# PROBABILISTIC WARING PROBLEMS FOR FINITE SIMPLE GROUPS

MICHAEL LARSEN, ANER SHALEV, AND PHAM HUU TIEP

Abstract. The probabilistic Waring problem for finite simple groups asks whether every word of the form $w_1w_2$, where $w_1$ and $w_2$ are non-trivial words in disjoint sets of variables, induces almost uniform distributions on finite simple groups with respect to the $L^1$ norm. Our first main result provides a positive solution to this problem.

We also provide a geometric characterization of words inducing almost uniform distributions on finite simple groups of Lie type of bounded rank, and study related random walks.

Our second main result concerns the probabilistic $L^\infty$ Waring problem for finite simple groups. We show that for every $l \geq 1$ there exists (an explicit) $N = N(l) = O(l^4)$, such that if $w_1, \ldots, w_N$ are non-trivial words of length at most $l$ in pairwise disjoint sets of variables, then their product $w_1 \cdots w_N$ is almost uniform on finite simple groups with respect to the $L^\infty$ norm. The dependence of $N$ on $l$ is genuine. This result implies that, for every word $w = w_1 \cdots w_N$ as above, the word map induced by $w$ on a semisimple algebraic group over an arbitrary field is a flat morphism.

Applications to representation varieties, subgroup growth, and random generation are also presented. In particular we show that, for certain one-relator groups $\Gamma$, a random homomorphism from $\Gamma$ to a finite simple group $G$ is surjective with probability tending to 1 as $|G| \to \infty$.

## Contents



1991 *Mathematics Subject Classification.* Primary 20P05; Secondary 11P05, 20C30, 20C33, 20D06, 20G40.

ML was partially supported by NSF grants DMS-1401419 and DMS-1702152. AS was partially supported by ISF grant 686/17 and the Vinik Chair of mathematics which he holds. PT was partially supported by NSF grant DMS-1840702. The authors were also partially supported by BSF grant 2016072.

The paper is partially based upon work supported by the NSF under grant DMS-1440140 while AS and PT were in residence at MSRI (Berkeley, CA), during the Spring 2018 semester. It is a pleasure to thank the Institute for the hospitality and support.

The authors are grateful to the referee for careful reading and insightful comments that helped greatly improve the paper.

## 1. Introduction

In the past two decades there has been much interest in word maps and related Waring type problems (see for instance [Sh2] and the references therein). Recall that a word is an element $w = w(x_1, \dots, x_d)$ of the free group $F_d$ on $x_1, \dots, x_d$. Given any group $G$, the word $w$ gives rise to a word map $w_G \colon G^d \to G$ induced by substitution. When the group $G$ is understood, we denote the map simply $w$.

Word maps on finite simple groups have attracted particular attention. Here and throughout this paper, by a finite simple group we mean a non-abelian finite simple group. Two words $w_1, w_2$ are said to be *disjoint* if they are words in disjoint sets of variables. If $w = w_1 w_2$ where $w_1, w_2 \neq 1$ are disjoint words, then it was shown in [LST] (following partial results from [LS1, LS2]) that the word map $w$ is surjective on all sufficiently large finite simple groups. This provides a best possible solution to the Waring problem for finite simple groups, inspired by the classical Waring problem in number theory.

The *probabilistic* Waring problem for finite simple groups asks whether, for $w = w_1 w_2$ as above, the push-forward distribution $p_{w,G} = w_* U_{G^d}$ on a finite simple group $G$ tends to the uniform distribution $U_G$ in the $L^1$ norm (see (1.1)) as the order of $G$ tends to infinity. That is, for a word $w$, a finite group $G$ and an element $g \in G$, we let $p_{w,G}(\{g\})$ denote the probability that $w(g_1, \dots, g_d) = g$ when $g_i \in G$ are chosen uniformly and independently:

$$p_{w,G}(\{g\}) = \frac{|w^{-1}(g)|}{|G|^d}.$$

It is conjectured that, for finite simple groups $G$, we have

$$\lim_{|G| \to \infty} \|p_{w,G} - U_G\|_{L^1} = 0$$

(see for instance [Sh2, 4.5]). When this holds we say that $w$ is *almost uniform* on finite simple groups.

This conjecture has already been established in some cases. In [GS, 7.1] it is proved for $w = x_1^2 x_2^2$ (and it is also shown in [GS] that the commutator word $[x_1, x_2]$ is almost uniform). In [LS4, 1.1] the conjecture is proved for $w = x_1^m x_2^n$ where $m, n$ are arbitrary non-zero integers. It is also shown in [LS4, 1.2] that admissible words, i.e., words in which each variable appears exactly twice, once as $x_i$ and once as $x_i^{-1}$, are almost uniform on finite simple groups. In [LS1] the conjecture is established for arbitrary $w_1 w_2$ for alternating groups $\mathsf{A}_n$ (see Theorem 1.18 there and the discussion following

it). In this paper we prove the conjecture in full by confirming it for simple groups of Lie type.

For a real function $f$ on a finite set $G$ and a real number $p > 0$, we define

$$\|f\|_{L^p} = (|G|^{p-1} \sum_{g \in G} |f(g)|^p)^{1/p}.$$

In particular,

$$\|f\|_{L^1} = \sum_{g \in G} |f(g)|, \ \|f\|_{L^\infty} = |G| \cdot \max_{g \in G} |f(g)|. \tag{1.1}$$

Our first main result is as follows.

**Theorem 1.** *Let $w_1, w_2 \neq 1$ be disjoint words and let $w = w_1 w_2$. Then*

$$\lim_{|G| \to \infty} \|p_{w,G} - U_G\|_{L^1} = 0,$$

*where $G$ ranges over all finite simple groups.*

Let $G$ and $w \in F_d$ be as in Theorem 1. Then it follows from the theorem that, as $|G| \to \infty$ and $S \subseteq G^d$ satisfies $|S|/|G|^d \to 1$, we have

$$|w(S)|/|G| \to 1,$$

namely, almost all elements $g \in G$ can be expressed in the form $g = w(g_1, \ldots, g_d)$ where $(g_1, \ldots, g_d) \in S$. Combining this observation with suitable known results we obtain some immediate applications of Theorem 1. For example, using [LiSh1, Theorem] we deduce that *almost all elements of $G$ have the form $w(g_1, \ldots, g_d)$ where $\langle g_i, g_j \rangle = G$ for all $1 \leq i < j \leq d$.* The same holds when we require that all $g_i$ are regular semisimple if $G$ is of Lie type and bounded rank; that $n^{(1/2-\epsilon)\log n} \leq o(g_i) \leq n^{(1/2+\epsilon)\log n}$ for $G = \mathsf{A}_n$ where $\epsilon > 0$ and $o(g)$ is the order $g$ (see [ET, Theorem]); that $|\mathbf{C}_G(g_i)| \leq q^{(1+\epsilon)r}$ where $\epsilon > 0$ and $G$ is classical of rank $r$ over $\mathbb{F}_q$ (this follows by combining Corollary 1.2(1) of [FG] with Lemma 5.3 of [Sh1]).

The proof of Theorem 1 makes use of the classification of finite simple groups. Since the result is asymptotic in nature, we do not need to consider sporadic groups at all, so it remains to deal with groups of Lie type. For groups of classical type of unbounded rank, we combine arguments of combinatorial flavor with essential use of strong new character estimates proved in [GLT2, GLT3]. For groups of Lie type of bounded rank (including exceptional groups) we provide two proofs: one character-theoretic, and the other geometric. The latter proof is based on the following characterization of almost uniform words in bounded rank, which is of independent interest.

**Theorem 2.** *Let $r$ and $d$ be positive integers, and $w \in F_d$ a non-trivial word. The following conditions are equivalent:*

(i) *As $G$ ranges over finite simple groups of Lie type of rank $\leq r$,*

$$\lim_{|G| \to \infty} \|p_{w,G} - U_G\|_{L^1} = 0.$$

(ii) *For every prime $p$ and every split simply connected semisimple group $\underline{G}$ over $\mathbb{F}_p$ of rank $\leq r$, there exists a power $q$ of $p$ such that*

$$\lim_{n\to\infty} \frac{|w(\underline{G}(\mathbb{F}_{q^n}))|}{|\underline{G}(\mathbb{F}_{q^n})|} = 1.$$

(iii) *For every simply connected semisimple group $\underline{G}$ of rank $\leq r$ over any field $\Bbbk$, the evaluation morphism $w\colon \underline{G}^d \to \underline{G}$ has geometrically irreducible generic fiber.*

If these equivalent conditions hold we say that $w$ is *almost uniform in rank* $\leq r$. If this is true for all $r$, we say that $w$ is *almost uniform in bounded rank*. Our geometric proof of Theorem 1 in bounded rank is based on Theorem 2 above and the fact that the generic fiber for any word of the form $w = w_1w_2$, where $w_1$ and $w_2$ are disjoint non-trivial words, is geometrically irreducible.

Theorem 2 shows, in particular, that words that are surjective on large enough finite simple groups of bounded rank are also almost uniform in bounded rank. This is by no means obvious. In Segal's monograph [Seg] a word $w \in F_d$ is said to be *silly* if $w \in x_1^{e_1} \dots x_d^{e_d} F_d'$ where $\gcd(e_1, \dots, e_d) = 1$. It is observed in [Seg, 3.1.1] that silly words are precisely the words that are surjective on all groups. It therefore follows that silly words are almost uniform in bounded rank. In our next result we estimate the probability that a random word has the above properties.

For any $d > 1$ and $n \geq 0$, we let $\mathsf{W}_{n,d}$ denote the random element of $F_d$ obtained from an $n$ step random walk on $F_d$ with steps uniformly distributed in $\{x_1^{\pm 1}, \dots, x_d^{\pm 1}\}$.

**Theorem 3.** (i) *For all $d > 1$ and all $n > 0$, the probability that $\mathsf{W}_{n,d}$ is surjective on all groups exceeds $1/3$ and tends to $1$ as $d \to \infty$.*

(ii) *For all $d > 1$ and all $n > 0$, the probability that $\mathsf{W}_{n,d}$ is almost uniform in bounded rank exceeds $1/3$ and tends to $1$ as $d \to \infty$.*

It has recently been shown in [CH, Theorem B] that a finite group $G$ is nilpotent if and only if all words $w$ which are surjective on $G$ induce the uniform distribution $U_G$ on $G$. Using this and part (i) of Theorem 3 it follows that *the probability that $\mathsf{W}_{n,d}$ is uniform on all finite nilpotent groups exceeds $1/3$ and tends to $1$ as $d \to \infty$.*

Next, we turn to almost uniformity results with respect to other norms.

For the groups $\mathrm{PSL}_2(q)$, we can strengthen Theorem 1, replacing the $L^1$ norm by the $L^2$ norm. Indeed, by Corollary 4.3 below, if $G = \mathrm{PSL}_2(q)$ where $q$ ranges over prime powers, then

$$\lim_{q\to\infty} \|p_{w,G} - U_G\|_{L^2} = 0.$$

It would be interesting to find out which families of finite simple groups satisfy this property. We note that if $w$ is $[x_1, x_2]$ or $x_1^2x_2^2$ then we have

$$\lim_{|G|\to\infty} \|p_{w,G} - U_G\|_{L^2} = 0$$

as $G$ ranges over finite simple groups; indeed, this follows from [GS, §2].

It is also interesting to obtain almost uniformity results in the $L^\infty$ norm or the $L^p$ norm for arbitrary $p > 1$. In this sense, for any fixed $k \geq 1$, the product $w_1 \cdots w_k$ of $k$ non-trivial pairwise disjoint words need not be almost uniform on all finite simple groups. Indeed we may take $w_i = x_i^n$ for $n \geq k+2$. If $G$ is an alternating group of large degree (compared to $n$), then it follows from Lemmas 2.17 and 2.18 of [LiSh2] that $p_{x_i^n,G}(1) \geq |G|^{-1/n}$. If $G$ is a classical group of large rank (compared to $n$) over a field with $q$ elements, then for some constant $c(n) > 0$ one has that

$$p_{x_i^n,G}(1) \geq q^{-c(n)}|G|^{-1/n} > |G|^{-1/(n-1)}$$

by [LiSh3, Theorem 4.3]. Hence,

$$p_{w,G}(\{1\}) \geq |G|^{-k/(n-1)},$$

and $w$ is not almost uniform in $L^\infty$. If moreover we take $n \geq kp/(p-1)+2$ then we see that $w$ is not almost uniform in $L^p$ whenever $p > 1$.

