

Silhouettes and generic properties of subgroups of the modular group

Frédérique Bassino, bassino@lipn.fr

Université Sorbonne Paris Nord, LIPN, CNRS UMR 7030, F-93430 Villetaneuse, France

Cyril Nicaud, cyril.nicaud@u-pem.fr

LIGM, Univ Gustave Eiffel, CNRS, ESIEE Paris, F-77454, Marne-la-Vallée, France

Pascal Weil, pascal.weil@labri.fr

Univ. Bordeaux, CNRS, Bordeaux INP, LaBRI, UMR 5800, F-33400 Talence, France*

CNRS, ReLaX, UMI 2000, Siruseri, India

November 15, 2023

Abstract

We show that the probability for a finitely generated subgroup of the modular group, of size n , to be almost malnormal or non-parabolic, tends to 0 as n tends to infinity — where the notion of the size of a subgroup is based on a natural graph-theoretic representation of the subgroup.

The proofs of these results rely on the combinatorial and asymptotic study of a natural map, which associates with any finitely generated subgroup of $\mathrm{PSL}_2(\mathbb{Z})$ a graph which we call its silhouette, which can be interpreted as a conjugacy class of free finite index subgroups of $\mathrm{PSL}_2(\mathbb{Z})$.

1 Introduction

The study of the modular group $\mathrm{PSL}_2(\mathbb{Z})$ and of its subgroups has played a central role in algebra, number theory and geometry since the late 19th century. This paper fits in this stream of research. We point out that, while a vast literature has concentrated on the finite index subgroups of $\mathrm{PSL}_2(\mathbb{Z})$, we deal instead with *all* finitely generated subgroups of $\mathrm{PSL}_2(\mathbb{Z})$.

Recall that a subgroup H is *almost malnormal* if $H \cap H^x$ (where H^x stands for $x^{-1}Hx$) is finite for every $x \notin H$; it is *parabolic* if it contains a parabolic element. Our main results

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

are the following: almost malnormality is negligible and parabolicity is generic for finitely generated subgroups of $\mathrm{PSL}_2(\mathbb{Z})$.

These results refer implicitly to a distribution of probabilities on the set of finitely generated subgroups of $\mathrm{PSL}_2(\mathbb{Z})$, that we now explain. With every finitely generated subgroup H of $\mathrm{PSL}_2(\mathbb{Z})$, one associates a (computable) finite edge-labeled graph $\Gamma(H)$, called its *Stallings graph*. The notion is an extension of the graphs Stallings introduced in 1983 [25] to represent finitely generated subgroups of free groups. The idea of using finite graphs to study subgroups of infinite groups has a long and distinguished history, with results of Gersten and Short [12, 24], Arzhantseva and Ol’shanskiĭ [2, 1], Gitik [14] and Kapovich [17] in the 1990s. The definition we use here finds its roots in Gitik’s work on hyperbolic groups [14] and in explicit constructions by Markus-Epstein [22] for subgroups of amalgamated products of finite groups. Its exact form was given by Kharlampovich, Miasnikov and Weil [18].

More precise statements are given in Section 2. At this point, we just want to point out the following. Say that the *size* of a finitely generated subgroup $H \leq \mathrm{PSL}_2(\mathbb{Z})$ is the number of vertices in its Stallings graph¹. Then, for each $n \geq 1$, $\mathrm{PSL}_2(\mathbb{Z})$ has only finitely many subgroups of size n , and we consider the uniform distribution on this finite set.² A property of subgroups is *negligible* (respectively, *generic*) if the proportion of size n subgroups with the property tends to 0 (respectively, 1) when n tends to infinity.

Note that this randomness model strongly differs from those considered by Gilman, Miasnikov and Osin in [13] for subgroups of hyperbolic groups and by Maher and Sisto [21] for subgroups of acylindrically hyperbolic groups. Roughly speaking, these models rely on fixing an integer k , randomly choosing k elements of G using an n -step Markovian mechanism, considering the subgroup generated by these k elements, and letting n tend to infinity. It is interesting to note that in this *few generators* model, almost malnormality is generic [21, Theorem 1.1(2)], in contrast with our model. A similar situation is already known to arise for subgroups of free groups: in the k -generated model, malnormality is generic (Jitsukawa [16]), whereas in the so-called graph-based model, it is negligible (Bassino, Martino, Nicaud and Weil [3]). Jitsukawa’s result was extended also to the Gromov density model, where the number k of generators is allowed to vary as an exponential function of n (Bassino, Nicaud and Weil [5]). One can argue that the model we consider in this paper is particularly natural since Stallings graphs are in bijection with subgroups.

The proof of our results relies on a natural construction which we call the *silhouetting* of the Stallings graph of a finitely generated subgroup of $\mathrm{PSL}_2(\mathbb{Z})$. This construction was introduced by the authors in [8], to give an efficient random generation algorithm for subgroups of $\mathrm{PSL}_2(\mathbb{Z})$ of a given size and isomorphism type. The silhouette of the Stallings graph of a subgroup H is, except in extremal cases, a uniform degree loop-free graph which captures essential features of the “shape” of $\Gamma(H)$. It is obtained by a sequence

¹For a finite index subgroup, this *size* coincides with the index of the subgroup.

²Exact and asymptotic enumeration results for the size n subgroups of $\mathrm{PSL}_2(\mathbb{Z})$ can be found in [7, 8].

of “simplifications” of the graph. Our result in this paper exploits a remarkable, and somewhat surprising property of the silhouetting construction, namely the fact that it preserves uniformity. More precisely, among the size n subgroups whose silhouette has size s , every size s silhouette graph is equally likely.

This property allows us to lift asymptotic properties of silhouette graphs, which are more easily understood, to all Stallings graphs of finitely generated subgroups of $\mathrm{PSL}_2(\mathbb{Z})$. The proof of the uniformity preservation result is of a combinatorial nature, and relies on a fine description of the simplification operations carried out in [8].

We would like to point out also an intermediate result which may be of independent interest. We show that, with high probability, in a finite group of permutations generated by a pair of fixpoint-free permutations (σ_2, σ_3) , of order 2 and 3 respectively, the composition $\sigma_2\sigma_3$ admits orbits of a certain, relatively small size (Proposition 5.7). Obtaining such results on the composition of two randomly chosen mappings is notoriously difficult. It is, for instance, a bottleneck in the study of the properties of random deterministic automata [23]. Most known results rely on a fine grain independent analysis of the mappings, but we know very little on their composition. Character theory has been used to tackle this kind of difficulties in the study of combinatorial maps [15, 11]: this approach yields enumeration results on triplets $(\sigma_1, \sigma_2, \tau)$ such that $\sigma_1 \circ \sigma_2 = \tau$, but for a fixed cyclic type of τ only, and it seems very difficult to exploit such results for our purposes (see also [9]).

We note that, while there are a good number of results in the literature about the genericity or negligibility of certain properties of subgroups of free groups ([16, 3, 4, 5, 6]), there are precious few such results for subgroups of other groups. We can cite in this direction, for their pioneering methods, the results of Arzhantseva and Ol’shanskiĭ [2] who show for instance that, for a very large (generic) class of r -generator, k -relator presentations (r, k fixed), all ℓ -generated subgroups ($\ell < r$) are free and quasi-convex; and the results mentioned above of Gilman, Miasnikov and Osin [13] and Maher and Sisto [21] on k -generated subgroups of a fixed hyperbolic or acylindrically hyperbolic group. To our knowledge, our results on almost malnormality and parabolicity for subgroups of the modular group are the first that are based on the distribution of subgroups given by Stallings graphs.

Organization of the paper Section 2 reviews the definitions of the Stallings graph of a subgroup of $\mathrm{PSL}_2(\mathbb{Z})$ and of the combinatorial type of a subgroup.

The silhouetting operation on Stallings graphs is introduced in Section 3. More precisely, we first introduce a number of local moves in a Stallings graph in Section 3.1. Iterating these moves turns out to be a confluent process, leading to the so-called *silhouette* of the given graph or subgroup. We show there that silhouetting preserves uniformity (Theorems 4.1 and 4.4), and that silhouetting only moderately decreases the size of a graph: with high probability (more precisely, super-polynomially generically), the silhouette of a size n subgroup has size at least $n - 3n^{\frac{2}{3}}$ (Propositions 4.2 and 4.5).

The last section, Section 5 contains the proof of our main results: subgroups of $\mathrm{PSL}_2(\mathbb{Z})$ generically contain parabolic elements (Proposition 5.1) and fail to be almost malnormal (Theorem 5.2). Both results exploit the statistical result on the existence of cycles reading non-trivial powers of ab (where a and b are the order 2 and order 3 generators of $\mathrm{PSL}_2(\mathbb{Z})$) in the Stallings graph of a subgroup (Theorem 5.3), whose proof reduces to proving the same (highly non-trivial) result on silhouette graphs (Proposition 5.7).

To conclude this introductory section, we note that, according to the results presented here and in [8], the silhouetting operation is combinatorially and asymptotically significant in the study of finitely generated subgroups of $\mathrm{PSL}_2(\mathbb{Z})$. We are able to lift a statistical property of silhouette graphs to the class of all $\mathrm{PSL}_2(\mathbb{Z})$ -reduced graphs, and it would be interesting to see what other properties can be lifted in that fashion. In addition, we think that the silhouetting operation also has a topological, or possibly a geometric interpretation, even for finite index subgroups, and we would be curious about its properties.

2 Preliminaries

We use the following presentation of the modular group, seen here as the free product of two cyclic groups:

$$\mathrm{PSL}_2(\mathbb{Z}) = \mathbb{Z}_2 * \mathbb{Z}_3 = \langle a, b \mid a^2 = b^3 = 1 \rangle.$$

The elements of $\mathrm{PSL}_2(\mathbb{Z})$ are represented by words over the alphabet $\{a, b, a^{-1}, b^{-1}\}$, or rather of $\{a, b, b^{-1}\}$ since $a^{-1} = a$. Each non-trivial element of $\mathrm{PSL}_2(\mathbb{Z})$ has a unique shortest (or *normal*, or *geodesic*) representative, which is a freely reduced word without factors in $\{a^2, b^2, b^{-2}\}$, that is, a word which either has length 1, or alternates letters a and letters in $\{b, b^{-1}\}$.

An element $g \neq 1$ in $\mathrm{PSL}_2(\mathbb{Z})$ is said to be *cyclically reduced* if it has length 1 or if its normal form starts with a and ends with $b^{\pm 1}$, or starts with $b^{\pm 1}$ and ends with a . It is immediate that every non trivial element of $\mathrm{PSL}_2(\mathbb{Z})$ is conjugated to a cyclically reduced element. Moreover [19, Theorem IV.2.8], two conjugated cyclically reduced elements are cyclic conjugates of one another (that is: their geodesic representatives are of the form, respectively, tt' and $t't$).

With each finitely generated subgroup H of $\mathrm{PSL}_2(\mathbb{Z})$, we associate its Stallings graph as in [18, 7, 8], which is the *interesting*, or *significant* part of the Schreier graph of H .

More precisely, first recall that the *Schreier graph*, or *coset graph* of H is the graph with vertex set $\{Hg \mid g \in \mathrm{PSL}_2(\mathbb{Z})\}$, with an a -labeled edge from Hg to Hga and a b -labeled edge from Hg to Hgb for every $g \in \mathrm{PSL}_2(\mathbb{Z})$. For every $g \in \mathrm{PSL}_2(\mathbb{Z})$, we also have a b^{-1} -labeled edge from Hgb to Hg . A *path* is a finite sequence of consecutive edges, and the word spelled out by the labels of these edges is the *label of the path*. In particular, if w is a word on the alphabet $\{a, b, b^{-1}\}$ and $g \in \mathrm{PSL}_2(\mathbb{Z})$, w labels a path from Hg to Hgw .

It follows that a word w is in H if and only if it labels a *cycle* at vertex $v_0 = H$ (that is: a path starting and ending at vertex H). The *Stallings graph* of H , written $(\Gamma(H), v_0)$,

is defined as the rooted subgraph of the Schreier graph spanned by the cycles at v_0 labeled by the geodesic representatives of the elements of H — that is, it consists of all the vertices and edges of the Schreier graph, which appear in these cycles.

It is immediately verified that, if H has finite index in $\mathrm{PSL}_2(\mathbb{Z})$, say, n , then $\Gamma(H)$ is the whole Schreier graph of H , and H has size n .

Example 2.1 Figure 1 shows examples of Stallings graphs. These graphs are *labeled graphs*, i.e. their vertices are labeled by an initial segment of \mathbb{N} . The definition of Stallings graphs does not entail labeling vertices — only designating a base vertex. \square

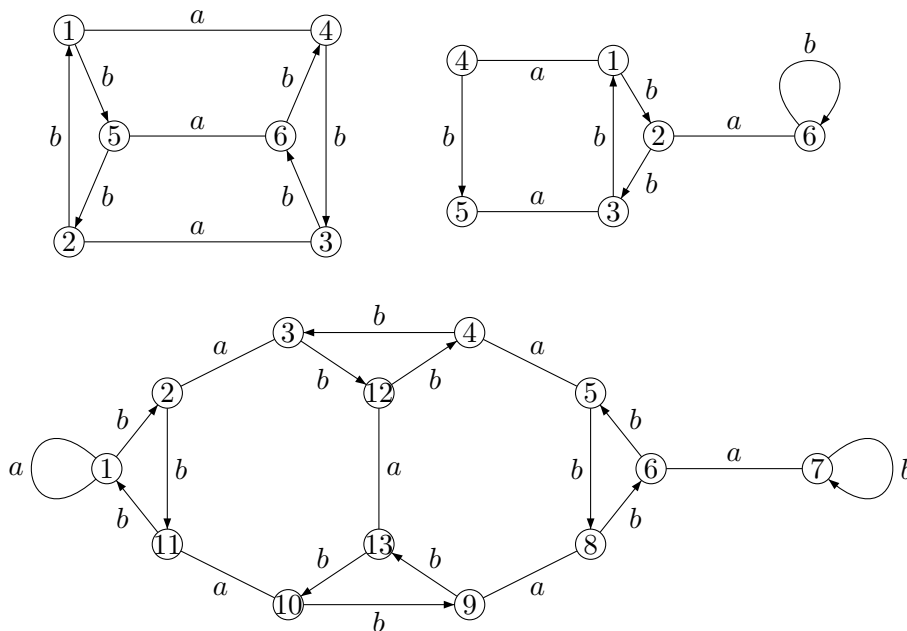


Figure 1: Top: The Stallings graphs of the subgroups $H = \langle abab^{-1}, babab \rangle$ and $K = \langle abab, babab^{-1} \rangle$ of $\mathrm{PSL}_2(\mathbb{Z})$. Bottom: the Stallings graph of $L = \langle a, (ba)^3b, (bab^{-1}a)^2b, (bab^{-1}ab^{-1}a)^2b \rangle$. In each case, the root is the vertex labeled 1.

