

AVERAGE-CASE COMPLEXITY AND DECISION PROBLEMS IN GROUP THEORY

ILYA KAPOVICH, ALEXEI MYASNIKOV, PAUL SCHUPP, AND VLADIMIR SHPILRAIN

ABSTRACT. We investigate the average-case complexity of decision problems for finitely generated groups, in particular the word and membership problems. Using our recent results on “generic-case complexity” we show that if a finitely generated group G has word problem solvable in subexponential time and has a subgroup of finite index which possesses a non-elementary word-hyperbolic quotient group, then the average-case complexity of the word problem of G is linear time, uniformly with respect to the collection of all length-invariant measures on G . This results applies to many of the groups usually studied in geometric group theory: for example, all braid groups B_n , all groups of hyperbolic knots, many Coxeter groups and all Artin groups of extra-large type.

1. INTRODUCTION

Formulated by Max Dehn in 1912, the word problem for finitely generated groups is one of the most classical of all decision problems. If G is a group with a finite generating set $A = \{x_1, \dots, x_k\}$, then the *word problem* $WP(G, A)$ of G relative to A consists of all those words w in the alphabet $X = A \cup A^{-1}$ which represent the identity element $1 \in G$. Thus $WP(G, A) \subseteq X^*$, where X^* is the set of all words in the alphabet X . For finitely generated groups the worst-case complexity of the word problem does not depend on the choice of a finite generating set and for that reason the reference to a generating set is often suppressed.

In general, group word problems may be very complicated with respect to worst-case complexity. The basic result of Novikov and of Boone (see for example [1, 42, 43]) states that there exists a finitely presented group G such that for any finite generating set A of G the word problem $WP(G, A)$ is undecidable. However, “most” finitely presented groups are word-hyperbolic and hence have the word problem solvable in linear time. This fact was first observed by Gromov [32] and made precise by Ol’shanskii [45, 8] and Champetier [17]. They defined the notion of a “generic” group-theoretic property and this notion is currently being successfully explored by many authors [5, 6, 7, 8, 18, 19, 20, 21, 28, 33, 34, 38, 39, 44, 55].

In our recent paper [37] we introduced the notion of *generic-case complexity* for decision problems (see Section 2 below for formal definitions) and showed that for most of the groups usually studied in combinatorial group theory, the generic-case complexity of the word problem is linear time in a very strong sense. This means that there is a partial algorithm which gives the correct answer for “most” inputs of the problem in linear time. We shall recall the precise definition of generic-case complexity later in Section 2. Such a result can hold for groups where the the word problem has very high worst-case complexity or is even undecidable, because generic-case complexity completely disregards the behavior of the algorithm on a “small” set of “difficult” inputs. Indeed, the result holds for the groups of Boone and of Novikov with undecidable word problem.

In the present article we show that our results regarding generic-case complexity can in fact be used to obtain precise average-case results on the expected value of complexity over the entire set of inputs, including the “difficult” ones. The basic idea here is very straightforward and is often used in practice. If we have a total algorithm Ω_1 solving a decision problem \mathcal{D} whose worst-case complexity is not “too high” and we also have a partial algorithm Ω_2 solving the problem with “strongly low” generic-case complexity, then by running Ω_1 and Ω_2 in parallel we have a total algorithm $\Omega_1 \parallel \Omega_2$ for which we can hope to prove low average-case complexity.

There are few average-case complexity results about decision problems for finitely presented groups. We mention here two papers by Wang [50, 52], whose results are very different from ours both in substance and in the technique used (for example, in Wang’s approach the set of instances of a problem involves all finite group presentations, rather than a fixed one).

Definition 1.1 (Subexponential functions). We say that a non-negative function $f(n)$ is *subexponential* if for any $r > 1$ we have

$$\lim_{n \rightarrow \infty} \frac{f(n)}{r^n} = 0.$$

Note that this implies that for every $r > 1$

$$\sum_{n=1}^{\infty} \frac{f(n)}{r^n} < \infty.$$

Definition 1.2. Let X be a finite alphabet with $k \geq 2$ elements. A *discrete probability measure* on X^* is a function $\mu : X^* \rightarrow [0, 1]$ such that $\sum_{w \in X^*} \mu(w) = 1$.

We say that μ is *length-invariant* if for any words $w, w' \in X^*$ with $|w| = |w'|$ we have $\mu(w) = \mu(w')$.

We say that a measure $\mu : X^* \rightarrow [0, 1]$ is *tame* if there exists a subexponential function $g(n) \geq 0$ such that for any $w \in X^*$ we have $\mu(w) \leq \frac{g(|w|)}{k^{|w|}}$.

It is easy to see that any length-invariant discrete probability measure is tame. Indeed, if $\mu : X^* \rightarrow [0, 1]$ is such a measure then, since there are

exactly k^n words of length n , there is a function $d(n) \geq 0$ such that for any $w \in X^*$ we have $\mu(w) = \frac{d(|w|)}{k^{|w|}}$ and such that $\sum_{n=0}^{\infty} d(n) = 1$. Hence $d(n) \leq 1$ and so $\mu(w) \leq \frac{1}{k^{|w|}}$ for any word w .

Requiring that a measure be length-invariant is a very natural assumption, since most complexity classes are defined in terms of the length of an input word.

Convention 1.3. We follow Papadimitriou [46] for our conventions on computational complexity. In particular, by an *algorithm* we shall always mean a deterministic multi-tape Turing machine. In this paper we restrict our consideration to deterministic time-complexity classes, although the methods of this paper can be applied to the analysis of average-case behavior in more general complexity classes. We also assume that for a time-complexity class \mathcal{C} the collection of functions bounding the time of a computation consists of proper complexity functions $f(n) \geq n$ and that for any function $f(n)$ in this collection and for any integer $C \geq 1$ the function $Cf(Cn + C) + C$ also belongs to this collection.

Definition 1.4 (Average-case complexity). Let X be a finite alphabet with $|X| \geq 2$ elements. Let $\mathcal{D} \subseteq X^*$ be a language and let Ω be an algorithm which for every $w \in X^*$ decides whether or not $w \in \mathcal{D}$ in time $T(w) < \infty$. Let $f(n)$ be a non-decreasing positive function. Let $\mu : X^* \rightarrow [0, 1]$ be a discrete probability measure.

