(uv)$, and the inverse of u is $\bar{u}$.

In the sequel, we fix a finite set A, with cardinality $r > 1$. If $n \in \mathbb{N}$, we denote by $\mathcal{R}_n$ (resp. $\mathcal{R}_{\leq n}$) the set of reduced words of length n (resp. at most n). A reduced word u is called *cyclically reduced* if $u^2 = uu$ is reduced, and we let $\mathcal{CR}_n$ (resp. $\mathcal{CR}_{\leq n}$) be the set of cyclically reduced words of length n (resp. at most n). If u is a reduced word, there exist uniquely defined words v, w such that w is cyclically reduced and $u = v^{-1}wv$. Then w is called the *cyclic reduction* of u, written $\kappa(u)$.

It is easily verified that

$$|\mathcal{R}_n| = 2r(2r-1)^{n-1} \quad \text{and} \quad 2r(2r-1)^{n-2}(2r-2) \leq |\mathcal{CR}_n| \leq 2r(2r-1)^{n-1},$$
$$|\mathcal{R}_{\leq n}| = \Theta((2r-1)^n) \quad \text{and} \quad |\mathcal{CR}_{\leq n}| = \Theta((2r-1)^n).$$

If $\vec{h} = (h_1, \cdots, h_k)$ is a tuple of elements of $F(A)$, we denote by $\langle \vec{h} \rangle$ the *subgroup of* $F(A)$ *generated by* $\vec{h}$: it is the set of all products of the elements of $\vec{h}$ and their inverses. And we denote by $\langle\langle \vec{h} \rangle\rangle$ the normal closure of $\langle \vec{h} \rangle$, namely the set of all products of conjugates $h_i^g = g^{-1}h_i g$ of the elements of $\vec{h}$ ($1 \leq i \leq k$, $g \in F(A)$) and

their inverses. The *group presented by the relators* $\vec{h}$, written $\langle A \mid \vec{h} \rangle$, is the quotient $F(A)/\langle\langle \vec{h} \rangle\rangle$.

If $\vec{h} = (h_1, \ldots, h_k)$ and $\kappa(\vec{h}) = (\kappa(h_1), \ldots, \kappa(h_k))$, then $\vec{h}$ and $\kappa(\vec{h})$ present the same group: that is, $\langle A \mid \vec{h} \rangle = \langle A \mid \kappa(\vec{h}) \rangle$. It is therefore customary, when considering group presentations, to assume that the relators are all cyclically reduced.

In general, if there exists a surjective morphism $\mu\colon F(A) \to G$, we say that G is *A-generated*. Then a word $u \in F(A)$ is called *geodesic* if it has minimum length in $\mu^{-1}(\mu(u))$.

Properties of interest for subgroups of G are, for instance, whether they are free or quasi-convex. Recall that a subgroup H of G is *quasi-convex* if there exists a constant $k > 0$ such that, for every geodesic word $u = a_1 \cdots a_n$ such that $\mu(u) \in H$, and for every $1 \leq i \leq n$ there exists a word v_i of length at most k such that $\mu(a_1 \cdots a_i v_i) \in H$. We observe that while being geodesic is a word property which depends on the chosen set of generators for the group, being quasi-convex is an intrinsic property of the subgroup, which is preserved when we consider a different finite set of generators for G.

We are also interested in malnormality and purity: a subgroup H is *almost malnormal* (resp. *malnormal*) if $H^g \cap H$ is finite (resp. trivial) for every $g \notin H$. Moreover, H is *almost pure* (resp. *pure*, also known as *isolated* or *closed under radical*) if $x^n \in H$ implies $x \in H$ for any $n \neq 0$ and any element $x \in G$ of infinite order (resp. any $x \in G$). Note that malnormality and almost malnormality (resp. purity and almost purity) are equivalent in torsion-free groups. It is easily verified that an almost malnormal (resp. malnormal) subgroup is almost pure (resp. pure).

It is a classical result that every finitely generated subgroup of a free group is free (Nielsen [44]) and quasi-convex (Gromov [27]). In addition, it is decidable whether a finitely generated subgroup of a free group is malnormal [8] and whether it is pure [10], see Section 3.1. In contrast, these properties are not decidable in a general finitely presented group, even if hyperbolic [12]. Quasi-convexity is also not decidable in general, even in hyperbolic or small cancellation groups [50]. Almost malnormality is however decidable for quasi-convex subgroups of hyperbolic groups [38, Corollary 6.8].

Finally, let us mention the property of Whitehead minimality for finitely generated subgroups of free groups: we say that H is *Whitehead minimal* if it has minimum size in its automorphic orbit, where the size of a subgroup is defined in terms of its Stallings graph, see Section 3.1 below. In the case of a cyclic subgroup $H = \langle u \rangle$, if $u = v^{-1}\kappa(u)v$, then the size of H is $|v| + |\kappa(u)|$. Whitehead minimality plays an important role in the solution of the automorphic orbit problem, to decide whether two subgroups are in the same orbit under the automorphism group of $F(A)$, see

[22, 31].

For group presentations, the emphasis can be on combinatorial properties of the presentation, such as small cancellation properties, or on the geometric properties of the given presented group, typically hyperbolicity. One of the main small cancellation properties is Property $C'(\lambda)$ (for some $0 < \lambda < 1$), which is defined as follows. A *piece* in a tuple $\vec{h}$ of cyclically reduced words is a word u which occurs as a prefix of two distinct elements of the set of cyclic conjugates of the elements of $\vec{h}$ and their inverses. For instance, $a\bar{b}a$ is a piece of $\vec{h} = (\bar{a}\bar{a}bb, b\underline{a\bar{b}a}b, \underline{\bar{a}}b\underline{\bar{a}b})$. A finite presentation $\langle A \mid \vec{h}\rangle$ *satisfies the small cancellation property* $C'(\lambda)$ if a piece u in $\vec{h} = (h_1, \ldots, h_k)$ satisfies $|u| < \lambda|h_i|$ for every i such that u is a prefix of a cyclic conjugate of h_i. This is an important property since it is well known that if $\vec{h}$ has Property $C'(\frac{1}{6})$, then the group $\langle A \mid \vec{h}\rangle$ is hyperbolic [27, 0.2.A]. An elegant generalization is due to Ollivier [47]. Other small cancellation properties are discussed in Section 2.1.

1.2 Models of randomness

In this chapter, the general model of randomness on a set S which we will consider, consists in the choice of a sequence $(\mathbb{P}_n)_n$ of probability laws on S. For instance, the set S could be the set of all k-relator presentations (for a fixed value of k), that is, the set of all k-tuples of cyclically reduced words, and the law $\mathbb{P}_n$ could be the uniform probability law with support the presentations where every relator has length at most n.

This general approach covers the classical models considered in the literature, such as the Arzhantseva-Ol'shanskiĭ model [3] or Gromov's density model [28]. It also allows us to consider probability laws that do not give equal weight to words of equal length, see Section 4 below.

A subset X of S is *negligible* if the probability for an element of S to be in X, tends to 0 when n tends to infinity; that is, if $\lim_n \mathbb{P}_n(X) = 0$. If this sequence converges exponentially fast (that is: $\mathbb{P}_n(X)$ is $\mathcal{O}(e^{-cn})$ for some $c > 0$), we say that X is *exponentially negligible*. The set X is *generic* (resp. *exponentially generic*) if its complement is negligible (resp. exponentially negligible).

2 Random finite presentations

2.1 The density model

The density model was introduced by Gromov [28]. Let $0 < d < 1$ be a real number. In the density d model, the set S (with reference to the notation in Section 1.2) is the set of all finite tuples of cyclically reduced words and $\mathbb{P}_n$ is the uniform probability law with support the set of $|\mathcal{CR}_n|^d$-tuples of elements of $\mathcal{CR}_n$. We say that a property is generic (resp. negligible) *at density d* if it is generic (resp. negligible) in the density d model.

In this model, small cancellation properties are generic at low enough density. For Property $C'(\lambda)$ ($0 < \lambda < 1$), we have an interesting, so-called *phase transition* statement ([28, 9.B], see also [46, Section I.2.a]).

Theorem 2.1 *Let $0 < d < 1$ and $0 < \lambda < \frac{1}{2}$. If $d < \frac{\lambda}{2}$, then at density d, a random finite presentation exponentially generically satisfies Property $C'(\lambda)$. If instead $d > \frac{\lambda}{2}$, then at density d, a random finite presentation exponentially generically does not satisfy $C'(\lambda)$.*

Proof. To lighten notation, we let $\rho = 2r - 1$, $\alpha = \frac{\rho-1}{\rho}$ and $\beta = \frac{\rho+1}{\rho}$. We saw in Section 1.1 that $|\mathcal{CR}_n| = c_n\rho^n$, with $\alpha\beta \leq c_n \leq \beta$. Let $\ell = \lambda n$.

A reduced word u of length ℓ is a prefix of a cyclic conjugate of $w \in \mathcal{CR}_n$ if either $w = u_2w_1u_1$ with $u_1u_2 = u$, or $w = w_1uw_2$ with $|w_1|, |w_2| > 0$. Let a and b be the first and last letters of u. For fixed values of u_1, u_2, the number of words w_1 such that $u_2w_1u_1 \in \mathcal{CR}_n$ (that is, w_1 is reduced, does not start with $\bar{b}$ and does not end with $\bar{a}$) is of the form $c'_{n,\ell}(a,b)\rho^{n-\ell}$, with $\alpha \leq c'_{n,\ell}(a,b) \leq 1$. Similarly, for $0 < \ell_1 < n - \ell$, the number of pairs (w_1, w_2) such that $|w_1| = \ell_1$ and $w_1uw_2 \in \mathcal{CR}_n$ is of the form $c''_{n,\ell_1,\ell}(a,b)\rho^{n-\ell}$, with $\alpha \leq c''_{n,\ell_1,\ell}(a,b) \leq 1$. Thus the probability $p_n(u)$ that a word of $\mathcal{CR}_n$ contains u as a piece is bounded above by

$$\begin{aligned} p_n(u) &\leq \frac{2}{|\mathcal{CR}_n|}\left(\sum_{\ell_1=1}^{n-\ell-1} c''_{n,\ell_1,\ell}(a,b)\rho^{n-\ell} + \ell\ c'_{n,\ell}(a,b)\rho^{n-\ell}\right) \\ &\leq \frac{2n\rho^{n-\ell}}{\alpha\beta\rho^{-n}} = \frac{2n}{\alpha\beta}\,\rho^{-\ell}. \end{aligned}$$

It follows that the probability that a word of length ℓ is a piece of at least two distinct words in a $|\mathcal{CR}_n|^d$-tuple of cyclically reduced words of length n is at most

$$\binom{|\mathcal{CR}_n|^d}{2}\sum_{u\in\mathcal{R}_\ell} p_n(u)^2 \leq \frac{\beta^2}{2}\rho^{2dn}\beta\rho^\ell\left(\frac{2n}{\alpha\beta}\,\rho^{-\ell}\right)^2 = \frac{2\beta}{\alpha^2}n^2\rho^{(2d-\lambda)n},$$

which vanishes exponentially fast if $2d < \lambda$.

Bounding the probability that u occurs twice as a piece in the same component as a tuple is technically more complicated, and we refer to [6, Theorem 3.20] for the details, discussed there in a more general situation. A brief summary is as follows: this double occurrence can arise because there are two non-overlapping occurrences of u, or one of u and one of u^{-1}, or because there are overlapping occurrences of u (u and u^{-1} cannot overlap). The non-overlapping situations lead to a probability with an upper bound of the form

$$|\mathcal{CR}_n| \sum_{u \in \mathcal{R}_\ell} \kappa n^2 \rho^{-2\ell} \leq \kappa' n^2 \rho^{(d-\lambda)n}$$

where κ, κ' are appropriate constants, see the proof of [6, Theorem 3.20] for a more general statement. The overlapping situation is more delicate to analyze, and it leads to an upper bound of the form

$$|\mathcal{CR}_n| \kappa'' n \rho^{\ell} \leq \kappa''' n \rho^{(d-\lambda)n}$$

for appropriate constants κ'', κ'''. At density $d < \frac{\lambda}{2}$, we have $d - \lambda < 0$, so both these probabilities vanish exponentially fast. Thus, at density less than $\frac{\lambda}{2}$, Property $C'(\lambda)$ holds exponentially generically.