However, we do show in Corollary 4.4 below, that if $w = w_1w_2w_3w_4$, a product of four pairwise disjoint non-trivial words, then $w$ is almost uniform on $\mathrm{PSL}_2(q)$ with respect to the $L^\infty$ norm. This result is best possible in the sense that it fails to hold for some products of three disjoint words. Indeed, it is shown in [Sh1, p. 1406] that $x_1^2x_2^2x_3^2$ is not almost uniform on $\mathrm{PSL}_2(q)$ with respect to the $L^\infty$ norm.

Our second main result concerns the probabilistic Waring problem for finite simple groups with respect to the $L^\infty$ norm.

**Theorem 4.** *For a positive integer $l$ define $N(l) = 2 \cdot 10^{18}l^4$. Let $N \geq N(l)$ be an integer and $w = w_1 \cdots w_N$ a product of pairwise disjoint non-trivial words of length at most $l$. Then*

$$\lim_{|G|\to\infty} \|p_{w,G} - U_G\|_{L^\infty} = 0,$$

*where $G$ runs over all finite simple groups.*

This theorem generalizes Proposition 8.5 of [Sh1] dealing with Lie type groups of bounded rank (where $N$ depends on the rank $r$ and not on $l$; cf. Proposition 6.10), and Theorem 2.8 of [Sh1] where $w_i$ are commutators and $N = 2$. Unlike Theorem 1, which was established long ago for alternating groups, Theorem 4 is new (and highly non-trivial) also for $\mathsf{A}_n$ – note that in this case the bound for $N(l)$ is substantially smaller, see Proposition 6.8. The proof of Theorem 4 is rather complicated, combining combinatorial and character methods. In particular it follows from the theorem that $x_1^l x_2^l \cdots x_{N(l)}^l$ is almost uniform in $L^\infty$ on finite simple groups, which may be regarded as a probabilistic non-commutative analogue of the Waring problem in number theory. The discussion prior to Theorem 4 shows that the conclusion of the theorem does not hold for $N \leq l-2$, so the dependence of $N$ on $l$ is genuine.

Our third main result concerns flatness of certain word maps on algebraic groups, representation varieties and subgroup growth of some one-relator groups, as well as random generation of finite simple groups. While parts (i)–(iv) below are applications of Theorem 4, parts (v) and (vi) are more challenging and require various additional tools.

Let us say that a word $w \in F_d$ is *even* if its image in the abelianization $F_d/[F_d, F_d]$ is a square, and that $w$ is *odd* otherwise (see also Definition 6.6 below).

**Theorem 5.** *For every positive integer $l$ there exists a positive integer $N^*(l)$ such that the following statement holds. Let $N \geq N^*(l)$ and $d$ be positive integers. Suppose $w = w_1 \cdots w_N \in F_d$ is a product of pairwise disjoint non-trivial words of length at most $l$, and $\Gamma = \langle x_1, \ldots, x_d \mid w(x_1, \ldots, x_d) = 1 \rangle$. Then all the following statements hold.*

(i) *For every field $\Bbbk$ and every semisimple algebraic group $\underline{G}$ over $\Bbbk$, the word morphism $w \colon \underline{G}^d \to \underline{G}$ is flat.*

(ii) *For every field $\Bbbk$ and every semisimple algebraic group $\underline{G}$ over $\Bbbk$, the dimension of the $\Bbbk$-variety $\mathrm{Hom}(\Gamma, \underline{G})$, i.e., the Krull dimension of its coordinate ring, is $(d-1)\dim \underline{G}$.*

(iii) *For every field $\Bbbk$ and every positive integer $n$, the dimension of the $\Bbbk$-variety $\mathrm{Hom}(\Gamma, \mathrm{GL}_n)$ is $(d-1)n^2 + a$, where $a = 0$ if $w \notin [F_d, F_d]$ and $a = 1$ otherwise.*

(iv) *The number $a_n(\Gamma)$ of index $n$ subgroups of $\Gamma$ satisfies*

$$a_n(\Gamma) \sim bn \cdot (n!)^{d-2},$$

*where $b = 1$ if $w$ is odd and $b = 2$ if $w$ is even. Thus $\dfrac{a_n(\Gamma)}{a_n(F_{d-1})} \to b$ as $n \to \infty$.*

(v) *The number $m_n(\Gamma)$ of maximal subgroups of $\Gamma$ of index $n$ satisfies*

$$m_n(\Gamma) \sim bn \cdot (n!)^{d-2},$$

*where $b$ is as above. Thus $\dfrac{m_n(\Gamma)}{a_n(\Gamma)} \to 1$ as $n \to \infty$.*

(vi) *The probability that a random homomorphism from $\Gamma$ to a finite simple group $G$ is an epimorphism tends to 1 as $|G| \to \infty$.*

Note that for statements (i)–(iv) of Theorem 5 to hold, it suffices to take $N^*(l) = N(l) = 2 \cdot 10^{18} l^4$ as in Theorem 4.

Some special cases of Theorem 5, where $w_i$ are commutators or squares, were already obtained in the past.

For example, in the case of surface words

$$w = x_1^{-1} x_2^{-1} x_1 x_2 \cdots x_{2g-1}^{-1} x_{2g}^{-1} x_{2g-1} x_{2g}$$

for $g \geq 2$, part (i) of Theorem 5 was obtained in [AA, 4.4]. In characteristic zero it is also shown in [AA, VIII] that, for $g \geq 374$, the fibers of $w_{\underline{G}}$ have rational singularities. It would be interesting to know whether the statement

about rational singularities holds in the generality of part (i) of Theorem 5, if $N$ is sufficiently large in terms of $l$.

Part (ii) of Theorem 5 for surface words (including non-oriented ones $w = x_1^2 \cdots x_g^2$ where $g \geq 3$) was obtained in [LiSh3, 1.11].

Part (iii) of Theorem 5 for (oriented and non-oriented) surface words was obtained in [RBC] and [BC] for fields of characteristic zero (see also [Go]), and in [LiSh3, 1.8] for arbitrary fields.

Parts (iv) and (v) for surface groups were obtained in [MP].

For Fuchsian groups of genus $g \geq 2$ ($g \geq 3$ in the non-oriented case), a result similar to part (vi) of Theorem 5 was obtained in Theorem 1.6 of [LiSh3].

We conclude the introduction with a result of independent interest, which plays an important role in this paper and might be useful for other purposes.

**Theorem 6.** *Let $w \in F_d$ be a non-trivial word, and let $G$ be a finite simple group. Choose $g_1, \ldots, g_d \in G$ uniformly and independently. Then, for every $\epsilon > 0$, the probability that*

$$|\chi(w(g_1, \ldots, g_d))| \leq \chi(1)^\epsilon \text{ for all } \chi \in \mathrm{Irr}(G)$$

*tends to 1 as $|G| \to \infty$.*

This result generalizes Proposition 4.2 of [LS4] dealing with the case $w = x_1$, and Theorem 7.4 of [LS1] dealing with alternating groups.

The rest of the paper is organized as follows. In Section 2 we use methods from algebraic geometry to prove Theorem 2 and deduce Theorem 1 for Lie type groups of bounded rank. In Section 3 we discuss random walks and prove Theorem 3. In Section 4 we use character methods to provide an alternative proof of Theorem 1 in bounded rank (as well as some stronger results for $\mathrm{PSL}_2(q)$). In Section 5 we discuss classical groups of large rank, and apply new character bounds obtained for them, and other tools, to complete the proof of Theorem 1. Theorem 6 is also proved in Section 5, and plays a key role in proving Theorem 1. The proof of Theorem 4 is given in Section 6, and Section 7 is devoted to the proof of Theorem 5.

## 2. Geometric methods

In this section we prove Theorem 2 and deduce Theorem 1 for Lie type groups of bounded rank. At the end of the section, we prove a result which will be needed below for Theorem 5. Note that by an $\mathbb{F}_q$-*variety*, we mean a separated, geometrically integral scheme of finite type over $\mathbb{F}_q$.

**Proposition 2.1.** *Let $\underline{X}$ be an $\mathbb{F}_q$-variety, $\underline{Y}$ a disjoint union of $\mathbb{F}_q$-varieties $\underline{Y}_i$ of equal dimension, and $f\colon \underline{Y} \to \underline{X}$ a morphism defined over $\mathbb{F}_q$. If $\underline{Y}$ is irreducible, then*

$$\lim_{n\to\infty} \frac{|f(\underline{Y}(\mathbb{F}_{q^n}))|}{|\underline{X}(\mathbb{F}_{q^n})|} = 1 \tag{2.1}$$

*if and only if $f$ is dominant and its generic fiber is geometrically irreducible. In general,*

$$\lim_{n\to\infty} \|f_* U_{\underline{Y}(\mathbb{F}_{q^n})} - U_{\underline{X}(\mathbb{F}_{q^n})}\|_{L^1} = 0, \tag{2.2}$$

*implies each restriction $f_i$ of $f$ to a component $\underline{Y}_i$ of $\underline{Y}$ is dominant and the generic fiber of each $f_i$ is geometrically irreducible.*

*Proof.* Let us first assume $\underline{Y}$ is irreducible. By the Lang-Weil estimate, we may replace $\underline{X}$, $\underline{Y}$, and $f$ by $\underline{X}'$, $\underline{Y}'$, and $f' = f|_{\underline{Y}'}$ respectively, for any open subvariety $\underline{X}'$ of $\underline{X}$ and any open subvariety $\underline{Y}'$ of $f^{-1}(\underline{X}')$. Thus, we are justified in assuming $\underline{X}$ and $\underline{Y}$ are affine and non-singular, and $f$ is dominant. We denote their coordinate rings $A$ and $A_Y$ respectively. As $\underline{X}$ and $\underline{Y}$ are varieties, these are integral domains. Let $K$ and $K_Y$ denote the fraction fields of $A$ and $A_Y$ respectively, and let $L$ denote the separable closure $L$ of $K$ in $K_Y$. As $K_Y$ is a finitely generated field, L is a finite extension of $K$. Our claim is that $L = K$ if and only if (2.1) holds.

Choose $\alpha \in L \cap A_Y$ to be a primitive element of $L/K$; after multiplying by a suitable element of $A$, we may assume it is also integral over $A$. Let $B = A[\alpha] \subset A_Y$, so $f$ factors through the finite morphism $\operatorname{Spec} B \to \operatorname{Spec} A$. By [EGA IV$_4$, Théorème 17.6.1], $\operatorname{Spec} B \to \operatorname{Spec} A$ is étale in a neighborhood of the generic point of $\operatorname{Spec} B$, so replacing $A$ by $A[1/a]$ for $\operatorname{Spec} A[1/a]$ small enough and $B$ and $A_Y$ by $B[1/a]$ and $A_Y[1/a]$ respectively, we may assume $\operatorname{Spec} B \to \operatorname{Spec} A$ is finite étale. In particular $\operatorname{Spec} B$ is non-singular [EGA IV$_4$, Théorème 17.11.1]. Both $A$ and $B$ are therefore integrally closed, and $B$ is module-finite over $A$ and hence integral. Thus $B$ is the integral closure of $A$ in $L$.

Let $M$ denote any finite extension of $L$ which is Galois over $K$ and $C$ the integral closure of $B$ in $M$. Thus $C^{\mathrm{Gal}(M/K)}$ contains $A$ and has fraction field $K$. It is contained in $K$ and integral over $A$, therefore equal to $A$. Thus $\underline{X} = \operatorname{Spec} A$ is the quotient of $\underline{Z} = \operatorname{Spec} C$ by $\mathrm{Gal}(M/K)$. Likewise, $B = C^{\mathrm{Gal}(M/L)}$, so $\underline{Z} \to \underline{X}$ factors through $\underline{Y} = \operatorname{Spec} B$. Let $m$ denote the common dimension of $\underline{X}$, $\underline{Y}$, and $\underline{Z}$. By the Lang-Weil estimate, $|\underline{X}(\mathbb{F}_{q^n})|$, $|\underline{Y}(\mathbb{F}_{q^n})|$, and $|\underline{Z}(\mathbb{F}_{q^n})|$ are all $(1 + O(q^{-n/2}))q^{mn}$.