(1) We say that Ω *solves* \mathcal{D} *with average case time-complexity bounded by* $f(n)$ *relative to* μ if

$$\int_{X^*} \frac{T(w)}{f(|w|)} \mu(w) = \sum_{w \in X^*} \frac{T(w)}{f(|w|)} \mu(w) < \infty.$$

If $f(n)$ satisfies the bound constraint of a time-complexity class \mathcal{C} we say that Ω *solves* \mathcal{D} *with average case time-complexity in* \mathcal{C} *relative to* μ .

(2) Let \mathfrak{R} be a family of discrete probability measures on X^* . We say that Ω *solves* \mathcal{D} *with average case time-complexity bounded by* $f(n)$ *uniformly relative to* \mathfrak{R} if there is $0 < C < \infty$ such that for any $\mu \in \mathfrak{R}$

$$\int_{X^*} \frac{T(w)}{f(|w|)} \mu(w) = \sum_{w \in X^*} \frac{T(w)}{f(|w|)} \mu(w) \leq C.$$

If $f(n)$ satisfies the constraint of a time-complexity class \mathcal{C} we say that Ω *solves* \mathcal{D} *with average case time-complexity in* \mathcal{C} *uniformly relative to* \mathfrak{R} .

Remark 1.5. Suppose Ω solves \mathcal{D} with average case time-complexity bounded by $f(n)$ relative to μ . Then

$$\sum_{w \in X^*} \frac{T(w)}{f(|w|)} \mu(w) = \sum_{n=0}^{\infty} \sum_{w \in X^*, |w|=n} \frac{T(w)\mu(w)}{f(n)} < \infty.$$

Therefore

$$\sup_{n \geq 0} \frac{\sum_{w \in X^*, |w|=n} T(w)\mu(w)}{f(n)} < \infty,$$

and

$$\sum_{w \in X^*, |w|=n} T(w)\mu(w) = o(f(n)).$$

The above inequality provides additional justification for Definition 1.4.

Convention 1.6. We shall denote by *SubExp* the class of languages decidable in deterministic subexponential time.

We are now ready to state the main results of this paper. Two classes of groups currently playing a major role in group theory are the class of word-hyperbolic groups defined by Gromov [32] and the class of automatic groups [24]. For this article we need only the facts that the word problem of a word-hyperbolic group is solvable in linear time and the word problem of an automatic group is solvable in quadratic time. Also, while the concept of a group being non-amenable plays an important role in our results, the reader need only have in mind that any group containing a free subgroup of rank two is non-amenable.

Theorem A. *Let G be a finitely presented group where the word problem is in *SubExp*. Suppose G has a subgroup of finite index which possesses a non-amenable quotient group \overline{G}_1 whose word problem is solvable in a complexity class \mathcal{C} , where $\mathcal{C} \subseteq \text{SubExp}$.*

Then the word problem $WP(G, A)$ of G is solvable with average-case complexity in \mathcal{C} , uniformly relative to the family of all length-invariant discrete probability measures $\mu : (A \cup A^{-1})^ \rightarrow [0, 1]$.*

Moreover, for any tame discrete probability measure μ on $(A \cup A^{-1})^$ the word problem $WP(G, A)$ of G is solvable with average-case complexity in \mathcal{C} relative to μ .*

The idea behind the proof of Theorem A is quite simple (see Proposition 3.2 below for a more technical statement) but applies to a surprising range of examples. We run in parallel the total subexponential algorithm for the word problem of G and the partial generic algorithm provided by the quotient group \overline{G}_1 . If the image of an element of g is nontrivial in G_1 then the element g is certainly nontrivial in G . On all but an exponentially negligible set of inputs this partial algorithm will actually terminate within the complexity bound \mathcal{C} and it turns out that the average-case complexity of the combined algorithm is in \mathcal{C} .

We can now formulate some more concrete corollaries of the above theorem.

Corollary B. *Let G be a finitely presented group where the word problem is solvable in subexponential time. Let A be a generating set of G and let*

$X = A \cup A^{-1}$. Let \mathfrak{R} be the family of all length-invariant discrete probability measures $\mu : X^* \rightarrow [0, 1]$.

- (1) Suppose that G has a subgroup of finite index that possesses a non-elementary word-hyperbolic quotient group. Then the word problem $WP(G, A) \subseteq X^*$ is solvable with linear average-case time-complexity uniformly relative to \mathfrak{R} .

Moreover, for any tame discrete probability measure μ on X^* the word problem $WP(G, A)$ for G is solvable with linear average-case time complexity relative to μ .

- (2) Suppose that G has a subgroup of finite index that possesses a non-amenable automatic quotient group. Then the word problem

$$WP(G, A) \subseteq X^*$$

is solvable with quadratic average-case time-complexity uniformly relative to \mathfrak{R} .

Moreover, for any tame discrete probability measure μ on X^* the word problem $WP(G, A)$ of G is solvable with quadratic average-case time complexity relative to μ .

We have observed before that assuming the measure to be length-invariant is very natural from the point of view of complexity theory. Therefore, when informally talking about average-case complexity of the word problem (or the membership problem) being linear (quadratic, polynomial, etc) for a group G , we will mean that for any finite generating set $X = A \cup A^{-1}$ of G the average-case complexity is of that type uniformly with respect to the family of all length-invariant measures on X^* .

From the point of view of group theory, one weakness of average-case complexity is that it says nothing about the complexity of the word problem for finitely generated subgroups of the group being considered. If G is a finitely presented group whose word problem is solvable in subexponential time, then an easy argument shows that the word problem for the direct product $G_1 = G \times F(a, b)$ is also solvable in subexponential time. The free group $F(a, b)$ is non-elementary word-hyperbolic and G_1 admits a homomorphism onto $F(a, b)$. Thus Corollary B implies that any finitely presented group whose word problem is solvable in subexponential time can be embedded into a finitely presented group whose word problem is solvable with linear time average-case complexity.

Example 1.7. Let B_n be the n -strand braid group where $n \geq 3$. We observed in [37] that the generic-case complexity of the word problem of the group B_n is strongly linear time. The reason for this is that the pure-braid group P_n , which has finite index in B_n , admits a homomorphism onto the group $P_3 \cong F(a, b) \times \mathbb{Z}$ (by pulling out all but the first three strands of a pure braid). The group P_3 in turn maps onto the free group $F(a, b)$ which is a non-elementary word-hyperbolic group. Since B_n is automatic [24], the

word problem for B_n is solvable in quadratic time. Hence by Corollary B the word problem in B_n is solvable with linear time average-case complexity.