Now let us assume that $d > \frac{\lambda}{2}$. In a variant of the birthday paradox, we show that, exponentially generically, two words in a random $|\mathcal{CR}_n|^d$-tuple of elements of $\mathcal{CR}_n$ have the same length ℓ prefix. Indeed, if u has length ℓ and first and last letters a and b, the number of words in $\mathcal{CR}_n$ which start with u is $c'_{n,\ell}(a,b)\rho^{n-\ell} \geq \alpha \rho^{n-\ell}$. If $u_1, \ldots, u_N$ are pairwise distinct reduced words of length ℓ, the number of elements $w \in \mathcal{CR}_n$ that starts with none of these words is greater than or equal to $N\alpha\rho^{n-\ell}$. It follows that the number of N-tuples of words in $\mathcal{CR}_n$ with pairwise distinct length ℓ prefixes is at most equal to

$$|\mathcal{CR}_n|(|\mathcal{CR}_n| - \alpha\rho^{n-\ell})(|\mathcal{CR}_n| - 2\alpha\rho^{n-\ell}) \ldots (|\mathcal{CR}_n| - (N-1)\alpha\rho^{n-\ell}).$$

Thus the probability p_N that an N-tuple of elements of $\mathcal{CR}_n$ all have distinct length ℓ prefixes satisfies

$$\begin{aligned} p_N &\leq (1 - \beta^{-1}\rho^{-\ell})(1 - 2\beta^{-1}\rho^{-\ell}) \ldots (1 - (N-1)\beta^{-1}\rho^{-\ell}) \\ &\leq \exp\left(-\beta^{-1}\rho^{-\ell} - 2\beta^{-1}\rho^{-\ell} - \cdots - (N-1)\beta^{-1}\rho^{-\ell}\right) \\ &\leq \exp\left(-\beta^{-1}\frac{N(N-1)}{2}\rho^{-\ell}\right). \end{aligned}$$

For $N = |\mathcal{CR}_n|^d$, we find that $N(N-1) \geq (\alpha\beta)^{2d}\rho^{2dn} - \beta^d\rho^{dn}$, which is greater than $\frac{(\alpha\beta)^{2d}}{2}\rho^{2dn}$ for n large enough. It follows that

$$p_N \leq \exp\left(-\frac{(\alpha\beta)^{2d}}{4\beta}\rho^{(2d-\lambda)n}\right),$$

which vanishes exponentially fast if $2d > \lambda$. □

As noted earlier, if $\vec{h}$ satisfies Property $C'(\frac{1}{6})$, then the group $\langle A \mid \vec{h}\rangle$ is hyperbolic but the condition is not necessary. Theorem 2.1 shows that, in the density model and up to density $\frac{1}{12}$, a finitely presented group is exponentially generically hyperbolic. Yet the property holds for higher densities, and we have another phase transition theorem.

Let us say that a finitely presentated group $G = \langle A \mid \vec{h}\rangle$, where $\vec{h}$ consists of cyclically reduced words of equal length n, is *degenerate* if G is trivial, or if G is the 2-element group and n is even. Then we have the following result, again a phase transition theorem, due to Ollivier [46].

Theorem 2.2 *Let $0 < d < 1$. If $d < \frac{1}{2}$, then at density d, a random finite presentation exponentially generically presents an infinite hyperbolic group. If instead $d > \frac{1}{2}$, then at density d, a random finite presentation exponentially generically presents a degenerate group.*

The proof of the statement in Theorem 2.2 about density greater than $\frac{1}{2}$ reduces to counting arguments on words in the spirit of the proof of Theorem 2.1 (see Section 4 for a generalization). The proof that hyperbolicity is generic at densities between $\frac{1}{12}$ and $\frac{1}{2}$ — that is: greater than the critical value for Property $C'(\frac{1}{6})$ —, is more complex and involves the combinatorics of van Kampen diagrams. An example of such a diagram is given on Figure 1; for a formal definition, the reader is referred to [40].

Remark 2.3 A natural variant of the density model considers tuples of words of length at most n, instead of words of length exactly n. More precisely, $\mathbb{P}_n$ is chosen to be the uniform probability law with support the set of $|\mathcal{CR}_{\leq n}|^d$-tuples of words in $\mathcal{CR}_{\leq n}$. Ollivier shows in [46] that Theorems 2.1 and 2.2 also hold for this model. □

Remark 2.4 The statement on hyperbolicity in Theorem 2.2 has an important predecessor. For a fixed number k of relators and a fixed k-tuple of lengths $(\ell_1, \ldots, \ell_k)$, consider the finite presentations with k relators of length, respectively, $\ell_1, \ldots, \ell_k$.

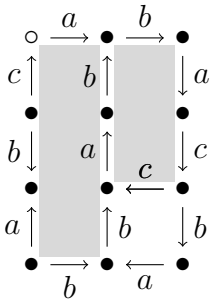

Figure 1: Informally, a van Kampen diagram is a planar finite cell complex with a specific embedding in the plane. Its edges (1-cells) are directed and labelled by letters in A, and the boundary of each face (2-cell) is cyclically labelled by a relator. There is one distinguished vertex (0-cell). On the left, an example of such a diagram for $\vec{h} = (b^2ac^2a, bab\bar{c}, b^2ab\bar{a}\bar{c}b\bar{a})$ of area 3, showing that $abacba\bar{a}\bar{b}c = 1$ in the group presented by $\vec{h}$. The two grey faces share the segment ab.

The probability that such a presentation presents an infinite hyperbolic group, tends exponentially fast to 1 when $\min(\ell_i)_{1\leq i\leq k}$ tends to infinity (while k remains fixed). This was originally stated by Gromov [27], and proved by Champetier [13] and Ol'shanskiĭ [49]. □

The small cancellation property $C'(\lambda)$ for a tuple of cyclically reduced words $\vec{h}$ can be interpreted geometrically as follows: in any reduced van Kampen diagram (w.r.t. the presentation $\langle A \mid \vec{h}\rangle$), a segment of consecutive edges in the boundary between two adjacent faces f and f' (namely, in $\partial f \cap \partial f'$) has length at most $\lambda \min(|\partial f|, |\partial f'|)$. *Greendlinger's property* (as interpreted by Ollivier [48]) is of the same nature: it states that in any reduced van Kampen diagram D with more than one face, there exist two faces f and f' for which there are segments of consecutive edges of $\partial f \cap \partial D$ (resp. $\partial f' \cap \partial D$) of length at least $\frac{1}{2}|\partial f|$ (resp. $\frac{1}{2}|\partial f'|$).

A closely related property of a tuple of relators $\vec{h}$ is whether Dehn's algorithm works for the corresponding presentation. More precisely, Dehn's algorithm is the following (non-deterministic) process applied to a reduced word w: if w if of the form $w = w_1uw_2$ for some word u such that uv is a cyclic permutation of a relator and $|v| < |u|$, then replace w by the reduction of $w_1v^{-1}w_2$ (which is a shorter word), and repeat. It is clear that this process always terminates, and that if it terminates with the empty word, then w is equal to 1 in the group $G = \langle A \mid \vec{h}\rangle$. The converse does not hold in general, but we say that $\vec{h}$ is a *Dehn presentation* if it does, that is, if every reduced word w that is trivial in G, contains a factor which is a prefix of some cyclic conjugate h' of a relator, of length greater than $\frac{1}{2}|h'|$. It is clear that the word problem (given a word u, is it equal to 1 in G) admits an efficient decision algorithm in groups given by a Dehn presentation. Note that a tuple $\vec{h}$ with the Greendlinger property provides a Dehn presentation. Moreover, every hyperbolic group has a computable Dehn presentation [1, Theorem 2.12] and [41].

Greendlinger [24] shows that a tuple $\vec{h}$ with property $C'(\frac{1}{6})$ also has (a stronger

version of) the Greendlinger property defined above and yields a Dehn presentation. Theorem 2.1 shows that this situation is exponentially generic at density $d < \frac{1}{12}$. Ollivier proved a phase transition result regarding this property, with critical density higher than $\frac{1}{12}$ [48].

Theorem 2.5 *Let $0 < d < 1$. If $d < \frac{1}{5}$, then at density d, a random finite presentation generically is Dehn and has the Greendlinger property. If instead $d > \frac{1}{5}$, then at density d, a random finite presentation generically fails both properties.*

Ollivier [45] also considered finite presentations based on a given, fixed hyperbolic A-generated group G, that is, quotients of G by the normal subgroup generated by a tuple $\vec{h}$ of elements of G, that can be taken randomly. There are actually two ways of generating $\vec{h}$. Let $\pi\colon F(A) \to G$ be the canonical onto morphism: then one can draw uniformly at random a tuple $\vec{h}$ of cyclically reduced words (that is, of elements of $F(A)$) and consider the quotient $G/\langle\langle\pi(\vec{h})\rangle\rangle$, or one can draw uniformly at random a tuple of cyclically reduced words that are geodesic for G. Here it is useful to remember that a hyperbolic group is geodesically automatic [17] — in particular its language of geodesics L is regular —, and there is a linear time algorithm to randomly generate elements of L of a given length [9].

To state Ollivier's result, let us recall the definition of the co-growth of G, relative to the morphism π, under the hypothesis that π is not an isomorphism, that is, G is not free over A. Let $r = |A|$. Then $\mathsf{cogrowth}(G) = \lim \frac{1}{n} \log_{2r-1}(|H_n|)$, where H_n is the set of reduced words of length n in $\ker \pi$ (and the limit is taken over all even values of n, to account for the situation where no odd length reduced word is in $\ker \pi$). This invariant of G (and π) was introduced by Grigorchuk [26], who proved that it is always greater than $\frac{1}{2}$ and less than or equal to 1 and, using a result of Kesten [37], that it is equal to 1 if and only if G is amenable (amenability is an important property which, in the case of discrete groups, is equivalent to the existence of a left-invariant, finitely additive probability measure on G). The above definition does not apply if G is free over A, but it is convenient to let $\mathsf{cogrowth}(F(A)) = \frac{1}{2}$ (see [45, Section 1.2] for a discussion). In particular, the following elegant phase transition statement generalizes Theorem 2.2.

Theorem 2.6 *Let G be a hyperbolic and torsion-free A-generated group and let $\pi\colon F(A) \to G$ be the canonical mapping. Let $0 < d < 1$. If $d < 1 - \mathsf{cogrowth}(G)$, then at density d, a random quotient $G/\langle\langle\pi(\vec{h})\rangle\rangle$ is exponentially generically hyperbolic. If $d > 1 - \mathsf{cogrowth}(G)$, then $G/\langle\langle\pi(\vec{h})\rangle\rangle$ is exponentially generically degenerate.*

If instead, we take a tuple of cyclically reduced words of length n that are geodesic for G, then the phase transition between hyperbolicity and degeneracy is at density $\frac{1}{2}$.

Remark 2.7 Theorem 2.6 above is for torsion-free hyperbolic groups G. It actually holds as well if G is hyperbolic and has *harmless torsion*, that is, if every torsion element either sits in the virtual center of G, or has a finite or virtually $\mathbb{Z}$ centralizer, see [45]. □

Finally we note another phase transition theorem, due to Żuk, about Kazhdan's property (T) — a property of the unitary representations of a group — for discrete groups [57].

Theorem 2.8 *Let $0 < d < \frac{1}{2}$. If $d < \frac{1}{3}$, then at density d, a random finite presentation generically does not present a group with Kazhdan's property (T). If instead $d > \frac{1}{3}$, then at density d, a random finite presentation generically presents a group satisfying this property.*

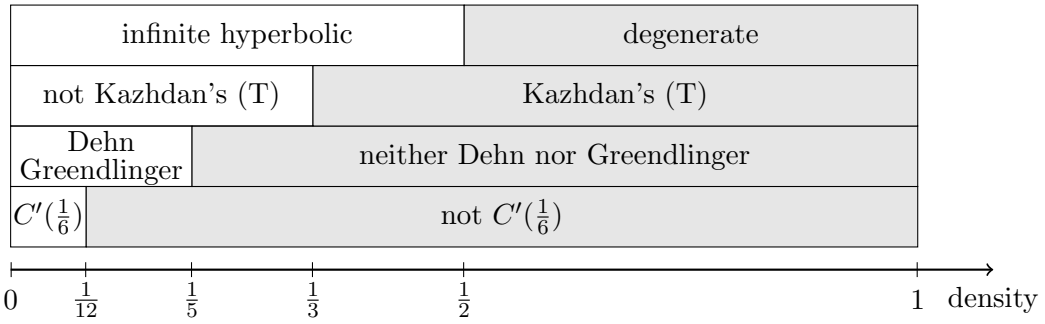

Figure 2: Phase transitions for properties of random presentations in the density model.

2.2 The few relators model

The few relators model was introduced by Arzhantseva and Ol'shanskiĭ [3]. In this model, the number of relators is fixed, say $k \geq 1$. Then the set S on which we define a model of randomness (see Section 1.2) is the set $\mathcal{CR}^k$ of all k-tuples of cyclically reduced words in $F(A)$ and $\mathbb{P}_n$ is the uniform probability law with support $(\mathcal{CR}_{\leq n})^k$.

Observe that if a tuple $\vec{h}$ of cyclically reduced words satisfies the small cancellation property $C'(\lambda)$ and if $\vec{g}$ is a sub-tuple of $\vec{h}$ (that is, the words in $\vec{g}$ are also in $\vec{h}$), then $\vec{g}$ satisfies Property $C'(\lambda)$ as well. From this observation and Theorem 2.1 (actually its variant in Remark 2.3) we deduce the following result, due to Arzhantseva and Ol'shanskii [3] (see also [33, Theorem B]).

Corollary 2.9 *In the few relator model, a random tuple exponentially generically satisfies Property $C'(\frac{1}{6})$ and presents an infinite hyperbolic group.*

Arzhantseva and Ol'shanskiĭ showed further that, in the few relator model, the finitely generated subgroups of a random k-relator subgroup are usually free or have finite index [3], statements (1) and (2) of Theorem 2.10 below. Statement (3) is due to Kapovich and Schupp [33, Theorem B]. Recall that a *Nielsen move* on a tuple $(x_1, \dots, x_k)$ of elements of a group G consists in replacing x_i by x_i^{-1}, exchanging x_i and x_j or replacing x_i by $x_i x_j$ for some $i \neq j$. We say that two k-tuples are *Nielsen-equivalent* if one can go from one to the other by a sequence of Nielsen moves.

Theorem 2.10 *Let $k, \ell \geq 1$ be integers. In the few relators model with k relators, exponentially generically,*

(1) every ℓ-generated subgroup of an A-generated group G, has finite index or is free;

(2) if $\ell < |A|$, every ℓ-generated subgroup of G is free and quasi-convex in G;

(3) an $|A|$-tuple which generates a non-free subgroup of G is Nielsen-equivalent to A in G. In particular, an $|A|$-tuple which generates a non-free subgroup generates G itself, and every automorphism of G is induced by an automorphism of $F(A)$.