Applying the Chebotarev density theorem for $\underline{Z} \to \underline{X}$ [Se], we see that in the limit $n \to \infty$, a positive proportion of points in $\underline{X}(\mathbb{F}_{q^n})$ split completely in $\underline{Z}$ and therefore in $\underline{Y}$. It follows that (2.1) implies $L = K$.

Conversely the condition $K = L$ is equivalent to the generic geometric irreducibility of $f$. By [EGA IV$_3$, Proposition 9.5.5, Théorème 9.7.7], we may assume without loss of generality that all fibers of $f$ are geometrically irreducible and of equal dimension. It is well known that the Lang-Weil theorem holds uniformly for families of varieties of the same dimension (see, e.g., [LS2, Lemma 2.2]), and this implies (2.1) and even the stronger (2.2).

Finally, we consider the case that $\underline{Y}$ has irreducible components $\underline{Y}_1, \dots, \underline{Y}_r$. We note first that Lang-Weil implies that as $n \to \infty$, the probability of a random element of $\underline{Y}(\mathbb{F}_{q^n})$ lying in any fixed $\underline{Y}_i(\mathbb{F}_{q^n})$ approaches $1/r$, so

(2.2) implies that the restriction of $f$ to each $\underline{Y}_i$ is dominant. Proceeding as before, we may assume that $\underline{X} = \operatorname{Spec} A$ is affine, each $\underline{Y}_i$ is affine and geometrically connected over $\operatorname{Spec} B_i$, $\operatorname{Spec} B_i$ is finite étale over $\underline{X}$, the fraction field $L_i$ of $B_i$ is a finite separable extension of the fraction field $K$ of $A$, and $M_i$ is a finite Galois extension of $K$ containing $L_i$. Let $e = \dim \underline{Y}_i - \dim \underline{X}$, the relative dimension of $\underline{Y}_i$ over $\underline{X}$, which is the same for all $i$ since the $\underline{Y}_i$ have the same dimension and the morphisms to $\underline{X}$ are all dominant. By the uniform version of the Lang-Weil theorem, for each $\mathbb{F}_{q^n}$-point of $\operatorname{Spec} B_i$, there are $(1 + o(1))q^{ne}$ elements of $\underline{Y}_i(\mathbb{F}_{q^n})$ lying over it.

Applying the Chebotarev density theorem for $M_1 \cdots M_r/K$, in the limit as $n \to \infty$, a positive proportion of points $x \in \underline{X}(\mathbb{F}_{q^n})$ split completely in each $L_i$, which means that there are $[L_i : K]$ $\mathbb{F}_{q^n}$-points of $\operatorname{Spec} B_i$ lying over $x$, therefore $(1 + o(1))[L_i : K]q^{ne}$ points of $\underline{Y}_i(\mathbb{F}_{q^n})$ lying over $x$, and, finally, $(1 + o(1))([L_1 : K] + \cdots + [L_r : K])q^{en}$ points of $\underline{Y}(\mathbb{F}_{q^n})$ lying over $x$. If any of $L_1, \ldots, L_r$ is of degree $\geq 2$ over $K$, then this sum of degrees strictly exceeds $r$. On the other hand, Lang-Weil implies

$$\lim_{n\to\infty} \frac{|\underline{Y}(\mathbb{F}_{q^n})|}{q^{en}|\underline{X}(\mathbb{F}_{q^n})|} = 1.$$

Thus, $\|f_* U_{\underline{Y}(\mathbb{F}_{q^n})} - U_{\underline{X}(\mathbb{F}_{q^n})}\|_{L^1}$ does not approach 0.

□

We now embark on the proof of Theorem 2.

*Proof.* If $G$ is any finite group and $H$ is contained in its center, then for all $g \in G$,

$$[G : H]^d |w_{G/H}^{-1}(gH)| = \sum_{h\in H} |w_G^{-1}(gh)|.$$

Defining $f\colon G^d \times H \to G$ by $f(g_1, \ldots, g_d, h) = w_G(g_1, \ldots, g_d)h$, we have

$$\|f_* U_{G^d\times H} - U_G\|_{L^1} = \|p_{w,G/H} - U_{G/H}\|_{L^1}. \tag{2.3}$$

On the other hand, the triangle inequality implies

$$\|p_{w,G/H} - U_{G/H}\|_{L^1} \leq \|p_{w,G} - U_G\|_{L^1}. \tag{2.4}$$

We specialize to the case that $H$ is the center of $G$, while $G$ is of the form $\underline{G}(\overline{\mathbb{F}}_p)^F$, where $F$ is a generalized Frobenius map and $\underline{G}$ is a simply connected, split, almost simple algebraic group of rank $\leq r$ over $\mathbb{F}_p$.

To prove (i) implies (ii), given $p$ and $\underline{G}$, we choose $q$ so that the center $Z$ of $\underline{G}(\overline{\mathbb{F}}_p)$ is contained in $\underline{G}(\mathbb{F}_q)$. Applying Proposition 2.1 to the morphism $\underline{G}^d \times Z \to \underline{G}$ given by $(g_1, \ldots, g_d, z) \mapsto w(g_1, \ldots, g_d)z$, condition (i) in the form given by (2.3) implies that each component of $\underline{G}^d \times Z$ maps to $\underline{G}$ with geometrically irreducible generic fiber. In particular this is true for the identity component, which is $\underline{G}^d$. A second application of Proposition 2.1 gives (ii).

To prove (ii) implies (iii), we first note that generic geometric irreducibility is stable under base change of $\Bbbk$, so we could assume without loss of

generality that $\Bbbk$ is algebraically closed and therefore that $\underline{G}$ is split. Since every split group is obtained by base change from a split group over a prime field, we assume instead that $\underline{G}$ is split and that $\Bbbk$ is either $\mathbb{Q}$ or $\mathbb{F}_p$ for some $p$. Let $\mathcal{G}$ denote a split semisimple group scheme over $\mathbb{Z}$ with the same root system as $\underline{G}$, and we denote by $w_{\mathcal{G}}$ the word morphism $\mathcal{G}^d \to \mathcal{G}$ of schemes of finite type over Spec $\mathbb{Z}$. By [EGA IV$_3$, Théorème 9.7.7], the set of points of $\mathcal{G}$ over which $w_{\mathcal{G}}$ is geometrically irreducible is constructible and contains the generic point. It therefore contains a non-empty open set $S$. By Chevalley's constructibility theorem [EGA IV$_1$, Corollaire 1.8.5], its image in Spec $\mathbb{Z}$ is constructible and therefore contains all but finitely many closed points. Thus $S$ contains the generic point of all but finitely many fibers of $\mathcal{G} \to$ Spec $\mathbb{Z}$, so it suffices to prove the geometric irreducibility in the case that $k = \mathbb{F}_p$ and $\underline{G}$ is split. This case follows from Proposition 2.1.

It remains to show that (iii) implies (i); by (2.4), it suffices to prove

$$\lim_{|G|\to\infty} \|p_{w,G} - U_G\|_{L^1} = 0,$$

where $G$ ranges over groups of the form $\underline{G}_0(\overline{\mathbb{F}}_p)^F$, where $F$ is a generalized Frobenius map and $\underline{G}_0$ is a simply connected, split, almost simple algebraic group of rank $\leq r$ over $\mathbb{F}_p$. We fix any root system $\Phi$ of rank $\leq r$ and prove the limit is zero as $G$ ranges over groups of this form with root system $\Phi$. In the case that $F$ is a standard Frobenius map, $\underline{G}_0(\overline{\mathbb{F}}_p)^F = \underline{G}(\mathbb{F}_q)$ for some simply connected $\underline{G}$ of rank $\leq r$ and some $q$. Thus, (i) follows from Proposition 2.1. In the case of Suzuki or Ree groups, it follows from the following lemma. □

**Lemma 2.2.** *Let $\underline{G}$ be a split simple algebraic group over $\mathbb{F}_p$ and $f\colon \underline{G}^d \to \underline{G}$ a morphism of schemes. There exists a constant $C$ such that if $\overline{x} \in \underline{G}(\overline{\mathbb{F}}_p)$ is a geometric point of $\underline{G}$ such that $w^{-1}(\overline{x})$ is irreducible of dimension $k$, and $F\colon \underline{G}_{\overline{\mathbb{F}}_p} \to \underline{G}_{\overline{\mathbb{F}}_p}$ an endomorphism which preserves $w^{-1}(\overline{x})$ and such that $F^2$ is a standard $p$-Frobenius endomorphism, and $s$ is a sufficiently large integer, then*

$$\left|\left|\, w^{-1}(\overline{x})(\overline{\mathbb{F}}_p)^{F^{2s+1}} \,\right| - p^{(2s+1)k/2}\right| \leq Cp^{(2s+1)k/2-1/4}.$$

*Proof.* We fix $\ell \neq p$. By the finiteness and proper base change theorems for étale cohomology over a field we see that for all $i$, $\dim H^i_c(w^{-1}(\overline{x}), \mathbb{Q}_\ell)$ is bounded as $\overline{x}$ varies.

We would like to apply the Lefschetz trace formula to count the $F$-fixed points of $w^{-1}(\overline{x})$. We use Fujiwara's theorem (formerly Deligne's conjecture) [Fu]. If $F$ is an endomorphism of $\underline{G}$ whose square is the $p$-Frobenius, then the naive Lefschetz trace formula applies to all sufficiently high odd powers of $F$:

$$|w^{-1}(\overline{x}) \cap (\underline{G}(\overline{\mathbb{F}}_p)^{F^{2s+1}})^d| = \sum_{i=0}^{2k} (-1)^i \mathrm{tr}(F^{2s+1} | H^i_c(w^{-1}(\overline{x}), \mathbb{Q}_\ell)). \tag{2.5}$$

Since $F^2$ is a standard $p$-Frobenius map, by [De, 3.3.1] the eigenvalues of $F$ on $H^i_c(w^{-1}(\overline{x}), \mathbb{Q}_\ell)$ have absolute value at most $p^{i/4} \leq \sqrt{p}^{\dim w^{-1}(\overline{x})}$. As $w^{-1}(\overline{x})$ is a variety, its top cohomology group, $H^{2k}(w^{-1}(\overline{X}), \mathbb{Q}_\ell)$, is 1-dimensional, and $F^2$ acts with eigenvalue $p^k$. Thus $F$ acts on the top cohomology with eigenvalue $\pm p^{k/2}$, and as left hand side of (2.5) is non-negative, for $f$ sufficiently large, the eigenvalue is $p^{k/2}$, and the number of $F^{2s+1}$-fixed points differs from $p^{(2s+1)k/2}$ by $O(q^{(2s+1)k/2-1/4})$. The lemma follows. □

An immediate consequence of Theorem 2 is the following.

**Corollary 2.3.** *If the image of $w \in F_d$ in the abelianization $\mathbb{Z}^d = F_d/[F_d, F_d]$ is primitive, then $w$ is almost uniform in bounded rank.*

*Proof.* It suffices to prove that $w(G) = G$ for all groups $G$. Indeed, as shown in [Seg, 3.1.1], if the image of $w$ in $\mathbb{Z}^d$ is a primitive $d$-tuple $(a_1, \ldots, a_d)$, we fix $b_1, \ldots, b_d \in \mathbb{Z}$ such that $\sum_i a_i b_i = 1$. Then $w(g^{b_1}, \ldots, g^{b_d}) = g$. □

We can now deduce Theorem 1 for Lie type groups of bounded rank. It follows immediately from Theorem 2 together with the following lemma.

**Lemma 2.4.** *Let $w = w_1 w_2 \in F_d$ where $w_1, w_2 \neq 1$ are disjoint words, and let $\underline{G}$ be a semisimple simply connected algebraic group. Then $w\colon \underline{G}^d \to \underline{G}$ has geometrically irreducible generic fiber.*

*Proof.* It suffices to prove this in the case that $\underline{G}$ is simple modulo its center. In the case, $G = \underline{G}(\mathbb{F}_q)$ is the universal central extension of a finite simple group if $q$ is sufficiently large. By Borel's theorem, $w_1$ and $w_2$ define dominant morphisms, so if $q$ is sufficiently large, there exist regular semisimple conjugacy classes $C_1$ and $C_2$ of $G$ lying in the image of $w_1$ and $w_2$ respectively. By [GT, Lemma 5.1], the image of $w$ contains all non-central semisimple elements of $G$ when $q$ is large, so condition (ii) of Theorem 2 is satisfied. Hence condition (iii) follows, as required. □

We conclude with a result which will be needed in §7.