Example 1.8. Let G be a finitely generated linear group over a field of characteristic zero. Then the result of Lipton and Zalcstein log-space and thus in polynomial time. If G has a finite index subgroup with a non-elementary word-hyperbolic quotient, then by Corollary B the word problem in G is solvable with linear time average-case complexity.

Example 1.9. Let $K \subseteq \mathbb{S}^3$ be a *hyperbolic knot*, that is, such that the complement $M = \mathbb{S}^3 - K$ admits a complete hyperbolic metric of finite volume. (A basic example of such a K is the figure-eight knot, and, as we note below, most prime knots are hyperbolic.) The group of K is the fundamental group, $G = \pi_1(M)$, of the knot complement.

A theorem of Thurston (see, for example, [13]) shows that all but finitely many Dehn fillings on the compactification M' of M (that is gluing a solid torus along the torus boundary of M') produce a closed hyperbolic 3-manifold M_1 . Then $G_1 = \pi_1(M_1)$ is a quotient of G (obtained, in fact, by adding a single extra relator) and G_1 is non-elementary word-hyperbolic. The group G has polynomial time word problem, as can be seen for two different reasons. First, G is linear over \mathbb{R} since it embeds into $O(3, 1)$ and hence has word problem solvable in log-space by the previous example. Second, G is automatic (see [24]) and hence the word problem in G is solvable in quadratic time. Therefore by Corollary B the word problem in G has linear time average-case complexity.

All prime alternating knots except the $(2, q)$ torus knots are hyperbolic ([13]). The group of the $(2, q)$ torus knot has a presentation $G\langle x, y \mid x^2 = y^q \rangle$ where $q \geq 2$ is a positive integer. That these groups have word problem solvable in linear time is an easy consequence of their structure as a free product with amalgamation. Thus the word problem of the group of any prime alternating knot has linear average-case complexity.

Example 1.10. Let G be a *Coxeter group*, that is, a finitely generated group admitting a presentation

$$G = \langle a_1, \dots, a_t \mid a_i^2 = 1, i = 1, \dots, t; (a_i a_j)^{m_{ij}} = 1, 1 \leq i < j \leq t \rangle$$

where $2 \leq m_{ij} \leq \infty$. (That $m_{ij} = \infty$ means that there is no defining relation involving a_i and a_j .)

Since a Coxeter group is realizable as a linear group over \mathbb{R} , it follows from the result of Lipton and Zalcstein that any Coxeter group G has word problem solvable in polynomial time. Moreover, Brink and Howlett [15] have shown that all Coxeter groups are automatic and thus have quadratic time word problem. If the Coxeter group G is of *extra-large type*, that is, $t \geq 3$ and all $m_{ij} \geq 4$, then G itself is a nonelementary word-hyperbolic group.

Let G be a Coxeter group with a presentation as above, let $I = \{1, \dots, t\}$. We say that G has a *hyperbolic core* if there is a subset $C \subseteq I$ with $|C| \geq 3$ with $m_{ij} \geq 4$ for all $i, j \in C$ and with m_{ij} even if $i \in I - C$ and $j \in C$. (Of

course, ∞ is even.) In this case let $G(C)$ be the subgroup of G generated by the generators in C . There is a natural retraction homomorphism from G onto $G(C)$ which sends all a_i , $i \notin C$, to 1 and fixes all a_j where $j \in C$. The group $G(C)$ is itself a Coxeter group with a presentation

$$G(C) = \langle a_i, i \in C \mid a_i^2 = 1, (a_i a_j)^{m_{ij}} = 1, i, j \in C \rangle$$

by the general theory of Coxeter groups and is thus a nonelementary word-hyperbolic group by the discussion above.

Hence by Corollary B the word problem for any Coxeter group with a hyperbolic core has word problem solvable with linear time average-case complexity.

Example 1.11. Let G be an *Artin group*, that is, G has a presentation

$$G = \langle a_1, \dots, a_t \mid u_{ij} = u_{ji}, \text{ for } 1 \leq i < j \leq t \rangle$$

where for $i \neq j$

$$u_{ij} := \underbrace{a_i a_j a_i \dots}_{m_{ij} \text{ terms}}$$

and where $m_{ij} = m_{ji}$ for each $i < j$. We allow $m_{ij} = \infty$ in which case the relation $u_{ij} = u_{ji}$ is omitted from the above presentation. Suppose also that G is of *extra large type*, that is $t \geq 3$ and all $m_{ij} \geq 4$. Appel and Schupp [4] solved the word problem for Artin groups of extra-large type. Peifer [47] proved that all such groups are automatic and hence have quadratic time word problem. An Artin group G has an associated Coxeter group G_1 which is the quotient group obtained by setting the squares of the generators a_i equal to the identity. If G is of extra-large type then so is its Coxeter quotient

$$G_1 = G/ncl_G(a_1^2, \dots, a_t^2).$$

Hence G_1 is a nonelementary (since $t > 2$) word-hyperbolic group. Thus by Corollary B the word problem of any Artin group G of extra-large type is solvable in linear time on average.

It is worth pointing out that if the Artin group G is not necessarily of extra-large type but is such that the associated Coxeter group G_1 has a hyperbolic core, then the results of [37] show that the word problem of G is strongly generically linear time. However, it is an important open problem to show that the word problem for arbitrary Artin groups (as opposed to those of extra-large type) is solvable, so that we cannot yet draw any conclusions regarding the average-case complexity of the word problem in such G .

Example 1.12. Although Magnus proved that the word problem for any one-relator group is solvable in the 1930's (see, for example, [42]), there are no bounds on complexity of the word problem over the class of one-relator groups. Indeed, we do not know whether the word problem for every one-relator group is solvable in polynomial time or whether for every m there is

a one-relator group whose word problem has time complexity greater than a tower of m exponentials.

Suppose G is a group which can be given by a finite presentation involving at least two more generators than defining relators. A beautiful theorem of Baumslag and Pride [10] shows that G has a subgroup of finite index that maps homomorphically onto the free group of rank two. In particular, the theorem applies to all one-relator groups with at least three generators.

Thus if G is any one-relator group on at least three generators whose word problem is solvable in subexponential time then by Corollary B the word problem of G is solvable in linear time on average.

Results similar to the conclusion of Theorem A also hold for the *subgroup membership problem*.