Sketch of proof. The core of the proof lies in the identification of a class $\mathcal{P}$ of k-relator presentations (over a fixed alphabet of size r), defined below, which is exponentially generic in the few relators model, and which is smooth enough for Properties (1), (2) and (3) to always hold. The class $\mathcal{P}$ was originally introduced by Arzhantseva and Ol'shanskiĭ [3], and revisited by Kapovich and Schupp [33].

This class, parametrized by positive real numbers λ and μ, consists in the tuples $(u_1, \dots, u_k)$ in $\mathcal{CR}_{\leq n}$ which satisfy Property $C'(\lambda)$, do not contain a proper power, and such that every prefix w of a cyclic conjugate of u_i, of length at least $\frac{1}{2}|u_i|$, satisfies the following negative property: there does not exist a subgroup H of F_r, whose Stallings graph (see Section 3.1) has at most $\mu|w|$ edges, such that there exists a reduced word in H containing w as a factor, and such that $\mathsf{rank}(H) \leq r-1$ or such that $\mathsf{rank}(H) \leq r$ and H has infinite index.

If $\lambda \leq \frac{\mu}{15r+3\mu} < \frac{1}{6}$, the class $\mathcal{P}$ is exponentially generic [3] and Properties (1), (2) and (3) hold for all tuples of relators in $\mathcal{P}$ [3, 33]. □

Arzhantseva also established the following related result [2], which refines in a sense Theorem 2.10 (1). Here a set of generators for the ℓ-generated subgroup of G is fixed in advance (as a tuple of words in $F(A)$), and it is assumed that it generates an infinite-index subgroup of $F(A)$.

Theorem 2.11 *Let H be a finitely generated, infinite index subgroup of $F(A)$. In the few relators model with k relators, exponentially generically, a finite presentation $G = \langle A \mid \vec{h} \rangle$ ($\vec{h} \in (\mathcal{CR}_{\leq n})^k$) is such that the canonical morphism $\varphi\colon F(A) \to G$ is injective on H (so $\varphi(H)$ is free) and $\varphi(H)$ has infinite index in G.*

We also note that Kapovich and Schupp extended Theorem 2.10 to the density model [34], with density bounds that depend on both parameters k and ℓ.

Theorem 2.12 *Let A be a fixed alphabet. For every $k, \ell \geq 1$, there exists $0 < d(k,\ell) < 1$ such that, at every density $d < d(k,\ell)$, generically, an ℓ-generated subgroup of an A-generated group presented by a random k-tuple of relators has finite index or is free.*

Also, for every $k \geq 1$, there exists $0 < d(k) < 1$ such that, at every density $d < d(k)$, every $(k-1)$-generated subgroup of an A-generated group presented by a random k-tuple of relators is free. But there is no single value of d such that this holds independently of k (that is: $\lim_{k\to\infty} d(k) = 0$).

2.3 1-relator groups

If u is a cyclically reduced word, let $G_u = \langle A \mid u \rangle$. 1-relator groups are of course covered by the few relators model, and the results of Section 2.2 apply to them. But more specific results are known for random 1-relator presentations.

Magnus showed that if $u, v \in F(A)$, then the normal closures of the subgroups generated by u and v, written $\langle\langle u \rangle\rangle$ and $\langle\langle v \rangle\rangle$, are equal if and only if u is a conjugate of v or v^{-1} (see [40, Prop. II.5.8]). Kapovich and Schupp combine this with Theorem 2.10 (3) to show the following [33, Theorem A].

Theorem 2.13 *There exists an exponentially generic (and decidable) class P of cyclically reduced words such that, if $u, v \in P$, then G_u and G_v are isomorphic if and only if there exists an automorphism φ of $F(A)$ such that $\varphi(u) \in \{v, v^{-1}\}$. In particular, the isomorphism problem for 1-relator groups with presentation in P is decidable.*

We now explain how this result gives access to generic properties of (isomorphism classes of) 1-relator groups and not just of 1-relator presentations. This is a more explicit rendering of arguments which can be found in particular in Ollivier [46, Section II.3], Kapovich, Schupp and Shpilrain [35] and Sapir and Špakulová [51]. For this discussion we consider probability laws $\mathbb{P}_n$ for 1-relator presentations and probability laws $\mathbb{Q}_n$ for isomorphism classes of 1-relator groups. More specifically, $\mathbb{P}_n$ is the uniform probability law with support $\mathcal{CR}_{\leq n}$ (that is: the probability law for the few relator model, with $k = 1$ relator); and $\mathbb{Q}_n$ is the uniform probability law with support the set T_n of isomorphism classes of groups G_u with $|u| \leq n$. We let $T = \bigcup_{n\geq 1} T_n$, that is, T is the set of isomorphism classes of 1-relator groups.

Let H be the group of length-preserving automorphisms of $F(A)$, that is, the automorphisms which permute $\tilde{A}$. Note that $|H| = 2^r r!$, where $r = |A|$. Let also W be the set of strictly Whitehead minimal words, that is, cyclically reduced words u such that $|\varphi(u)| > |u|$ for every automorphism $\varphi \in \mathsf{Aut}(F(A)) \setminus H$. Kapovich *et al.* [35] show that W is exponentially generic (see [7] and Section 3.4 for a more general result).

Fix an arbitrary order on $\tilde{A}$. For each word $u \in \mathcal{CR}$, we let $\tau(u)$ be the lexicographically least element of the set of all cyclic permutations of images of u and u^{-1} by an automorphism in H. A set P as in Theorem 2.13 can be assumed to be closed under taking inverses and cyclic conjugation (see for instance the description of P in [33, Section 4]). In that case, a word u is in P if and only if $\tau(u) \in P$. The same clearly holds for W, and we have $2r \leq |\tau^{-1}(\tau(u))| \leq 2^{r+1}|u|r!$ — where the lower bound corresponds to a word of the form $u = a^{|u|}$. It is immediate that $G_u = G_{\tau(u)}$. Moreover, in view of Theorem 2.13, if $u, v \in P \cap W$, then G_u and G_v are isomorphic if and only if $\tau(u) = \tau(v)$.

Proposition 2.14 *Let X be a property of isomorphism classes of 1-relator groups, that is, X is a subset of T. Let $Y = \{u \in \mathcal{CR} \mid G_u \in X\}$. If $\mathbb{P}_n(Y) = o(n^{-1})$ (resp. Y is exponentially negligible), then X is negligible (resp. exponentially negligible). The same statement holds for genericity instead of negligibility.*

Proof. Let Z be the set of 1-relator groups G_u such that $u \in W \cap P$, where P is a set as in Theorem 2.13 and W is the set of strictly Whitehead minimal cyclically reduced words. Since $W \cap P$ is exponentially generic in $\mathcal{CR}$, there exist constants $C, c > 0$ such that $\mathbb{P}_n(W \cap P) \geq 1 - Ce^{-cn}$. We have

$$\mathbb{Q}_n(X) = \mathbb{Q}_n(X \cap Z) + \mathbb{Q}_n(X \setminus Z) \leq \mathbb{Q}_n(X \cap Z) + \mathbb{Q}_n(T \setminus Z).$$

We first deal with $\mathbb{Q}_n(T \setminus Z)$. Let $\alpha_n = |T_n \cap Z|$ and $\beta_n = |T_n \setminus Z|$. Then $\mathbb{Q}_n(T \setminus Z) = \frac{\beta_n}{\alpha_n+\beta_n}$. Note that $T_n \setminus Z \subseteq \{G_u \mid u \in \mathcal{CR}_{\leq n} \setminus (W \cap P)\}$. So $\beta_n \leq |\mathcal{CR}_{\leq n} \setminus (W \cap P)| \leq Ce^{-cn}|\mathcal{CR}_{\leq n}|$.

On the other hand, $T_n \cap Z$ is in bijection with $\{\tau(u) \mid u \in \mathcal{CR}_{\leq n} \cap (W \cap P)\}$, and it follows that

$$\alpha_n \geq \frac{1}{2^{r+1}nr!}|\mathcal{CR}_{\leq n} \cap (W \cap P)| \geq \frac{1 - Ce^{-cn}}{2^{r+1}nr!}|\mathcal{CR}_{\leq n}|.$$

Therefore

$$\mathbb{Q}_n(T \setminus Z) = \frac{\beta_n}{\alpha_n + \beta_n} \leq \frac{\beta_n}{\alpha_n} \leq \frac{Ce^{-cn}}{1 - Ce^{-cn}} 2^{r+1}nr!,$$

which vanishes exponentially fast.

Let us now consider $\mathbb{Q}_n(X \cap Z)$. We have

$$\begin{aligned}
\mathbb{Q}_n(X \cap Z) &= \frac{|X \cap Z \cap T_n|}{\alpha_n + \beta_n} \\
&\leq \frac{|\{G_u \mid u \in \mathcal{CR}_{\leq n} \cap W \cap P, u \in Y\}|}{\alpha_n} \\
&\leq \frac{2^{r+1}nr!\ \mathbb{P}_n(Y \cap W \cap P)\ |\mathcal{CR}_{\leq n}|}{2r(1 - Ce^{-cn})\ |\mathcal{CR}_{\leq n}|} \\
&\leq 2^r(r-1)!\ \frac{n\mathbb{P}_n(Y)}{1 - Ce^{-cn}},
\end{aligned}$$

and this concludes the proof. □

Then the results of Section 2.2 (Corollary 2.9, Theorem 2.10), together with Proposition 2.14 yield the following.

Corollary 2.15 *Exponentially generically, a 1-relator group G is infinite hyperbolic, every automorphism of G is induced by an automorphism of $F(A)$, and every ℓ-generated subgroup is free and quasi-convex if $\ell < |A|$.*

Kapovich, Schupp and Shpilrain use the ideas behind Proposition 2.14 to compute an asymptotic equivalent of the number of (isomorphism classes of) 1-relator groups in T_n [35].

Theorem 2.16 *Let $I_n(A)$ be the number of isomorphism classes of 1-relator groups of the form $\langle A \mid u \rangle$ with $|u| \leq n$. If $|A| = r$, then $I_n(A)$ is asymptotically equivalent to $\frac{1}{2^{r+1}r!}\ \frac{(2r-1)^n}{n}$.*

Finally we note the following result of Sapir and Špakulová [51]. Recall that a group G is *residually* $\mathcal{P}$ (for some property $\mathcal{P}$) if for all distinct elements $x, y \in G$, there exists a morphism φ from G to a group having property $\mathcal{P}$, such that $\varphi(x) \neq \varphi(y)$. We let *finite-p* be the property of being a finite p-group. A group G is *coherent* if every finitely generated subgroup is finitely presented.

Theorem 2.17 *Suppose that $|A| \geq 3$. Then an A-generated 1-relator group is generically residually finite, residually finite-p and coherent.*

2.4 Rigidity properties

Theorem 2.13 above gives a generic rigidity property: at least on a large (exponentially generic) set of words, the isomorphism class of the 1-relator group $G_u = \langle A \mid u \rangle$ is uniquely determined by u, up to inversion and an automorphism of $F(A)$. That is, the only words v such that G_v is isomorphic to G_u are those that come immediately to mind. As indicated, this result follows from a theorem of Magnus which states that the normal closure $\langle\langle u \rangle\rangle$ has essentially only one generator as a normal subgroup: $\langle\langle u \rangle\rangle = \langle\langle v \rangle\rangle$ if and only if u is a conjugate of v or v^{-1}.

There is no such general statement for normal subgroups generated by a k-tuple with $k \geq 2$. A closely related result due to Greendlinger generalizes Magnus's statement, but only for tuples that satisfy the small cancellation property $C'(\frac{1}{6})$ [25]: if $\vec{g}$ and $\vec{h}$ are such tuples, respectively a k-tuple and an ℓ-tuple, and if $\langle\langle \vec{g} \rangle\rangle = \langle\langle \vec{h} \rangle\rangle$, then $k = \ell$ and there is a re-ordering $\vec{g}'$ of $\vec{g}$ such that, for each i, h_i is a cyclic permutation of g'_i or g'^{-1}_i. The restriction to tuples satisfying $C'(\frac{1}{6})$ prevents us from proceeding as in Section 2.3 to prove a more general analogue of Theorem 2.13. Whether Theorem 2.13 can be extended to m-tuples of cyclically reduced words, is essentially the *Stability Conjecture* formulated by Kapovich and Schupp [32, Conjecture 1.2].

Nevertheless, Kapovich and Schupp show that one can circumvent this obstacle when considering the quotients of the *modular* group $M = \mathsf{PSL}(2, \mathbb{Z}) = \langle a, b \mid a^2, b^3 \rangle$. If $\vec{h}$ is a tuple of cyclically reduced words in $F(a, b)$, we denote by $M_{\vec{h}}$ the quotient of M by the images of the elements of $\vec{h}$ in M, that is, $M_{\vec{h}} = \langle a, b \mid a^2, b^3, \vec{h} \rangle$. Let η be the automorphism of M which fixes a and maps b to $b^{-1} = b^2$. Then the following holds [36, Theorem A and Corollary 2.5].