Remark 2.2 As we noted in the introduction, finite graphs have long been used to discuss properties of arbitrary index subgroups of infinite groups. The connection with quasi-convexity and with the Howson property (which deals with question whether the intersection of finitely generated subgroups is again finitely generated) was observed by Gersten and Short in the early 1990s [12, 24]. Arzhantseva and Ol’shanskiĭ [2] developed the usage of graphs to investigate the subgroups of an exponentially generic class of k -relator groups. The earliest usage of a graphical representation of subgroups to derive algorithmic results, may be Kapovich’s work on detecting quasi-convexity in automatic groups [17] (note that

$\mathrm{PSL}_2(\mathbb{Z})$ is automatic). For a systematic algorithmic approach, it is however very convenient to have a unique graphical representation, such as the Stallings graph defined above. This definition was first introduced³ by Gitik [14] for quasi-convex subgroups of hyperbolic groups (and every finitely generated subgroup of $\mathrm{PSL}_2(\mathbb{Z})$ is quasi-convex). It was generalized by Kharlampovich, Miasnikov and Weil [18], who showed that the quasi-convex subgroups of automatic groups have a finite and computable Stallings graph, which led to a unified approach to a number of algorithmic decidability problems. See [18] for more historical details. \square

Remark 2.3 The computation of the Stallings graph of a quasi-convex subgroup of an automatic group is a high complexity problem, and may not be practical in general. It is simpler when the ambient group is an amalgamated product of finite groups, as showed by Markus-Epstein [22] (although she used slightly different graphs than our Stallings graphs). In the particular case of $\mathrm{PSL}_2(\mathbb{Z})$, the computation of Stallings graphs is particularly straightforward, and can be achieved in time $\mathcal{O}(n \log^* n)$, where n is the sum of the length of the generators of the given subgroup, see [7]: given a tuple of words $\vec{h} = (h_1, \dots, h_r)$ on alphabet $\{a, b, b^{-1}\}$, one first computes the (classical) Stallings graph of the subgroup generated by \vec{h} in the free group $F(a, b)$, as in [25, 26]. This is done starting with a wedge of cycles each labeled by one of the h_i , and then repeatedly applying Stallings foldings (in which two edges with the same label and the same start (resp. end) vertex are identified). The next step is to add an a -labeled edge from vertex v to vertex w whenever there is an a -labeled edge from w to v , add a b -labeled edge from v to w whenever there is a b^2 -labeled path from w to v , and then apply another round of repeated Stallings foldings. \square

It is immediate from the definition of Stallings graphs that $\Gamma(H)$ is connected and that its a -edges (respectively, b -edges) form a partial, injective map on the vertex set of the graph. Moreover, because $a^2 = b^3 = 1$, distinct a -edges are never adjacent to the same vertex: we distinguish therefore a -loops and so-called *isolated a -edges*. Similarly, if we have two consecutive b -edges, say, from v_1 to v_2 and from v_2 to v_3 , then $\Gamma(H)$ also has a b -edge from v_3 to v_1 . Thus each b -edge is either a loop, or an *isolated b -edge*, or a part of a b -triangle. Finally, every vertex except maybe the root vertex is adjacent to an a - and to a b -edge.

A rooted edge-labeled graph satisfying these conditions is called $\mathrm{PSL}_2(\mathbb{Z})$ -*reduced* and it is not difficult to see that every finite $\mathrm{PSL}_2(\mathbb{Z})$ -reduced graph is the Stallings graph of a unique finitely generated subgroup of $\mathrm{PSL}_2(\mathbb{Z})$ [7]. That is, the mapping $H \mapsto (\Gamma(H), v_0)$ is a bijection between finitely generated subgroups of $\mathrm{PSL}_2(\mathbb{Z})$ and $\mathrm{PSL}_2(\mathbb{Z})$ -reduced graphs. The asymptotic results that are at the heart of this paper rely on this bijection: enumeration

³under the name of *geodesic core*.

results for subgroups are equivalent to enumeration results on $\mathrm{PSL}_2(\mathbb{Z})$ -reduced graphs (their Stallings graphs).

An edge-labeled graph is said to be $\mathrm{PSL}_2(\mathbb{Z})$ -*cyclically reduced* if every vertex is adjacent to an a - and a b -edge. This is equivalent to asking that Γ be $\mathrm{PSL}_2(\mathbb{Z})$ -reduced when rooted at every one of its vertices. We also say that a finitely generated subgroup of $\mathrm{PSL}_2(\mathbb{Z})$ is $\mathrm{PSL}_2(\mathbb{Z})$ -cyclically reduced if its Stallings graph is. We note that a subgroup is $\mathrm{PSL}_2(\mathbb{Z})$ -cyclically reduced if and only if it has minimum size in its conjugacy class (see, *e.g.*, [8, Section 2.2].)

As is classical in combinatorics, it is actually more convenient to work with *labeled* and *weakly labeled graphs* (see [10, Section II.1]). A graph is said to be *weakly labeled* if its vertex set is equipped with a (labeling) one-to-one map to $\mathbb{N} \setminus \{0\}$.⁴ It is *labeled* if the range of the labeling map is an initial segment of $\mathbb{N} \setminus \{0\}$ (a set of the form $[n] = \{1, \dots, n\}$). It is immediate that a weakly labeled graph Γ can be relabeled in a unique fashion into a labeled graph, in such a way that the order of vertices is preserved: the resulting labeled graph is called the *normalization*⁵ of Γ , written $\mathrm{norm}(\Gamma)$.

To lighten up notation, we often abusively identify the vertices of a weakly labeled graph with their labels. The graphs in Figure 1 are in fact labeled graphs.

Example 2.4 The $\mathrm{PSL}_2(\mathbb{Z})$ -cyclically reduced graphs Γ with 1 or 2 vertices are represented in Figure 2. There is only one with 1 vertex, and three with 2 vertices. Note that Δ_2 and Δ_3 can be labeled in two different ways while Δ_4 admits only one labeling. \square

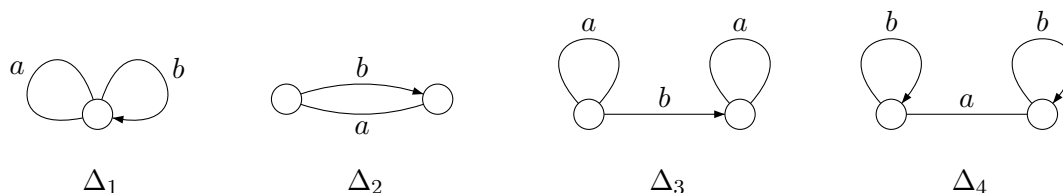


Figure 2: All $\mathrm{PSL}_2(\mathbb{Z})$ -cyclically reduced graphs with at most 2 vertices.

Remark 2.5 Counting of graphs is done up to isomorphism. To be more precise, an isomorphism $\Gamma \rightarrow \Gamma'$ between graphs is a pair of bijections from the vertex set of Γ to the vertex set of Γ' , and from the edge set of Γ to the edge set of Γ' , which preserves the incidence relation. Isomorphisms between rooted graphs must also map the root of one graph to the root of the other. Finally, isomorphisms between edge-labeled graphs must also preserve these labels. \square

⁴This notion of labeling, which injectively assigns an integer to each vertex, must be distinguished from the edge labeling used so far, where each edge is labeled by letter a or b and each path is labeled by a word.

⁵An operation called *reduction* in [10].

If Γ is a $\mathrm{PSL}_2(\mathbb{Z})$ -reduced graph, its *combinatorial type* is the tuple $(n, k_2, k_3, \ell_2, \ell_3)$ where n is the number of vertices of Γ , k_2 and k_3 are the numbers of isolated a - and b -edges, and ℓ_2 and ℓ_3 are the numbers of a - and b -loops. We may also talk of the combinatorial type of a subgroup to mean the combinatorial type of its Stallings graph. The integer n is called the *size* of the graph or the subgroup. One can find in [7, Section 2.3.1] a discussion of the constraints on tuples that arise as combinatorial types. The combinatorial type information on a subgroup refines algebraic information such as freeness or finite index (see, e.g., [7, Propositions 2.7, 2.9, 8.18 and Section 8.2]).

Proposition 2.6 *A subgroup $H \leq \mathrm{PSL}_2(\mathbb{Z})$ has finite index if and only if its Stallings graph is $\mathrm{PSL}_2(\mathbb{Z})$ -cyclically reduced and has combinatorial type of the form $(n, k_2, 0, \ell_2, \ell_3)$. It is free if and only if its combinatorial type is of the form $(n, k_2, k_3, 0, 0)$.*

Free $\mathrm{PSL}_2(\mathbb{Z})$ -cyclically reduced subgroups have even size, and free and finite index subgroups have size a positive multiple of 6.

3 Moves on $\mathrm{PSL}_2(\mathbb{Z})$ -reduced graphs and silhouette of a subgroup

In this section, we review a combinatorial construction on labeled $\mathrm{PSL}_2(\mathbb{Z})$ -cyclically reduced graphs first introduced in [8]. It consists in applying a number of moves on such a graph Γ , depending on its geometry, see Section 3.1. These moves constitute a confluent and terminating graph rewriting system (see Section 3.2): this means in particular that iteratively applying these moves to a graph Γ leads to a uniquely defined $\mathrm{PSL}_2(\mathbb{Z})$ -cyclically reduced graph $\mathrm{silh}(\Gamma)$, which we call the *silhouette* of Γ .

This silhouetting operation was studied in [8] from a combinatorial (and, to a lesser extent, algebraic) point of view, towards exact enumeration and random generation results. We investigate in Section 4 asymptotic properties of this operation. These properties, in turn, are used to establish certain asymptotic properties of subgroups of $\mathrm{PSL}_2(\mathbb{Z})$, see Section 5 below.

3.1 Moves on a labeled $\mathrm{PSL}_2(\mathbb{Z})$ -cyclically reduced graph

Except for the so-called exceptional moves defined further down, the moves we define on weakly labeled $\mathrm{PSL}_2(\mathbb{Z})$ -cyclically reduced graphs delete vertices without changing their labels: they result in smaller weakly labeled $\mathrm{PSL}_2(\mathbb{Z})$ -cyclically reduced graphs.

We start with the λ_3 -moves, which decrease the number of b -loops. Let Γ be a weakly labeled $\mathrm{PSL}_2(\mathbb{Z})$ -cyclically reduced graph and let v, w be distinct vertices such that Γ has a b -loop at v and an a -edge between v and w , see Figure 3. The corresponding λ_3 -move deletes vertex v and the edges adjacent to it, and adds an a -loop at vertex w .

Next we define the two kinds of λ_2 -moves, which decrease the number of a -loops. Let Γ be a weakly labeled $\mathrm{PSL}_2(\mathbb{Z})$ -cyclically reduced graph and let v be a vertex such that Γ

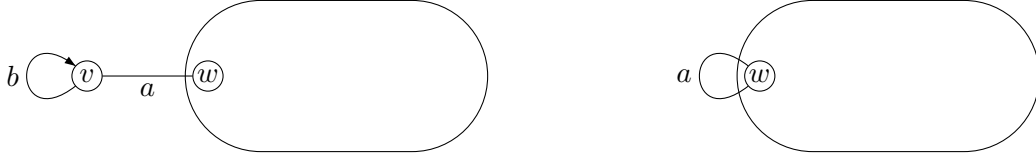


Figure 3: A λ_3 -move

has an a -loop at v and v is part of a b -triangle, see Figure 4. The corresponding $\lambda_{2,1}$ -move deletes vertex v and the edges adjacent to it.

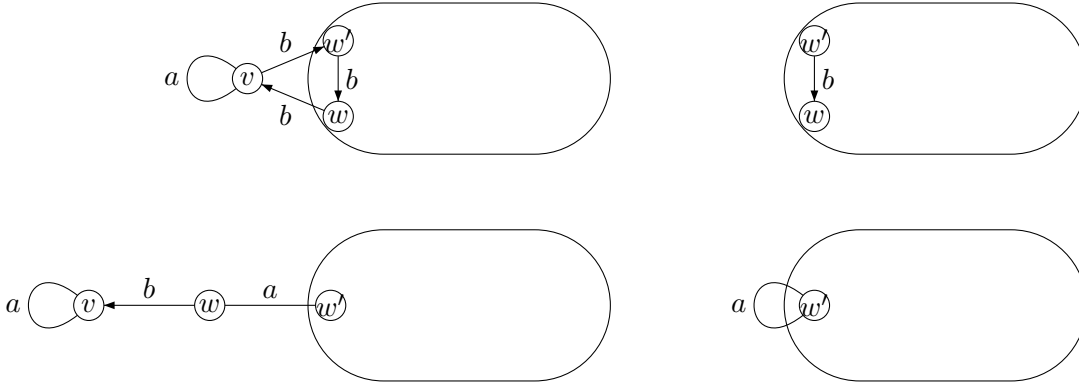


Figure 4: Above: A $\lambda_{2,1}$ -move. Below: a $\lambda_{2,2}$ -move

The second kind of λ_2 -moves is as follows. Let Γ be a weakly labeled $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graph and let v, w, w' be distinct vertices such that Γ has an a -loop at v , an isolated b -edge between v and w (in either direction) and an a -edge between w and w' , see Figure 3. The corresponding $\lambda_{2,2}$ -move deletes vertices v and w and the edges adjacent to them, and adds an a -loop at vertex w' .

Our next kind of moves deletes isolated b -edges connecting vertices that sit on distinct isolated a -edges. More precisely, let Γ be a weakly labeled $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graph and let v, w, v', w' be distinct vertices such that Γ has isolated a -edges between v and v' , and between w and w' , and an isolated b -edge between v and w (in either direction), see Figure 5. The corresponding κ_3 -move deletes vertices v and w and the edges adjacent to them, and adds an a -edge between v' and w' .