Recall that if G is a group with a finite generating set $A = \{x_1, \dots, x_k\}$ and $H \leq G$ is a subgroup, then the *membership problem* $MP(G, H, A)$ for H in G relative to A consists of all those words w in the alphabet $X = A \cup A^{-1}$ which represent elements of H . Thus $MP(G, H, A) \subseteq X^*$. As in the case of the word problem, the complexity of the membership problem does not depend on the choice of A .

Our main result regarding the membership problem is:

Theorem C. *Let G be a finitely presented group and let $H \leq G$ be a subgroup where the membership problem for H in G is in $SubExp$. Let $G_1 \leq G$ be a subgroup of finite index in G such that $H \leq G_1$ and let $\phi : G \rightarrow \overline{G}$ be an epimorphism.*

Suppose there is a subgroup $\phi(H) \leq K \leq \overline{G}$ such that the Schreier coset graph for \overline{G} over K is non-amenable and such that the membership problem for K in \overline{G} is solvable in complexity class $\mathcal{C} \subseteq SubExp$.

Then for any finite generating set A of G the membership problem

$$MP(G, H, A)$$

for H in G is solvable with average-case complexity in \mathcal{C} uniformly relative to the family of all length-invariant discrete probability measures $\mu : (A \cup A^{-1})^ \rightarrow [0, 1]$.*

Moreover, for any tame discrete probability measure μ on $(A \cup A^{-1})^$ the membership problem $MP(G, H, A)$ is solvable with average-case complexity in \mathcal{C} relative to μ .*

Corollary D. *Let G be a finitely presented group and let $H \leq G$ be a subgroup where the membership problem for H in G is solvable in subexponential time. Let $G_1 \leq G$ be a subgroup of finite index in G such that $H \leq G_1$. Let $\phi : G_1 \rightarrow \overline{G}$ be an epimorphism and let $\phi(H) \leq K \leq \overline{G}$.*

- (1) *Suppose \overline{G} is a non-elementary word-hyperbolic group and that $K \leq \overline{G}$ is a rational subgroup of infinite index.*

Then for any finite generating set A of the group G the membership problem $MP(G, H, A)$ for H in G is solvable with linear time

average-case complexity uniformly relative to the family of all length-invariant discrete probability measures $\mu : (A \cup A^{-1})^* \rightarrow [0, 1]$.

Moreover, for any tame discrete probability measure μ on $(A \cup A^{-1})^*$ the membership problem $MP(G, H, A)$ is solvable with linear time average-case complexity relative to μ .

- (2) Suppose \overline{G} is automatic and that $K \leq \overline{G}$ is a rational subgroup such that the Schreier coset graph for \overline{G} over K is non-amenable (and hence K has infinite index in \overline{G}). Then for any finite generating set A of G and for any length-invariant discrete probability measure $\mu : (A \cup A^{-1})^* \rightarrow [0, 1]$ the membership problem $MP(G, H, A)$ for H in G is solvable with quadratic time average-case complexity uniformly relative to the family of all length-invariant discrete probability measures $\mu : (A \cup A^{-1})^* \rightarrow [0, 1]$.

Moreover, for any tame discrete probability measure μ on $(A \cup A^{-1})^*$ the membership problem $MP(G, H, A)$ is solvable with quadratic time average-case complexity relative to μ .

We should make an important disclaimer. The notion of average-case complexity used in this paper does not have many of the robustness properties that are desirable at finer levels of the complexity theory. For example, while our results are independent of the choice of a finite generating set for a group, more delicate issues such as model independence still have to be addressed (see [40, 52] for a more detailed discussion).

Nevertheless, we believe that our results constitute a valuable step in studying the largely unexplored field of the average-case complexity of algorithmic problems in group theory.

2. GENERIC-CASE COMPLEXITY

We need to recall some definitions from our earlier paper [37].

Definition 2.1 (Asymptotic density). Let X be a finite alphabet with at least two elements and let X^* denote the set of all words in X . Let S be a subset of X^* . For every $n \geq 0$ let B_n be the set all words in A^* of length at most n .

The *asymptotic density* $\rho(S)$ for S in X^* is defined as

$$\rho(S) := \limsup_{n \rightarrow \infty} \rho_n(S)$$

where

$$\rho_n(S) := \frac{|S \cap B_n|}{|B_n|},$$

If the actual limit $\lim_{n \rightarrow \infty} \rho_n(S)$ exists, we write $\hat{\rho}(S) := \rho(S)$.

The above notion was first suggested in the paper of Borovik, Myasnikov and Shpilrain [14].

If $\lim_{n \rightarrow \infty} a_n = a$, where $a_n, a \in \mathbb{R}$, we say that the convergence is *exponentially fast* provided there exist $0 \leq \sigma < 1$ and $C > 0$ such that for every $n \geq 1$ we have $|a_n - a| \leq C\sigma^n$.

Definition 2.2 (Generic performance of a partial algorithm). Let X be a finite alphabet with at least two letters and let $\mathcal{D} \subseteq X^*$ be a language in X .

Let Ω be a correct partial algorithm for \mathcal{D} which accepts as inputs words from X^* . That is, whenever Ω reaches a decision (either positive or negative) on whether a word w belongs to \mathcal{D} , that decision is correct. Let \mathcal{C} be a complexity class (e.g. linear time, quadratic time, linear space, etc).

We say that Ω *solves \mathcal{D} with generic-case complexity \mathcal{C}* if there is a subset $S \subseteq X^*$ with $\hat{\rho}(S) = 1$ such that for every $w \in S$ the partial algorithm Ω decides whether or not w is an element of \mathcal{D} within the complexity bound \mathcal{C} (in terms of w). If in addition $\lim_{n \rightarrow \infty} \rho_n(S) = 1$ converges exponentially fast, we say that Ω *solves \mathcal{D} with generic-case complexity strongly \mathcal{C}* .

We again stress that, unlike the average-case complexity, generic-case complexity totally disregards the complexity of the algorithm on the “small” set of inputs $X^* - S$.

3. PROOFS OF THE MAIN RESULTS

Convention 3.1. Let G be a group and let A be a finite alphabet equipped with a map $\pi : A \rightarrow G$ such that $\pi(A)$ generates G . In this case we say that (A, π) (or, by abuse of notation, just A) is a *finite generating set* for G . (Different letters of A may represent the same element of G and some letters of A may represent the identity $1 \in G$.) Every word w in the alphabet $A \cup A^{-1}$ represents an element of G which we will still denote by $\pi(w)$.