Theorem 2.18 *For each $k \geq 1$, there exists an exponentially generic (in the k-relator model), decidable subset Q_k of $\mathcal{CR}^k$ such that the following holds.*

- *If $\vec{h} \in Q_k$, then the group $M_{\vec{h}}$ is hyperbolic and one-ended, the generators a and b have order 2 and 3, respectively, in $M_{\vec{h}}$, and all the automorphisms of $M_{\vec{h}}$ are inner.*

- *If $\vec{g}, \vec{h} \in Q_k$ and $M_{\vec{g}}$ and $M_{\vec{h}}$ are isomorphic, then there is a re-ordering $\vec{g}'$ of $\vec{g}$ and a value $\varepsilon \in \{0,1\}$ such that, for each $1 \leq i \leq k$, h_i is a cyclic permutation of $\eta^\varepsilon(g'_i)$ or $\eta^\varepsilon(g_i'^{-1})$.*
- *If $\vec{g} \in Q_k$, $\vec{h} \in Q_\ell$ are such that the g_i and the h_j all have the same length, and if $M_{\vec{g}}$ and $M_{\vec{h}}$ are isomorphic, then $k = \ell$.*

In the k-relator model, the isomorphism problem for quotients of M is exponentially generically solvable in time $\mathcal{O}(n^4)$.

The last statement of this theorem is all the more interesting as the isomorphism problem, and even the triviality problem, for quotients of M is undecidable in general (Schupp [52]).

As in Section 2.3, Theorem 2.18 can be used to discuss asymptotic properties of k-relator quotients of the modular group, rather than of k-tuples of relators. The few-relator model for the quotients of M considers the set T of isomorphism classes of k-relator quotients of M, and the probability laws $\mathbb{Q}_n$ which are uniform on the set T_n of isomorphism classes of groups $M_{\vec{h}}$ with $\vec{h} \in (\mathcal{CR}_{\leq n})^k$. We can reason as for Proposition 2.14, modifying the map τ in such a way that $\tau(h)$ is the lexicographically least element of h, h^{-1}, $\eta(h)$ and $\eta(h^{-1})$. Then, with essentially the same proof as Proposition 2.14, we get the following result.

Proposition 2.19 *Let $k \geq 1$ and let X be a property of isomorphism classes of k-relator quotients of the modular group, that is, X is a subset of T. Let $Y = \{\vec{h} \in (\mathcal{CR})^k \mid M_{\vec{h}} \in X\}$. If $\mathbb{P}_n(Y) = o(n^{-k})$ (resp. Y is exponentially negligible), then X is negligible (resp. exponentially negligible). The same statement holds for genericity instead of negligibility.*

As in Section 2.3 again, one can derive from Theorem 2.18 an asymptotic equivalent of the number of isomorphism classes of k-relator quotients of the modular group [36, Theorem C].

Corollary 2.20 *Let $k \geq 1$. The number of isomorphism classes of quotients of M by k relators which are cyclically reduced words of length n, is asymptotically equivalent to*

$$\frac{(2^{\frac{n}{2}+1})^k}{2k!(2n)^k}.$$

Kapovich and Schupp go on to give further generic rigidity properties of homomorphisms between quotients of M, which are proved to be generically hopfian and co-hopfian (that is, every surjective (resp. injective) endomorphism is an isomorphism), and on the generic incompressibility of the presentations by k relators [36, Theorems B and D].

2.5 Nilpotent groups

We conclude this section with recent results on random groups in a particular class, that of nilpotent groups. If G is a group, the *lower central series* of G is defined by letting $G_1 = G$ and, for $n \geq 1$, $G_{n+1} = [G_n, G]$. That is: G_{n+1} is the subgroup generated by the commutators $[g, h] = g^{-1}h^{-1}gh$, with $g \in G_n$ and $h \in G$. Then each G_n is normal in G and G_{n+1} is contained in G_n. The group G is said to be *nilpotent of class* s if $G_{s+1} = 1$. In particular, G_2 is the derived subgroup of G and the class 1 nilpotent groups are exactly the abelian groups. Nilpotent groups of class 2 are those in which the derived subgroup lies in the center of the group.

Let us extend the commutator notation by letting, for $s \geq 2$, $[x_1, \ldots, x_{s+1}] = [[x_1, \ldots, x_s], x_{s+1}]$. One can show that the class of nilpotent groups of class s is defined by the identity $[x_1, \ldots, x_{s+1}] = 1$. As a result, this class constitutes a variety (in the sense of universal algebra) and we denote by $\mathsf{N}_s(A)$ its free object over the finite alphabet A: $\mathsf{N}_s(A) = F(A)/F(A)_{s+1}$.

Note that a torsion-free non-cyclic nilpotent group contains a free abelian group of rank 2, a standard obstacle for hyperbolicity: so torsion-free non-cyclic nilpotent groups are not hyperbolic. In particular, they form a negligible set in the few-relator as well as in the density models discussed in the previous sections, and we can not use earlier results to discuss random nilpotent groups. This difficulty was circumvented in several different ways in the literature.

Cordes et al. view finitely presented nilpotent groups as quotients of free nilpotent groups (of a fixed class and rank) by a random tuple of relators whose length tends to infinity [14]. In this model, relators are words over the symmetrized alphabet $\tilde{A}$. Depending on the number of relators, this extends the few relator and the density models. Garreta et al. extend in [19, 20] the study initiated in [14]. The following result is a summary of [14, Theorem 29, Proposition 30 and Corollaries 32 and 35] and of [20, Theorems 3.7 and 4.1].

Theorem 2.21 *Let $s \geq 1$, $r \geq 2$, let A be an alphabet of cardinality r, let $\mathsf{N}_{s,r} = \mathsf{N}_s(A)$ be the free nilpotent group of class s over A, and let π be the canonical morphism from $\tilde{A}^*$ onto $\mathsf{N}_{s,r}$.*

In the density model, at any density $d > 0$, a random quotient $\mathsf{N}_{s,r}/\langle\langle \pi(\vec{h}) \rangle\rangle$ is generically trivial. In fact, this holds in any model where the size of the tuple of relators is not bounded.

In the few relator model with k relators with $k \leq r - 2$, a random quotient $\mathsf{N}_{s,r}/\langle\langle \pi(\vec{h}) \rangle\rangle$ is generically non abelian and regular (that is: every element of the center of G has a non-trivial power in the derived subgroup).

If $k = r-1$, then such a quotient is generically virtually abelian (it has an abelian finite index subgroup), and if $k = r$, then it is generically finite. In either case, it is abelian if and only if it is cyclic. Finally, if $k \geq r+1$, then it is generically finite and abelian.

In the particular case where $r = 2$, $k = 1$ and $s \geq 2$, the probability that a random 1-relator quotient of $\mathsf{N}_{s,2}$ is cyclic (and hence, abelian) tends to $\frac{6}{\pi^2}$.

Cordes et al. also give a full classification of the 1-relator quotients of $\mathsf{N}_{2,2}$ (the Heisenberg group) [14, Section 3]. Moreover, they deduce from Theorem 2.21 the following result on random finitely presented groups [14, Corollary 36].

Corollary 2.22 *In the density model, at any density $d > 0$, and in any model where the size of the tuple of relators is not bounded, a random tuple $\vec{h}$ generically presents a perfect group (that is: a group G such that $[G, G] = G$, or equivalently, a group whose abelian quotient is trivial).*

Delp *et al.* use a different view of nilpotent groups [15]: it is well known that every torsion-free nilpotent group embeds in $U_n(\mathbb{Z})$ for some $n \geq 2$, where $U_n(\mathbb{Z})$ is the group of upper-triangular matrices with entries in $\mathbb{Z}$ and diagonal elements equal to 1. If $1 \leq i < n$, let $a_{i,n}$ be the matrix in $U_n(\mathbb{Z})$ with coefficients 1 on the diagonal and on row i and column $i+1$, and all other coefficients 0. Then $A_n = \{a_{1,n}, \ldots, a_{n-1,n}\}$ generates $U_n(\mathbb{Z})$. Let $\ell\colon \mathbb{N} \to \mathbb{N}$ be a function such that $\lim \ell(n) = \infty$ when n tends to infinity. We let $G_{\ell,n}$ be the subgroup of $U_n(\mathbb{Z})$ generated by a random pair of words of length $\ell(n)$ on alphabet $\tilde{A}_n$: in the language of Section 1.2, S is the set of pairs of words on an alphabet of the form $\tilde{A}_n$ for some $n \geq 2$, S_n is the set of all pairs of length $\ell(n)$ words on alphabet $\tilde{A}_n$, and $\mathbb{P}_n$ is the uniform probability law with support S_n. Then we have the following result, a combination of [15, Theorems 1 and 2]. Note that $U_n(\mathbb{Z})$ is nilpotent of class $n-1$: we say that a subgroup of $U_n(\mathbb{Z})$ has full class if it is nilpotent of class $n-1$.

Theorem 2.23 *Let $\ell\colon \mathbb{N} \to \mathbb{N}$ be a function such that $\lim \ell(n) = \infty$.*

- *If $\ell = o(\sqrt{n})$, then $G_{\ell,n}$ is generically abelian (that is: of class 1). If $\sqrt{n} = o(\ell(n))$, then $G_{\ell,n}$ is generically non abelian. And if $\ell(n) = c\sqrt{n}$, then the probability that $G_{\ell,n}$ is abelian tends to e^{-2c^2}.*
- *If $\ell = o(n^2)$, then $G_{\ell,n}$ generically does not have full class; and if $n^3 = o(\ell(n))$, then $G_{\ell,n}$ generically has full class.*

Garreta *et al.* use yet another representation of nilpotent groups [19, 21], the polycyclic presentation. A group G is *polycyclic* if it admits a sequence of subgroups $1 = H_n \leq H_{n-1} \leq \cdots \leq H_1 = G$ such that, for every $1 < i \leq n$, H_i is normal in H_{i-1} and H_{i-1}/H_i is cyclic. It is elementary to verify that every finitely generated nilpotent group is polycyclic. Polycyclic groups admit presentations of a particular form, the so-called *polycyclic presentations* (see [29] for a precise description), which can be characterized by a k-tuple of integers, where k is a function of the number of generators in the presentation.

In the case of torsion-free nilpotent groups, polycyclic presentations with generators $x_1, \ldots, x_r$ have relators of the following form, called a *torsion-free nilpotent presentation*:

$$[x_j, x_i] = x_{j+1}^{b_{i,j,j+1}} \cdots x_r^{b_{i,j,r}}$$
$$[x_j, x_i^{-1}] = x_{j+1}^{c_{i,j,j+1}} \cdots x_r^{c_{i,j,r}},$$

for all $1 \leq i < j \leq r$, where the $b_{i,j,h}$ and $c_{i,j,h}$ $(1 \leq i < j < h \leq r)$ are integers. Garreta *et al.* introduce a notion of random torsion-free nilpotent presentations as follows [19, 21]: with the number r of generators fixed, they let S be the set of tuples $(b_{i,j,h}, c_{i,j,h})_{1 \leq i < j < h \leq r}$ of integers, S_n be the set of those tuples whose components sit in the interval $[-n, n]$ and $\mathbb{P}_n$ be the uniform law with support S_n. They then show the following result [19, 21, Lemma 8].

Proposition 2.24 *Let $r \geq 2$. The group presented by a random torsion-free nilpotent presentation is generically finite.*

The situation becomes more interesting if one restricts one's attention to torsion-free nilpotent groups of class 2, also known as *τ_2-groups*. Recall that these are the torsion-free groups where the derived subgroup is contained in the center. In particular, the derived subgroup and the center are both free abelian groups. In the case of τ_2-groups, torsion-free nilpotent presentations can be simplified to the following *τ_2-presentations*:

$$\langle A, C \mid [a_i, c_h] = 1, [c_h, c_k] = 1, [a_i, a_j] = \prod_{1 \leq h \leq m} c_h^{\alpha_{i,j,h}},\ 1 \leq i < j \leq \ell \rangle,$$

for some $A = \{a_1, \ldots, a_\ell\}$, $C = \{c_1, \ldots, c_m\}$ $(\ell, m \geq 0)$ and some choice of $\alpha_{i,j,h} \in \mathbb{Z}$ $(1 \leq i, j \leq \ell$ and $1 \leq h \leq m)$. If $\vec{\alpha} = (\alpha_{i,j,h})_{i,j,h}$, we denote the group thus presented by $G(A, C, \vec{\alpha})$. Note that C generates a free abelian group of rank m, contained in the center of $G(A, C, \vec{\alpha})$.

For a fixed choice of A and C, a natural notion of randomness is given by letting S be the set of all the tuples $\vec{\alpha}$ of the appropriate size, S_n be the set of these tuples where every element has absolute value at most n, and $\mathbb{P}_n$ be the uniform probability law with support S_n. In this situation, Garreta *et al.* show the following [19, 21, Theorems 4 and 5].

Theorem 2.25 *Let $\ell, m \geq 0$ and let G be the group presented by a random τ_2-presentation on the pair of alphabets $A = \{a_1, \ldots, a_\ell\}$ and $C = \{c_1, \ldots, c_m\}$.*

If $\ell - 1 \leq m$, then generically C generates $Z(G)$, the center of G, and G is directly indecomposable into non-abelian factors.

If $m \leq \frac{\ell(\ell-1)}{2}$, then generically the derived subgroup G' of G has finite index in $Z(G)$. If $\ell - 1 \leq m \leq \frac{\ell(\ell-1)}{2}$, then G is generically regular.

If $m > \frac{\ell(\ell-1)}{2}$, then G is not regular, G' is freely generated (as an abelian group) by the $[a_i, a_j]$ $(1 \leq i < j \leq \ell)$, and G' generically has infinite index in $Z(G)$.

Garreta *et al.* also discuss whether the diophantine problem is generically decidable in a random τ_2-group [21].

3 Random subgroups

Before we discuss asymptotic properties of finitely generated subgroups, let us introduce a privileged tool to describe and reason about subgroups of free groups. Most of this section is devoted to this type of subgroups, only Section 3.5 below goes beyond the free group case.