**Proposition 2.5.** *Let $w \in F_d$ be a word such that, as $G$ ranges over all finite simple groups of Lie type, we have*

$$\lim_{|G| \to \infty} \|p_{w,G} - U_G\|_{L^\infty} = 0. \tag{2.6}$$

*Then for every field $\Bbbk$ and every semisimple algebraic group $\underline{G}$ over $\Bbbk$, the word map $w_{\underline{G}}\colon \underline{G}^d \to \underline{G}$ associated to $w$ is a flat morphism.*

*Proof.* As flatness is not affected by faithfully flat base change [EGA $\mathrm{IV}_2$, Cor. 2.2.11 (iii)], we can proceed as in Proposition 2.1, observing that it suffices to consider the case that $\Bbbk$ is prime and $\underline{G}$ is split. Suppose we can prove flatness for $\Bbbk = \mathbb{F}_p$ for all $p$ and therefore for $\Bbbk$ any finite field. Let $\mathcal{G}$ denote the split semisimple group scheme over Spec $\mathbb{Z}$ with the same root datum as $\underline{G}$, and let $w_{\mathcal{G}}$ denote the word map $\mathcal{G}^d \to \mathcal{G}$. Every non-empty closed set of $\mathcal{G}^d$ contains a closed point. By [EGA $\mathrm{IV}_3$, Théorème 11.1.1],

the flat locus of $w_{\mathcal{G}}$ is open, so if it contains all closed points of $\mathcal{G}^d$, it must be all of $\mathcal{G}^d$. Thus, we assume that $\Bbbk = \mathbb{F}_p$.

By "miracle flatness" [EGA IV$_2$, Proposition 6.1.5], it suffices to prove that every fiber of $w_{\underline{G}}$ has dimension $(d-1)\dim \underline{G}$, the inequality

$$\dim w_{\underline{G}}^{-1}(g) \geq (d-1)\dim \underline{G}$$

being automatic [EGA IV$_2$, (5.5.2.1)]. If there exists a point on $\underline{G}$ over which the inequality is strict, then by Chevalley's semicontinuity theorem [EGA IV$_3$, Théorème 13.1.3], there exists a closed point $x$ with this property. If $\underline{G}^{\mathrm{ad}}$ denotes the adjoint quotient of $\underline{G}$, then the image of $x$ in $\underline{G}^{\mathrm{ad}}$ has the same property for the word map $w_{\underline{G}^{\mathrm{ad}}}$. Thus, we may assume $\underline{G}$ is adjoint, and since it is also split, it suffices to consider the case that it is absolutely simple. The closed point $x$ corresponds to a $\mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p)$-orbit of points of $\underline{G}(\mathbb{F}_q)$ for some $q$, and we let $x_0$ denote a point in this orbit. Thus, the fiber $F_{x_0}$ of $w_{\underline{G}_{\mathbb{F}_q}}$ over $x_0$ has dimension at least $(d-1)\dim \underline{G} + 1$.

Replacing $\mathbb{F}_q$ by a finite extension field if necessary, we may assume that the fiber $F_{x_0} \subset \underline{G}^d_{\mathbb{F}_q}$ has the property that all of its irreducible components are geometrically irreducible. We may further assume the same for the inverse image $F^{\mathrm{sc}}_{x_0}$ of $F_{x_0}$ in $(\underline{G}^{\mathrm{sc}}_{\mathbb{F}_q})^d$:

$$\begin{array}{ccc}
F^{\mathrm{sc}}_{x_0} & \longrightarrow & (\underline{G}^{\mathrm{sc}})^d \\
\downarrow & & \downarrow \\
F_{x_0} & \longrightarrow & \underline{G}^d \\
\downarrow & & \downarrow{\scriptstyle w_{\underline{G}}} \\
\mathrm{Spec}\ \mathbb{F}_q & \xrightarrow{x_0} & \underline{G}
\end{array}$$

Let $G$ denote the derived group of $\underline{G}(\mathbb{F}_q)$. The image of $F^{\mathrm{sc}}_{x_0}(\mathbb{F}_q)$ in $F_{x_0}(\mathbb{F}_q) \subset \underline{G}(\mathbb{F}_q)^d$ lies in

$$\mathrm{im}(\underline{G}^{\mathrm{sc}}(\mathbb{F}_q)^d \to \underline{G}(\mathbb{F}_q)^d) = G^d,$$

and by the Lang-Weil estimate, if $q$ is sufficiently large,

$$|F^{\mathrm{sc}}_{x_0}(\mathbb{F}_q)| > \frac{q^{1+(d-1)\dim \underline{G}}}{2},$$

so

$$|\mathrm{im}(F^{\mathrm{sc}}_{x_0}(\mathbb{F}_q) \to F_{x_0}(\mathbb{F}_q))| \geq \frac{q^{1+(d-1)\dim \underline{G}}}{2|\mathbf{Z}(\underline{G}^{\mathrm{sc}}(\mathbb{F}_q))|^d}.$$

The denominator does not depend on $q$, so when $q$ is sufficiently large, this is inconsistent with (2.6). □

## 3. Random words

This section is devoted to the proof of Theorem 3.

For $n \geq 0$ and $d \geq 1$ let $\mathsf{X}_{n,d}$ denote the random variable associated with the standard random walk with $n$ steps in $\mathbb{Z}^d$. We also set $\mathsf{X}_n = \mathsf{X}_{n,2}$. Thus $\mathsf{X}_n$ is the probability distribution in $\mathbb{Z}^2$ corresponding to a random walk of length $n$ in which each step in the set $\{(\pm 1, 0), (0, \pm 1)\}$ has probability $1/4$.

**Lemma 3.1.** *Let $(a, b), (a', b') \in \mathbb{N}^2$, with $a + b \equiv a' + b' \pmod 2$. Then*

$$\mathbf{P}[\mathsf{X}_n = (a, b)] \leq \mathbf{P}[\mathsf{X}_n = (a', b')]$$

*for all $n \in \mathbb{N}$ if either of the following conditions holds:*

(3.1.1) $a - a', b - b' \in \mathbb{N}$.

(3.1.2) $a + b = a' + b'$, *and* $|a - b| \geq |a' - b'|$

*Proof.* We proceed by induction on $n$, the claim being trivial for $n = 0$. For any $a, b \in \mathbb{Z}$, we abbreviate $\mathbf{P}[\mathsf{X}_n = (a, b)]$ by $(a, b)_n$. Thus,

$$\begin{aligned}(a, b)_{n+1} &= \frac{(a-1, b)_n + (a+1, b)_n + (a, b-1)_n + (a, b+1)_n}{4} \\ &= \frac{(|a-1|, b)_n + (a+1, b)_n + (a, |b-1|)_n + (a, b+1)_n}{4}.\end{aligned}$$

We write $(a', b') \preceq_* (a, b)$ if $a, b, a', b' \in \mathbb{N}$ and (3.1.$*$) holds for $* \in \{1, 2\}$. If $(a', b') \preceq_1 (a, b)$, then

$$(a' + 1, b') \preceq_1 (a + 1, b),\ (a', b' + 1) \preceq_1 (a, b + 1).$$

Moreover,

$$(|a' - 1|, b') \preceq_1 (|a - 1|, b)$$

unless $a' = 0$ and $a = 1$. In this case, the parity condition implies $b \geq b' + 1$, so

$$(|a' - 1|, b') = (1, b') \preceq_2 (0, b' + 1) \preceq_1 (0, b) = (|a - 1|, b).$$

Likewise,

$$(a', |b' - 1|) \preceq_1 (a, |b - 1|)$$

unless $b' = 0$ and $b = 1$, in which case

$$(a', |b' - 1|) = (a', 1) \preceq_2 (a' + 1, 0) \preceq_1 (a, 0) = (a, |b - 1|),$$

so (3.1.1) follows by induction.

Suppose, on the other hand, that $(a', b') \preceq_2 (a, b)$. It suffices to consider the case that $a > a' \geq b' > b$, so $a$ and $a'$ are positive. It follows that

$$\begin{aligned}(|a' - 1|, b') = (a' - 1, b') &\preceq_2 (a - 1, b) = (|a - 1|, b), \\ (a' + 1, b') &\preceq_2 (a + 1, b), \\ (a', b' + 1) &\preceq_2 (a, b + 1).\end{aligned}$$

If $b > 0$, then

$$(a', |b' - 1|) = (a', b' - 1) \preceq_2 (a, b - 1) = (a, |b - 1|),$$

and we are done by induction. If $b = 0$, then

$$(a', |b'-1| = (a', b'-1) \preceq_1 (a'+2, b'-1) \preceq_2 (a, 1) = (a, |b-1|),$$

and we are done by induction. □

**Proposition 3.2.** *If $p > 2$ is prime and $n > 0$, then*

$$\mathbf{P}[\mathsf{X}_n \in p\mathbb{Z}^2 \setminus \{0,0\}] < \frac{4}{(p+1)^2}\ .$$

*Proof.* Let $\mathcal{Z} = \mathbb{Z}_{>0} \times \mathbb{N}$. If $R$ is the group of automorphisms of $\mathbb{Z}^2$ generated by rotation by $\pi/2$, then $\mathbb{Z}^2 \setminus \{0,0\}$ is the disjoint union of $\rho(\mathcal{Z})$ for all $\rho \in R$. Let $\mathcal{Z}_p = p\mathbb{Z}^2 \cap \mathcal{Z}$. As $|R| = 4$, it suffices to prove that

$$\mathbf{P}[\mathsf{X}_n \in \mathcal{Z}_p] < \frac{1}{(p+1)^2}. \tag{3.1}$$

For $(a,b) \in \mathcal{Z}_p$, we define subsets $\mathcal{Y}(a,b)$ of $\mathcal{Z}$ as follows. For $b = 0$,

$$\mathcal{Y}(a,0) = \mathcal{Z} \cap \bigcup_{\rho \in R} \rho\{(x,y) \in \mathbb{Z}^2 : |x-a| + |y| \in 2\mathbb{Z} \cap [0,p],\ x + |y| \le a\},$$

and for $b > 0$,

$$\mathcal{Y}(a,b) = \{(x,y) \in \mathcal{Z} : |x-a| + |y-b| \in 2\mathbb{Z} \cap [0,p],\ |x| \le a,\ |y| \le b\}.$$

By Lemma 3.1, $(x,y) \in \mathcal{Y}(a,b)$ implies $(x,y)_n \ge (a,b)_n$ for all $n$.

We claim the sets $\{\mathcal{Y}(a,b) \mid (a,b) \in \mathcal{Z}_p\}$ are pairwise disjoint. Indeed, for $\mathcal{Y}(a_1,b_1) \cap \mathcal{Y}(a_2,b_2)$ to be non-empty for distinct elements $(a_1,b_1)$ and $(a_2,b_2)$ of $\mathcal{Z}_p$, it is necessary that $a_1 + b_1 \equiv a_2 + b_2 \pmod 2$, and this together with the fact that $a_1, b_1, a_2, b_2 \in p\mathbb{Z}$ implies that the the $L^1$ distance between any point in the $R$-orbit of $(a_1,b_1)$ and any point in the $R$-orbit of $(a_2,b_2)$ is at least $2p$. On the other hand, all the elements of $\mathcal{Y}(a,b)$ are within distance $p-1$ in the $L^1$ norm of some element of the $R$-orbit of $(a,b)$.

Whether $b$ is 0 or not, $|\mathcal{Y}(a,b)| = (p+1)^2/4$. Thus,

$$(a,b)_n \le \frac{4}{(p+1)^2}\mathbf{P}[\mathsf{X}_n \in \mathcal{Y}(a,b)].$$

By symmetry, $\mathbf{P}[\mathsf{X}_n \in \mathcal{Z}] \le 1/4$, so

$$\begin{aligned}\mathbf{P}[\mathsf{X}_n \in \mathcal{Z}_p] = \sum_{(a,b) \in \mathcal{Z}_p} (a,b)_n &\le \frac{4}{(p+1)^2}\mathbf{P}[\mathsf{X}_n \in \bigcup_{(a,b) \in \mathcal{Z}_p} \mathcal{Y}(a,b)] \\ &< \frac{4}{(p+1)^2}\mathbf{P}[\mathsf{X}_n \in \mathcal{Z}] = \frac{\mathbf{P}[\mathsf{X}_n \ne (0,0)]}{(p+1)^2},\end{aligned}$$

implying (3.1). □

*Proof of Theorem 3.* Part (ii) of the theorem follows from part (i) and Theorem 2, so it suffices to prove the two assertions in part (i).