Finally, we introduce three *exceptional* moves, which can modify the labeling of vertices. One transforms an improperly labeled version of Δ_1 to the (unique) labeled version of Δ_1 (that is: if the unique vertex of Δ_1 is not labeled by 1, the label of that vertex is made 1. Another exceptional move can be applied only to a weakly labeled version of Δ_3 , turning it into the labeled version of Δ_1 . This move can be seen as a degenerate version of a $\lambda_{2,2}$ -move. The last exceptional move can be applied to any weakly labeled version of

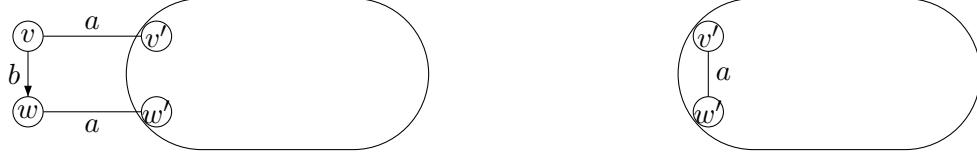


Figure 5: A κ_3 -move

Δ_2 different from the particular labeling showed in Figure 6, turning it to that preferred labeling of Δ_2 .

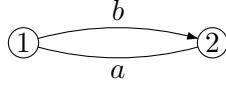


Figure 6: The preferred labeling of Δ_2

It is easily verified that the λ_3 -, λ_2 - and κ_3 -moves modify the combinatorial type of a $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graph Γ as follows. Let τ be the combinatorial type of Γ .

- If Δ is obtained from Γ by a λ_3 -move, then Δ has combinatorial type $\tau + \lambda_3$, where $\lambda_3 = (-1, -1, 0, 1, -1)$.
- If Δ is obtained from Γ by a $\lambda_{2,1}$ -move, then Δ has combinatorial type $\tau + \lambda_{2,1}$, where $\lambda_{2,1} = (-1, 0, 1, -1, 0)$.
- If Δ is obtained from Γ by a $\lambda_{2,2}$ -move, then Δ has combinatorial type $\tau + \lambda_{2,2}$, where $\lambda_{2,2} = (-2, -1, -1, 0, 0)$.
- If Δ is obtained from Γ by a κ_3 -move, then Δ has combinatorial type $\tau + \kappa_3$, where $\kappa_3 = (-2, -1, -1, 0, 0)$.

The combinatorial study of these moves carried out in [8, Propositions 4.4, 4.5, 4.6] establishes the following enumeration results.

Proposition 3.1 *Let $\tau = (n, k_2, k_3, \ell_2, \ell_3)$ be a combinatorial type.*

- *If $n \geq 2$, $\ell_3 > 0$ and Δ is a $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graph with combinatorial type $\tau + \lambda_3$, then the set of labeled $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graphs Γ of combinatorial type τ , such that a λ_3 -move takes Γ to Δ , has $\frac{n(\ell_2+1)}{\ell_3}$ elements.*
- *If $n \geq 3$, $\ell_2 > 0$ and Δ has combinatorial type $\tau + \lambda_{2,1}$, then the set of labeled $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graphs Γ of combinatorial type τ , such that a $\lambda_{2,1}$ -move takes Γ to Δ , has $\frac{n(k_3+1)}{\ell_2}$ elements.*

- If $n \geq 3$, $\ell_2 > 0$ and Δ has combinatorial type $\tau + \lambda_{2,2}$, then the set of labeled $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graphs Γ of combinatorial type τ , such that a $\lambda_{2,2}$ -move takes Γ to Δ , has $2n(n-1)$ elements.
- If $n \geq 4$, $\ell_2 = 0$, $k_3 > 0$ and Δ has combinatorial type $\tau + \kappa_3$, then the set of labeled $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graphs Γ of combinatorial type τ , such that a κ_3 -move takes Γ to Δ , has $\frac{2n(n-1)(k_2-1)}{k_3}$ elements.

3.2 Silhouette graphs, silhouette of a labeled $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graph

Say that a $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graph Γ (with or without a weak labeling function) is a *silhouette graph* if it is equal to Δ_1 or Δ_2 , or if it has combinatorial type $(n, n/2, 0, 0, 0)$. By Proposition 2.6, the latter are the Stallings graphs of free and finite index subgroups of $\text{PSL}_2(\mathbb{Z})$ where we forget which vertex is the base vertex, and their size n is a positive multiple of 6. It is clear that, under any labeling, no move is defined on a silhouette graph of size $n > 2$.

The rewriting system on weakly labeled $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graphs, given by the moves defined in Section 3.1, was showed to be *confluent* in the following sense [8, Proposition 3.4].

Proposition 3.2 *Let Γ be a weakly labeled $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graph Γ and let Δ and Δ' be obtained from Γ after maximal sequences of λ_3 -, $\lambda_{2,1}$ -, $\lambda_{2,2}$ - and κ_3 - and exceptional moves. Then $\Delta = \Delta'$.*

In view of this result, we define the *quasi-silhouette* and the *silhouette* of a $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graph Γ as $\text{q-silh}(\Gamma) = \Delta$ and $\text{silh}(\Gamma) = \text{norm}(\Delta)$, where Δ is obtained from Γ after a maximal sequence of moves. Proposition 3.2 establishes that the quasi-silhouette (resp. silhouette) of Γ is a well defined weakly labeled (resp. labeled) $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graph, which does not depend on the choice of a particular maximal sequence of moves.

Example 3.3 The first Stallings graph in Example 2.1 is a silhouette graph: it is therefore its own silhouette and quasi-silhouette. The silhouette of the second Stallings graph is Δ_2 with its preferred labeling (see Figure 6). Finally, the quasi-silhouette and silhouette of the last graph are given in Figure 7. \square

Remark 3.4 A useful consequence of Proposition 3.2 is the following: given a $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graph Γ , its quasi-silhouette can always be computed by iteratively performing λ_3 -moves until there are no b -loops left, then iteratively performing λ_2 -moves until there are no a -loops left (note that these moves do not add b -loops), then iteratively performing κ_3 -moves until we reach a graph of size at most 2, or there are no more loops or isolated b -edges. In the first case, one may have to perform an exceptional move; and

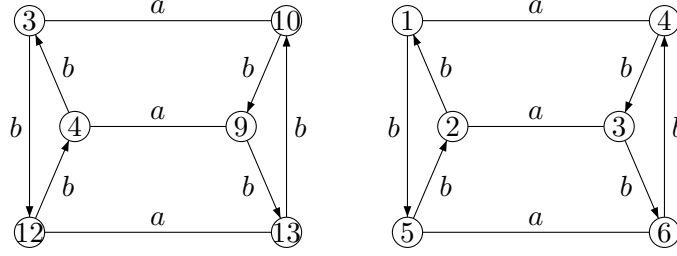


Figure 7: The quasi-silhouette and the silhouette of the labeled Stallings graph of $L = \langle a, (ba)^3b, (bab^{-1}a)^2b, (bab^{-1}ab^{-1}a)^2b \rangle$ given in Figure 1.

in the second case, a normalization step is required to obtain $\text{silh}(\Gamma)$. Several of the proofs below use this possibility of choosing the “path” to compute $\text{q-silh}(\Gamma)$ and $\text{silh}(\Gamma)$ in this fashion. \square

The silhouetting operation captures algebraic properties of subgroups of $\text{PSL}_2(\mathbb{Z})$ (see [8, Proposition 3.4], but we concentrate here on its probabilistic properties.

4 Probabilistic properties of the silhouetting operation

In this section, we first explore two properties of the silhouetting operation which are central to the proof of our main results. The first is that the map $\Gamma \mapsto \text{silh}(\Gamma)$ preserves uniformity on the set of labeled $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graphs of a given combinatorial type, and whose silhouette have a prescribed size, see Section 4.1 for the precise statement.

The second property of the silhouetting operation is that the expected number of vertices of the silhouette of a size n $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graph is close to n , and is concentrated around this expected value, again see Section 4.2 for a precise statement.

Finally, we show how the silhouetting operation can be extended to all finitely generated subgroups of $\text{PSL}_2(\mathbb{Z})$ (not just the $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced subgroups), and that the properties mentioned above also hold in that case.

4.1 Silhouetting preserves uniformity on labeled $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graphs

If $\tau = (n, k_2, k_3, \ell_2, \ell_3)$ and $1 \leq s \leq n$, we let $\mathfrak{G}_\tau(s)$ (resp. $\mathfrak{G}_n(s)$) be the set of labeled $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graphs of type τ (resp. of size n) whose silhouette has size s . Note that $\mathfrak{G}_s(s)$ is the set of labeled silhouette graphs of size s . Let $\mathbb{P}_{\tau,s}$ (resp. $\mathbb{P}_{n,s}$) denote the uniform probability on $\mathfrak{G}_\tau(s)$ (resp. $\mathfrak{G}_n(s)$). We show that the map $\text{silh}: \Gamma \mapsto \text{silh}(\Gamma)$, from $\mathfrak{G}_\tau(s)$ (resp. $\mathfrak{G}_n(s)$) to $\mathfrak{G}_s(s)$, preserves uniformity: that is, if X is a subset of $\mathfrak{G}_s(s)$ and $X_{\tau,s}$ (resp. $X_{n,s}$) is the set of elements Γ of $\mathfrak{G}_\tau(s)$ (resp. $\mathfrak{G}_n(s)$) such that $\text{silh}(\Gamma) \in X$, then $\mathbb{P}_{\tau,s}(X_{\tau,s}) = \mathbb{P}_{n,s}(X_{n,s}) = \mathbb{P}_{s,s}(X)$.

Theorem 4.1 *Let $\tau = (n, k_2, k_3, \ell_2, \ell_3)$ be a combinatorial type and let $1 \leq s \leq n$. If Γ is an element of $\mathfrak{G}_\tau(s)$ (resp. $\mathfrak{G}_n(s)$) taken uniformly at random, then $\text{silh}(\Gamma)$ is a uniformly random element of the set $\mathfrak{G}_s(s)$ of size s labeled silhouette graphs. That is: for any $\Delta, \Delta' \in \mathfrak{G}_s(s)$, we have*

$$\begin{aligned} \mathbb{P}_{\tau,s}(\Gamma \in \mathfrak{G}_\tau(s) \text{ and } \text{silh}(\Gamma) = \Delta) &= \mathbb{P}_{\tau,s}(\Gamma \in \mathfrak{G}_\tau(s) \text{ and } \text{silh}(\Gamma) = \Delta') \\ \mathbb{P}_{n,s}(\Gamma \in \mathfrak{G}_n(s) \text{ and } \text{silh}(\Gamma) = \Delta) &= \mathbb{P}_{n,s}(\Gamma \in \mathfrak{G}_n(s) \text{ and } \text{silh}(\Gamma) = \Delta'). \end{aligned}$$

The proof of Theorem 4.1 uses the following lemma, a “weakly labeled version” of the general statement. For $s > 2$, we let $\mathfrak{D}_n(s)$ be the set of size s silhouette graphs weakly labeled with elements of $[n]$ (in particular, $\mathfrak{D}_s(s) = \mathfrak{G}_s(s)$). If $\Delta \in \mathfrak{D}_n(s)$, we let $\mathbf{q}\mathfrak{G}_\tau(\Delta)$ (resp. $\mathbf{q}\mathfrak{G}_n(\Delta)$) be the set of labeled graphs in $\mathfrak{G}_\tau(s)$ (resp. $\mathfrak{G}_n(s)$) whose quasi-silhouette is Δ .

Lemma 4.2 *Let $\tau = (n, k_2, k_3, \ell_2, \ell_3)$ be a combinatorial type and let $3 \leq s \leq n$. If $\Delta, \Delta' \in \mathfrak{D}_n(s)$, then the sets $\mathbf{q}\mathfrak{G}_\tau(\Delta)$ and $\mathbf{q}\mathfrak{G}_\tau(\Delta')$ have the same cardinality.*

Proof. We show by induction on n that the cardinality of $\mathbf{q}\mathfrak{G}_\tau(\Delta)$ depends only on τ , not on Δ .

If $n = s$, then $\tau = (s, s/2, 0, 0, 0)$ and $\mathbf{q}\mathfrak{G}_\tau(\Delta) = \{\Delta\}$ has cardinality 1.

Now suppose that $n > s$. We first deal with the case where $\ell_3 > 0$. Let $\Gamma \in \mathbf{q}\mathfrak{G}_\tau(\Delta)$ and let Γ' be obtained from Γ by a λ_3 -move. Then $\mathbf{q}\text{-silh}(\Gamma') = \mathbf{q}\text{-silh}(\Gamma) = \Delta$ by Proposition 3.2 and $\Gamma' \in \mathbf{q}\mathfrak{G}_{\tau+\lambda_3}(\Delta)$. Proposition 3.1 then shows that $|\mathbf{q}\mathfrak{G}_\tau(\Delta)| = \frac{n(\ell_2+1)}{\ell_3} |\mathbf{q}\mathfrak{G}_{\tau+\lambda_3}(\Delta)|$, which only depends on τ .

If $\ell_3 = 0$ and $\ell_2 > 0$, let v be a vertex carrying an a -loop in Γ and let Γ' be the $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graph obtained from Γ by the corresponding λ_2 -move: a $\lambda_{2,1}$ -move if v sits on a b -triangle of Γ , a $\lambda_{2,2}$ -move otherwise. Again, $\mathbf{q}\text{-silh}(\Gamma') = \mathbf{q}\text{-silh}(\Gamma) = \Delta$, so $\Gamma' \in \mathbf{q}\mathfrak{G}_{\tau+\lambda_{2,1}}(\Delta)$ if v is adjacent to a b -triangle and $\Gamma' \in \mathbf{q}\mathfrak{G}_{\tau+\lambda_{2,2}}(\Delta)$ otherwise. Proposition 3.1 then establishes that

$$|\mathbf{q}\mathfrak{G}_\tau(\Delta)| = \frac{n(k_3+1)}{\ell_2} |\mathbf{q}\mathfrak{G}_{\tau+\lambda_{2,1}}(\Delta)| + 2n(n-1) |\mathbf{q}\mathfrak{G}_{\tau+\lambda_{2,2}}(\Delta)|,$$

again showing the only dependence on τ of $|\mathbf{q}\mathfrak{G}_\tau(\Delta)|$.