If $H \leq G$ is a subgroup, we define the *Schreier coset graph* $\Gamma(G, H, A)$ as follows. The vertices of Γ are the cosets $\{Hg \mid g \in G\}$. The graph Γ is oriented. The edge-set of Γ is partitioned as $E(\Gamma) = E^+ \sqcup E^-$, where the set E^+ is in one-to-one correspondence with $V(\Gamma) \times A$. Namely for each coset Hg and each $a \in A$ there is an edge e from Hg to Hga with label a in E^+ . The set E^- consists of formal inverses of the edges from E^+ , where the inverse of an the edge e above is the edge e^{-1} from Hga to Hg with label a^{-1} . Thus the edges of Γ are labeled by letters of $A \cup A^{-1}$. Note that if $k = |A|$ then Γ is connected and $2k$ -regular.

Intuitively, a regular connected graph Γ is *non-amenable* if the graph “grows rapidly in all directions”. We refer the reader to [9, 16, 22, 31, 36, 49, 53, 54] for a detailed discussion of non-amenability for groups and graphs, including the many equivalent definitions (such as the Følner condition, the growth-rate criterion for regular graphs, the doubling condition etc). A finitely generated group G is *non-amenable* if for some (and therefore for any) finite generating set A of G the Cayley graph $\Gamma(G, A)$ is non-amenable. This is equivalent to the standard definition of non-amenability for finitely generated groups. (A finitely generated group G is amenable if and only if

for any action of G by homeomorphisms on a compact space Q there exists a G -invariant probability measure on Q .)

Recall that a subgroup H of an automatic group G is said to be *rational* if for some automatic language L for G the pre-image L_H of H in L is itself a regular language. If G is word-hyperbolic then rational subgroups are also often called *quasiconvex*. We refer the reader to [2, 11, 23, 25, 26, 27, 30, 32, 29, 24, 48] for background information on hyperbolic and automatic groups and their rational subgroups. For the moment we need only recall that the word problem is solvable in quadratic time for any automatic group and in linear time for any hyperbolic group [2, 3, 24, 26, 35]. Also, if H is a rational subgroup of a hyperbolic (automatic) group G then the membership problem for H in G is solvable in linear (quadratic) time.

Our main technical tool is:

Proposition 3.2. *Let X be a finite alphabet with at least two elements and let $\mathcal{D} \subseteq X^*$. Suppose that \mathcal{D} is decidable in time $f(n)$ and strongly generically decidable in time $f_1(n) \leq f(n)$ where the function $f(n)/f_1(n)$ is subexponential. Then the language \mathcal{D} is decidable with average-case time-complexity bounded by $f_1(n)$ uniformly relative to the family of all length-invariant discrete probability measures $\mu : (A \cup A^{-1})^* \rightarrow [0, 1]$.*

Proof. Let Ω' be an algorithm solving \mathcal{D} in time $f(n)$. Let Ω'' be a partial algorithm which solves \mathcal{D} strongly generically in time $f_1(n)$. Define Ω to be the algorithm consisting of running Ω' and Ω'' concurrently. Let $\mu : X^* \rightarrow [0, 1]$ be a length-invariant discrete probability measure.

Denote $k := |X| \geq 2$. Let B_n be the set of all words in X^* of length at most n . Thus $|B_n| = 1 + k + k^2 + \dots + k^n = \frac{k^{n+1}-1}{k-1}$. Let $S \subseteq X^*$ be the set of inputs such that for each $w \in S$ the algorithm Ω'' terminates in time $f_1(|w|)$ and $\lim_{n \rightarrow \infty} \frac{|S \cap B_n|}{|B_n|} = 1$ exponentially fast. Let $K = X^* - S$. Then there is $C > 0$ and $1 \leq q < k$ such that for every $n \geq 1$ we have $|K \cap B_n| \leq Cq^n$. For each $w \in X^*$ denote by $T(w)$ the time required for algorithm Ω to reach a decision on input w . Thus for every $w \in S$ we have $T(w) \leq f_1(|w|)$ and for every $w \in X^*$ we have $T(w) \leq f(|w|)$.

Since μ is length-invariant, there is a function $d(n) \geq 0$ such that

$$\sum_{n=0}^{\infty} d(n) = 1$$

and such that for every $w \in X^*$ with $|w| = n$ we have $\mu(w) = \frac{d(n)}{k^n}$ where k^n is the number of all words of length n in X^* . In particular, $d(n) \leq 1$ for each n . We have:

$$\int_{X^*} \frac{T(w)}{f_1(|w|)} \mu(w) = \int_S \frac{T(w)}{f_1(|w|)} \mu(w) + \int_K \frac{T(w)}{f_1(|w|)} \mu(w).$$

Since for all $w \in S$ we have $T(w) \leq f_1(|w|)$, then

$$\int_S \frac{T(w)}{f_1(|w|)} \mu(w) \leq \int_S \frac{f_1(|w|)}{f_1(|w|)} \mu(w) = \mu(S) \leq 1 < \infty.$$

On the other hand for $w \in K$ with $|w| = n$ we have $T(w) \leq f(n)$ and $\mu(w) = \frac{d(n)}{k^n}$. Hence $\frac{f(|w|)}{f_1(|w|)} d(|w|) \leq \frac{f(n)}{f_1(n)}$ where the function $\frac{f(n)}{f_1(n)}$ is subexponential. Let K_n be the set of all $w \in K$ with $|w| = n$. Then $|K_n| \leq |K \cap B_n| \leq Cq^n$.