Let $\mathsf{W}_{n,d} = w_1 \cdots w_n$, where the $w_i$ are chosen independently from the standard generating set $\{x_1^{\pm 1}, \dots, x_d^{\pm 1}\}$, with all elements equally likely, and

$n > 0$. Let $\phi\colon F_d \to \mathbb{Z}^d$ be the abelianization map. Thus $\phi(\mathsf{W}_{n,d})$ is exactly $\mathsf{X}_{n,d}$.

We first assume $d = 2$, so $\mathsf{X}_{n,d}$ is just $\mathsf{X}_n$, and the probability that $\phi(\mathsf{W}_{n,d})$ is primitive is the probability $P_n$ that an $n$ step random walk in $\mathbb{Z}^2$ gives a primitive element.

By [Seg, 3.1.1], if $\phi(w)$ is primitive, then $w$ is surjective on all groups, and by Theorem 2, it is almost uniform in bounded rank. Thus, to prove the theorem for $d = 2$, it suffices to prove that $P_n > 1/3$ for all $n > 0$. Now, if $a_{n,m}$ denotes the probability that $\mathsf{X}_n \neq (0,0)$ and the g.c.d. of $m$ and the two coordinates of $\mathsf{X}_n$ is $> 1$, then

$$P_n \geq 1 - a_{n,6} - \sum_p a_{n,p} - (0,0)_n,$$

where $p$ ranges over primes $\geq 5$, so

$$\inf_{n\geq 1} P_n \geq 1 - \sup_n a_{n,6} - \sum_p \sup_n a_{n,p} - (0,0)_n.$$

To estimate $\sup_n a_{n,6}$, we fix a cutoff $N$ and calculate $a_{n,6}$ for $n \leq N$ (using interval arithmetic to get a rigorous upper bound). To bound $a_{n,6}$ for $n \geq N$, we consider the $n$ step random walk on $(\mathbb{Z}/6\mathbb{Z})^2$ in which the steps $(\pm 1, 0)$, $(0, \pm 1)$ each have probability $1/4$. An upper bound for the probability of any state occurring in $n \geq N$ steps is given by the maximum over all states of the probability of occurrence in $N$ steps. Since the image of $(2\mathbb{Z})^2 \cup (3\mathbb{Z})^2$ in $(\mathbb{Z}/6\mathbb{Z})^2$ has 12 elements of which 10 have even coordinate sum and 2 have odd coordinate sum, the probability of landing in $(2\mathbb{Z})^2 \cup (3\mathbb{Z})^2$ after $n \geq N$ steps is at most 10 times the maximum probability at time $N$ of any state in the (mod 6) Markov chain.

Likewise, for any given $p \geq 5$, to estimate $\sup_n a_{n,p}$, we can fix a cutoff $N$ and proceed as before. In practice, to obtain a good bound, $N$ should be chosen of order $p^2$. We use this method for small $p$, while for large $p$, we use the estimate $\sup_n a_{n,p} < 4/(p+1)^2$ given by Proposition 3.2. For $n \geq N$, $(0,0)_n$ is bounded above by the maximum of $(a,b)_N$ over pairs $(a,b) \in \mathbb{N}$. Implementing these calculations by computer using $N = 1000$,

$$(0,0)_n \leq .0006,\ a_{n,6} < .5556,\ a_{n,5} < .0401,\ a_{n,7} < .0205,\ \cdots\ a_{n,59} < .0007$$

for all $n \geq 1000$, so

$$\begin{aligned}\inf_{n\geq 1000} P_n &> 1 - .5556 - .0401 - .0205 - .0083 - .\cdots - .0007 \\ &\quad - \sum_{p>60} \frac{4}{(p-1)^2} - .0006 \\ &> .3535 - \sum_{60<p<10003} \frac{4}{(p-1)^2} - \int_{10000}^{\infty} \frac{2\,dx}{x^2} \\ &= .3535 - .0132 - .0005 - .0006 > \frac{1}{3}.\end{aligned}$$

This proves the the theorem for $n \geq 1000$; and for $1 \leq n < 1000$, machine computation shows that the probability that the coordinates of $\mathsf{X}_n$ are relatively prime is greater than .4.

We now consider the general case $d \geq 2$. Recall that $\mathsf{X}_{n,d}$ denotes the random variable associated with the standard random walk with $n$ steps in $\mathbb{Z}^d$. It suffices to prove that the probability that $\mathsf{X}_{n,d}$ is primitive always exceeds $1/3$ and tends to 1 as $d \to \infty$. For $1 \leq i < j \leq d$, let $\pi_{i,j} \colon \mathbb{Z}^d \to \mathbb{Z}^2$ denote the projection map onto the $i$th and $j$th coordinates.

For $\mathsf{X}_{n,d}$ to be primitive, it suffices that $\pi_{i,j}(\mathsf{X}_{n,d})$ is primitive for some $i, j$. Let $n_{i,j}$ denote the number of terms in the sequence $w_1, \ldots, w_n$ which belong to $\{x_i^{\pm 1}, x_j^{\pm 1}\}$; conditioning on $n_{i,j}$, $\pi_{i,j}(\phi(\mathsf{W}_{n,d}))$ has the same probability distribution as $\mathsf{X}_{n_{i,j}}$. Since there is always at least one pair $(i, j)$ for which $n_{i,j} > 0$ it follows that $\mathsf{X}_{n,d}$ is primitive with probability greater than $1/3$.

If $n$ is fixed and $d \to \infty$, the probability approaches 1 that $\phi(w_1), \ldots, \phi(w_n)$ are linearly independent, which implies that $\mathsf{X}_{n,d}$ is primitive. On the other hand, for any $k > 0$, as $n$ and $d$ both grow without bound

$$\mathbf{P}[\mathrm{Span}(\phi(w_1), \ldots, \phi(w_n)) \geq k]$$

goes to 1. Assuming the span has dimension $\geq k$ and $d \geq 2k$, there exist $k$ disjoint pairs of coordinates such that each projection of the random walk associated to one of the $k$ pairs $(i, j)$ satisfies $n_{i,j} > 0$, and therefore, conditioning on the choice of the $k$ pairs, the probability that each of the $k$ projections of $\mathsf{X}_{n,d}$ is imprimitive is less than $(2/3)^k$. Thus, $\mathsf{X}_{n,d}$ is primitive with probability greater than $1 - (2/3)^k$. Taking $k \to \infty$, this implies the second assertion in part (i) of the theorem and completes the proof. □

*Remark* 3.3. For any odd number $m$, the Markov chain on $(\mathbb{Z}/m\mathbb{Z})^2$ given by our $\pmod m$ random walk is irreducible and aperiodic, since the set of possible steps does not lie in a single coset of any proper subgroup of $(\mathbb{Z}/m\mathbb{Z})^2$. Therefore, it converges to the unique invariant distribution, which is the uniform distribution. It follows that

$$\lim_{n\to\infty} a_{n,m} = 1 - \prod_{p|m}(1 - p^{-2}).$$

For $m$ even, the situation is slightly more complicated, since for $n$ odd, $a_{n,2} = 0$ and for $n > 0$ even, $a_{n,2} = 1/2$. Thus,

$$\lim_{n\to\infty} a_{2n,m} = 1 - (2/3)\prod_{p|m}(1 - p^{-2})$$

while

$$\lim_{n\to\infty} a_{2n+1,m} = 1 - (4/3)\prod_{p|m}(1 - p^{-2}).$$

From this together with Proposition 3.2 it is easy to deduce that

$$\limsup_{n\to\infty} P_n = \frac{4}{\pi^2},$$

and it follows, without any necessity for computer calculation, that there exists a positive lower bound for $P_n$ for all $n > 0$. We do not know whether $P_n > 4/\pi^2$ for all $n > 0$.

## 4. Character methods

In this section we provide an alternative proof of Theorem 1 for Lie type groups of bounded rank using character theory. We also prove a stronger $L^2$ result in the case $G = \mathrm{PSL}_2(q)$ by studying the non-commutative Fourier expansion of the probability distribution $p_{w,G}$.

**Lemma 4.1.** *Let $w \in F_d$ be a non-trivial word. Let $G(q)$ be a finite simple group of Lie type of rank $r$ over a field with $q$ elements. Let $S$ be the set of regular semisimple elements of $G(q)$. Then we have*

$$p_{w,G(q)}(S) \geq 1 - cq^{-1},$$

*where $c > 0$ depends on $w$ and $r$ but not on $q$.*

*Proof.* At the level of the algebraic group $G$, the regular semisimple elements form an open dense subset, and its complement is a proper subvariety. By Borel's theorem [Bor] the inverse image of this subvariety under the word map induced by $w$ on $G^d$ is a proper subvariety of $G^d$. By the Lang-Weil estimate,

$$p_{w,G(q)}(G(q) \setminus S) \leq cq^{-1},$$

yielding the desired conclusion. □

Next, let $w_1, w_2$ be non-trivial disjoint words, and let $G = G(q)$ be as above. Let $C_1, C_2$ be conjugacy classes of regular semisimple elements of $G$, and let $g$ be a regular semisimple element of $G$. For $i = 1, 2$ choose $x_i \in C_i$ uniformly and independently. It is well known that the probability $p(C_1, C_2, g)$ that $x_1 x_2 = g$ satisfies

$$p(C_1, C_2, g) = |G|^{-1} \sum_{\chi \in \mathrm{Irr}(G)} \frac{\chi(C_1)\chi(C_2)\chi(g^{-1})}{\chi(1)}. \tag{4.1}$$

It is known that there exists a constant $b$ depending only on $r$ such that $|\chi(s)| \leq b$ for all regular semisimple elements $s \in G$ (see for instance [Sh1, 4.4]). This yields

$$|p(C_1, C_2, g) - |G|^{-1}| \leq |G|^{-1} \sum_{1 \neq \chi \in \mathrm{Irr}(G)} b^3/\chi(1) = b^3 |G|^{-1}(\zeta_G(1) - 1),$$

where $\zeta_G(s) = \sum_{\chi \in \mathrm{Irr}(G)} \chi(1)^{-s}$ is the Witten zeta function of $G$. Suppose $G \neq \mathrm{PSL}_2(q)$. Then we have $\zeta_G(1) \to 1$ as $|G| \to \infty$ by [LiSh3, 1.1]. This yields

$$p(C_1, C_2, g) = |G|^{-1}(1 + o(1)),$$

for all $C_1, C_2, g$ as above. Summing up over $C_1, C_2$ and applying Lemma 4.1 we see that for every $\epsilon > 0$ and large enough $G$, for at least $(1-\epsilon)|G|$ elements $g \in G$ we have $p_{w_1w_2,G}(\{g\}) \geq (1-\epsilon)|G|^{-1}$. This easily yields

$$\|p_{w_1w_2,G} - U_G\|_{L^1} \to 0$$

as $|G| \to \infty$. This proves Theorem 1 for bounded rank Lie-type groups $G \neq \mathrm{PSL}_2(q)$.

In the case $G = \mathrm{PSL}_2(q)$ we obtain a somewhat stronger result, see Corollary 4.3 below. We need some preparations.

Let $G$ be a finite group, $w \in F_d$ a word, and $p_{w,G}$ its induced probability distribution on $G$. We express the class function $P_{w,G}$ as a linear combination of irreducible characters

$$P_{w,G} = |G|^{-1} \sum_{\chi \in \mathrm{Irr}(G)} a_{w,\chi}\chi.$$

It is well known (see for instance [Sh1, §8]) that if $w_1, w_2$ are disjoint words, then we have

$$a_{w_1w_2,\chi} = a_{w_1,\chi}a_{w_2,\chi}/\chi(1)$$

for all $\chi \in \mathrm{Irr}(G)$. Using an inverse Fourier transform one obtains

$$a_{w,\chi} = |G|^{-d} \sum_{g_1,\ldots,g_d \in G} \chi(w(g_1,\ldots,g_d)^{-1}) = \sum_{g \in G} P_{w,G}(g)\chi(g^{-1}).$$

The following result, which may be of some independent interest, will be useful in this section.