If $\ell_2 = \ell_3 = 0$ and $k_3 > 0$, Let $\Gamma \in \mathbf{q}\mathfrak{G}_\tau(\Delta)$ and let Γ' be obtained from Γ by a κ_3 -move. Here too $\mathbf{q}\text{-silh}(\Gamma') = \mathbf{q}\text{-silh}(\Gamma) = \Delta$ and $\Gamma' \in \mathbf{q}\mathfrak{G}_{\tau+\kappa_3}(\Delta)$. Proposition 3.1 then shows that $|\mathbf{q}\mathfrak{G}_\tau(\Delta)| = 2 \frac{n(n-1)(k_2-1)}{k_3} |\mathbf{q}\mathfrak{G}_{\tau+\kappa_3}(\Delta)|$, concluding the proof by induction. \square

Proof of Theorem 4.1. If $s \leq 2$, $\mathfrak{G}_s(s)$ has only one element, so the result holds trivially. We now assume that $s \geq 3$. It suffices to show that, if $\Delta \in \mathfrak{G}_s(s)$, the cardinality of the set $\{\Gamma \in \mathfrak{G}_\tau(s) \mid \text{silh}(\Gamma) = \Delta\}$ depends on τ only, not on Δ .

If Γ' is a weakly labeled graph with labels in $[n]$ and σ is a permutation on $[n]$, we denote by $\sigma(\Gamma')$ the graph obtained from Γ' by relabeling each vertex v by $\sigma(v)$. Observe that Γ' and $\sigma(\Gamma')$ have the same combinatorial type.

Moreover, if $\Gamma \in \mathfrak{G}_\tau(s)$ and $\Delta \in \mathfrak{G}_s(s)$, then $\text{silh}(\Gamma) = \Delta$ if and only if $\text{norm}(\mathbf{q}\text{-silh}(\Gamma)) = \Delta$, if and only if $\mathbf{q}\text{-silh}(\Gamma) = \sigma(\Delta)$ for some permutation σ on $[n]$ which is increasing on $[s]$. The number of such permutations depends only on n and s (it is exactly $\frac{n!}{s!}$).

Thus, if τ is a combinatorial type, then the set $\{\Gamma \in \mathfrak{G}_\tau(s) \mid \text{silh}(\Gamma) = \Delta\}$ is the disjoint union of the $\mathbf{q}\mathfrak{G}_\tau(\sigma(\Delta))$, where σ runs over the permutations on $[n]$ which are increasing on $[s]$. By Lemma 4.2, all these sets have the same cardinality, which does not depend on the choice of $\Delta \in \mathfrak{G}_s(s)$. Therefore the cardinality

$$|\{\Gamma \in \mathfrak{G}_\tau(s) \mid \text{silh}(\Gamma) = \Delta\}| = \frac{n!}{s!} |\mathbf{q}\mathfrak{G}_\tau(\Delta)|$$

depends on τ and s , not on Δ .

The corresponding result on the cardinality of $\{\Gamma \in \mathfrak{G}_n(s) \mid \text{silh}(\Gamma) = \Delta\}$ follows, since this set is the disjoint union of the $\{\Gamma \in \mathfrak{G}_\tau(s) \mid \text{silh}(\Gamma) = \Delta\}$ when τ runs over the combinatorial types of size n . \square

4.2 Size of the silhouette of a uniform random $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graph

Let Γ be a labeled $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graph. The computation of $\text{silh}(\Gamma)$ shows how vertices are deleted until there are no more loops or isolated b -edges, or until the process has led to Δ_1 or Δ_2 . Statistically however, only a small number of vertices are deleted. Proposition 4.3, a large deviation result, quantifies this statement. It follows that most vertices and edges of Γ are untouched in the reduction to $\text{silh}(\Gamma)$. This observation will play an important role in our discussion of generic properties of subgroups in Section 5, especially Theorem 5.3. The proof of Proposition 4.3 relies in an essential way on results in [7].

Proposition 4.3 *If Γ is a labeled $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graph of size n , the average number of vertices of $\text{silh}(\Gamma)$ is greater than $n - 2n^{\frac{2}{3}} + o(n^{\frac{2}{3}})$.*

Moreover, if $\mu > 0$, there exists $0 < \gamma < 1$ such that the probability that $\text{silh}(\Gamma)$ has less than $n - (2 + \mu)n^{\frac{2}{3}}$ vertices is $\mathcal{O}(\gamma^{n^{\frac{1}{3}}})$.

Proof. Let Γ be a labeled $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graph with combinatorial type $(n, k_2, k_3, \ell_2, \ell_3)$. By Proposition 3.2, $\text{silh}(\Gamma)$ can be computed as follows: first a sequence of ℓ_3 λ_3 -moves, which delete ℓ_3 vertices (and all ℓ_3 b -loops) and add ℓ_3 a -loops.

The graph obtained after these moves has combinatorial type $(n - \ell_3, k_2 - \ell_3, k_3, \ell_2 + \ell_3, 0)$, so that $\ell_2 + \ell_3$ λ_2 -moves must be performed. Every $\lambda_{2,1}$ -move deletes 1 vertex and 1 a -loop, and adds one isolated b -edge. Every $\lambda_{2,2}$ -move deletes 2 vertices, 1 a -loop and

1 isolated b -edge. Therefore the $\ell_2 + \ell_3$ λ_2 -moves delete at most $2(\ell_2 + \ell_3)$ vertices. They also create at most $\ell_2 + \ell_3$ isolated b -edges.

The graph obtained after these moves now has combinatorial type $(n', k'_2, k'_3, 0, 0)$, where $n' \leq n - \ell_2 - 2\ell_3$, $k'_2 \leq k_2 - \ell_2 - 2\ell_3$ and $k'_3 \leq k_3 + \ell_2 + \ell_3$. At that point, k'_3 κ_3 -moves must be performed. Each deletes 2 vertices. Finally, if an exceptional move must be performed, it deletes at most 1 vertex.

Therefore the total number of vertices deleted in the computation of $\text{silh}(\Gamma)$ is at most

$$\ell_3 + 2(\ell_2 + \ell_3) + 2(k_3 + \ell_2 + \ell_3) = 2k_3 + 4\ell_2 + 5\ell_3.$$

The statement on average values is then a direct consequence of this computation and of [7, Proposition 5.3]. Moreover, [7, Theorem 5.5 and Section 7.2] establish that, for each $\mu > 0$, there exists $0 < \gamma < 1$ such that

$$\mathbb{P}\left(k_3 \geq \left(1 + \frac{\mu}{6}\right)n^{\frac{2}{3}}\right) = \mathcal{O}\left(\gamma^{n^{\frac{2}{3}}}\right), \quad \mathbb{P}\left(\ell_2 \geq 2n^{\frac{1}{2}}\right) = \mathcal{O}\left(\gamma^{n^{\frac{1}{2}}}\right), \quad \mathbb{P}\left(\ell_3 \geq 2n^{\frac{1}{3}}\right) = \mathcal{O}\left(\gamma^{n^{\frac{1}{3}}}\right).$$

For n large enough, $2n^{\frac{1}{3}} < 2n^{\frac{1}{2}} < \frac{\mu}{15}n^{\frac{2}{3}}$. So $\mathbb{P}\left(2k_3 + 4\ell_2 + 5\ell_3 \geq (2 + \mu)n^{\frac{2}{3}}\right) = \mathcal{O}\left(\gamma^{n^{\frac{1}{3}}}\right)$, which concludes the proof. \square

4.3 Silhouetting all finitely generated subgroups of $\text{PSL}_2(\mathbb{Z})$

We have, until now, discussed the silhouetting operation for labeled $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graphs. It can be naturally extended to labeled $\text{PSL}_2(\mathbb{Z})$ -reduced graphs, and the same probabilistic results holds in that extended framework.

Let (Γ, v) be a labeled $\text{PSL}_2(\mathbb{Z})$ -reduced graph. Recall that (Γ, v) is not $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced if and only v is not adjacent to both an a - and a b -edge, and is the only such vertex. Let Γ° be the graph obtained from Γ by adding an a -loop (resp. a b -loop) at v if v is not adjacent to an a -edge (resp. a b -edge). Then Γ° is $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced⁶ and $\Gamma^\circ = \Gamma$ if Γ is $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced. We define the *silhouette* $\text{silh}(\Gamma, v)$ and the *quasi-silhouette* $\mathbf{q}\text{-silh}(\Gamma, v)$ of the rooted graph (Γ, v) to be equal to $\text{silh}(\Gamma^\circ)$ and $\mathbf{q}\text{-silh}(\Gamma^\circ)$.

If $1 \leq s, \ell \leq n$, we let $\mathfrak{R}_{n,\ell}(s)$ be the set of size n labeled $\text{PSL}_2(\mathbb{Z})$ -reduced graphs (Γ, v) whose silhouette has s vertices, and such that Γ° has ℓ loops (a single counter for both the a - and the b -loops). Observe that the set $\mathfrak{D}(n, s)$ of quasi-silhouettes of the elements of $\mathfrak{G}_n(s)$, is also the set of quasi-silhouettes of the elements of $\bigcup_\ell \mathfrak{R}_{n,\ell}(s)$. A statement parallel to Theorem 4.1 holds for rooted graphs.

⁶The only case where two loops must added is when Γ is the graph with one vertex and no edge (the Stallings graph of the trivial subgroup), which we do not need to consider in the rest of this section.

Theorem 4.4 *Let $n \geq 3$ and $1 \leq s, \ell \leq n$. If (Γ, v) is an element of $\mathfrak{R}_{n,\ell}(s)$ taken uniformly at random, then $\text{silh}(\Gamma, v)$ is a uniformly random element of $\mathfrak{G}_s(s)$. That is: for any $\Delta, \Delta' \in \mathfrak{G}_s(s)$, we have*

$$\mathbb{P}(\text{silh}(\Gamma, v) = \Delta) = \mathbb{P}(\text{silh}(\Gamma, v) = \Delta'),$$

where \mathbb{P} denotes the uniform probability on $\mathfrak{R}_{n,\ell}(s)$.

Proof. If $\Delta \in \mathfrak{D}_n(s)$, let $\mathfrak{qR}_{n,\ell}(\Delta)$ be the set of all $(\Gamma, v) \in \mathfrak{qR}_{n,\ell}(s)$ such that $\text{q-silh}(\Gamma, v) = \Delta$. This set is the disjoint union of the sets $\mathfrak{qR}_{n,\ell}(\Delta, \alpha)$ ($\alpha \in \{a, b, 0\}$), where $\mathfrak{qR}_{n,\ell}(\Delta, 0)$ consists of the elements $(\Gamma, v) \in \mathfrak{qR}_{n,\ell}(\Delta)$ such that Γ is a labeled $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graph, and $\mathfrak{qR}_{n,\ell}(\Delta, a)$ (resp. $\mathfrak{qR}_{n,\ell}(\Delta, b)$) consists of the $(\Gamma, v) \in \mathfrak{qR}_{n,\ell}(\Delta)$ such that v is not adjacent to an a -loop (resp. a b -loop).

We have $|\mathfrak{qR}_{n,\ell}(\Delta, 0)| = n \sum_{\tau} |\mathfrak{qG}_{\tau}(\Delta)|$ and $|\mathfrak{qR}_{n,\ell}(\Delta, a)| = \sum_{\ell_i=0}^{\ell} \ell_i \sum_{\tau} |\mathfrak{qG}_{\tau}(\Delta)|$ for $i = 2, 3$. hence $|\mathfrak{qR}_{n,\ell}(\Delta)| = (n + \ell) \sum_{\tau} |\mathfrak{qG}_{\tau}(\Delta)|$, where the summations over τ are over all combinatorial types $(n, k_2, k_3, \ell_2, \ell_3)$ such that $\ell_2 + \ell_3 = \ell$.

By Lemma 4.2, $|\mathfrak{qG}_{\tau}(\Delta)|$ does not depend on the choice of Δ in $\mathfrak{D}_n(s)$, so the quantities above depend only on n and ℓ . As a result, if $\Delta, \Delta' \in \mathfrak{D}_n(s)$, then $\mathbb{P}(\text{q-silh}(\Gamma, v) = \Delta) = \mathbb{P}(\text{q-silh}(\Gamma, v) = \Delta')$ (over $\mathfrak{qR}_{n,\ell}(s)$).

Next, as in the proof of Theorem 4.1, we verify that this implies that, if $\Delta, \Delta' \in \mathfrak{G}_s(s)$, then $\mathbb{P}(\text{silh}(\Gamma, v) = \Delta) = \mathbb{P}(\text{silh}(\Gamma, v) = \Delta')$ (over $\mathfrak{R}_{n,\ell}(s)$), which concludes the proof. \square

We also establish a result similar to Proposition 4.3.

Proposition 4.5 *If (Γ, v) is a random labeled $\text{PSL}_2(\mathbb{Z})$ -reduced graph of size n , the average number of vertices of $\text{silh}(\Gamma, v)$ is greater than $n - 2n^{\frac{2}{3}} + o(n^{\frac{2}{3}})$.*

Moreover, if $\mu > 0$, there exists $0 < \gamma < 1$ such that the probability that $\text{silh}((\Gamma, v))$ has less than $n - (2 + \mu)n^{\frac{2}{3}}$ vertices is $\mathcal{O}(\gamma^{n^{\frac{1}{3}}})$.

Proof. Let \mathfrak{G}_n and \mathfrak{R}_n denote the set of size n labeled $\text{PSL}_2(\mathbb{Z})$ -cyclically reduced graphs and $\text{PSL}_2(\mathbb{Z})$ -reduced graphs, respectively.

Let $(\Gamma, v) \in \mathfrak{R}_n$ be a $\text{PSL}_2(\mathbb{Z})$ -reduced graph. We note that Γ° has the same number of vertices and isolated a - and b -edges as Γ . If Γ has ℓ_2 a -loops and ℓ_3 b -loops, then Γ° has ℓ'_2 and ℓ'_3 such loops, with $(\ell'_2, \ell'_3) = (\ell_2, \ell_3)$, $(\ell_2 + 1, \ell_3)$ or $(\ell_2, \ell_3 + 1)$.

Since we need to manipulate simultaneously random variables defined on different probability spaces (namely the uniform distributions on \mathfrak{G}_n and \mathfrak{R}_n), we introduce the following notation.

- $\mathbb{P}_{\mathfrak{G}_n}$ and $\mathbb{P}_{\mathfrak{R}_n}$ denote the uniform probabilities on \mathfrak{G}_n and \mathfrak{R}_n , respectively.
- L_2 (respectively, L_3 , K_3) is the random variable defined on \mathfrak{G}_n such that $L_2(\Gamma)$ (respectively, $L_3(\Gamma)$, $K_3(\Gamma)$) is the number of a -loops (respectively, b -loops, isolated b -edges) in Γ .