3.1 Stallings graph of a subgroup

It is classical to represent the finitely generated subgroups of a free group by a finite labeled graph, subject to certain combinatorial constraints. An *A-graph* is a finite graph Γ whose edges are labeled by elements of A. It can be seen also as a transition system on alphabet $\tilde{A}$, with the convention that every a-edge from p to q represents an a-transition from p to q and an $\bar{a}$-transition from q to p. Note that this transition system is strongly connected as soon as (the underlying undirected graph of) Γ is connected. Say that Γ is *reduced* if it is connected and if no two edges with the same label start (resp. end) at the same vertex: this is equivalent to stating that the corresponding transition system is deterministic and co-deterministic. If 1 is a vertex of Γ, we say that $(\Gamma, 1)$ is *rooted* if every vertex, except possibly 1, has valency at least 2.

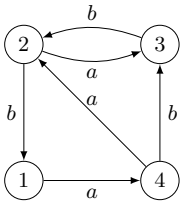

Figure 3: The Stallings graph of $H = \langle aab, ab\bar{a}b, abbb\rangle$. The reduced word $u = aa\bar{b}\bar{a}b$ is in H as it is accepted by $\Gamma(H)$: it labels a path starting from 1 and ending at 1, with edges being used backward when reading a negative letter. Since every vertex has valency at least 2, this graph is cyclically reduced.

We say that Γ is *cyclically reduced* if it is reduced and every vertex has valency at least 2. The A-graph in Fig. 3 is cyclically reduced. If Γ is reduced, we denote by $\kappa(\Gamma)$ the *cyclic reduction of* Γ, namely the cyclically reduced A-graph obtained from Γ by iteratively removing vertices of valency 1 and the edges adjacent to them.

If H is a finitely generated subgroup of $F(A)$, there exists a unique reduced rooted graph $(\Gamma(H), 1)$, called the *Stallings graph* of H, such that H is exactly the set of reduced words accepted by $(\Gamma(H), 1)$: a reduced word is accepted when it labels a loop starting and ending at 1. Moreover, this graph can be effectively computed given a tuple of reduced words generating H, in time $\mathcal{O}(n \log^* n)$ (that is to say: almost linear) [54, 55].

Conversely, every rooted reduced A-graph $(\Gamma, 1)$ is the Stallings graph of a (computable) finitely generated subgroup of $F(A)$. Moreover, two subgroups H_1 and H_2 are conjugated if and only if $\kappa(\Gamma(H_1)) = \kappa(\Gamma(H_2))$. Thus the reduced A-graph $\kappa(\Gamma(H))$ can be seen as a representative of the conjugacy class of H.

Many interesting properties of a subgroup H can be characterized – and often decided – in terms of the Stallings graph $\Gamma(H)$. We list here a couple that will be used in the sequel:

- the rank of H is equal to $|E| - |V| + 1$, where E (resp. V) is the number of edges (resp. vertices) of $\Gamma(H)$ [54];
- H has finite index in $F(A)$ if and only if $\Gamma(H)$ has the maximal number of possible edges, namely $|V||A|$ (one edge starting from every vertex, labeled by every letter of A), and if this is the case, then the index of H in $F(A)$ is $|V|$ [54];
- H is malnormal if and only if no non-empty reduced word u labels a loop in $\Gamma(H)$ at two different vertices, if and only if every non-diagonal connected component of the direct product $\Gamma(H) \times \Gamma(H)$ (in the category of A-graphs) is a tree [8];
- H is pure if and only if $\Gamma(H)$ never has a loop labeled u^n (u a non-empty reduced word, $n \geq 0$) at a vertex v, without having in fact a u-labeled loop at

that vertex [10].

3.2 The central tree property and its consequences

Let $\vec{h} = (h_1, \ldots, h_k)$ be a a tuple of reduced words in $F(A)$ and let $\vec{h}^{\pm}$ be the $2k$-tuple consisting of the elements of $\vec{h}$ and their inverses. Let $\min(\vec{h}) = \min_i |h_i|$ and let $\mathsf{lcp}(\vec{h})$ be the length of the longest common prefix of the elements of $\vec{h}^{\pm}$.

We say that $\vec{h}$ has the *central tree property* if $2\,\mathsf{lcp}\,\vec{h} < \min \vec{h}$. This property is identified explicitly by Bassino *et al.* [6, Section 1.3], but it is underlying the reasoning in the work of Arzhantseva and Ol'shanskiĭ [3], Jitsukawa [30] and several others. The central tree property, which could also be termed a small initial cancellation property, has the following interesting consequences. All are easily verified, except perhaps the last one, see [6, Proposition 1.3].

Proposition 3.1 *Let $\vec{g}$ and $\vec{h}$ be tuples of reduced words in $F(A)$, with the central tree property.*

(1) The Stallings graph $\Gamma(\langle\vec{h}\rangle)$ consists of a central tree*, which can be identified with the tree of prefixes of length $t = \mathsf{lcp}(\vec{h})$ of the elements of $\vec{h}^{\pm}$, and of pairwise edge-disjoint paths, one for each h_i, from the length t prefix of h_i to the length t prefix of h_i^{-1}.*

(2) $\Gamma(\langle\vec{h}\rangle)$ can be computed in linear time and $\vec{h}$ freely generates $\langle\vec{h}\rangle$.

(3) If $\langle\vec{h}\rangle = \langle\vec{g}\rangle$, then $\vec{h}^{\pm}$ and $\vec{g}^{\pm}$ differ only by the order of their components.

(4) If, in addition, $3\,\mathsf{lcp}(\vec{h}) < \min(\vec{h})$ and every word of length at most $\frac{1}{2}(\min(\vec{h}) - 2\,\mathsf{lcp}(\vec{h}))$ has at most one occurrence as a factor of a word in $\vec{h}^{\pm}$, then $\langle\vec{h}\rangle$ is malnormal and pure.

Let us now turn to asymptotic properties. Just like we were dealing with properties of finite presentations in Section 2 and not with properties of finitely presented groups (with the exception of 1-relator groups and k-relator quotients of the modular group, see Propositions 2.14 and 2.19), we first discuss asymptotic properties of tuples of generators of a subgroup (see however Proposition 3.4 and Corollary 3.5 below). We will see in Section 3.3 another way of approaching the asymptotic properties of finitely generated subgroups of $F(A)$.

In analogy with the tuples of cyclically reduced words used as relators, we can distinguish here:

- the density model, where at density d, $\mathbb{P}_n$ is the uniform probability law with support the $|\mathcal{R}_{\leq n}|^d$-tuples of words in $\mathcal{R}_{\leq n}$;
- and the few generators model, where an integer $k \geq 1$ is fixed, and $\mathbb{P}_n$ is the uniform probability law with support $\mathcal{R}^k_{\leq n}$.

Then we have the following result [6, Propositions 3.21 and 3.22].

Theorem 3.2 *Let $0 < d < 1$.*

If $d < \frac{1}{4}$, then at density d, a tuple of reduced words $\vec{h}$ exponentially generically has the central tree property and in particular, it freely generates $\langle \vec{h} \rangle$.

If instead $d > \frac{1}{4}$, then at density d, $\vec{h}$ exponentially generically does not have the central tree property.

If $d < \frac{1}{16}$, then at density d, a tuple of reduced words $\vec{h}$ exponentially generically generates a malnormal and pure subgroup.

It is immediate that, if every element of $\vec{g}$ is also an element of $\vec{h}$, and $\vec{h}$ has the central tree property, then so does $\vec{g}$. In that case, it is not hard to show also that $\langle \vec{g} \rangle$ is malnormal if $\langle \vec{h} \rangle$ is (see for instance [6, Proposition 1.5]). Then Theorem 3.2 yields the following corollary, which was already observed by Arzhantseva and Ol'shanskiĭ [3] for the free generation statement, and Jitsukawa [30] for the malnormality statement.

Corollary 3.3 *In the few generators model, a tuple of reduced words exponentially generically has the central tree property, it is a basis of the subgroup it generates, and this subgroup is malnormal and pure.*

We now see how to use the rigidity property in Proposition 3.1 (3) to discuss asymptotic properties of subgroups themselves, and not of tuples of generators, at least in the few generators model. This is in the same spirit as in Propositions 2.14 and 2.19 above.

Fix $k \geq 1$. In the k-generator model for tuples, the set S (in the terminology of Section 1.2) is $\mathcal{R}^k$ and $\mathbb{P}_n$ is the uniform probability law with support $S_n = \mathcal{R}^k_{\leq n}$. Now consider the set T of all k-generated subgroups of $F(A)$, the set T_n of subgroups of the form $\langle \vec{h} \rangle$ for some $\vec{h} \in S_n$ and the probability law $\mathbb{Q}_n$ which is uniform on T_n. We call this the *k-generator model for subgroups.*

Proposition 3.4 *Let X be a property of k-generator subgroups of $F(A)$, that is, X is a subset of T. Let $Y = \{\vec{h} \in \mathcal{R}^k \mid \langle \vec{h} \rangle \in X\}$. If Y is negligible (resp. exponentially negligible) in the k-generator model for tuples, then so is X, in the k-generator model for subgroups. The same statement holds for genericity instead of negligibility.*

Proof. Let P be the set of tuples with the central tree property. By Corollary 3.3, there exist $C, d > 0$ such that $\mathbb{P}_n(P) > 1 - Ce^{-dn}$. Moreover, by Proposition 3.1 (3), if $\vec{h} \in P$, there are at most $2^k k!$ elements of P which generate the subgroup $\langle \vec{h} \rangle$.

If Z is the set of subgroups of $F(A)$ of the form $\langle \vec{h} \rangle$ such that $\vec{h} \in P$, one shows as in the proof of Proposition 2.14 that $\mathbb{Q}_n(X) \leq \mathbb{Q}_n(X \cap Z) + \mathbb{Q}_n(T \setminus Z)$, and that both terms of this sum vanish exponentially fast. □

The following corollary immediately follows from Corollary 3.3.

Corollary 3.5 *Let $k \geq 1$. In the k-generator model for subgroups, malnormality and purity are exponentially generic.*

Remark 3.6 The proof of Proposition 3.4 does not extend to the density model: if the number of elements of a tuple $\vec{h}$ is a function $k(n)$ that tends to infinity, the multiplying fact $2^k k!$ is not a constant anymore, and neglibibility for X is obtained only if $\mathbb{P}_n(Y)$ vanishes very fast (namely, if $\mathbb{P}_n(Y) = o(2^{k(n)}k(n)!)$). □

We conclude this section with a discussion of the height of the central tree of the Stallings graph of $\langle \vec{h} \rangle$ (that is: the parameter $\mathsf{lcp}(\vec{h})$) for a random choice of $\vec{h}$. Arzhantseva and Ol'shanskiĭ [3] showed that in the few generators model, the height of the central tree (namely the parameter $\mathsf{lcp}(\vec{h})$) is exponentially generically at most αn, for any $\alpha > 0$. It is in fact generically much smaller, see [6, Proposition 3.24].

Proposition 3.7 *Let f be an unbounded non-decreasing integer function and let $k \geq 1$. The following inequality holds generically for a tuple $\vec{h}$ chosen randomly in the k-generator model: $\mathsf{lcp}(\vec{h}) \leq f(n)$.*

This implies that, generically in the few generators model, for tuples as well as for subgroups, the proportion of vertices of $\Gamma(\langle \vec{h} \rangle)$ that lie in the central tree (at most $2r(2r-1)^{\mathsf{lcp}(\vec{h})-1}$) tends to 0 (apply Proposition 3.7 with, say, $f(n) = \log \log n$).

3.3 Random Stallings graphs

Another point of view on random subgroups of $F(A)$ relies on the observation that each finitely generated subgroup corresponds to a unique Stallings graph, and that these graphs admit an intrinsic combinatorial characterization, as reduced rooted A-graphs (see Section 3.1). The problem of drawing a random subgroup can therefore be reduced to the problem of drawing a random reduced rooted A-graph.

When considering such graphs, it is natural to measure their size by their number of vertices (the number of edges of such a graph of size n lies between $n-1$ and $2|A|n$). By extension, we say that the size of a subgroup H, written $|H|$, is the size of its Stallings graph $\Gamma(H)$. Then we consider the set S of all Stallings graphs over alphabet A (that is: of all the reduced rooted A-graphs), and the uniform probability law $\mathbb{P}_n$ with support the Stallings graphs with n vertices. This is called the *graph-based model* for subgroups of $F(A)$.

Implementation of the graph-based model The problem of drawing a tuple of reduced words uniformly at random is easily solved: one draws each word independently, one letter at a time, with $2r = |\tilde{A}|$ choices for the first letter, and $2r-1$ choices for each of the following letters.

Drawing (a tuple of) cyclically reduced words uniformly at random is also done in a simple way. Indeed, the probability that a random reduced word of length n is cyclically reduced tends to $\frac{2r-1}{2r}$ when n tends to infinity, and we can use a rejection algorithm: repeatedly draw a reduced word until that word is cyclically reduced. The expected number of draws tends to $\frac{2r}{2r-1} = 1 + \frac{1}{2r-1}$.

Drawing a Stallings graph with n vertices is a less immediate task. Bassino *et al.* [5] use a recursive method and the tools of analytic combinatorics to solve it in an efficient manner: they give a rejection algorithm with expected number of draws $1 + o(1)$, which requires a linear time precomputation, and takes linear time for each draw. These linear time bounds are evaluated in the RAM model; in the bit complexity model, the precomputation is done in time $\mathcal{O}(n^2 \log n)$ and each draw is done in time $\mathcal{O}(n^2 \log^2 n)$ (see [5, Section 3]).