Hence

$$\begin{aligned} \int_K \frac{T(w)}{f_1(|w|)} \mu(w) &= \sum_{n=0}^{\infty} \sum_{w \in K_n} \frac{T(w)}{f_1(|w|)} \mu(w) = \sum_{n=0}^{\infty} \sum_{w \in K_n} \frac{f(n)}{f_1(n)} d(n) \frac{1}{k^n} \leq \\ &\leq C \sum_{n=0}^{\infty} \frac{f(n)}{f_1(n)} d(n) \frac{q^n}{k^n} \leq C \sum_{n=0}^{\infty} \frac{f(n)}{f_1(n)} \frac{q^n}{k^n} = C_0 < \infty \end{aligned}$$

since $k > q$ and the function $\frac{f(n)}{f_1(n)}$ is subexponential. Therefore

$$\int_{X^*} \frac{T(w)}{f_1(|w|)} \mu(w) = \int_S \frac{T(w)}{f_1(|w|)} \mu(w) + \int_K \frac{T(w)}{f_1(|w|)} \mu(w) \leq 1 + C_0 < \infty.$$

Since C_0 does not depend on the choice of μ , the statement of Proposition 3.2 follows. \square

Remark 3.3. The proof of Proposition 3.2 is sufficiently robust to accommodate some other notions of average-case complexity. One of such definitions uses the Cauchy density $d(n) = \frac{1}{n^2}$ to define the notion of average-case complexity being at most polynomial of degree p . Namely, we say that, using the notations of Definition 1.4, *the algorithm Ω solves \mathcal{D} with Cauchy average-case time-complexity bounded by polynomial of degree p* if for any $\epsilon > 0$

$$\int_{X^*} \frac{T(w)}{|w|^{p-1+\epsilon}} \mu(w) = \sum_{w \in X^*} \frac{T(w)}{|w|^{p-1+\epsilon}} \mu(w) < \infty,$$

where $\mu(w) = \frac{1}{|w|^{2k^{|w|}}}$.

The proof of Proposition 3.2 easily goes through to show that if \mathcal{D} is decidable in subexponential time and strongly generically decidable in polynomial of degree p time, then \mathcal{D} is decidable with Cauchy average-case time-complexity bounded by polynomial of degree p .

Indeed, we have to analyze the same integral with $f_1(n) = n^{p-1+\epsilon}$. As before, we decompose this integral into two parts corresponding to K and S accordingly. For the K -part the functions $T(w)$ and $T(w)/|w|^{p-1+\epsilon}$ are subexponential and hence the K -integral is finite by the same argument as in the proof of Proposition 3.2. For the S -part we have $T(w) \leq C|w|^p$ for any $w \in S$. Let s_n be the number of words of length n in S . Hence $s_n \leq k^n$ and

$$\begin{aligned} \int_S \frac{T(w)}{|w|^{p-1+\epsilon}} \mu(w) &\leq C \int_S \frac{|w|^p}{|w|^{p-1+\epsilon} |w|^{2k|w|}} = \\ &= C \sum_{n=1}^{\infty} \frac{1}{n^{1+\epsilon}} \frac{s_n}{k^n} \leq C \sum_{n=1}^{\infty} \frac{1}{n^{1+\epsilon}} < \infty. \end{aligned}$$

This implies that the whole integral over X^* is finite, as required.

The requirement that measures in Proposition 3.2 be length-invariant can be considerably weakened at the cost of sacrificing the “uniform” conclusion of Proposition 3.2.

Proposition 3.4. *Let X be a finite alphabet with $k \geq 2$ elements and let $\mathcal{D} \subseteq X^*$. Suppose that \mathcal{D} is decidable in time $f(n)$ and strongly generically decidable in time $f_1(n) \leq f(n)$ where the function $f(n)/f_1(n)$ is subexponential. Then for any tame measure μ on X^* the language \mathcal{D} is decidable with μ -average-case time-complexity bounded by $f_1(n)$.*

Proof. Recall that since μ is tame, there exists a subexponential function $g(n) \geq 0$ such that for any $w \in X^*$ we have $\mu(w) = \frac{g(|w|)}{k^{|w|}}$.

The argument is almost exactly the same as in the proof of Proposition 3.2 and the only change required is in the estimate of the integral over the strongly negligible set K with $|K \cap B_n| \leq Cq^n$, $0 < q < k$, and with $K_n = \{w \in K : |w| = n\}$:

$$\begin{aligned} \int_K \frac{T(w)}{f_1(|w|)} \mu(w) &= \sum_{n=0}^{\infty} \sum_{w \in K_n} \frac{T(w)}{f_1(|w|)} \mu(w) \leq \sum_{n=0}^{\infty} \sum_{w \in K_n} \frac{f(n)}{f_1(n)} \frac{g(n)}{k^n} \leq \\ &\leq C \sum_{n=0}^{\infty} \frac{f(n)}{f_1(n)} g(n) \frac{q^n}{k^n} < \infty. \end{aligned}$$

The last inequality holds since $\frac{f(n)}{f_1(n)}$ and $g(n)$ are subexponential functions and hence so is their product. \square

Proposition 3.2 together with the results of our previous paper [37] immediately implies the main results stated in the Introduction.

Proof of Theorem A and Theorem C. Suppose the assumptions of Theorem A are satisfied. By the results of [37] since \overline{G} is non-amenable, for any finite generating set A of G the word problem $WP(G, A)$ for G is solvable strongly generically with complexity \mathcal{C} . Together with Proposition 3.2 and Proposition 3.4 this implies the statement of Theorem A.

Any non-elementary hyperbolic group contains a free subgroup of rank two and thus is non-amenable. Moreover, by [3, 2, 35] the word problem in a hyperbolic group is solvable in linear time. Additionally, the word problem in an automatic group is solvable in quadratic time [24]. The statement of Corollary B now follows from Theorem A.

Suppose now that the assumptions of Theorem C hold. Since the Schreier graph of \overline{G} over \overline{H} is non-amenable, the result of [37] implies that for any finite generating set A of G the membership problem $MP(G, H, A)$ is solvable strongly generically with complexity \mathcal{C} . The statement of Theorem C now follows from Proposition 3.2 and Proposition 3.4.

Recall that the membership problem for a rational subgroup of a hyperbolic group is solvable in linear time. Also, the membership problem for a rational subgroup of an automatic group is solvable in quadratic time. Since non-elementary hyperbolic groups are non-amenable, Corollary D now follows from Theorem C. \square

Remark 3.5. One may argue that the natural set of inputs for the word problem of a group $G = \langle A \rangle$ is the set of elements of the free group $F(A)$ rather than the set of all words (including those which are not freely reduced) in the alphabet $A \cup A^{-1}$. Then one would need to talk about the average-case complexity of the word or the membership problem with respect to a length-invariant discrete probability measure $\mu : F(A) \rightarrow [0, 1]$. For an element $w \in F(A)$, let $|w|$ be the length of the unique freely reduced word in $A \cup A^{-1}$ representing w . Here a “length-invariant measure” would mean a measure such that for any $w_1, w_2 \in F(A)$ with $|w_1| = |w_2|$ we have $\mu(w_1) = \mu(w_2)$. Then the definition of average-case complexity would need to be modified so that the summation occurs over all elements of $F(A)$ rather than over all words in the language $(A \cup A^{-1})^*$.