- L'_2 , L'_3 and K'_3 are the random variables defined on \mathfrak{R}_n by $L'_2(\Gamma, v) = L_2(\Gamma^\circ)$, $L'_3(\Gamma, v) = L_3(\Gamma^\circ)$ and $K'_3(\Gamma, v) = K_3(\Gamma^\circ)$.

We know from the proof of Proposition 4.3, that the number of vertices of Γ deleted when silhouetting $(\Gamma, v) \in \mathfrak{R}_n$, is at most $2K'_3(\Gamma, v) + 4L'_2(\Gamma, v) + 5L'_3(\Gamma, v)$. So we want to show that there exists $0 < \gamma < 1$ such that

$$\mathbb{P}_{\mathfrak{R}_n} \left(2K'_3(\Gamma, v) + 4L'_2(\Gamma, v) + 5L'_3(\Gamma, v) \geq (2 + \mu)n^{\frac{2}{3}} \right) = \mathcal{O} \left(\gamma^{n^{\frac{1}{3}}} \right).$$

Let $\Gamma' \in \mathfrak{G}_n$ be a size n labeled $\mathrm{PSL}_2(\mathbb{Z})$ -cyclically reduced graph. Then $\Gamma' = \Gamma^\circ$ for exactly $n + L_2(\Gamma') + L_3(\Gamma')$ values of $(\Gamma, v) \in \mathfrak{R}_n$. Moreover $n \leq n + L_2(\Gamma') + L_3(\Gamma') \leq 2n$ and $|\mathfrak{R}_n| \geq n|\mathfrak{G}_n|$, hence

$$\mathbb{P}_{\mathfrak{R}_n}(L'_2(\Gamma, v) = i) = \frac{|\{(\Gamma, v) \in \mathfrak{R}_n \mid L'_2(\Gamma, v) = i\}|}{|\mathfrak{R}_n|} \leq 2 \mathbb{P}_{\mathfrak{G}_n}(L_2(\Delta) = i),$$

where the first probability is in \mathfrak{R}_n and the second in \mathfrak{G}_n . Similarly we have

$$\mathbb{P}_{\mathfrak{R}_n}(L'_3(\Gamma, v) = i) \leq 2 \mathbb{P}_{\mathfrak{G}_n}(L_3(\Delta) = i) \text{ and } \mathbb{P}_{\mathfrak{R}_n}(K'_3(\Gamma, v) = i) \leq 2 \mathbb{P}_{\mathfrak{G}_n}(K_3(\Delta) = i).$$

Summing over successive values of i , this yields

$$\begin{aligned} \mathbb{P}_{\mathfrak{R}_n} \left(L'_2(\Gamma, v) \geq 2n^{\frac{1}{2}} \right) &\leq 2 \mathbb{P}_{\mathfrak{G}_n} \left(L_2(\Delta) \geq 2n^{\frac{1}{2}} \right), \\ \mathbb{P}_{\mathfrak{R}_n} \left(L'_3(\Gamma, v) \geq 2n^{\frac{1}{3}} \right) &\leq 2 \mathbb{P}_{\mathfrak{G}_n} \left(L_3(\Delta) \geq 2n^{\frac{1}{3}} \right) \\ \text{and } \mathbb{P}_{\mathfrak{R}_n} \left(K'_3(\Gamma, v) \geq \left(1 + \frac{\mu}{6}\right)n^{\frac{2}{3}} \right) &\leq 2 \mathbb{P}_{\mathfrak{G}_n} \left(K_3(\Delta) \geq \left(1 + \frac{\mu}{6}\right)n^{\frac{2}{3}} \right). \end{aligned}$$

We saw in the proof of Proposition 4.3 that the probabilities in the right-hand side terms are small, and they remain small even when multiplied by 2. In particular, there exists $0 < \gamma < 1$ such that

$$\begin{aligned} \mathbb{P}_{\mathfrak{R}_n} \left(L'_2(\Gamma, v) \geq 2n^{\frac{1}{2}} \right) &= \mathcal{O} \left(\gamma^{n^{\frac{1}{2}}} \right), \quad \mathbb{P}_{\mathfrak{R}_n} \left(L'_3(\Gamma, v) \geq 2n^{\frac{1}{3}} \right) = \mathcal{O} \left(\gamma^{n^{\frac{1}{3}}} \right) \\ \text{and } \mathbb{P}_{\mathfrak{R}_n} \left(K'_3(\Gamma, v) \geq \left(1 + \frac{\mu}{6}\right)n^{\frac{2}{3}} \right) &= \mathcal{O} \left(\gamma^{n^{\frac{2}{3}}} \right). \end{aligned}$$

As in the proof of Proposition 4.3, this leads to

$$\mathbb{P}_{\mathfrak{R}_n} \left(2K'_3(\Gamma, v) + 4L'_2(\Gamma, v) + 5L'_3(\Gamma, v) \geq (2 + \mu)n^{\frac{2}{3}} \right) = \mathcal{O} \left(\gamma^{n^{\frac{1}{3}}} \right),$$

which concludes the proof. \square

5 Generic properties of subgroups of $\mathrm{PSL}_2(\mathbb{Z})$

Recall that the *parabolic elements* of $\mathrm{PSL}_2(\mathbb{Z})$ are the conjugates of non-trivial powers of ab and that a subgroup H of $\mathrm{PSL}_2(\mathbb{Z})$ is *non-parabolic* if it contains no parabolic element.

Recall also, H is *almost malnormal* if, for every $x \notin H$, $H \cap H^x$ is finite. It is *malnormal* if each of these intersections is trivial: malnormality coincides with almost malnormality if H is torsion-free (e.g. a free subgroup of $\mathrm{PSL}_2(\mathbb{Z})$).

We show that, generically, a finitely generated subgroup of $\mathrm{PSL}_2(\mathbb{Z})$ contains parabolic elements (Proposition 5.1) and fails to be almost malnormal (Theorem 5.2). More precisely, we prove the following statements.

Proposition 5.1 *Let $0 < \alpha < \frac{1}{6}$. A random size n subgroup (respectively, cyclically reduced subgroup) of $\mathrm{PSL}_2(\mathbb{Z})$ is non-parabolic with probability $\mathcal{O}(n^{-\alpha})$.*

Theorem 5.2 *Let $0 < \alpha < \frac{1}{6}$. The probability that a size n subgroup (respectively, cyclically reduced subgroup) of $\mathrm{PSL}_2(\mathbb{Z})$ is almost malnormal is $\mathcal{O}(n^{-\alpha})$.*

The proofs of both statements rely on Theorem 5.3 below. This technical statement deals with the presence of cycles in a $\mathrm{PSL}_2(\mathbb{Z})$ -reduced graph labeled by a power of ab . In Theorem 5.3 and in the rest of the paper, a cycle labeled by $(ab)^m$ ($m \geq 1$) is called an *ab-cycle of size m* . It is clear that such a cycle has length $2m$.

Theorem 5.3 *Let $0 < \alpha < \frac{1}{6}$. Then a random size n labeled $\mathrm{PSL}_2(\mathbb{Z})$ -reduced (respectively, $\mathrm{PSL}_2(\mathbb{Z})$ -cyclically reduced) graph fails to have an ab-cycle of size at least 2 and at most n^α with probability $\mathcal{O}(n^{-\alpha})$.*

The proof of Proposition 5.1 is a direct application of Theorem 5.3, and is given in Section 5.1. The proof of Theorem 5.2 is less direct and relies on a graph-theoretic characterization of almost malnormality, see Section 5.2. Finally, the complex proof of Theorem 5.3 is given in Section 5.3 except for a technical but deep lemma, which is interesting in its own right and is established in Section 5.4.

5.1 Non-parabolic subgroups of $\mathrm{PSL}_2(\mathbb{Z})$

Recall from Section 2, the notion of a cyclically reduced element of $\mathrm{PSL}_2(\mathbb{Z})$.

Lemma 5.4 *Let H be a finitely generated subgroup of $\mathrm{PSL}_2(\mathbb{Z})$ and $(\Gamma(H), v_0)$ its Stallings graph, let g be a non-trivial element of H . If x is a word with minimal length such that $u = x^{-1}gx$ is cyclically reduced, then $\Gamma(H)$ has an x -labeled path from v_0 to some vertex v_1 and a u -labeled cycle at v_1 .*

Proof. The result is trivial if g is cyclically reduced since, in that case, $x = 1$. Let us now assume that g is not cyclically reduced, so that $x \neq 1$.

If $|u| = 1$, then xux^{-1} is the normal form of g . Indeed, let c be the last letter of x , say, $x = yc$. If $u = a$ and $c = b^{\pm 1}$, or $u = b^{\pm 1}$ and $c = a$, then xux^{-1} is clearly in normal form. If instead $u = c$ or $u = c^{-1}$, then $yuy^{-1} = xux^{-1} = g$, and $y^{-1}gy = u$ is cyclically reduced, contradicting the minimality of $|x|$. It follows that $\Gamma(H)$ has a cycle at v_0 labeled xux^{-1} and therefore, as announced, it has a u -cycle at v_1 , where v_1 is the extremity of the x -labeled path starting at v_0 .

If $|u| > 1$, without loss of generality, set $u = au'b^\varepsilon$ (for some $\varepsilon = \pm 1$), where u' is either empty or ends with an a . If $x = 1$, then $\Gamma(H)$ has a u -cycle at $v_1 = v_0$. If $x \neq 1$ and x ends with an a , say, $x = ya$, we have $g = xux^{-1} = yu'b^\varepsilon ay^{-1}$ and $u'b^\varepsilon a$ is cyclically reduced, contradicting the minimality of $|x|$. So x cannot end with an a . The same reasoning shows that it cannot end with b^ε . So $x = yb^{-\varepsilon}$. The word $xux^{-1} = yb^{-\varepsilon}au'b^{-\varepsilon}y^{-1}$ is therefore the normal form of g . Let v_1, v_2 and v_3 be the vertices reached from v_0 after reading $x = yb^{-\varepsilon}$, y and yb^ε , respectively. Then v_1, v_2 and v_3 constitute a b -triangle and there is a path labeled au' from v_1 to v_3 , so that considering the cycle at v_0 labeled by $yb^{-\varepsilon}au'b^{-\varepsilon}y^{-1}$, we see that v and v_1 sit on the same b -triangle, and that $u = au'b^\varepsilon$ labels a cycle at v_1 . \square

Proof of Proposition 5.1. Recall that the parabolic elements of $\mathrm{PSL}_2(\mathbb{Z})$ are the conjugates of the non-trivial powers of ab . Such an element is a product of the form $g = x^{-1}(ab)^m x$ ($m \in \mathbb{Z}$, $m \neq 0$), and we observe that $(ab)^m$ is cyclically reduced. Lemma 5.4 shows that, if H is a subgroup of $\mathrm{PSL}_2(\mathbb{Z})$, then $g \in H$ if and only if $(ab)^m$ labels a cycle in $\Gamma(H)$.

And from Theorem 5.3: for $0 < \alpha < 1/6$, $\Gamma(H)$ contains a cycle labeled $(ab)^m$ for some $2 \leq m \leq n^\alpha$ with probability $1 - \mathcal{O}(n^\alpha)$. As a result, H contains a parabolic element with probability $1 - \mathcal{O}(n^\alpha)$. \square

5.2 Almost malnormal subgroups of $\mathrm{PSL}_2(\mathbb{Z})$

We start with the following characterization of almost malnormality, very similar to the classical characterization of malnormality for subgroups of free groups due to Jitsukawa [16].

Proposition 5.5 *Let H be a finitely generated subgroup of $\mathrm{PSL}_2(\mathbb{Z})$ and let (Γ, v_0) be its Stallings graph. Then H is almost malnormal if and only if there does not exist distinct vertices p and q in $\Gamma(H)$ and an element $g \in \mathrm{PSL}_2(\mathbb{Z})$ (in normal form) of infinite order (that is: g is not a conjugate of a or b) which labels cycles in $\Gamma(H)$ at both p and q .*

Proof. Let $h \notin H$ such that $H \cap H^h$ is infinite. It is well known that $\mathrm{PSL}_2(\mathbb{Z})$ is locally quasi-convex and that such groups satisfy the Howson property [24]. As a result, $H \cap H^h$ is finitely generated. In particular it is a quasi-convex subgroup and hence itself a hyperbolic

group. As a result, $H \cap H^h$ contains an element g of infinite order. Let w be its normal form. By definition, w labels a cycle in $\Gamma(H)$ at v_0 . In addition note that the finite order elements of $\mathrm{PSL}_2(\mathbb{Z})$ are exactly the conjugates of a and b .

If $g = y^p$ for some reduced word y and $p > 1$, let x be a minimum length word such that $u = x^{-1}yx$ is cyclically reduced. Then g has normal form xu^px^{-1} , and u^p is cyclically reduced. In particular, x labels a path in $\Gamma(H)$ from v_0 to a vertex v_1 , and u^p labels a cycle C at v_1 . It follows that u^p labels a cycle at p distinct vertices along C . We now assume that g is not a proper power.

Let x be a word of minimal length such that $u = x^{-1}gx$ is cyclically reduced. Lemma 5.4 shows that x labels a path in $\Gamma(H)$ from v_0 to a vertex v_1 , and that u labels a cycle at v_1 .

Since $hgh^{-1} \in H$, Lemma 5.4 shows that $\Gamma(H)$ also has a y -path from v_0 to a vertex v_2 and a cycle at v_2 labeled by a cyclically reduced word w such that $hgh^{-1} = ywy^{-1}$. Moreover, since hgh^{-1} is conjugated to g , the words u and w are cyclic conjugates, say, $u = tt'$ and $w = t't$. Let then v_3 be the vertex reached from v_2 reading t' . There are tt' -cycles at v_1 and v_3 and we need to show that $v_1 \neq v_3$.

If $v_1 = v_3$, then $tx^{-1} \in H$. Moreover

$$hgh^{-1} = xux^{-1} = xt'tx^{-1} = xt^{-1}tt'tx^{-1} = (tx^{-1})^{-1}g(tx^{-1})$$

so $tx^{-1}h$ and g commute. By the classical characterization of commuting elements in free products [20, Theorem 4.5], it follows that either $tx^{-1}h$ and g sit in the same conjugate of $\langle a \rangle$ or $\langle b \rangle$, or $tx^{-1}h$ and g sit in the same cyclic subgroup.