**Proposition 4.2.** *For every word $1 \neq w \in F_d$ there exists a positive number $c(w)$ such that for every group $G = \mathrm{PSL}_2(q)$ and every character $\chi \in \mathrm{Irr}(G)$ we have $|a_{w,\chi}| \leq c(w)$.*

*Proof.* Inspecting the well known character table of $G$, we see that, if $1 \neq g \in G$ and $\chi \in \mathrm{Irr}(G)$, then $|\chi(g)| \leq 2$ except if $g$ is unipotent. In this case we have $|\chi(g)| > 2$ for at most two irreducible characters $\chi$, and in any case, $|\chi(g)| \leq q^{1/2}$. Let $S \subset G$ be the set of (regular) semisimple elements and let $U$ be the set of (regular) unipotent elements. Then, at the level of algebraic groups, $U$ is contained in a proper subvariety, and it follows from Borel's theorem [Bor] and the Lang-Weil theorem that $p_{w,G}(U) \leq eq^{-1}$ for some constant $e = e(w)$.

We have

$$|a_{w,\chi}| \leq \sum_{g \in G} P_{w,G}(g)|\chi(g)| \leq 2p_{w,G}(S) + p_{w,G}(U)q^{1/2} + P_{w,G}(1)\chi(1).$$

Since $P_{w,G}(1) \leq f(w)q^{-1}$ and $\chi(1) \leq q+1$ this yields

$$|a_{w,\chi}| \leq 2 + e(w)q^{-1}q^{1/2} + f(w)q^{-1}(q+1) \leq 2 + e(w)q^{-1/2} + 2f(w) \leq c(w)$$

for a suitable $c(w)$. □

The above result has some applications, as follows.

**Corollary 4.3.** *Let $w = w_1w_2$ where $w_1, w_2 \in F_d$ are non-trivial disjoint words. If $G = \mathrm{PSL}_2(q)$ where $q$ ranges over prime powers, then we have*

$$\lim_{q\to\infty} \|p_{w,G} - U_G\|_{L^2} = 0.$$

*Proof.* This follows easily using non-commutative Fourier methods. For $\chi \in \mathrm{Irr}(G)$ we have

$$|a_{w,\chi}| = \frac{|a_{w_1,\chi}||a_{w_2,\chi}|}{\chi(1)} \leq \frac{c(w_1)c(w_2)}{\chi(1)}$$

by Proposition 4.2. Applying [GS, Lemma 2.2] we obtain

$$(\|p_{w,G} - U_G\|_{L^2})^2 \leq \sum_{1\neq\chi\in\mathrm{Irr}(G)} |a_{w,\chi}|^2 \leq c(w_1)^2c(w_2)^2(\zeta_G(2) - 1), \tag{4.2}$$

where $\zeta_G$ is as before. By [LiSh2, Theorem 1.1] the RHS of (4.2) tends to 0 as $|G| \to \infty$ for all finite simple groups $G$. This completes the proof. □

Note that the above result completes the proof of Theorem 1 for Lie-type groups of bounded rank. We also obtain an $L^\infty$ result as follows.

**Corollary 4.4.** *Let $w_1, w_2, w_3, w_4$ be pairwise disjoint non-trivial words. Let $w = w_1w_2w_3w_4$ and $G = \mathrm{PSL}_2(q)$. Then*

$$\lim_{q\to\infty} \|p_{w,G} - U_G\|_{L^\infty} = 0.$$

*Proof.* Note that

$$a_{w,\chi} = \frac{a_{w_1,\chi}a_{w_2,\chi}a_{w_3,\chi}a_{w_4,\chi}}{\chi(1)^3}.$$

Combining this with Proposition 4.2 we obtain

$$|a_{w,G}| \leq \frac{C}{\chi(1)^3},$$

where $C = c(w_1)c(w_2)c(w_3)c(w_4)$.

Proposition 8.1 of [Sh1] shows that

$$\|p_{w,G} - U_G\|_{L^\infty} \leq \sum_{1\neq\chi\in\mathrm{Irr}(G)} |a_{w,\chi}|\chi(1).$$

This yields

$$\|p_{w,G} - U_G\|_{L^\infty} \leq \sum_{1\neq\chi\in\mathrm{Irr}(G)} C\chi(1)^{-2} = C(\zeta_G(2) - 1).$$

As noted above, the right hand side tends to 0 as $|G| \to \infty$, completing the proof. □

A similar statement for three words is false. Indeed, it is shown in [Sh1, p. 1406] that for $w = x_1^2x_2^2x_3^2$ and $G = \mathrm{PSL}_2(q)$, $p_{w,G}$ is not almost uniform in $L^\infty$.

Finally, it is easy to see that the bound on the Fourier coefficients in Proposition 4.2 cannot hold for all finite simple groups; indeed words of the kind $w = x_1^n$ give counter-examples. However, we conjecture that, for every

non-trivial word $w$ there exist a real number $\epsilon(w) > 0$ and a positive integer $N(w)$ such that, for all finite simple groups $G$ of order at least $N(w)$ and for all characters $\chi \in \mathrm{Irr}(G)$ we have

$$|a_{w,\chi}| \leq \chi(1)^{1-\epsilon(w)}.$$

## 5. The $L^1$ Waring problem

In this section we prove Theorem 6 and use it to complete the proof of Theorem 1.

Recall that Theorem 6 is known for alternating groups, see [LS1, Theorem 7.4]. For groups of Lie type $G = G(q)$ of bounded rank, it follows easily from Lemma 4.1 above. (Indeed, when $|G(q)| \to \infty$, $q$ tends to infinity, and Lemma 4.1 then shows that the probability that $w(g_1, \ldots, g_d)$ is regular semisimple tends to 1. The character values of regular semisimple elements are bounded in term of the rank $r$ of $G$, whereas $\chi(1)$ is at least of the magnitude of $q^r$, whence Theorem 6 follows.) Hence it remains to prove Theorem 6 for simple classical groups of arbitrarily high rank (which, in particular, can be assumed to be of type $A_r$, ${}^2A_r$, $B_r$, $C_r$, $D_r$, or ${}^2D_r$.) Thus we may (and do) assume that $G$ is a simple classical group of Lie type whose rank can be taken as large as we wish.

Let $H$ be a group satisfying for some $n$ and $q$ one of the following conditions: $\mathrm{SL}_n(q) \lhd H \leq \mathrm{GL}_n(q)$, $\mathrm{SU}_n(q) \lhd H \leq \mathrm{U}_n(q)$, $\mathrm{Sp}_n(q) \lhd H \leq \mathrm{CSp}_n(q)$ (with $2|n$), or $\Omega_n^{\pm}(q) \lhd H \leq \mathrm{O}_n^{\pm}(q)$. We will consider the natural action of $H$ on $V = \mathbb{F}_q^n$, $\mathbb{F}_{q^2}^n$, $\mathbb{F}_q^n$, $\mathbb{F}_q^n$, which in the last three cases is endowed with a non-degenerate $H$-invariant Hermitian, symplectic, or quadratic form $\langle\ ,\ \rangle$. In the unitary and orthogonal cases, the form is preserved; in the symplectic case, it only needs to be preserved up to a multiplier. We set $f = 2$ if $H$ is unitary; otherwise $f = 1$. In what follows, we will write $H = \mathrm{Cl}(V)$ to specify that $H$ is one of the described groups.

By a *classical group* $G$ (*in dimension* $n$, if we wish to specify), we mean henceforth a group which is the quotient of some group $H = \mathrm{Cl}(V)$ as above by a central subgroup $Z$ of $H$. Note that $|Z| \leq \max(q+1, 2) < 2q$ and that $|H/[H,H]| < 2q$. For such a group $G$, the *rank* of $G$ is the semisimple rank of the algebraic group underlying $H$. The finite simple groups $G$ with which we are concerned are of this type, but for the purposes of §7 it will be useful to do things in this slightly greater generality. Note that any classical group in our sense is a classical group in the sense of [GLT3, Definition 1.2]; hence the results of [GLT3] apply.

Theorem 6 is obtained by combining recent estimates for values of irreducible characters of classical groups with the following result, which may be of independent interest.

**Theorem 5.1.** *For every non-trivial word $w \in F_d$ there exists a constant $c = c(w)$ such that, if $G$ is a classical group of rank $r$ over the field with $q$*

*elements, and $g_1, \dots, g_d \in G$ are chosen uniformly and independently, then*

$$\mathbf{P}[|\mathbf{C}_G(w(g_1, \dots, g_d))| \leq q^{cr}] \to 1 \text{ as } |G| \to \infty.$$

We now embark on the proof of Theorem 5.1. This result is trivial if the rank $r$ is bounded, so we may assume $G$ is classical of arbitrarily high rank. We follow [LS3] closely.

**Lemma 5.2.** *If $h \in H$ maps to $g \in G$, with $G = H/Z$ as above, then $|\mathbf{C}_G(g)| \leq |\mathbf{C}_H(h)|$.*

*Proof.* Let

$$J = \{j \in H \mid j^{-1}hj \in hZ\}.$$

Then $J$ is a group containing $Z$, and $x \mapsto j^{-1}hjh^{-1}$ defines a homomorphism $J \to Z$ whose kernel is $\mathbf{C}_H(h)$. It follows that $|J| \leq |\mathbf{C}_H(h)| \cdot |Z|$. The restriction of the quotient map $H \to G$ to $J$ has kernel $Z$ and image $\mathbf{C}_G(g)$. Thus,

$$|\mathbf{C}_G(g)| = |Z|^{-1}|J| \leq |\mathbf{C}_H(h)|.$$

□

**Lemma 5.3.** *For any $A \in \mathbb{R}_{>0}$,*

$$\frac{|\{(g_1, \dots, g_d) \in G^d : \ |\mathbf{C}_G(w(g_1, \dots, g_d))| > A\}|}{|G|^d} \leq \frac{|\{(h_1, \dots, h_d) \in H^d : \ |\mathbf{C}_H(w(h_1, \dots, h_d))| > A\}|}{|H|^d}.$$

*Proof.* Indeed, any preimage in $H^d$ of an element $(g_1, \dots, g_d)$ in the left-hand side numerator belongs to the set in the right-hand side numerator. The lemma follows. □

Equivalently,

$$\mathbf{P}[|\mathbf{C}_G(w(g_1, \dots, g_d))| > A] \leq \mathbf{P}[|\mathbf{C}_H(w(h_1, \dots, h_d))| > A].$$

Therefore, to prove that there exists $c > 0$ such that

$$\limsup_{|G| \to \infty} \mathbf{P}[|\mathbf{C}_G(w(g_1, \dots, g_d))| > q^{cr}] = 0$$

it suffices to prove that there exists $c > 0$ such that

$$\limsup_{|H| \to \infty} \mathbf{P}[|\mathbf{C}_H(w(h_1, \dots, h_d))| > q^{cr}] = 0,$$

so it is certainly enough to prove there exists $c > 0$ such that

$$\limsup_{|H| \to \infty} \mathbf{P}[|\mathbf{C}_{\mathrm{GL}(V)}(w(h_1, \dots, h_d))| > q^{cr}] = 0.$$

If $\mathbb{F}$ is a finite field and $h \in \mathrm{GL}_n(\mathbb{F})$, we define for each monic irreducible polynomial $P(x) \in \mathbb{F}[x]$

$$a_{P,1} \geq a_{P,2} \geq \dots$$

to be the descending sequence giving the sizes of Jordan blocks for any root $\lambda$ of $P(x)$. (Clearly, this sequence does not depend on the choice of root $\lambda$.) Clearly,

$$\sum_{P,m} a_{P,m} \deg P = n. \tag{5.1}$$

It is well known [Hu, §1.3] that the centralizer of $h$ in $M_n(\mathbb{F})$ is a vector space over $\mathbb{F}$ of dimension

$$\sum_P \sum_m (2m-1) a_{P,m} \deg P.$$

Thus,

$$|\mathbf{C}_{\mathrm{GL}_n(\mathbb{F})}(h)| < |\mathbb{F}|^{\sum_P \sum_m (2m-1) a_{P,m} \deg P}.$$