Since spheres and balls of radius n in $F(A)$ grow as $(2k - 1)^n$ (where k is the number of letters in A), the definition of a *tame* discrete probability measure μ on $F(A)$ needs to be adjusted by requiring that for any $w \in F(A)$ we have $\mu(w) \leq \frac{g(|w|)}{(2k-1)^{|w|}$, where $g(n) \geq 0$ is some subexponential function.

One can also define asymptotic density for a subset $S \subseteq F(A)$ similarly to Definition 2.1. The only difference is that the denominator in the fraction $\rho_n(S)$ would be the total number of elements $f \in F$ with $|w| \leq n$. All the results of [37] regarding the generic-case complexity of the word and the membership problem are proved in parallel for both approaches.

Thus all the results (together with exactly the same proofs) of this paper remain true in this alternative approach.

We give here a sample version of such an alternatively stated result:

Theorem E. *Let G be a finitely presented group where the word problem is in deterministic time-complexity class \mathcal{C} . Suppose G has a subgroup of finite index that possesses a non-amenable quotient group \overline{G} where the word problem belongs to a time-complexity class $\mathcal{C} \subseteq \text{SubExp}$.*

Then for any finite generating set A of G the word problem for G is solvable with average-case time-complexity in \mathcal{C} uniformly relative to the family of all length-invariant discrete probability measures $\mu : F(A) \rightarrow [0, 1]$.

Moreover, for any tame discrete probability measure μ on $F(A)$ the word problem for G is solvable with average-case complexity in \mathcal{C} relative to μ .

REFERENCES

- [1] S. I. Adian and V. G. Durnev, *Algorithmic problems for groups and semigroups*, Uspekhi Mat. Nauk **55** (2000), no. 2, 3–94; translation in Russian Math. Surveys **55** (2000), no. 2, 207–296
- [2] J. Alonso, T. Brady, D. Cooper, V. Ferlini, M. Lustig, M. Mihalik, M. Shapiro and H. Short, *Notes on hyperbolic groups*, In: " Group theory from a geometrical viewpoint", Proceedings of the workshop held in Trieste, É. Ghys, A. Haefliger and A. Verjovsky (editors). World Scientific Publishing Co., 1991
- [3] M. Anshel and B. Domanski, *The complexity of Dehn's algorithm for word problems in groups*, J. Algorithms **6** (1985), no. 4, 543–549
- [4] K. Appel and P. Schupp, *Coxeter groups, Artin groups and small cancellation theory*, Invent. Math. **72** (1983), no. 2, 201–220
- [5] G. Arzhantseva, *On groups in which subgroups with a fixed number of generators are free*, (Russian) Fundam. Prikl. Mat. **3** (1997), no. 3, 675–683
- [6] G. Arzhantseva, *Generic properties of finitely presented groups and Howson's theorem*, Comm. Algebra **26** (1998), no. 11, 3783–3792
- [7] G. Arzhantseva, *A property of subgroups of infinite index in a free group*, Proc. Amer. Math. Soc. **128** (2000), no. 11, 3205–3210
- [8] G. Arzhantseva and A. Olshanskii, *Genericity of the class of groups in which subgroups with a lesser number of generators are free*, (Russian) Mat. Zametki **59** (1996), no. 4, 489–496
- [9] L. Bartholdi, *Counting paths in graphs*, Enseign. Math. (2) **45** (1999), no. 1-2, 83–131
- [10] B. Baumslag and S. J. Pride, *Groups with two more generators than relators*, J. London Math. Soc. (2) **17** (1978), no. 3, 425–426
- [11] G. Baumslag, S. M. Gersten, M. Shapiro and H. Short, *Automatic groups and amalgams*. J. Pure Appl. Algebra **76** (1991), no. 3, 229–316
- [12] G. Baumslag and C. F. Miller, III, *Experimenting and computing with infinite groups*. Groups and computation, II (New Brunswick, NJ, 1995), 19–30, DIMACS Ser. Discrete Math. Theoret. Comput. Sci., **28**, Amer. Math. Soc., Providence, RI, 1997.
- [13] R. Benedetti, and C. Petronio, *Lectures on hyperbolic geometry*. Universitext. Springer-Verlag, Berlin, 1992
- [14] A. Borovik, A. G. Myasnikov, V. Shpilrain, *Measuring sets in infinite groups*, Contemp. Math., Computational and statistical group theory (Las Vegas, NV/Hoboken, NJ, 2001), 21–42, Contemp. Math., 298, Amer. Math. Soc., Providence, RI, 2002.
- [15] B. Brink and R. Howlett, *A finiteness property and an automatic structure for Coxeter groups*. Math. Ann. **296** (1993), no. 1, 179–190.
- [16] T. Ceccherini-Silberstein, R. Grigorchuck and P. de la Harpe, *Amenability and paradoxical decompositions for pseudogroups and discrete metric spaces*, (Russian) Tr. Mat. Inst. Steklova **224** (1999), Algebra. Topol. Differ. Uravn. i ikh Prilozh., 68–111; translation in Proc. Steklov Inst. Math., **224** (1999), no. 1, 57–97.
- [17] C. Champetier, *Petite simplification dans les groupes hyperboliques*, Ann. Fac. Sci. Toulouse Math. (6) **3** (1994), no. 2, 161–221.
- [18] C. Champetier, *Propriétés statistiques des groupes de présentation finie*, Adv. Math. **116** (1995), no. 2, 197–262.
- [19] C. Champetier, *The space of finitely generated groups*, Topology **39** (2000), no. 4, 657–680
- [20] P.-A. Cherix and A. Valette, *On spectra of simple random walks on one-relator groups*, With an appendix by Paul Jolissaint. Pacific J. Math. **175** (1996), no. 2, 417–438