The first case is impossible since we assumed that g has infinite order. Therefore there exist $z \in \mathrm{PSL}_2(\mathbb{Z})$ and $p, q \in \mathbb{Z}$ such that $g = z^p$ and $tx^{-1}h = z^q$. We assumed that g is not a proper power, so $p = \pm 1$. Thus $tx^{-1}h = g^{pq} \in H$ and hence $h \in H$, a contradiction. \square

The following is a very convenient corollary of Proposition 5.5.

Corollary 5.6 *Let H be a subgroup of $\mathrm{PSL}_2(\mathbb{Z})$. If there exist a word w in normal form which is not a conjugate of a or b , and an integer $m \geq 2$ such that w^m labels a cycle at some vertex p in the Stallings graph $\Gamma(H)$ but w does not label a cycle at p , then H fails to be almost malnormal.*

Proof. The word w^m labels cycles at every vertex reached from p reading w^i , for $0 \leq i < m$ and we conclude by Proposition 5.5. \square

Our result on the probability for a subgroup to not be almost malnormal follows directly.

Proof of Theorem 5.2. Theorem 5.3 shows that for $0 < \alpha < \frac{1}{6}$ the Stallings graph of a size n random subgroup H contains a cycle labeled $(ab)^m$ (with $2 \leq m \leq n^\alpha$) with probability $1 - \mathcal{O}(n^{-\alpha})$. The result then follows from Corollary 5.6. \square

5.3 ab -cycles in $\mathrm{PSL}_2(\mathbb{Z})$ -reduced graphs

Let $0 < \alpha < 1$. We want to show that a size n $\mathrm{PSL}_2(\mathbb{Z})$ -reduced (respectively, $\mathrm{PSL}_2(\mathbb{Z})$ -cyclically reduced) graph has an ab -cycle of size $m \in [2, n^\alpha]$ with high probability. More precisely, we want to show that the probability that this is not the case, is $\mathcal{O}(n^{-\alpha})$.

The structure of the proof is the following: we first show that Theorem 5.3 holds when restricted to labeled silhouette graphs (this is Proposition 5.7 below); we then lift this result to all labeled $\mathrm{PSL}_2(\mathbb{Z})$ -cyclically reduced graphs and $\mathrm{PSL}_2(\mathbb{Z})$ -reduced graphs using the randomness preservation results of Sections 4.1 and 4.3, and the results on the expected size of the silhouette proved in Sections 4.2 and 4.3.

Proposition 5.7 *Let $0 < \alpha < \frac{1}{6}$ and let $n > 0$ be a multiple of 6. Then a random size n labeled silhouette graph fails to have an ab -cycle of size at least 2 and at most n^α , with probability $\mathcal{O}(n^{-\alpha})$.*

Besides its importance in the proof of Theorem 5.3, Proposition 5.7 is of independent interest, as it deals with the probability of the presence of short cycles in certain permutation groups, see the discussion in the introduction. The proof of Proposition 5.7, given in Section 5.4, is tricky and uses different techniques than what has been used so far in this paper.

Proof of Theorem 5.3. Recall that \mathfrak{G}_n and \mathfrak{R}_n denote the sets of size n labeled $\mathrm{PSL}_2(\mathbb{Z})$ -cyclically reduced graphs and $\mathrm{PSL}_2(\mathbb{Z})$ -reduced graphs, respectively. We first deal with $\mathrm{PSL}_2(\mathbb{Z})$ -cyclically reduced graphs.

Fix $0 < \alpha < \frac{1}{6}$. Let $\mathfrak{p}_\alpha(n)$ be the probability that a graph in \mathfrak{G}_n has no ab -cycle of size in $[2, n^\alpha]$ and let $\mathfrak{p}'_\alpha(n)$ be the probability that an element of \mathfrak{G}_n has a silhouette of size at least $n - 3n^{\frac{2}{3}}$ and has no ab -cycle of length in $[2, n^\alpha]$.

Proposition 4.3 shows that there exists $0 < \gamma < 1$ such that the probability that an element of \mathfrak{G}_n has a silhouette with size less than $n - 3n^{\frac{2}{3}}$ is $\mathcal{O}(\gamma^{n^{\frac{2}{3}}})$. This yields

$$\mathfrak{p}_\alpha(n) \leq \mathfrak{p}'_\alpha(n) + \mathcal{O}(\gamma^{n^{\frac{2}{3}}}).$$

Let now $s \geq n - 3n^{\frac{2}{3}}$ and $\mathfrak{p}'_\alpha(n, s)$ be the probability that an element of $\mathfrak{G}_n(s)$ (the set of elements of \mathfrak{G}_n whose silhouette has size s) has no ab -cycle of length in $[2, n^\alpha]$. Then

$$\mathfrak{p}'_\alpha(n) = \sum_{s=n-3n^{2/3}}^n \mathbb{P}_n(\mathfrak{G}_n(s)) \mathfrak{p}'_\alpha(n, s),$$

where \mathbb{P}_n is the uniform probability on \mathfrak{G}_n .

Finally, for Γ taken uniformly at random in $\mathfrak{G}_n(s)$ let $\mathfrak{q}_\alpha(n, s)$ be the probability that both $\mathrm{silh}(\Gamma)$ and Γ have no ab -cycle of size in $[2, n^\alpha]$, and $\mathfrak{q}'_\alpha(n, s)$ the probability that $\mathrm{silh}(\Gamma)$ has such a cycle but Γ does not. Then

$$\mathfrak{p}'_\alpha(n, s) \leq \mathfrak{q}_\alpha(n, s) + \mathfrak{q}'_\alpha(n, s).$$

Observe that $q_\alpha(n, s)$ is at most equal to the probability that Γ or $\text{silh}(\Gamma)$ has no ab -cycle of size in $[2, n^\alpha]$. Theorem 4.1 shows that the latter is equal to the probability that a size s silhouette graph has no ab -cycle of size in $[2, s^\alpha]$, which is $\mathcal{O}(s^{-\alpha})$ according to Proposition 5.7. Thus, there exist positive constants C (independent of s and n) such that

$$q_\alpha(n, s) \leq Cs^{-\alpha} \leq C(n - 3n^{\frac{2}{3}})^{-\alpha}$$

Turning our attention to $q'_\alpha(n, s)$, we note that

$$q'_\alpha(n, s) = \sum \mathbb{P}_{n,s}(\text{silh}(\Gamma) = \Delta) q_\Delta,$$

where $\mathbb{P}_{n,s}$ is the uniform probability on $\mathfrak{G}_n(s)$, q_Δ is the probability that a graph in $\mathfrak{G}_n(s)$ having silhouette Δ has no ab -cycle of size in $[2, n^\alpha]$, and the sum is taken over all size s silhouette graphs Δ who do have an ab -cycle of such a size.

Let $\Delta = \text{silh}(\Gamma)$. By Proposition 3.2, Δ can be obtained from Γ after a maximal sequence of λ_3 -, $\lambda_{2,1}$ -, $\lambda_{2,2}$ - and κ_3 -moves (since $s \geq n - 3n^{\frac{2}{3}} > 2$). In particular, the a -edges of Δ that were not already edges in Γ arose during a κ_3 -move, which deleted two vertices, two a -edges and a b -edge, and added a new a -edge.

Since Γ has at most $3n^{\frac{2}{3}}$ vertices more than Δ , the path from Γ to Δ has at most $\frac{3}{2}n^{\frac{2}{3}}$ κ_3 -moves, and the majority of a -edges of Δ are also a -edges of Γ . More precisely, at most $\frac{3}{2}n^{\frac{2}{3}}$ of the $\frac{s}{2}$ a -edges of Δ are not a -edges of Γ . Thus the probability that an a -edge of Δ fails to be an edge of Γ is at most

$$\frac{3n^{\frac{2}{3}}}{s} \leq \frac{3n^{\frac{2}{3}}}{n - 3n^{\frac{2}{3}}} = 3n^{-\frac{1}{3}} \left(1 + \mathcal{O}\left(n^{-\frac{1}{3}}\right)\right).$$

For each of the silhouette graphs Δ under consideration (of size s , with an ab -cycle of size in $[2, n^\alpha]$), fix an ab -cycle in Δ of size, say, λ . The probability that at least one of the λ a -edges in this ab -cycle is not an edge in Γ (that is: that this cycle does not exist in Γ), is bounded above by

$$3\lambda n^{-\frac{1}{3}} \left(1 + \mathcal{O}\left(n^{-\frac{1}{3}}\right)\right) \leq 3n^{\alpha-\frac{1}{3}} \left(1 + \mathcal{O}\left(n^{-\frac{1}{3}}\right)\right).$$

As a result, $q_\Delta \leq 3n^{\alpha-\frac{1}{3}} \left(1 + \mathcal{O}\left(n^{-\frac{1}{3}}\right)\right)$, $q'_\alpha(n, s) \leq 3n^{\alpha-1/3} \left(1 + \mathcal{O}\left(n^{-1/3}\right)\right)$, and

$$p'_\alpha(n, s) \leq q_\alpha(n, s) + q'_\alpha(n, s) \leq C'n^{-\alpha} + 3n^{\alpha-1/3} \left(1 + \mathcal{O}\left(n^{-1/3}\right)\right).$$

Since $\alpha < \frac{1}{6}$, $\alpha - \frac{1}{3} < -\alpha$ and there exists a constant $C'' > 0$ such that $p'_\alpha(n, s) \leq C''n^{-\alpha}$, and hence $p'_\alpha(n) \leq C''n^{-\alpha}$ which, in turn, yields $p_\alpha(n) = \mathcal{O}(n^{-\alpha})$.

The same proof holds for $\text{PSL}_2(\mathbb{Z})$ -reduced graphs, reasoning within \mathfrak{R}_n instead of \mathfrak{G}_n , and using Proposition 4.5 and Theorem 4.4 instead of Proposition 4.3 and Theorem 4.1. \square

5.4 Proof of Proposition 5.7

In this section, we give a proof of Proposition 5.7 which states that, with high probability, in a finite group of permutations generated by a pair of permutations without fixed points (σ_2, σ_3) , of order 2 and 3 respectively, the composition $\sigma_2\sigma_3$ admits orbits of a certain, relatively small size. This type of result, on the composition of two randomly chosen mappings, is notoriously difficult to obtain and is therefore interesting *per se*.

Let $0 < \alpha < \frac{1}{6}$ and let $n > 0$ be a multiple of 6. We want to show that a random size n silhouette graph fails to have an ab -cycle of size at least 2 and at most n^α with probability $\mathcal{O}(n^{-\alpha})$.

Let \mathcal{G} be the set of labeled graphs that are disjoint unions of silhouette graphs. Recall that the set of n -vertex elements of \mathcal{G} (n a multiple of 6) is in bijection with the set of pairs (σ_2, σ_3) of fixpoint-free permutations of $[n]$, the first of order 2 (the a -edges) and the second of order 3 (the b -edges).

If we fix σ_3 , the set \mathcal{G}_{σ_3} of elements of \mathcal{G} characterized by pairs of the form (σ_2, σ_3) has cardinality the number of fixpoint-free, order 2 permutations on $[n]$, namely $(n-1)!!$, where the double factorial of an odd integer q is given by $q!! = q(q-2)(q-4)\dots 1$. Since this value does not depend on σ_3 , it is sufficient to establish the restriction of Proposition 5.7 to an arbitrary \mathcal{G}_{σ_3} , provided that the constants in the \mathcal{O} -notation do not depend on σ_3 .

For convenience, we fix $\sigma_3 = (1\ 2\ 3)(4\ 5\ 6)\dots(n-2\ n-1\ n)$, and we write \mathcal{G}_n instead of \mathcal{G}_{σ_3} . We now concentrate on a particular set of ab -cycles. Say that an ab -cycle in a graph $G \in \mathcal{G}$ is *simple* if it visits at most one vertex in each b -triangle in G . Equivalently, walking along the corresponding $(ab)^m$ -labeled path does not require traveling through an a -edge in both directions.

Now let $M = \lfloor n^\alpha \rfloor$ and $\mathbb{M} = \{2, \dots, M\}$. In the context of this proof, we say that a set is *small* if its cardinality is in \mathbb{M} . Then a *small* ab -cycle in an element of \mathcal{G}_n visits a *small* set of b -triangles. We need to prove that, only with probability $\mathcal{O}(n^{-\alpha})$, a random size n silhouette graph fails to have a small simple ab -cycle.

Let \mathcal{C} be the set of graphs in \mathcal{G}_n with at least a small simple ab -cycle. Let also \mathcal{S} be the set of all small sets of b -triangles in a graph of \mathcal{G}_n (recall that all the graphs of \mathcal{G}_n have the same b -triangles). If $I \in \mathcal{S}$, let \mathcal{C}_I be the set of graphs in \mathcal{G}_n containing a simple ab -cycle visiting exactly the b -triangles in I . Then $\mathcal{C} = \bigcup_{I \in \mathcal{S}} \mathcal{C}_I$.