Let us sketch a more precise description of this random generation algorithm and its justification. The central idea is the observation that a size n Stallings graph Γ defines an A-tuple $(f_a)_{a\in A}$ of partial injections (partial, one-to-one maps) from $\{1,\dots,n\}$ into itself: $f_a(i) = j$ if and only if there is an a-labeled edge in Γ from vertex i to vertex j. Conversely, such a tuple of partial injections defines an A-labeled graph with vertex set $\{1,\dots,n\}$, which is a Stallings graph (rooted at vertex 1) if and only if it is connected and every vertex $i > 2$ is adjacent to at least 2 edges.

Drawing an n-vertex Stallings graph uniformly at random can therefore be done by drawing independently $|A|$ partial injections uniformly at random, checking whether the resulting graph is a Stallings graph and, if it isn't, rejecting this draw and drawing a fresh one, repeating the operation until a Stallings graph has been drawn. The justification of the efficiency of such a *rejection algorithm* relies on the proof of the following statement: with probability tending to 1 when n tends to infinity, the graph defined by an A-tuple of randomly chosen partial injections on $\{1,\dots,n\}$ is a

Stallings graph. Once this is established, it is elementary that the expected number of draws in the rejection algorithm is $1 + o(1)$.

We refer the reader to Theorems 2.4 and 2.6 and Corollary 2.7 in [5] for a proof of this assertion. This proof relies on a combinatorial understanding of partial injections which we discuss below. We first note that drawing uniformly at random a size n partial injection can be done by the following elementary method: there are $\mathsf{PI}_{n,k} = \binom{n}{k}^2 k!$ partial injections with domain size k (choose a size k domain, a size k codomain, and a bijection between them). For n fixed, if the $\mathsf{PI}_{n,k}$ are precomputed, one can draw a size n partial injection as follows: first draw the domain size k according to the distribution given by the $\mathsf{PI}_{n,k}$, draw two size k subsets to be the domain and codomain, and draw a permutation of $\{1, \ldots, k\}$. This method has shortcomings: it does not give us a handle to prove the genericity of connectedness, which is essential to justifying the rejection algorithm, or to easily estimate such parameters as, say, the expected value of the domain size of a partial injection, which is essential in the proof of Statements (2) to (6) of Theorem 3.8 below. Note also that some care needs to be exercised to obtain the linear complexity bounds mentioned above: the binomial coefficients $\binom{n}{k}$ (or the ratios $\frac{\mathsf{PI}_{n,k}}{\mathsf{PI}_n}$, where PI_n is the number of size n partial injections) must be computed by a linear recurrence (based on Pascal's triangle) and random permutations must be generated in linear time.

The algorithm used in [5] to efficiently draw a partial injection uniformly at random, is an instance of the *recursive method* (see [18]). A size n partial injection f (or rather its functional graph) is analyzed as follows: it is a disjoint union of its maximal orbits which are either cycles (as in permutations) or linear graphs (or *sequences*), that is, subsets $\{i_1, \ldots, i_\ell\}$ of $\{1, \ldots, n\}$ ($\ell \geq 1$) such that $f(i_j) = i_{j+1}$ for $1 \leq j < \ell$, i_1 has no pre-image and i_ℓ has no image. The exponential generating sequences (EGSs) of these simple combinatorial structures (cycles and sequences) are easily computed and the calculus of EGSs inherent to the recursive method yields an explicit formula for the EGS of partial injections. This formula, together with a healthy dose of complex analysis, allows us to justify our rejection algorithm and to establish a number of asymptotic properties of Stallings graphs, see Theorem 3.8 below.

The resulting efficient random generator uses the explicit computation of the coefficients of the EGS for partial injections, and the fact that this EGS is the result of specific algebraic operations applied to the EGSs of cycles and sequences. This reduces the random generation of a size n partial injection to a 2-step algorithm: first we draw the *profile* of a random permutation, that is, the sequence of sizes and types (cycle or sequence) of its maximal orbit, and second we draw a random size n permutation to label the objects in the profile we just drew. Drawing the

profile uniformly at random consists in determining the size k of a maximal orbit (according to the distribution of the sizes of these orbits, which is obtained along the way), determining whether this orbit is a cycle or a sequence (the distribution of these two types of size k orbits was also obtained along the way) and completing the profile by randomly generating the profile of a size $n-k$ partial injection (this is the recursion in the *recursive* method), see [5] for more details.

Asymptotic properties of subgroups in the graph-based model The following is a combination of [5, Section 2.4, Corollary 4.1] and [4, Corollary 4.8 and Theorems 5.1 and 6.1]. We say that a property X is *super-polynomially negligible* (resp. *generic*) if $\mathbb{P}_n(X)$ is $\mathcal{O}(n^{-k})$ (resp. $1-\mathcal{O}(n^{-k})$) for every positive integer k.

Theorem 3.8 *Let $r = |A|$.*

(1) The number of subgroups of $F(A)$ of size n is asymptotically equivalent to

$$\frac{(2e)^{-r/2}}{\sqrt{2\pi}}\ e^{-(r-1)n+2r\sqrt{n}}\ n^{(r-1)n+\frac{r+2}{4}}.$$

(2) The expected rank of a size n subgroup of $F(A)$ is $(r-1)n - r\sqrt{n}+1$, with standard deviation $o(\sqrt{n})$.

(3) In the graph-based model, a random subgroup of $F(A)$ of size n is generically neither malnormal nor pure: it is malnormal (resp. pure) with vanishing probability $\mathcal{O}(n^{-\frac{r}{2}})$.

(4) The probability that a subgroup of $F(A)$ of size n avoids all the conjugates of the elements of A tends to e^{-r}.

(5) The probability that a subgroup of $F(A)$ of size n has finite index admits an $\mathcal{O}(n^{\frac{r}{4}}e^{-2r\sqrt{n}})$ upper bound. In particular, this class of subgroups is super-polynomially negligible.

(6) In the graph-based model, the quotient of $F(A)$ by the normal closure of a random subgroup is generically trivial.

Theorem 3.8 (1) is a direct consequence of the previous discussion: generically, an r-tuple of partial injections drawn independently defines a Stallings graph, so an asymptotic estimate of the number of size n subgroups can be derived from an asymptotic estimage of the number of size n partial injections.

Theorem 3.8 (2) uses the fact that the rank of a subgroup H is equal to $e-v+1$, where e and v are the numbers of edges and vertices, respectively, of $\Gamma(H)$ (see Section 3.1). For a size n subgroup, $v = n$. As for the number of a-labeled edges, it is the difference $n - s_a$, where s_a is the number of sequences among the maximal orbits of the partial injection f_a determined by a. Thus the proof of Theorem 3.8 (2) reduces to the study of the asymptotic behavior of the random variable which counts the number of sequences in a random partial injection of size n. This relies on the saddle point analysis of the bivariate EGS which counts partial injections by size and by the number of their sequences (see [5, Section 2.3]). The counting of partial injections by the seemingly indirect recursive method is crucial for this purpose.

It is interesting to contrast Theorem 3.8 (2) with the results reported in Section 3.2. As discussed at the very end of that section, in the Stallings graph of a subgroup taken at random in the few generators model, the immense majority of vertices are on the outer loops, adjacent to exactly two edges. In fact, since the rank of a subgroup is the difference between the number of edges and the number of vertices plus 1, the ratio between the number of edges and vertices tends to 1 in the few generators model (Proposition 3.1 (2) and Corollary 3.3), and it tends to $|A|-1$ in the graph-based model (Theorem 3.8 (2)). Observe that the minimum and maximum possible values for this ratio are 1 and $|A|$: in intuitive terms, the Stallings graph of a random group is sparse in the few generators model, and rather full in the graph based model. In other words, there are many more loops, including short loops, in the latter model, whereas in the k relator model, there are only k loops, and they are all very long: using close to a $\frac{1}{k}$ proportion of the edges. This is the feature that is exploited in [4] to show that the property in Theorem 3.8 (4) is exponentially negligible in the few generators model, and indeed in the density model at densities $d < \frac{1}{4}$. Similarly, generically in the graph based model, a Stallings graph has a cycle labeled by a power of a letter, and hence the corresponding subgroup is neither malnormal nor pure (Theorem 3.8 (3)). This is a very rough sufficient reason for a subgroup to fail being malnormal or pure, and the probability of this property may well vanish faster than stated above. A refinement of this result (namely the fact that for each letter a, the lengths of the cycles labeled by a power of a are relatively prime) leads to Theorem 3.8 (6). In this respect, we see that drawing uniformly at random the Stallings graph of the subgroup generated by a tuple of relators is not a fruitful avenue, to discuss 'typical' properties of finite presentations.

Finally, we note that the estimates in Theorem 3.8 (1) and (5) can be seen as an extension of the study of subgroup growth, see in particular Lubotzky and Segal [39].

3.4 Whitehead minimality

The following property of a subgroup H of $F(A)$ has already been mentioned: we say that H is *Whitehead minimal* (resp. *strictly Whitehead minimal*) if $|\varphi(H)| \geq |H|$ (resp. $|\varphi(H)| > |H|$) for every non length-preserving automorphism φ of $F(A)$, where $|H|$ is the number of vertices of its Stallings graph $\Gamma(H)$. This property plays an important role in the solution of the automorphic orbit problem, to decide whether two subgroups are in the same orbit under the automorphism group of $F(A)$, as shown by Gersten [22, Corollary 2], in an extension of the famous Whitehead peak reduction theorem [56] (see also [40, Section 1.4]) from elements of $F(A)$ to finitely generated subgroups.

Note that a cyclic subgroup $H = \langle u \rangle$ is (strictly) Whitehead minimal if and only if the word u is (strictly) Whitehead minimal in the sense discussed in Section 2.3. As mentioned there, Kapovich *et al.* proved that strictly Whitehead minimal cyclically reduced words are exponentially generic in $F(A)$ [35, Theorem A].

This can be generalized to all finitely generated subgroups. Since the Stallings graph of a Whitehead minimal subgroup must be cyclically reduced, the graph based model must be restricted (in the natural way) to these graphs. If we consider instead the few generators model, we note that being cyclically reduced is not a generic property (see [7, Proposition 4.6]): here too, the few generators model must be restricted to tuples of cyclically reduced words, that is, to the few relator model of Section 2. Under these restrictions, Bassino *et al.* proved that strict Whitehead minimality is generic both in the graph based and in the few generators models [7, Theorems 3.1 and 4.1].

Theorem 3.9 *Strict Whitehead minimality is super-polynomially generic for the uniform distribution of cyclically reduced Stallings graphs.*

The same property is exponentially generic in the few relator model, restricted to tuples of cyclically reduced words.

Remark 3.10 The reasons for genericity are different for the two models, due to the very different expected geometry of a random Stallings graph: in the few generator models, it is very sparse and most of its vertices are on very long loops, whereas the graph is fuller and has many short loops in the graph-based model. See [7] for more details. □

3.5 Random subgroups of non-free groups

Let us first return to the few generators model, but for subgroups of some fixed, non-free A-generated group G. Here, the probability laws $\mathbb{P}_n$ we consider are the

uniform probability laws with support $(\tilde{A}^{\leq n})^k$ for some fixed $k \geq 1$: that is, we draw uniformly at random k-tuples of words of length at most n, that are not necessarily reduced.

Gilman *et al.* show the following proposition [23, Theorem 2.1]. Recall that a group is *non-elementary hyperbolic* if it is hyperbolic and does not have a cyclic, finite index subgroup.

Proposition 3.11 *Let G be a non-elementary hyperbolic group and let $k \geq 1$. Then for any choice of generators A of G and any onto morphism $\pi\colon F(A) \to G$, exponentially generically in the k-generator model, a tuple $\vec{h}$ of elements of $F(A)$ is such that $\pi(\vec{h})$ freely generates a free, quasi-convex subgroup of G.*

Note that a free group $F(A)$ is non-elementary hyperbolic if $|A| \geq 2$: thus Proposition 3.11 generalizes part of Corollary 3.3, since the latter is only relative to the standard set of generators of $F(A)$.

Say that a group G has the *(exponentially) generic free basis property* if, for every choice of generators A of G and every onto morphism $\pi\colon F(A) \to G$, for every integer $k \geq 1$, the π-image of a k-tuple $\vec{h}$ of elements of $\tilde{A}^*$ (exponentially) generically freely generates a free subgroup of G (in the k-generator model). Proposition 3.11 states that non-elementary hyperbolic groups have the exponentially generic free basis property. Gilman *et al.* [23] and Myasnikov and Ushakov [43] note that this property is preserved as follows: if $\varphi\colon G_1 \to G_2$ is an onto morphism and G_2 has the (exponentially) generic free basis property, then so does G_1. For instance, non abelian right-angled Artin groups and pure braid groups PB_n ($n \geq 3$) have the exponentially generic free basis property, since they admit morphisms onto a rank 2 free group (see *e.g.* [11] for PB_n).

Proposition 3.11 can be used also to show the following result [23, Theorem 2.2] on the membership problem in subgroups – a problem which is, in general, undecidable in hyperbolic groups [50].

Corollary 3.12 *Let G be a non-elementary hyperbolic A-generated group, let π be a surjective morphism from A^* onto G and let $k \geq 1$. There exists an exponentially generic set X of k-tuples of words in $\tilde{A}^*$ and a cubic time algorithm which, on input a k-tuple $\vec{h}$ and an element $x \in \tilde{A}^*$, decides whether $\vec{h} \in X$, and if so, solves the membership problem for $\pi(x)$ and $\pi(\vec{h})$, that is, decides whether $\pi(x) \in \langle \pi(\vec{h}) \rangle$.*

There is no study as yet of asymptotic properties of subgroups of non free groups using a graph based model, in the spirit of Section 3.3. Let us however mention that recent results may open the way towards such a study: Kharlampovich *et al.*

[38] effectively construct Stallings graphs which are uniquely associated with each quasi-convex subgroup of a geodesically automatic group, e.g. hyperbolic groups, right-angled Artin groups. Like in the free group case, this has a large number of algorithmic consequences. It may be difficult to combinatorially characterize these graphs in general, and to design random generation algorithms or to explore their asymptotic properties. But it may be possible to tackle this task for specific groups or classes of groups.