For later use, we note that by (5.1), if

$$|\mathbf{C}_{\mathrm{GL}_n(\mathbb{F})}(h)| > |\mathbb{F}|^{2\delta n^2},$$

then

$$\text{there exist some } P \text{ and some } m_0 > \delta n \text{ such that } a_{P,m_0} \neq 0, \tag{5.2}$$

i.e., some eigenspace of $h$ has dimension greater than $\delta n$. For immediate use, we note that

$$|\mathbf{C}_{\mathrm{GL}_n(\mathbb{F})}(h)| > |\mathbb{F}|^{6cn}$$

implies

$$\begin{aligned}
\sum_P \sum_{m>c} (m-c) a_{P,m} \deg P &= -cn + \sum_P \sum_{m>c} m a_{P,m} \deg P \\
&> -cn + \sum_P \sum_{m\geq 1} m a_{P,m} \deg P - \sum_P \sum_{m=1}^{c} \sum_{k=m}^{c} a_{P,k} \deg P \\
&\geq -cn + \sum_P \sum_{m\geq 1} m a_{P,m} \deg P - c \sum_P \sum_{k=m}^{c} a_{P,k} \deg P \\
&> -2cn + \frac{1}{2} \sum_P \sum_{m\geq 1} (2m-1) a_{P,m} \deg P > cn
\end{aligned}$$

As $a_{P,m}$ is non-increasing in $m$,

$$\sum_{m>c} (m-c) a_{P,m} \deg P \leq \left( \max_{\{(m,P) | a_{P,m} > 0\}} (m-c) \right) \sum_P a_{P,c+1} \deg P.$$

Thus, at least one of the following conditions holds:

$$\sum_P a_{P,c+1} \deg P > \sqrt{cn}, \tag{5.3}$$

or

$$\text{for some polynomial } P \text{ and some } m_0 > \sqrt{cn}, \text{ we have } a_{P,m_0} \neq 0. \tag{5.4}$$

**Lemma 5.4.** *Condition* (5.2) *implies that there exists a non-constant polynomial* $Q(x) \in \mathbb{F}[x]$ *such that*

$$\dim_{\mathbb{F}} \ker Q(g) > \delta n \deg Q.$$

*For any positive integer* $t > 0$*, if* $c$ *is sufficiently large in terms of* $t$ *and* $n$ *is sufficiently large in terms of* $c$*, then the conditions* (5.3) *and* (5.4) *each imply that there exists a non-zero polynomial* $Q(x) \in \mathbb{F}[x]$ *such that*

$$\dim_{\mathbb{F}} \ker Q(g) > 2t \deg Q + \sqrt{n}.$$

*Proof.* If (5.2) holds, setting $Q = P$, we have that

$$\dim \ker Q(g) \geq m_0 \deg P > \delta n \deg P.$$

In case of (5.3), $a_{P,1} \geq a_{P,2} \geq \ldots \geq a_{P,c+1}$ for all $P$ implies

$$\sum_P a_{P,c+1} \deg P \leq \frac{n}{c+1}.$$

Assuming $c \geq 2t$ and $n > (c+1)^2$, we set $Q = \prod P^{a_{P,c+1}}$ and obtain $\deg Q > \sqrt{n}$. Regarding $\mathbb{F}^n$ as $\mathbb{F}[x]$-module, where $x$ acts as $g$, the kernel of $Q(g)$ is isomorphic to

$$\bigoplus_P \Big[ (\mathbb{F}[x]/(P(x)^{a_{P,c+1}}))^{c+1} \oplus \bigoplus_{m>c+1} \mathbb{F}[x]/(P(x)^{a_{P,m}}) \Big],$$

whose dimension is $\geq (c+1) \deg Q > 2t \deg Q + \sqrt{n}$.

In case of (5.4), if $c > (2t+1)^2 \geq 1$ then we have $\deg P \leq n/m_0 < \sqrt{n}$. Setting $Q = P$, we obtain

$$\dim \ker Q(g) \geq m_0 \deg P > \sqrt{cn} \deg P > (2t+1)\sqrt{n} \deg P > 2t \deg P + \sqrt{n}.$$

□

**Proposition 5.5.** *Let* $H = \mathrm{Cl}(V)$ *be a finite classical group as described, where* $\mathbb{F} = \mathbb{F}_{q^f}$ *and* $n = \dim_{\mathbb{F}} V$*. Let* $w \in F_d$ *be a word of length* $l > 0$*. If* $k$ *and* $D$ *are positive integers and* $(h_1, \ldots, h_d)$ *is chosen uniformly from* $H^d$*, the probability that there exists a polynomial* $Q(x) \in \mathbb{F}[x]$ *of degree* $D$ *such that*

$$\dim \ker Q(w(h_1, \ldots, h_d)) \geq 2lDk$$

*is at most* $q^{-fk((k-1)lD-2.5)}$.

*Proof.* We choose an ordered $k$-tuple $(v_1, \ldots, v_k)$ uniformly from $V^k$ and a $d$-tuple $(h_1, \ldots, h_d)$ uniformly from $H^d$. It suffices to prove that the probability that $Q(w(h_1, \ldots, h_d))(v_i) = 0$ for all $i \in [1, k]$ is less than $q^{f(2.5k+k(k+1)lD-kn)}$. Indeed, the probability that a uniformly chosen random $k$-tuple $(v_1, \ldots, v_k)$ of vectors belongs to any particular subspace of dimension $2lDk$ is $q^{2flDk^2-fkn}$, so this implies that the probability that $\dim Q(w(h_1, \ldots, h_d)) \geq 2lDk$ is at most $q^{-fk((k-1)lD-2.5)}$, as claimed.

We write $w^D$ as a reduced word $y_m y_{m-1} \dots y_1$, where $m \le lD$ and each $y_i$ belongs to $\{x_1^{\pm 1}, \dots, x_d^{\pm 1}\}$. Let $z_j = y_j y_{j-1} \cdots y_1$ and for $j \ge 0$, let

$$e_{i,j} = z_j(h_1, \dots, h_d)(v_i).$$

If $\{v_1, \dots, v_k\} \subset \ker Q(w(h_1, \dots, h_d))$, then for each $i \in [1,k]$, the set $\{e_{i,0}, e_{i,1}, \dots, e_{i,m}\}$ is linearly dependent.

We endow the set of integer pairs in $[1,k] \times [0,m]$ with the lexicographic ordering. Let the event $X_{i,j}$ be the condition

$$e_{i,j} \notin \operatorname{Span}\{e_{i',j'} \mid (i',j') < (i,j)\}.$$

Let

$$Y_{i,j} = X_{i,0} X_{i,1} \cdots X_{i,j-1} X_{i,j}^c.$$

Let $Z_i$ be the event that $e_{i,0}, e_{i,1}, \dots, e_{i,m}$ is a linearly independent sequence. If $Z_i^c$ occurs, then $Y_{i,j}$ occurs for some $j \in [0,m]$. Thus,

$$\mathbf{P}[Z_i^c | Z_1^c Z_2^c \cdots Z_{i-1}^c] \le \sum_{j=0}^{m} \mathbf{P}[Y_{i,j} | Z_1^c Z_2^c \cdots Z_{i-1}^c].$$

We find an upper bound for each term on the right hand side by giving an upper bound on the conditional probability of $Y_{i,j}$ with respect to any possible set of data $e_{i',j}$ for $i' \in [1,i)$ and $j' \in [0,m]$.

Given this data, the event $Y_{i,0}$, or, equivalently, $X_{i,0}^c$, is the condition that $v_i$ belongs to the span of $\{e_{i',j'} \mid i' < i, 0 \le j' \le m\}$. As

$$\dim \operatorname{Span}\{e_{i',0}, \dots, e_{i',m}\} \le m,$$

the probability of $X_{i,0}^c$ is at most $q^{f((i-1)m-n)}$. For $j \in [1,m]$, we further condition on $e_{i,0}, e_{i,1}, \dots, e_{i,j-1}$ consistent with $X_{i,j'}$ for $j' < i$. Either $y_i = x_t$ or $y_i = x_t^{-1}$ for some $t$, and $e_{i,j}$ is determined by the specified value $e_{i,j-1}$ and the random variable $h_t$, so the conditional probability in question depends only $h_t$.

If $h$ (taken to be $h_t$ or $h_t^{-1}$ depending on whether $y_i$ is $x_t$ or $x_t^{-1}$), is a uniformly distributed random variable on $H$, and for some linearly independent sequence of $r \le (i-1)m + j$ vectors, $w_1, \dots, w_r$ and a second linearly independent sequence $w_1', \dots, w_r'$, we condition on $h(w_j) = w_j'$ for $j = 1, \dots, r-1$, then the probability that $h(w_r) = w_r'$ is the reciprocal of the number of possibilities for $h(w_r)$ given $h(w_j) = w_j'$ for $j = 1, \dots, r-1$.

If $H$ is of linear type, it contains $\mathrm{SL}(V)$ and therefore acts transitively on $r$-tuples of linearly independent vectors of $V$ for $r < n$. It follows that any $w'$ not in the span of $w_1', \dots, w_{r-1}'$ is possible. Otherwise $H$ contains $\mathrm{SU}(V)$, $\mathrm{Sp}(V)$, or $\Omega(V)$, respectively. In each of these cases, Witt's extension theorem [KlL, Proposition 2.1.6] applied to $\tilde{H} = \mathrm{U}(V)$, $\mathrm{Sp}(V)$, or $\mathrm{O}(V)$, respectively, implies that the number of possibilities for $h(w_r)$ is at least $1/\alpha$ of the number of solutions in $w'$ of the system of equations

$$\langle w_j', w' \rangle = \langle w_j, w_r \rangle,\ j = 1, \dots, r-1;\ \langle w', w' \rangle = \langle w_r, w_r \rangle, \tag{5.5}$$

where $\alpha = q+1$ in the U-case, 1 in the Sp-case, and $\alpha = 2$ or 4 in the O-case, depending on whether $2|q$ or not.

The equations (5.5) are $\mathbb{F}_{q^f}$-linear except for the last, in which the left hand side is a quadratic form over $\mathbb{F}_q$. Since a quadratic form in $k$ variables over a field of cardinality $q$ takes on each possible value at least $q^{k-2}$ times, we conclude that the probability of any single possible value $w'$ for $h(w_r)$ is at most $\alpha q^{f(1+r-n)} \le \alpha q^{f(1+(i-1)m+j-n)}$. Since

$$\dim \mathrm{Span}\{e_{i',j'} \mid (i',j') < (i,j)\} \le (i-1)m + j,$$

we conclude that

$$\mathbf{P}[Y_{i,j}|Z_1^c Z_2^c \cdots Z_{i-1}^c] \le \alpha q^{f((i-1)m+j)} q^{f(1+(i-1)m+j-n)} = \alpha q^{f(1+2(i-1)m+2j-n)}.$$

It follows that

$$\begin{aligned}\mathbf{P}[Z_i^c \mid Z_1^c Z_2^c \cdots Z_{i-1}^c] &\le \sum_{j=0}^{m} \alpha q^{f(1+2(i-1)m+2j-n)} \\ &< \frac{\alpha}{1-q^{-2f}} q^{f(1+2im-n)} < q^{f(2.5+2im-n)}.\end{aligned}$$

This implies

$$\mathbf{P}[Z_1^c \cdots Z_k^c] \le q^{f(3k+k(k+1)m-kn)} \le q^{f(2.5k+k(k+1)lD-kn)},$$

as claimed. □

*Proof of Theorem 5.1.* This follows by combining Lemma 5.4 with Proposition 5.5. Indeed, by the discussion preceding Lemma 5.4, we need to bound from above the probability $\mathbf{P}'$ that either (5.3) or (5.4) holds for $h = w(h_1, \ldots, h_d)$. We may assume that the rank $r$ of $H$ is as large as we wish; in particular, we may assume that

$$r_0 = \lfloor \sqrt[4]{r}/\sqrt{2l} \rfloor \ge 4,$$

where $l$ is the length of $w$. First we apply Lemma 5.4 with $t = 2l$ to see that either of (5.3), (5.4) for $h$ implies the existence of non-constant $Q \in \mathbb{F}[x]$ such that

$$\dim_{\mathbb{F}} \ker Q(h) > 2t \deg Q + \sqrt{n} > \max(4l \deg Q, 2lr_0^2).$$