Finally, let $\bar{\mathcal{C}}$ denote the complement of \mathcal{C} in \mathcal{G}_n (the elements of \mathcal{G}_n without a small simple ab -cycle). We want to show that $\frac{|\bar{\mathcal{C}}|}{|\mathcal{G}_n|} = \frac{|\bar{\mathcal{C}}|}{(n-1)!!}$ is of the form $\mathcal{O}(n^{-\alpha})$. By the inclusion-exclusion principle, we have

$$|\bar{\mathcal{C}}| = |\mathcal{G}_n| - \sum_{I \in \mathcal{S}} |\mathcal{C}_I| + \sum_{\substack{\{I_1, I_2\} \subseteq \mathcal{S} \\ I_1 \neq I_2}} |\mathcal{C}_{I_1} \cap \mathcal{C}_{I_2}| - \dots = \sum_{\mathcal{I} \subseteq \mathcal{S}} (-1)^{|\mathcal{I}|} \left| \bigcap_{I \in \mathcal{I}} \mathcal{C}_I \right| = \sum_k (-1)^k \sum_{\substack{\mathcal{I} \subseteq \mathcal{S} \\ |\mathcal{I}|=k}} \left| \bigcap_{I \in \mathcal{I}} \mathcal{C}_I \right|.$$

Truncating the inclusion-exclusion formula on even or odd cardinalities for \mathcal{I} , yields

upper and lower bounds for $|\bar{\mathcal{C}}|$. For any $\kappa \geq 0$ we have

$$\sum_{k=0}^{2\kappa+1} (-1)^k \sum_{\substack{\mathcal{I} \subseteq S \\ |\mathcal{I}|=k}} \left| \bigcap_{I \in \mathcal{I}} \mathcal{C}_I \right| \leq |\bar{\mathcal{C}}| \leq \sum_{k=0}^{2\kappa} (-1)^k \sum_{\substack{\mathcal{I} \subseteq S \\ |\mathcal{I}|=k}} \left| \bigcap_{I \in \mathcal{I}} \mathcal{C}_I \right| \quad (1)$$

It turns out to be more convenient to work with tuples of b -triangles rather than sets. If $\mathcal{I} = \{I_1, \dots, I_k\}$ is a small set of b -triangles with cardinality k , there are $k!$ tuples $\mathbf{J} = (J_1, \dots, J_k) \in \mathcal{S}^k$ such that $\{J_1, \dots, J_k\} = \mathcal{I}$, so

$$\left| \bigcap_{I \in \mathcal{I}} \mathcal{C}_I \right| = \frac{1}{k!} \sum_{\substack{(J_1, \dots, J_k) \in \mathcal{S}^k \\ \{J_1, \dots, J_k\} = \mathcal{I}}} |\mathcal{C}_{J_1} \cap \dots \cap \mathcal{C}_{J_k}|. \quad (2)$$

Say that two simple ab -cycles *overlap* if they visit a same b -triangle (of course, on different vertices). We now distinguish the tuples $\mathbf{J} = (J_1, \dots, J_k)$ in Equation (2) according to the cardinality of their components and to their overlaps. More precisely, if $\mathbf{d} = (d_1, \dots, d_k) \in \mathbb{M}^k$, we let

$$\begin{aligned} \mathcal{P}^{\circ\circ}(\mathbf{d}) &= \left\{ \mathbf{J} = (J_1, \dots, J_k) \in \mathcal{S}^k \mid \forall i, |J_i| = d_i \text{ and the } J_i \text{ are pairwise disjoint} \right\} \\ \mathcal{P}^{\circ}(\mathbf{d}) &= \left\{ \mathbf{J} = (J_1, \dots, J_k) \in \mathcal{S}^k \mid \forall i, |J_i| = d_i \text{ and the } J_i \text{ are pairwise distinct} \right. \\ &\quad \left. \text{but not pairwise disjoint} \right\}. \end{aligned}$$

Returning to the summands in the estimation of the inclusion-exclusion bounds (Equation (1)), we now have, for every k ,

$$\sum_{\substack{\mathcal{I} \subseteq S \\ |\mathcal{I}|=k}} \left| \bigcap_{I \in \mathcal{I}} \mathcal{C}_I \right| = \frac{1}{k!} \left(\underbrace{\sum_{\mathbf{d} \in \mathbb{M}^k} \sum_{\mathbf{J} \in \mathcal{P}^{\circ\circ}(\mathbf{d})} \left| \bigcap_{i=1}^k \mathcal{C}_{J_i} \right|}_{A_k} + \underbrace{\sum_{\mathbf{d} \in \mathbb{M}^k} \sum_{\mathbf{J} \in \mathcal{P}^{\circ}(\mathbf{d})} \left| \bigcap_{i=1}^k \mathcal{C}_{J_i} \right|}_{B_k} \right) \quad (3)$$

We now study successively the quantities A_k and B_k in Equation (3). If $q \geq 1$, we let H_q denote the partial sum $\sum_{i=1}^q \frac{1}{i}$ of the harmonic series. It is well known that $H_q = \log q + \gamma + o(1)$, where γ is Euler's constant.

Lemma 5.8 *Let n be a positive multiple of 6. Let $0 < \alpha < \frac{1}{2}$, $M = \lfloor n^\alpha \rfloor$, $0 < \beta < \frac{1}{2} - \alpha$ and $1 \leq k \leq n^\beta$. Finally, let $\delta = \alpha + \beta$ and let A_k be as in Equation (3). Then*

$$\frac{1}{(n-1)!!} A_k = (H_M - 1)^k \left(1 + \mathcal{O}\left(n^{2\delta-1}\right) \right),$$

uniformly in k (that is: the constants intervening in the \mathcal{O} notation do not depend on n or k).

Proof. Let $\mathbf{d} = (d_1, \dots, d_k) \in \mathbb{M}^k$, $d = d_1 + \dots + d_k$ and $\mathbf{J} = (J_1, \dots, J_k) \in \mathcal{P}^{\circ\circ}(\mathbf{d})$. To construct a graph in $\bigcap_i \mathcal{C}_{J_i}$, that is, a graph in \mathcal{G}_n with (non-overlapping) simple ab -cycles over the sets of b -triangles J_1, \dots, J_k respectively, we must

- select for each b -triangle in J_1, \dots, J_k a vertex belonging to the collection of simple ab -cycles, *i.e.* a vertex reached after reading an occurrence of ab in an ab -cycle; there are 3 possibilities for each b -triangle, and therefore a total of 3^d choices;
- cyclically order the triangles in each J_i ; there are $(d_1 - 1)! \dots (d_k - 1)!$ possibilities to do so; note that this second step fully determines the a -edges adjacent to the d vertices chosen in the first step;
- choose the missing a -edges arbitrarily: they connect the $n - 2d$ vertices not yet adjacent to an a -edge, and there are $(n - 2d - 1)!!$ ways to do so.

Thus, for every $\mathbf{J} = (J_1, \dots, J_k) \in \mathcal{P}^{\circ\circ}(\mathbf{d})$, there are $3^d(d_1 - 1)! \dots (d_k - 1)!(n - 2d - 1)!!$ graphs in $\bigcap_{i=1}^k \mathcal{C}_{J_i}$ (which is independent of the choice of \mathbf{J} in $\mathcal{P}^{\circ\circ}(\mathbf{d})$).

Moreover, since a graph in \mathcal{G}_n has $\frac{n}{3}$ b -triangles and the components of a tuple in $\mathcal{P}^{\circ\circ}(\mathbf{d})$ are pairwise disjoint, we have $|\mathcal{P}^{\circ\circ}(\mathbf{d})| = \binom{n/3}{d_1, \dots, d_k, n/3-d}$. Thus

$$\sum_{\mathbf{J} \in \mathcal{P}^{\circ\circ}(\mathbf{d})} \left| \bigcap_{i=1}^k \mathcal{C}_{J_i} \right| = \binom{n/3}{d_1, \dots, d_k, n/3-d} 3^d (d_1 - 1)! \dots (d_k - 1)! (n - 2d - 1)!!.$$

Note that $\frac{(n-2d-1)!!}{(n-1)!!} = \prod_{i=0}^{d-1} \frac{1}{n-1-2i}$ and $\frac{(n/3)! 3^d}{(n/3-d)!} = \prod_{i=0}^{d-1} (n-3i)$. Therefore

$$\frac{1}{(n-1)!!} \sum_{\mathbf{J} \in \mathcal{P}^{\circ\circ}(\mathbf{d})} \left| \bigcap_{i=1}^k \mathcal{C}_{J_i} \right| = \frac{Q_d}{d_1 \dots d_k}, \quad \text{with } Q_d = \prod_{i=0}^{d-1} \frac{n-3i}{n-1-2i}.$$

Now observe that, for n sufficiently large,

$$Q_d \leq \frac{n}{n-1} \quad \text{and} \tag{4}$$

$$\begin{aligned} Q_d &= \prod_{i=0}^{d-1} \left(1 - \frac{i-1}{n-1-2i} \right) \geq \left(1 - \frac{d}{n-2d} \right)^d = \exp \left(d \log \left(1 - \frac{d}{n-2d} \right) \right) \\ &\geq 1 + d \log \left(1 - \frac{d}{n-2d} \right). \end{aligned} \tag{5}$$

The lower bound in Equation (5) for Q_d is a decreasing function of d and the possible values of d satisfy $1 \leq d \leq kM \leq n^{\alpha+\beta} = n^\delta$, so

$$1 + n^\delta \log \left(1 - \frac{n^\delta}{n-2n^\delta} \right) \leq Q_d \leq \frac{n}{n-1}. \tag{6}$$

Let U and L be the upper and lower bounds of Q_d in Equation (6). Note that $\sum_{d \in \mathbb{M}^k} \frac{1}{d_1 \dots d_k} = (H_M - 1)^k$. Therefore

$$L(H_M - 1)^k = L \sum_{d \in \mathbb{M}^k} \frac{1}{d_1 \dots d_k} \leq \frac{1}{(n-1)!!} A_k \leq U \sum_{d \in \mathbb{M}^k} \frac{1}{d_1 \dots d_k} = U(H_M - 1)^k,$$

which concludes the proof of Lemma 5.8 since both U and L equal $1 + \mathcal{O}(n^{2\delta-1})$. \square

Lemma 5.9 *Let n be a positive multiple of 6. Let $0 < \alpha < \frac{1}{6}$, $M = \lfloor n^\alpha \rfloor$, $0 < \beta < \frac{1}{6} - \alpha$ and $1 \leq k \leq n^\beta$. Finally, let $\delta = \alpha + \beta$ and let B_k be as in Equation (3). Then*

$$\frac{1}{(n-1)!!} B_k \leq \mathcal{O}\left(n^{6\delta-1}\right) (H_M - 1)^k$$

uniformly in k .

Proof. Let $\mathbf{d} = (d_1, \dots, d_k) \in \mathbb{M}^k$ and $d = d_1 + \dots + d_k$. Since every vertex of a silhouette graph occurs in exactly one ab -cycle, a b -triangle can occur in at most 3 ab -cycles. Let $\mathbf{t} = (t_1, t_2, t_3)$ be a tuple of non-negative integers such that $d = t_1 + 2t_2 + 3t_3$. We denote by $\mathcal{P}_k^\circ(\mathbf{d}; \mathbf{t})$ the set of elements $\mathbf{J} = (J_1, \dots, J_k) \in \mathcal{P}_k^\circ(\mathbf{d})$ such that there are t_1 (respectively, t_2, t_3) b -triangles belonging to exactly 1 (respectively, 2, 3) of the components of \mathbf{J} . We talk of b -triangles of *type 1* (respectively, 2, 3).

Claim 5.10 *If $\mathbf{J} \in \mathcal{P}_k^\circ(\mathbf{d}; \mathbf{t})$ and $\bigcap_{i=1}^k \mathcal{C}_{J_i} \neq \emptyset$, then $t_2 + t_3$ is even.*

Proof of Claim 5.10. Let $G \in \bigcap_{i=1}^k \mathcal{C}_{J_i}$. Consider the a -edges occurring in the k small simple ab -cycles visiting, respectively, J_1, \dots, J_k . It is convenient at this point to think of the a -edges in our ab -cycles as matched pairs of half-edges, which we denote (T, e) , where e is an a -edge and T is a b -triangle adjacent to e . Since no a -edge in a simple ab -cycle may connect vertices from the same b -triangle, each such a -edge e corresponds to a pair of distinct, matched half-edges (T, e) and (T', e) .

Even though a -edges are undirected (or can be traversed in both directions), considering one in a simple ab -cycle uniquely defines a direction and we can talk of matched outgoing and incoming half-edges along a simple ab -cycle. Clearly, an a -edge is used in only one direction in a given ab -cycle, but it can be used in different directions by distinct ab -cycles. So we say that a half-edge (T, e) such that e occurs in the union of the k ab -cycles under consideration is *outgoing* (respectively, *incoming*, *2-way*) if it only occurs as outgoing (respectively, it only occurs as incoming, it occurs both as outgoing and incoming).

If T is a b -triangle in $\bigcup J_i$, occurring in just one of the k small simple ab -cycles, then T is a component of exactly 1 incoming and 1 outgoing half-edges; if T occurs in two of these ab -cycles, it is a component of 1 incoming, 1 outgoing and 1 2-way half-edges; finally, if T occurs in three ab -cycles, it is a component of 3 2-way half-edges.

The result follows since a 2-way half-edge must be matched with another, distinct, 2-way half-edge, and we have $t_2 + 3t_3$ such half-edges. \square

Let $\mathbf{J} = (J_1, \dots, J_k) \in \mathcal{P}_k^\oplus(\mathbf{d}; \mathbf{t})$. Then $|\bigcup_i J_i| = t_1 + t_2 + t_3$ whereas $d = \sum_i |J_i| = t_1 + 2t_2 + 3t_3$. The overlaps between the J_i determine which b -triangles in $\bigcup_i J_i$ are of type 1, 2 or 3. To construct an n -vertex graph in $\bigcap_i \mathcal{C}_{J_i}$, that is, a graph in \mathcal{G}_n with (overlapping) simple ab -cycles over the sets of b -triangles J_1, \dots, J_k respectively, we must

- select for each b -triangle of type 1 a vertex belonging to the collection of simple ab -cycles (recall that it is a vertex reached after reading an occurrence of ab in an ab -cycle); there are 3^{t_1} choices. Similarly, for each b -triangle of type 2, select two vertices belonging to the collection of simple ab -cycles; there are 3^{t_2} choices. Note that every vertex of a type 3 b -triangle belongs to the collection of simple ab -cycles. As discussed in the proof of Claim 5.10, the choice of these $t_1 + 2t_2 + 3t_3$ vertices in the ab -cycles implies the presence of $2t_1 + 3t_2 + 3t_3$ vertices along the corresponding $(ab)^{d_i}$ -labeled cycles.
- cyclically order the triangles in each J_i ; there are $(d_1 - 1)! \dots (d_k - 1)!$ possibilities to do so; note that this second step fully determines the a -edges adjacent to the $2t_1 + 3t_2 + 3t_3$ vertices determined in the first step;
- choose the missing a -edges arbitrarily: they connect the $n - (2t_1 + 3t_2 + 3t_3)$ vertices not yet selected, and there are $(n - (2t_1 + 3t_2 + 3t_3) - 1)!!$ ways to do so.

Thus, for every $\mathbf{J} = (J_1, \dots, J_k) \in \mathcal{P}_k^\oplus(\mathbf{d}; \mathbf{t})$, we have

$$\left| \bigcap_{i=1}^k \mathcal{C}_{J_i} \right| = 3^{t_1+t_2} (d_1 - 1)! \dots (d_k - 1)! (n - 2t_1 - 3t_2 - 3t_3 - 1)!!.$$

Now, to construct a tuple in $\mathcal{P}_k^\oplus(\mathbf{d}; \mathbf{t})$, we must

- choose t_i b -triangles of type i ($i = 1, 2, 3$); there are $\binom{n/3}{t_1, t_2, t_3, n/3 - t_1 - t_2 - t_3}$ choices;
- allocate the $t_1 + t_2 + t_3$ selected b -triangles to sets J_1, \dots, J_k respecting multiplicities and the required cardinality of the J_i ; there are at most $\binom{t_1 + 2t_2 + 3t_3}{d_1, \dots, d_k} = \binom{d}{d_1, \dots, d_k}$ choices (this is an upper bound as some choices may be unrealizable or produce cycles that are not simple).