In fact, somewhat earlier results already gave more efficient and more combinatorially luminous constructions, for amalgams of finite groups (Markus-Epstein [42]) and for virtually free groups (Silva *et al.* [53]). Note that both classes of groups are locally quasi-convex, and these constructions apply to all their finitely generated subgroups.

4 Non-uniform distributions

In this final section, we introduce non-uniform distributions, both for relators and generators, as explored by Bassino *et al.* [6]. We keep the idea of randomly drawing tuples of words by independently drawing the elements of the tuple, but we relax the distribution on the lengths of the tuples and on the lengths of the words, and we use non-uniform probability laws of probability on each $\mathcal{R}_n$ (resp. $\mathcal{CR}_n$).

More precisely, the model of randomness is the following [6]. For each $n \geq 0$, let $\mathbb{R}_n$ be a law of probability on $\mathcal{R}_n$ (or $\mathcal{CR}_n$ if we are dealing with presentations) and let $\mathbb{T}_n$ be a law of probability on the set of tuples of positive integers. If $\vec{h} = (h_1, \ldots, h_k)$ is a tuple of words, let $|\vec{h}| = (|h_1|, \ldots, |h_k|)$. Together, $(\mathbb{R}_n)_n$ and $(\mathbb{T}_n)_n$ define a sequence of probability laws $\mathbb{P}_n$ on the set of tuples of (cyclically) reduced words as follows:

$$\mathbb{P}_n(\vec{h}) = \mathbb{T}_n(|\vec{h}|) \prod_i \mathbb{R}_{|h_i|}(h_i).$$

Note that this includes the density and the few generators (relators) models discussed in Sections 2 and 3: for instance, in the k-generator model, $\mathbb{R}_n$ is the uniform distribution on $\mathcal{R}_n$ and $\mathbb{T}_n$ is the distribution with support the k-tuples of integers between 0 and n, each with probability

$$\mathbb{T}_n(\ell_1, \ldots, \ell_k) = \prod_{i=1}^{k} \frac{|\mathcal{R}_{\ell_i}|}{|\mathcal{R}_{\leq n}|}.$$

4.1 Prefix-heavy distributions

For each word $u \in \mathcal{R}$, denote by $\mathcal{P}(u)$ the set of reduced words starting with u, that is, $\mathcal{P}(u) = u\tilde{A}^* \cap \mathcal{R}$. For $C \geq 1$ and $0 < \alpha < 1$, say that the sequence of probability laws $(\mathbb{R}_n)_n$ (each with support in $\mathcal{R}_n$) is *prefix-heavy with parameters* (C, α) if, for all $u, v \in \mathcal{R}$, we have

$$\mathbb{R}_n(\mathcal{P}(uv) \mid \mathcal{P}(u)) \ \leq \ C\alpha^{|v|}.$$

This definition captures the idea that the probability of a prefix-defined set (a set of the form $\mathcal{P}(u)$) decreases exponentially fast with the length of u. It is satisfied by the sequence of uniform probability laws on the $\mathcal{R}_n$ ($n \geq 0$).

If $(\mathbb{P}_n)_n$ is a sequence of laws of probability on tuples of reduced words, defined as above by sequences $(\mathbb{R}_n)_n$ and $(\mathbb{T}_n)_n$ of probability laws on words and on tuples of integers, and if $(\mathbb{R}_n)_n$ is prefix-heavy with parameters (C, α), then we say that $(\mathbb{P}_n)_n$ is *prefix-heavy* as well, with the same parameters.

Under this hypothesis, Bassino *et al.* obtain a series of general results [6, Theorems 3.18, 3.19 and 3.20], summarized as follows. If $\vec{h} = (h_1, \ldots, h_k)$ is a tuple of reduced words, we let $\mathsf{size}(\vec{h}) = k$, $\min(\vec{h}) = \min\{|h_i| \mid 1 \leq i \leq k\}$ and $\max(\vec{h}) = \max\{|h_i| \mid 1 \leq i \leq k\}$. Let us say, also, that $(\mathbb{R}_n)_n$ and $(\mathbb{P}_n)_n$ *do not ignore cyclically reduced words* if $\liminf \mathbb{R}_n(\mathcal{CR}_n) > 0$.

Theorem 4.1 *Let $(\mathbb{P}_n)_n$ be a sequence of probability laws on tuples of reduced words, which is prefix-heavy with parameters (C, α), with $C \geq 1$ and $0 < \alpha < 1$. Let $0 < \lambda < \frac{1}{2}$.*

- *If the random variable $\mathsf{size}^2 \alpha^{\frac{\min}{2}}$ is increasingly small — more precisely, if there exists a sequence $(\eta_n)_n$ tending to 0, such that $\mathbb{P}_n(\mathsf{size}^2 \alpha^{\frac{\min}{2}} > \eta_n)$ tends to 0 —, then a random tuple of reduced words generically satisfies the central tree property, and freely generates a subgroup of $F(A)$.*

- *If there exists a sequence $(\eta_n)_n$ tending to 0, such that $\mathbb{P}_n(\mathsf{size}^2 \max^2 \alpha^{\frac{\min}{8}} > \eta_n)$ tends to 0, then a random tuple of reduced words generically generates a malnormal subgroup of $F(A)$.*

- *Let $0 < \lambda < \frac{1}{2}$. If the sequence $(\mathbb{P}_n)_n$ does not ignore cyclically reduced words and if there exists a sequence $(\eta_n)_n$ tending to 0, such that $\mathbb{P}_n(\mathsf{size}^2 \max^2 \alpha^{\lambda \min} > \eta_n)$ tends to 0, then a random tuple of cyclically reduced words generically satisfies the small cancellation property $C'(\frac{1}{6})$.*

In all three statements, exponential genericity is guaranteed if the vanishing sequences converge exponentially fast to 0.

The technical aspect of these statements is due to the very general nature of the random model considered. In the next section, we discuss a more specific model, where the $\mathbb{R}_n$ are generated by a Markovian scheme.

4.2 Markovian automata

When it comes to drawing words at random, an automaton-theoretic model comes naturally to mind. Bassino *et al.* introduce the following notion: a *Markovian automaton* $\mathcal{A}$ over a finite alphabet X consists in a finite deterministic transition system $(Q, \cdot)$ (that is: an action of the free monoid X^* on the finite set Q, or seen otherwise, a deterministic finite state automaton over alphabet X without initial or terminal states), an initial probability vector $\gamma_0 \in [0,1]^Q$ (that is: $\sum_{p\in Q} \gamma_0(p) = 1$), and a stochastic matrix $M \in [0,1]^{Q\times X}$ (that is, a matrix where each column is a probability vector) such that $M(p,x) > 0$ if and only if $p \cdot x$ is defined.

Such a scheme defines a sequence $(\mathbb{R}_n)_n$ of laws of probability, over each set X^n ($n \geq 0$), as follows:

$$\mathbb{R}_n(x_1 \cdots x_n) = \sum_{p\in Q} \gamma_0(p) M(p, x_1) M(p \cdot x_1, x_2) \cdots M(p \cdot (x_1 \cdots x_{n-1}), x_n).$$

Note that the union over n of the support sets of the $\mathbb{R}_n$ is always a prefix-closed rational language: that accepted by the transition system $(Q, \cdot)$, with initial states the support of γ_0 and all states final.

Example 4.2 For instance, if $Q = \tilde{A}$, if for each $a, b \in \tilde{A}$, $a \cdot b$ is defined whenever $b \neq a^{-1}$, and equal to b when defined, if the entries of γ_0 are all equal to $\frac{1}{2r}$ and if the non-zero entries of M are all equal to $\frac{1}{2r-1}$, then $\mathbb{R}_n$ is the uniform probability law on $\mathcal{R}_n$.

The Markovian automata in Figure 4 also yield the uniform probability law (at fixed length) on two languages which both provide unique representatives for the elements of the modular group (see Section 2.4): the support of $\mathcal{A}$ is the set of words over alphabet $\{a, b, b^{-1}\}$ without occurrences of the factors a^2, b^2, $(b^{-1})^2$, bb^{-1} and $b^{-1}b$ (the shortlex geodesics of the modular group), and the support of $\mathcal{A}'$ consists of the words on alphabet $\{a, b\}$, without occurrences of a^2 or b^3. □

A first set of results is obtained by specializing Theorem 4.1 to the case where the sequence $(\mathbb{P}_n)_n$ is induced by a Markovian automaton $\mathcal{A}$. If $0 < \alpha < 1$, we introduce *the α-density model with respect to $\mathcal{A}$*, in analogy with Sections 2.1 and 3.2: at density $d < 1$, the sequence $(\mathbb{P}_n)_n$ is induced by the sequences $(\mathbb{R}_n)_n$, induced by $\mathcal{A}$,

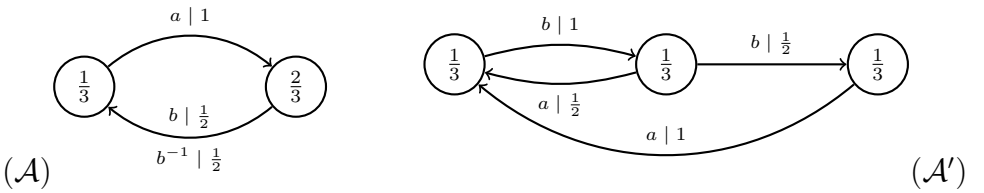

Figure 4: Markovian automata $\mathcal{A}$ and $\mathcal{A}'$. Transitions are labeled by a letter and a probability, and each state is decorated with the corresponding initial probability.

and $(\mathbb{T}_n)_n$, where the support of $\mathbb{T}_n$ is reduced to the α^{dn}-tuple $(n, \dots, n)$. The usual density model corresponds to $\alpha = \frac{1}{2r-1}$.

The following is a generalization of the results in Section 3.2 [6, Proposition 4.3 and Corollary 4.5].

Theorem 4.3 *Let $\mathcal{A}$ be a Markovian automaton.*

If $\mathcal{A}$ does not have a cycle with probability 1, then the induced sequence of probability laws on $\mathcal{R}$ is prefix-heavy, with computable parameters (C, α).

If that is the case, then in the density model with respect to $\mathcal{A}$, at α-density $d < \frac{1}{4}$, a tuple of reduced words exponentially generically has the central tree property.

And at α-density $d < \frac{1}{16}$, a tuple of reduced words exponentially generically generates a malnormal subgroup.

Sketch of proof. Let Q, γ_0 and M be, respectively, the state set, the initial probability vector and the stochastic matrix of $\mathcal{A}$. If $p \in Q$ and $x_1 \cdots x_n \in A^*$, let

$$\gamma(p, x_1 \cdots x_n) = M(p, x_1) M(p \cdot x_1, x_2) \cdots M(p \cdot (x_1 \cdots x_{n-1}), x_n),$$

so that $\mathbb{R}_n(u) = \sum_{p \in Q} \gamma_0(p) \gamma(p, u)$ for every word u. Let ℓ be the maximum length of an elementary cycle in $\mathcal{A}$ and let δ be the maximal value of $\gamma(q, \kappa)$ when κ is an elementary cycle at state q. By assumption, $\delta < 1$. Then, for every cycle u (elementary or not) at a vertex q, we have $\gamma(q, u) \leq \delta^{\frac{|u|}{\ell}}$. Since a path starting from q can be seen as a sequence of cycles interspersed with at most $|Q|$ transitions, we find that $\gamma(q, w) \leq \delta^{\frac{|w|-|Q|}{\ell}}$ for every word w, and hence $\mathbb{R}_n(w) \leq \delta^{\frac{|w|-|Q|}{\ell}}$. Letting

$C = \delta^{\frac{-|Q|}{\ell}}$ and $\alpha = \delta^{\frac{1}{\ell}}$, we find that, if $n \geq |uv|$,

$$\begin{aligned}\mathbb{R}_n(\mathcal{P}(uv)) = \mathbb{R}_{|uv|}(uv) &= \sum_{p\in Q} \gamma_0(p)\gamma(p,u)\gamma(p\cdot u, v) \\ &\leq \left(\sum_{p\in Q} \gamma_0(p)\gamma(p,u)\right) C\alpha^{|v|} \\ &= \mathbb{R}_{|u|}(u)C\alpha^{|v|} = \mathbb{R}_n(\mathcal{P}(u))C\alpha^{|v|}.\end{aligned}$$

We now consider the probability P that an α^{dn}-tuple $\vec{h}$ of reduced words in $\mathcal{R}_n$ fails to satisfy the central tree property (see Section 3.2), that is, some word of length $t = \frac{1}{2}n$ occurs as a prefix of h_i or h_i^{-1}, and of h_j or h_j^{-1}, for some $i < j$ (it is not possible for a word of that length to occur as a prefix of both h_i and h_i^{-1}). It is easily seen that $P \leq 4\ \Sigma_{i<j}\Sigma_{w\in\mathcal{R}_t}\mathbb{R}_n(\mathcal{P}(w))^2$. Since $\mathbb{R}_n(\mathcal{P}(w)) \leq C\alpha^t$ for each w, we have $\Sigma_{w\in\mathcal{R}_t}\mathbb{R}_n(\mathcal{P}(w))^2 \leq (\Sigma_{w\in\mathcal{R}_t}\mathbb{R}_n(\mathcal{P}(w)))C\alpha^t = C\alpha^t$ and hence $P \leq 4\alpha^{(2d-\frac{1}{2})n}$. If $d < \frac{1}{4}$, this vanishes exponentially fast, as announced.