- [21] P.-A. Cherix and G. Schaeffer, *An asymptotic Freiheitssatz for finitely generated groups*, Enseign. Math. (2) **44** (1998), no. 1-2, 9–22
- [22] J. M. Cohen, *Cogrowth and amenability of discrete groups*, J. Funct. Anal. **48** (1982), 301–309.
- [23] M. Coornaert, T. Delzant, and A. Papadopoulos, *Géométrie et théorie des groupes. Les groupes hyperboliques de Gromov*. Lecture Notes in Mathematics, 1441; Springer-Verlag, Berlin, 1990
- [24] D. Epstein, J. Cannon, D. Holt, S. Levy, M. Paterson, W. Thurston, *Word Processing in Groups*, Jones and Bartlett, Boston, 1992
- [25] B. Farb, *Automatic groups: a guided tour*, Enseign. Math. (2) **38** (1992), no. 3-4, 291–313
- [26] S. Gersten, *Introduction to hyperbolic and automatic groups*. Summer School in Group Theory in Banff, 1996, 45–70, CRM Proc. Lecture Notes, **17**, Amer. Math. Soc., Providence, RI, 1999
- [27] S. Gersten and H. Short, *Rational subgroups of bi-automatic groups*, Ann. of Math. (2) **134** (1991), no. 1, 125–158
- [28] E. Ghys, *Groupes Aléatoires*. Seminar Bourbaki (March 2003), to appear in *Asterisque*
- [29] E. Ghys and P. de la Harpe (editors), *Sur les groupes hyperboliques d'après Mikhael Gromov*, Birkhäuser, Progress in Mathematics series, vol. **83**, 1990.
- [30] R. Gilman, *Groups with a rational cross-section*. Combinatorial group theory and topology (Alta, Utah, 1984), 175–183, Ann. of Math. Stud., 111, Princeton Univ. Press, Princeton, NJ, 1987
- [31] R. Grigorchuk, *Symmetrical random walks on discrete groups*, Multicomponent random systems, pp. 285–325, Adv. Probab. Related Topics, **6**, Dekker, New York, 1980
- [32] M. Gromov, *Hyperbolic groups*, Essays in group theory, Springer, New York, 1987, pp. 75–263.
- [33] M. Gromov, *Random walks in random groups*, Geom. Funct. Analysis **13** (2003), no. 1, 73–146
- [34] M. Gromov, *Asymptotic invariants of infinite groups*, Geometric group theory, Vol. 2 (Sussex, 1991), Cambridge Univ. Press, Cambridge, 1993, pp. 1–295.
- [35] D. Holt, *Word-hyperbolic groups have real-time word problem*, Internat. J. Algebra Comput. **10** (2000), no. 2, 221–227
- [36] I. Kapovich, *The non-amenability of Schreier graphs for infinite index quasiconvex subgroups of hyperbolic groups*, Enseign. Math. **48** (2002), no. 3-4, 359–375
- [37] I. Kapovich, A. Myasnikov, V. Shpilrain and P. E. Schupp, *Generic-case complexity and decision problems in group theory and random walks*, J. Algebra **264** (2003), no. 2, 665–694
- [38] I. Kapovich and P. Schupp, *Genericity, the Arzhantseva-Ol'shanskii method and the isomorphism problem for one-relator groups*, preprint, 2002; <http://front.math.ucdavis.edu/math.GR/0210307>
- [39] I. Kapovich, P. Schupp and V. Shpilrain, *Generic properties of Whitehead's algorithm, stabilizers in $\text{Aut}(F_k)$ and one-relator groups*, preprint, 2003; <http://front.math.ucdavis.edu/math.GR/0303386>
- [40] L. Levin, *Average case complete problems*, SIAM J. Comput. **15** (1986), 285–286
- [41] R. Lipton and Y. Zalcstein, *Word problems solvable in logspace*, J. Assoc. Comput. Mach. **24** (1977), no. 3, 522–526
- [42] R. C. Lyndon and P. E. Schupp, *Combinatorial Group Theory*, Ergebnisse der Mathematik, band 89, Springer 1977. Reprinted in the Springer Classics in Mathematics series, 2000.

- [43] C. F. Miller III, *Decision problems for groups- Survey and reflections.* in *Algorithms and Classification in Combinatorial Group Theory*, G. Bamuslag and C.F. Miller III, editors, (1992), Springer, 1–60.
- [44] Y. Ollivier, *Critical densities for random quotients of hyperbolic groups.* C. R. Math. Acad. Sci. Paris **336** (2003), no. 5, 391–394
- [45] A. Yu. Ol’shanskii, *Almost every group is hyperbolic*, Internat. J. Algebra Comput. **2** (1992), no. 1, 1–17.
- [46] C. Papadimitriou, *Computation Complexity*, Addison-Wesley, Reading, 1994
- [47] D. Peifer, *Artin groups of extra-large type are biautomatic*, J. Pure Appl. Algebra **110** (1996), no. 1, 15–56
- [48] H. Short, *An introduction to automatic groups.* Semigroups, formal languages and groups (York, 1993), 233–253, NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci., **466**, Kluwer Acad. Publ., Dordrecht, 1995
- [49] A. M. Vershik, *Dynamic theory of growth in groups: entropy, boundaries, examples* Uspekhi Mat. Nauk **55** (2000), no. 4(334), 59–128; translation in Russian Math. Surveys **55** (2000), no. 4, 667–733
- [50] J. Wang, *Average-case completeness of a word problem in groups*, Proc. of the 27-th Annual Symposium on Theory of Computing, ACM Press, New York, 1995, 325–334
- [51] J. Wang, *Average-case computational complexity theory*, Complexity Theory Retrospective, II. Springer-Verlag, New York, 1997, 295–334
- [52] J. Wang, *Distributional word problem for groups*, SIAM J. Comput. **28** (1999), no. 4, 1264–1283
- [53] W. Woess, *Cogrowth of groups and simple random walks*, Arch. Math. **41** (1983), 363–370.
- [54] W. Woess, *Random walks on infinite graphs and groups - a survey on selected topics*, Bull. London Math. Soc. **26** (1994), 1–60.
- [55] A. Zuk, *On property (T) for discrete groups.* Rigidity in dynamics and geometry (Cambridge, 2000), 473–482, Springer, Berlin, 2002

Department of Mathematics, University of Illinois at Urbana-Champaign, 1409 West Green Street, Urbana, IL 61801, USA

E-mail address: kapovich@math.uiuc.edu

Department of Mathematics, The City College of New York, New York, NY 10031

E-mail address: alexeim@att.net

Department of Mathematics, University of Illinois at Urbana-Champaign, 1409 West Green Street, Urbana, IL 61801, USA

E-mail address: schupp@math.uiuc.edu

Department of Mathematics, The City College of New York, New York, NY 10031

E-mail address: shpil@groups.sci.ccnycunyu.edu