Therefore we have

$$|\mathcal{P}_k^\oplus(\mathbf{d}; \mathbf{t})| \leq \binom{n/3}{t_1, t_2, t_3, n/3 - t_1 - t_2 - t_3} \binom{d}{d_1, \dots, d_k}$$

and $\sum_{\mathbf{J} \in \mathcal{P}_k^\oplus(\mathbf{d}; \mathbf{t})} \left| \bigcap_{i=1}^k \mathcal{C}_{J_i} \right|$ is bounded above by

$$\binom{n/3}{t_1, t_2, t_3, n/3 - t_1 - t_2 - t_3} \binom{d}{d_1, \dots, d_k} 3^{t_1+t_2} (d_1 - 1)! \dots (d_k - 1)! (n - 2t_1 - 3t_2 - 3t_3 - 1)!!.$$

Dividing by $(n-1)!!$, we get

$$\frac{\sum_{\mathbf{J} \in \mathcal{P}^\circ(\mathbf{d}; \mathbf{t})} \left| \bigcap_{i=1}^k \mathcal{C}_{J_i} \right|}{(n-1)!!} \leq \frac{d!}{d_1 \cdots d_k t_1! t_2! t_3! 3^{t_3}} \frac{n(n-3) \cdots (n-3(t_1+t_2+t_3-1))}{(n-1)(n-3) \cdots (n-2t_1-3t_2-3t_3+1)}.$$

Let $t = t_1 + t_2 + t_3$, so that $t_1 = d - 2t_2 - 3t_3$. Recall that $Q_t = \prod_{i=0}^{t-1} \frac{n-3i}{n-1-2i} \leq \frac{n}{n-1}$ (Equation (4)). Then, as $t_2! t_3! 3^{t_3} \geq 1$,

$$\begin{aligned} \frac{\sum_{\mathbf{J} \in \mathcal{P}^\circ(\mathbf{d}; \mathbf{t})} \left| \bigcap_{i=1}^k \mathcal{C}_{J_i} \right|}{(n-1)!!} &\leq \frac{1}{d_1 \cdots d_k} \cdot \frac{d!}{(d-2t_2-3t_3)!} \cdot \frac{Q_t}{\prod_{i=t}^{t_1+\frac{3}{2}(t_2+t_3)-1} (n-1-2i)} \\ &\leq \frac{n d^{2t_2+3t_3}}{(n-1) d_1 \cdots d_k (n-2t_1-3t_2-3t_3+1)^{\frac{1}{2}(t_2+t_3)}} \end{aligned} \quad (7)$$

Since $d = t_1 + 2t_2 + 3t_3$ and $d \leq kM \leq n^{\alpha+\beta} = n^\delta$, we have $2t_1 + 3t_2 + 3t_3 \leq 2d \leq 2n^\delta$, yielding

$$\frac{d^{2t_2+3t_3}}{(n-2t_1-3t_2-3t_3+1)^{\frac{1}{2}(t_2+t_3)}} \leq \frac{n^{\delta(2t_2+3t_3)}}{(n-2n^\delta)^{\frac{1}{2}(t_2+t_3)}}.$$

Moreover, $\frac{1}{2}(t_2+t_3) \leq \frac{1}{4}d \leq \frac{1}{4}n^\delta$ and we have

$$\begin{aligned} (n-2n^\delta)^{\frac{1}{2}(t_2+t_3)} &= n^{\frac{1}{2}(t_2+t_3)} \exp \left(\left(\frac{1}{2}(t_2+t_3) \right) \log \left(1-2n^{\delta-1} \right) \right) \\ &\geq n^{\frac{1}{2}(t_2+t_3)} \exp \left(\frac{1}{4}n^\delta \log \left(1-2n^{\delta-1} \right) \right) \end{aligned}$$

The term under the exponential is $\frac{1}{4}n^\delta \log(1-2n^{\delta-1}) \sim -\frac{1}{2}n^{3\delta-1}$, which tends to 0 (since $\delta < \frac{1}{6} < \frac{1}{3}$). Therefore, for n sufficiently large,

$$(n-2n^\delta)^{\frac{1}{2}(t_2+t_3)} \geq \frac{1}{2}n^{\frac{1}{2}(t_2+t_3)}.$$

Going back to Equation (7), we have

$$\frac{\sum_{\mathbf{J} \in \mathcal{P}^\circ(\mathbf{d}; \mathbf{t})} \left| \bigcap_{i=1}^k \mathcal{C}_{J_i} \right|}{(n-1)!!} \leq \frac{1}{d_1 \cdots d_k} \frac{n^{\delta(2t_2+3t_3)}}{n^{\frac{1}{2}(t_2+t_3)}} = \frac{1}{d_1 \cdots d_k} n^{\delta(2t_2+3t_3)-\frac{1}{2}(t_2+t_3)} \quad (8)$$

Equation (8) must be summed over \mathbf{t} and \mathbf{d} . Let us first fix \mathbf{d} and recall that $\mathbf{t} = (t_1, t_2, t_3)$ with $(t_2, t_3) \neq (0, 0)$ and $t_2 + t_3$ even (Claim 5.10).

Consider first the \mathbf{t} of the form $(t_1, t_2, 0) = (t_1, 2s, 0)$ ($s \geq 1$). The corresponding subsum of powers of n is bounded above (for $n \geq 2$) by

$$\sum_{s \geq 1} \left(n^{2\delta-\frac{1}{2}} \right)^{2s} = \frac{n^{4\delta-1}}{1-n^{4\delta-1}} \leq 2n^{4\delta-1}.$$

Similarly, the subsum of powers of n corresponding to the \mathbf{t} of the form $(t_1, 0, t_3)$ is bounded above by $2n^{6\delta-1}$.

Next, the subsum corresponding to the \mathbf{t} where both t_2 and t_3 are non-zero is bounded above by

$$\sum_{t_2 \geq 1} \sum_{t_3 \geq 1} \left(n^{2\delta-\frac{1}{2}}\right)^{t_2} \left(n^{3\delta-\frac{1}{2}}\right)^{t_3} \leq \left(\sum_{t_2 \geq 1} \left(n^{2\delta-\frac{1}{2}}\right)^{t_2}\right) \left(\sum_{t_3 \geq 1} \left(n^{3\delta-\frac{1}{2}}\right)^{t_3}\right) \leq 2n^{5\delta-1}$$

It follows that, for a fixed tuple \mathbf{d} ,

$$\frac{\sum_{\mathbf{J} \in \mathcal{P}^\oplus(\mathbf{d}; \mathbf{t})} \left| \bigcap_{i=1}^k \mathcal{C}_{J_i} \right|}{(n-1)!!} \leq \frac{2(n^{4\delta-1} + n^{5\delta-1} + n^{6\delta-1})}{d_1 \cdots d_k}.$$

Now, summing over \mathbf{d} , we get

$$\begin{aligned} \frac{\sum_{\mathbf{J} \in \mathcal{P}^\oplus(\mathbf{d}; \mathbf{t})} \left| \bigcap_{i=1}^k \mathcal{C}_{J_i} \right|}{(n-1)!!} &\leq 2(n^{4\delta-1} + n^{5\delta-1} + n^{6\delta-1}) \sum_{\mathbf{d}} \frac{1}{d_1 \cdots d_k} \\ &\leq 2(n^{4\delta-1} + n^{5\delta-1} + n^{6\delta-1})(H_M - 1)^k \end{aligned}$$

Since $n^{4\delta-1} + n^{5\delta-1} + n^{6\delta-1} = \mathcal{O}(n^{6\delta-1})$, we have, for n large enough,

$$\frac{\sum_{\mathbf{J} \in \mathcal{P}^\oplus(\mathbf{d}; \mathbf{t})} \left| \bigcap_{i=1}^k \mathcal{C}_{J_i} \right|}{(n-1)!!} = (H_M - 1)^k \mathcal{O}(n^{6\delta-1}),$$

thus concluding the proof of Lemma 5.9. \square

We can now conclude the proof of Proposition 5.7. By Equations (1) and (3), we want to show that, for $\kappa = \lfloor n^\beta \rfloor$ and $\kappa = \lfloor n^\beta \rfloor - 1$,

$$\frac{1}{(n-1)!!} \sum_{k=0}^{\kappa} (-1)^k \sum_{\substack{\mathcal{I} \subseteq S \\ |\mathcal{I}|=k}} \left| \bigcap_{I \in \mathcal{I}} \mathcal{C}_I \right| = \sum_{k=0}^{\kappa} \frac{(-1)^k}{k!} \frac{A_k}{(n-1)!!} + \sum_{k=0}^{\kappa} \frac{(-1)^k}{k!} \frac{B_k}{(n-1)!!}$$

is $\mathcal{O}(n^{-\alpha})$. By Lemma 5.8, the absolute value of the first sum is bounded above by

$$(1 + \mathcal{O}(n^{2\delta-1})) \exp(-(H_M - 1)) = e^{1-\gamma} n^{-\alpha} (1 + \mathcal{O}(n^{2\delta-1})).$$

And by Lemma 5.9, the absolute value of the second sum is bounded above by

$$\exp(-(H_M - 1)) \mathcal{O}(n^{6\delta-1}) = e^{1-\gamma} n^{-\alpha} \mathcal{O}(n^{6\delta-1}).$$

Thus the whole sum is $\mathcal{O}(n^{-\alpha})$, establishing the expected bound for disjoint unions of silhouette graphs.

Such a union, of size n (a multiple of 6), is connected (and hence silhouette) with probability $1 - \frac{5}{6}n^{-1} + o(n^{-1})$ by [7, Proof of Proposition 8.18]. Thus the probability that a silhouette graph has no small simple ab -cycle is, again, $\mathcal{O}(n^{-\alpha})$.

References

- [1] G. N. Arzhantseva. Generic properties of finitely presented groups and Howson’s theorem. *Comm. Algebra*, 26(11):3783–3792, 1998.
- [2] G. N. Arzhantseva and A. Y. Ol’shanskii. Generality of the class of groups in which subgroups with a lesser number of generators are free. *Mat. Zametki*, 59(4):489–496, 1996.
- [3] F. Bassino, A. Martino, C. Nicaud, E. Ventura, and P. Weil. Statistical properties of subgroups of free groups. *Random Struct. Algorithms*, 42(3):349–373, 2013.
- [4] 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.
- [5] 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.
- [6] F. Bassino, C. Nicaud, and P. Weil. On the genericity of Whitehead minimality. *J. Group Theory*, 19(1):137–159, 2016.
- [7] F. Bassino, C. Nicaud, and P. Weil. Statistics of subgroups of the modular group. *Int. J. Algebra Comput.*, 31(8):1691–1751, 2021.
- [8] F. Bassino, C. Nicaud, and P. Weil. Random generation of subgroups of the modular group with a fixed isomorphism type. [arXiv:2310.18923](https://arxiv.org/abs/2310.18923), 2023.
- [9] T. Budzinski, N. Curien, and B. Petri. The diameter of random Belyi surfaces. *Algebraic & Geometric Topology*, 21:2929–2957, 2021.
- [10] P. Flajolet and R. Sedgewick. *Analytic combinatorics*. Cambridge University Press, 2009.
- [11] A. Gamburd. Poisson-Dirichlet distribution for random Belyi surfaces. *Ann. Probab.*, 34(5):1827–1848, 2006.
- [12] S. M. Gersten and H. B. Short. Rational subgroups of biautomatic groups. *Ann. of Math.*, 134(1):125–158, 1991.
- [13] R. Gilman, A. Miasnikov, and D. Osin. Exponentially generic subsets of groups. *Illinois J. Math.*, 54(1):371–388, 2010.
- [14] R. Gitik. Nielsen generating sets and quasiconvexity of subgroups. *J. Pure Appl. Algebra*, 112(3):287–292, 1996.
- [15] D. M. Jackson and T. I. Visentin. A character-theoretic approach to embeddings of rooted maps in an orientable surface of given genus. *Trans. Am. Math. Soc.*, 322(1):343–363, 1990.
- [16] 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.
- [17] I. Kapovich. Detecting quasiconvexity: algorithmic aspects. In *Geometric and computational perspectives on infinite groups (Minneapolis, MN and New Brunswick, NJ, 1994)*, volume 25 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 91–99. Amer. Math. Soc., Providence, RI, 1996.

- [18] O. Kharlampovich, A. Miasnikov, and P. Weil. Stallings graphs for quasi-convex subgroups. *J. Algebra*, 488:442–483, 2017.
- [19] R. C. Lyndon and P. E. Schupp. *Combinatorial group theory*. Springer-Verlag, 1977.
- [20] W. Magnus, A. Karrass, and D. Solitar. *Combinatorial group theory*. Dover, 1976.
- [21] J. Maher and A. Sisto. Random subgroups of acylindrically hyperbolic groups and hyperbolic embeddings. *Int. Math. Res. Not. IMRN*, 13:3941–3980, 2019.
- [22] L. Markus-Epstein. Stallings foldings and subgroups of amalgams of finite groups. *Internat. J. Algebra Comput.*, 17(8):1493–1535, 2007.
- [23] C. Nicaud. Random deterministic automata. In E. Csuhaj-Varjú, M. Dietzfelbinger, and Z. Ésik, editors, *Mathematical Foundations of Computer Science 2014 — 39th International Symposium, MFCS 2014, Budapest, Hungary, August 25-29, 2014. Proceedings, Part I*, volume 8634 of *Lecture Notes in Computer Science*, pages 5–23. Springer, 2014.
- [24] H. Short. Quasiconvexity and a theorem of Howson’s. In *Group theory from a geometrical viewpoint (Trieste, 1990)*, pages 168–176. World Sci. Publ., River Edge, NJ, 1991.
- [25] J. R. Stallings. Topology of finite graphs. *Invent. Math.*, 71(3):551–565, 1983.
- [26] N. W. M. Touikan. A fast algorithm for Stallings’ folding process. *Internat. J. Algebra Comput.*, 16(6):1031–1045, 2006.