The proof of the statement on malnormality is established in the same spirit, using Proposition 3.1 (4). □

Remark 4.4 Theorem 4.3 holds also for the variant of the α-density model where the words are picked in $\mathcal{R}_{\leq n}$ instead of $\mathcal{R}_n$. Its proof then requires the more technical statements in Theorem 4.1. □

We get more precise results if the Markovian automaton $\mathcal{A}$ is *ergodic*, that is, if its underlying graph is strongly connected and if, for every large enough n, there are paths of length n from every state to every other one. In that situation, it is well known that $\mathcal{A}$ has a stationary vector $\tilde{\gamma} \in [0,1]^Q$, and we let $(\tilde{\mathbb{R}}_n)_n$ be the sequence of probability laws defined by $\mathcal{A}$ with initial vector $\tilde{\gamma}$ instead of γ_0. We say that $\mathcal{A}$ is *non-degenerate* if $\sum_{a\in\tilde{A}} \mathbb{R}_n(a)\tilde{\mathbb{R}}_n(a^{-1}) \neq 1$. Finally we define the *coincidence probability* $\alpha_{[2]}$ of $\mathcal{A}$ as follows: let $M_{[2]}$ be the $((Q\times\tilde{A})\times(Q\times\tilde{A}))$-matrix with entries

$$M_{[2]}((p,a),(q,b)) = \begin{cases}\gamma(p,b)^2 & \text{if } p\cdot b = q,\\ 0 & \text{otherwise.}\end{cases}$$

Then $\alpha_{[2]}$ is the largest eigenvalue of $M_{[2]}$. Bassino *et al.* proved the following phase transition result, which generalizes Theorem 2.1 and part of Theorem 2.2, see [6, Propositions 4.9 and 4.14, Theorem 4.15].

Theorem 4.5 *Let $\mathcal{A}$ be a non-degenerate ergodic Markovian automaton. Then the induced sequence of probability laws on $\mathcal{R}$ is prefix-heavy with parameters $(C, \sqrt{\alpha_{[2]}})$, for some computable $C \geq 1$, and it does not ignore cyclically reduced words.*

In particular, in the density model with respect to $\mathcal{A}$, at $\alpha_{[2]}$-density $d < \frac{1}{8}$, a tuple of reduced words exponentially generically has the central tree property; and at $\alpha_{[2]}$-density $d < \frac{1}{32}$, a tuple of reduced words exponentially generically generates a malnormal subgroup.

Moreover, let $0 < \lambda < \frac{1}{2}$. Then at $\alpha_{[2]}$-density $d < \frac{\lambda}{2}$, a tuple of cyclically reduced words exponentially generically satisfies Property $C'(\lambda)$. And at $\alpha_{[2]}$-density $d > \frac{\lambda}{2}$, it exponentially generically does not satisfy Property $C'(\lambda)$.

Finally, at $\alpha_{[2]}$-density $d > \frac{1}{2}$, a tuple $\vec{h}$ of cyclically reduced words exponentially generically presents a degenerate group, in the following sense: let $B \subseteq \tilde{A}$ be the set of letters which label a transition in $\mathcal{A}$ and let $D = A \setminus (B \cup B^{-1})$. Then $\langle A \mid \vec{h} \rangle$ is equal to the free group of rank $|D| + 1$ if $B \cap B^{-1} = \emptyset$, and otherwise to $F(D) * \mathbb{Z}/2\mathbb{Z}$ if n is even, $F(D)$ if n is odd.

References

[1] J. M. Alonso, T. Brady, D. Cooper, V. Ferlini, M. Lustig, M. Mihalik, M. Shapiro, and H. Short. Notes on word hyperbolic groups. In *Group theory from a geometrical viewpoint (Trieste, 1990)*, pages 3–63. World Sci. Publ., River Edge, NJ, 1991. Edited by Short.

[2] G. N. Arzhantseva. A property of subgroups of infinite index in a free group. *Proc. Amer. Math. Soc.*, 128(11):3205–3210, 2000.

[3] G. N. Arzhantseva and A. Y. Ol'shanskiĭ. Generality of the class of groups in which subgroups with a lesser number of generators are free. *Mat. Zametki*, 59(4):489–496, 638, 1996.

[4] F. Bassino, A. Martino, C. Nicaud, E. Ventura, and P. Weil. Statistical properties of subgroups of free groups. *Random Structures Algorithms*, 42(3):349–373, 2013.

[5] F. Bassino, C. Nicaud, and P. Weil. Random generation of finitely generated subgroups of a free group. *Internat. J. Algebra Comput.*, 18(2):375–405, 2008.

[6] F. Bassino, C. Nicaud, and P. Weil. Generic properties of subgroups of free groups and finite presentations. In *Algebra and computer science*, volume 677 of *Contemp. Math.*, pages 1–43. Amer. Math. Soc., Providence, RI, 2016.

[7] F. Bassino, C. Nicaud, and P. Weil. On the genericity of Whitehead minimality. *J. Group Theory*, 19(1):137–159, 2016.

[8] G. Baumslag, A. Myasnikov, and V. Remeslennikov. Malnormality is decidable in free groups. *Internat. J. Algebra Comput.*, 9(6):687–692, 1999.

[9] O. Bernardi and O. Giménez. A linear algorithm for the random sampling from regular languages. *Algorithmica*, 62(1-2):130–145, 2012.

[10] J.-C. Birget, S. Margolis, J. Meakin, and P. Weil. PSPACE-complete problems for subgroups of free groups and inverse finite automata. *Theoret. Comput. Sci.*, 242(1-2):247–281, 2000.

[11] J. S. Birman. *Braids, links, and mapping class groups*. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1974. Annals of Mathematics Studies, No. 82.

[12] M. R. Bridson and D. T. Wise. Malnormality is undecidable in hyperbolic groups. *Israel J. Math.*, 124:313–316, 2001.

[13] C. Champetier. Propriétés statistiques des groupes de présentation finie. *Journal of Advances in Mathematics*, 116(2):197–262, 1995.

[14] M. Cordes, M. Duchin, Y. Duong, M.-C. Ho, and A. P. Sánchez. Random nilpotent groups I. *International Mathematics Research Notices*, to appear, 2016. arXiv:1506.01426.

[15] K. Delp, T. Dymarz, and A. Schaffer-Cohen. A matrix model for random nilpotent groups. arXiv:1602.01454, 2016.

[16] J. D. Dixon. Probabilistic group theory. *C. R. Math. Acad. Sci. Soc. R. Can.*, 24(1):1–15, 2002.

[17] D. B. A. Epstein, J. W. Cannon, D. F. Holt, S. V. F. Levy, M. S. Paterson, and W. P. Thurston. *Word Processing in Groups*. Jones and Bartlett, Boston, 1992.

[18] P. Flajolet, P. Zimmerman, and B. Van Cutsem. A calculus for the random generation of labelled combinatorial structures. *Theoret. Comput. Sci.*, 132(1-2):1–35, 1994.

[19] A. Garreta. *The diophantine problem over random diophantine groups*. PhD thesis, Stevens Institute of Technology, 2016.

[20] A. Garreta, A. Miasnikov, and D. Ovchinnikov. Properties of random nilpotent groups. Technical report, `arXiv:1612.01242`, 2016.

[21] A. Garreta, A. Miasnikov, and D. Ovchinnikov. Random nilpotent groups, polycyclic presentations, and Diophantine problems. *Groups Complex. Cryptol.*, 9(2):99–115, 2017.

[22] S. M. Gersten. On Whitehead's algorithm. *Bull. Amer. Math. Soc. (N.S.)*, 10(2):281–284, 1984.

[23] R. Gilman, A. Miasnikov, and D. Osin. Exponentially generic subsets of groups. *Illinois J. Math.*, 54(1):371–388, 2010.

[24] M. Greendlinger. Dehn's algorithm for the word problem. *Comm. Pure Appl. Math.*, 13:67–83, 1960.

[25] M. Greendlinger. An analogue of a theorem of Magnus. *Arch. Math.*, 12:94–96, 1961.

[26] R. I. Grigorchuk. Symmetrical random walks on discrete groups. In *Multicomponent random systems*, volume 6 of *Adv. Probab. Related Topics*, pages 285–325. Dekker, New York, 1980.

[27] M. Gromov. Hyperbolic groups. In *Essays in group theory*, volume 8 of *Math. Sci. Res. Inst. Publ.*, pages 75–263. Springer, New York, 1987.

[28] M. Gromov. Asymptotic invariants of infinite groups. In *Geometric group theory, Vol. 2 (Sussex, 1991)*, volume 182 of *London Math. Soc. Lecture Note Ser.*, pages 1–295. Cambridge Univ. Press, Cambridge, 1993.

[29] D. F. Holt, B. Eick, and E. A. O'Brien. *Handbook of computational group theory*. Discrete Mathematics and its Applications (Boca Raton). Chapman & Hall/CRC, Boca Raton, FL, 2005.

[30] T. Jitsukawa. Malnormal subgroups of free groups. In *Computational and statistical group theory (Las Vegas, NV/Hoboken, NJ, 2001)*, volume 298 of *Contemp. Math.*, pages 83–95. Amer. Math. Soc., Providence, RI, 2002.

[31] S. Kalajdžievski. Automorphism group of a free group: centralizers and stabilizers. *J. Algebra*, 150(2):435–502, 1992.

[32] I. Kapovich and P. Schupp. Delzant's T-invariant, Kolmogorov complexity and one-relator groups. *Comment. Math. Helv.*, 80(4):911–933, 2005.

[33] I. Kapovich and P. Schupp. Genericity, the Arzhantseva-Ol'shanskii method and the isomorphism problem for one-relator groups. *Math. Ann.*, 331(1):1–19, 2005.

[34] I. Kapovich and P. Schupp. On group-theoretic models of randomness and genericity. *Groups Geom. Dyn.*, 2(3):383–404, 2008.

[35] I. Kapovich, P. Schupp, and V. Shpilrain. Generic properties of Whitehead's algorithm and isomorphism rigidity of random one-relator groups. *Pacific J. Math.*, 223(1):113–140, 2006.

[36] I. Kapovich and P. E. Schupp. Random quotients of the modular group are rigid and essentially incompressible. *J. Reine Angew. Math.*, 628:91–119, 2009.

[37] H. Kesten. Symmetric random walks on groups. *Trans. Amer. Math. Soc.*, 92:336–354, 1959.

[38] O. Kharlampovich, A. Miasnikov, and P. Weil. Stallings graphs for quasi-convex subgroups. *J. Algebra*, 488:442–483, 2017.

[39] A. Lubotzky and D. Segal. *Subgroup growth*, volume 212 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 2003.

[40] R. C. Lyndon and P. E. Schupp. *Combinatorial group theory*. Springer-Verlag, Berlin, 1977. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 89.

[41] I. G. Lysënok. Some algorithmic properties of hyperbolic groups. *Izv. Akad. Nauk SSSR Ser. Mat.*, 53(4):814–832, 912, 1989.

[42] L. Markus-Epstein. Stallings foldings and subgroups of amalgams of finite groups. *Internat. J. Algebra Comput.*, 17(8):1493–1535, 2007.

[43] A. G. Myasnikov and A. Ushakov. Random subgroups and analysis of the length-based and quotient attacks. *J. Math. Cryptol.*, 2(1):29–61, 2008.

[44] J. Nielsen. Die Isomorphismen der allgemeinen, unendlichen Gruppe mit zwei Erzeugenden. *Mathematische Annalen*, 78, 1918.

[45] Y. Ollivier. Sharp phase transition theorems for hyperbolicity of random groups. *Geom. Funct. Anal.*, 14(3):595–679, 2004.

[46] Y. Ollivier. *A January 2005 invitation to random groups*, volume 10 of *Ensaios Matemáticos [Mathematical Surveys]*. Sociedade Brasileira de Matemática, 2005.

[47] Y. Ollivier. On a small cancellation theorem of Gromov. *Bull. Belg. Math. Soc. Simon Stevin*, 13(1):75–89, 2006.

[48] Y. Ollivier. Some small cancellation properties of random groups. *Internat. J. Algebra Comput.*, 17(1):37–51, 2007.

[49] A. Y. Ol'shanskiĭ. Almost every group is hyperbolic. *Internat. J. Algebra Comput.*, 2(1):1–17, 1992.

[50] E. Rips. Subgroups of small cancellation groups. *Bull. London Math. Soc.*, 14(1):45–47, 1982.

[51] M. Sapir and I. Špakulová. Almost all one-relator groups with at least three generators are residually finite. *J. Eur. Math. Soc. (JEMS)*, 13(2):331–343, 2011.

[52] P. E. Schupp. Embeddings into simple groups. *J. London Math. Soc. (2)*, 13(1):90–94, 1976.

[53] P. Silva, X. Soler-Escriva, and E. Ventura. Finite automata for Schreier graphs of virtually free groups. *J. Group Theory*, to appear, 2015.

[54] J. R. Stallings. Topology of finite graphs. *Invent. Math.*, 71(3):551–565, 1983.

[55] N. W. M. Touikan. A fast algorithm for Stallings' folding process. *Internat. J. Algebra Comput.*, 16(6):1031–1045, 2006.

[56] J. H. C. Whitehead. On equivalent sets of elements in a free group. *Ann. of Math. (2)*, 37(4):782–800, 1936.

[57] A. Żuk. Property (T) and Kazhdan constants for discrete groups. *Geom. Funct. Anal.*, 13(3):643–670, 2003.