By Proposition 5.5 applied to $k = 2$, the probability $\mathbf{P}'_1$ that $D = \deg Q$ is at least $r_0$ is

$$\mathbf{P}'_1 < q^{5f} \sum_{D=r_0}^{\infty} q^{-2flD} < \frac{q^{5f}}{q^{2r_0 f}(1-q^{-2f})} < q^{-r_0 f/2}.$$

On the other hand, by Proposition 5.5 applied to $k = r_0 \ge 4$, the probability $\mathbf{P}'_2$ that $D = \deg Q \ge 1$ is $< r_0$ is

$$\mathbf{P}'_2 < q^{2.5r_0 f} \sum_{D=1}^{\infty} q^{-3r_0 flD} < \frac{q^{2.5r_0 f}}{q^{3r_0 f}-1} < 2q^{-r_0 f/2}.$$

Note that when $|G| \to \infty$, we have that $q^{r_0} \to \infty$, and so

$$\mathbf{P}' = \mathbf{P}'_1 + \mathbf{P}'_2 < 3q^{-r_0 f/2}$$

tends to 0, as desired. □

*Proof of Theorem 6.* By [GLT2, Theorem 1.3] and [GLT3, Theorem 1.3], for all $c$ and $\epsilon > 0$, increasing $r$ if necessary, $|\mathbf{C}_G(g)| < q^{cr}$ implies

$$|\chi(g)| \leq \chi(1)^\epsilon$$

for every irreducible character $\chi$ of $G$. Now apply Theorem 5.1. □

*Proof of Theorem 1.* It remains to show that, given any two disjoint words $w_1, w_2 \neq 1$, there exists a positive constant $R$ such that if $S$ is any set of finite simple groups of rank $r \geq R$, and $w = w_1 w_2$, then

$$\lim_{G \in S, |G| \to \infty} \|p_{w,G} - U_G\|_{L^1} = 0.$$

Fix any $0 < \epsilon < 1/3$. We say that an element $g \in G$ is *$\epsilon$-good* if $\chi(g)| \leq \chi(1)^\epsilon$ for all $\chi \in \mathrm{Irr}(G)$; a conjugacy class $C$ is *$\epsilon$-good* if it consists of $\epsilon$-good elements. By Theorem 6, if $R$ is chosen sufficiently large, $G$ is of rank $r > R$, and $g_1, \ldots, g_d \in G$ are chosen uniformly and independently, then the probability that $w_1(g_1, \ldots, g_d)$ and $w_2(g_1, \ldots, g_d)$ are both $\epsilon$-good approaches 1 as $|G| \to \infty$. For proving $L^1$ convergence to the uniform distribution, we may therefore assume that both $w_1(g_1, \ldots, g_d)$ and $w_2(g_1, \ldots, g_d)$ belong to $\epsilon$-good conjugacy classes.

For $i \in \{1, 2\}$ and any conjugacy class $C_i$, the conditional distribution of $w_i(g_1, \ldots, g_d)$ given that it belongs to $C_i$ is the uniform distribution on $C_i$. Thus, it suffices to prove that the convolution of the uniform distribution on an $\epsilon$-good $C_1$ with the uniform distribution on an $\epsilon$-good $C_2$ approaches the uniform distribution on $G$ in the $L^1$ norm uniformly in $C_1$ and $C_2$ as the order of $G$ (of sufficiently high rank) grows without bound. This would follow if we knew that there exist at least $(1 - o(1))|G|$ elements $g \in G$ for which the probability that $x_1 x_2 = g$ as $x_i \in C_i$ are chosen uniformly and independently is $(1 + o(1))|G|^{-1}$, where $o(1) = o_{|G|}(1)$.

By Proposition 4.2 of [LS4], if $\epsilon > 0$ and the rank of $G$ is sufficiently large in terms of $\epsilon$, then the proportion of $\epsilon$-good elements in $G$ tends to 1 as $|G| \to \infty$. Hence we may assume that $g$ is $\epsilon$-good. By (4.1), the probability $p(C_1, C_2, g)$ that $x_1 x_2 = g$ satisfies

$$p(C_1, C_2, g) = |G|^{-1} \sum_{\chi \in \mathrm{Irr}(G)} \frac{\chi(C_1)\chi(C_2)\chi(g^{-1})}{\chi(1)},$$

so

$$|p(C_1, C_2, g) - |G|^{-1}| \leq |G|^{-1} \sum_{1_G \neq \chi \in \mathrm{Irr}(G)} \frac{\chi(1)^{3\epsilon}}{\chi(1)} = |G|^{-1}(\zeta_G(1 - 3\epsilon) - 1),$$

where $\zeta_G(s) = \sum_{\chi \in \mathrm{Irr}(G)} \chi(1)^{-s}$ is the Witten zeta function of $G$. Since $1 - 3\epsilon > 0$, we may choose $R$ sufficiently large so that $\zeta_G(1 - 3\epsilon) \to 1$ as

$|G| \to \infty$; indeed, this follows from Theorem 1.1 and 1.2 of [LiSh2]. This yields

$$p(C_1, C_2, g) = |G|^{-1}(1 + o(1)),$$

for all $C_1, C_2, g$ as above. This completes the proof of Theorem 1. □

## 6. The $L^\infty$ Waring problem

In this section we prove Theorem 4.

**Proposition 6.1.** *Fix any $0 < \epsilon < 1$. There exists some $A(\epsilon) > 0$ such that, for any $n \geq A(\epsilon)$, any element $g$ in $G \in \{\mathsf{A}_n, \mathsf{S}_n\}$, and any $\chi \in \mathrm{Irr}(G)$, the following two statements hold.*

(i) *If $\mathsf{fix}(g) \leq n^{1-\epsilon}$, then $|\chi(g)| \leq \chi(1)^{1-\epsilon/3}$.*
(ii) *If $\mathsf{fix}(g) = k$, then $|\chi(g)| \leq n^{k/4}\chi(1)^{1/2+\epsilon}$.*

*Proof.* (a) First we consider the case $G = \mathsf{S}_n$. Then [LS1, Theorem 1.2] implies (i) immediately. For (ii), it implies that there exists $C(\epsilon) > 0$ such that, if $n - k \geq C(\epsilon)$ and $x \in \mathsf{S}_{n-k}$ is fixed-point-free, then

$$|\chi(x)| \leq \chi(1)^{1/2+\epsilon} \tag{6.1}$$

for all $\chi \in \mathrm{Irr}(\mathsf{S}_{n-k})$. On the other hand,

$$\frac{|\chi(g)|}{n^{k/4}\chi(1)^{1/2}} \leq \frac{\chi(1)^{1/2}}{n^{k/4}} \leq \frac{|\mathsf{S}_n|^{1/4}}{n^{k/4}} < \frac{(n/2)^{n/4}}{n^{k/4}} = \left(\frac{n^{n-k}}{2^n}\right)^{1/4}.$$

In particular, the desired bound in (ii) holds if $n-k < C(\epsilon)$ is bounded, but $n$ is large enough.

Hence we may assume that $n - k \geq C(\epsilon)$ is sufficiently large, and also $\epsilon < 1/2$. We use the branching rule from $\mathsf{S}_m$ to $\mathsf{S}_{m-1}$ for $n \geq m \geq n-k+1$ consecutively and write

$$\chi|_{\mathsf{S}_{n-k}} = \chi_1 + \cdots + \chi_N$$

of irreducible characters of $\mathsf{S}_{n-k}$, with repetition allowed. The number of terms $N$ is at most the $k^{\text{th}}$ power of the maximum number of removable boxes from any Young diagram of size $\leq n$, and so $N < (2n)^{k/2}$.

Let $h \in \mathsf{S}_{n-k}$ map to an element of $\mathsf{S}_n$ conjugate to $g$. Then $\chi(g) = \sum_i \chi_i(h)$. Since $\mathsf{fix}(g) = k$, $h$ has no fixed point; also, $n - k \geq C(\epsilon)$. Hence $|\chi_i(h)| \leq \chi_i(1)^{1/2+\epsilon}$ by (6.1). As $\epsilon < 1/2$, we obtain

$$\begin{aligned}|\chi(g)| &\leq \sum_i |\chi_i(h)| \leq \sum_i \chi_i(1)^{1/2+\epsilon} \leq N\left(\frac{\sum_i \chi_i(1)}{N}\right)^{1/2+\epsilon} \\ &= N\left(\frac{\chi(1)}{N}\right)^{1/2+\epsilon} < (2n)^{k/4-k\epsilon/2}\chi(1)^{1/2+\epsilon} \leq n^{k/4}\chi(1)^{1/2+\epsilon}\end{aligned}$$

when $n > 2^{1/2\epsilon}$.

(b) Now we consider the case $G = \mathsf{A}_n$. We are certainly done by (a) if $\chi$ extends to $G$. Hence we may assume that there is a self-associated partition

$$\lambda = (\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_r \geq 1)$$

of $n$ such that the character $\chi^\lambda$ of $\mathsf{S}_n$ labeled by $\lambda$ restricts to $G$ as $\chi^+ + \chi^-$, with $\chi^\pm \in \mathrm{Irr}(G)$ and $\chi = \chi^+$. Let $h_{11} > \ldots > h_{tt} \geq 1$ denote the hook lengths of the Young diagram of $\lambda$ at the diagonal nodes. By [JK, Theorem 2.5.13], if the cycle type of $g$ is not $(h_{11}, h_{22}, \ldots, h_{tt})$, then

$$|\chi(g)| = \frac{|\chi^\lambda(g)|}{2} \leq \frac{\chi^\lambda(1)^{1-\epsilon/3}}{2} \leq \chi(1)^{1-\epsilon/3}$$

if $\mathsf{fix}(g) \leq n^{1-\epsilon}$; and the same argument applies to (ii). On the other hand, if the cycle type of $g$ is $(h_{11}, h_{22}, \ldots, h_{tt})$, then

$$|\chi(g)| \leq \left(1 + \sqrt{\prod_{i=1}^{t} h_{ii}}\right)/2.$$

Since $\lambda$ is self-associated, we have that $\lambda_1 \leq (n+1)/2$, and so

$$\chi(1) = \chi^\lambda(1)/2 \geq 2^{(n-5)/4} \tag{6.2}$$

by [GLT1, Theorem 5.1]. Also, all $h_{ii}$ are odd integers. Note that if $m \geq 3$ is any odd integer, then

$$m < 2^{m/4}$$

if and only if $m \geq 17$; furthermore,

$$\prod_{j=1}^{7} \frac{2j+1}{2^{(2j+1)/4}} < 37.$$

It follows that

$$\prod_{i=1}^{t} h_{ii} < 37 \cdot 2^{\sum_{i=1}^{t} h_{ii}/4} = 37 \cdot 2^{n/4}$$

and so

$$|\chi(g)| < (1 + 6.1 \cdot 2^{n/8})/2.$$

Together with (6.2), this implies that

$$|\chi(g)| < \min\{\chi(1)^{2/3}, \chi(1)^{1/2+\epsilon}\} \leq \min\{\chi(1)^{1-\epsilon/3}, n^{k/4}\chi(1)^{1/2+\epsilon}\}$$

when $n$ is large enough. □

For use in §7, we will also need an imprimitive version of Proposition 6.1:

**Lemma 6.2.** *For any $0 < \epsilon < 1$, let $A(\epsilon) > 0$ be the constant in Proposition 6.1. Let $m|n$ and consider the subgroup $H = \mathsf{S}_{n/m} \wr \mathsf{S}_m$ of $G = \mathsf{S}_n$. If $n/m \geq A(\epsilon)$, then for any $h \in H$ and $\chi \in \mathrm{Irr}(H)$ we have*

(i) *If $\mathsf{fix}(h) \leq (n/m)^{1-\epsilon}$, then $|\chi(h)| \leq m!\chi(1)^{1-\epsilon/3}$.*
(ii) *If $\mathsf{fix}(h) = k$, then $|\chi(h)| \leq m!(n/m)^{k/4}\chi(1)^{1/2+\epsilon}$.*

*Proof.* Let $K = (\mathsf{S}_{n/m})^m \lhd H$. The restriction of any irreducible representation $V$ of $H$ to $K$ is a direct sum of representations of the form $V_1 \boxtimes \cdots \boxtimes V_m$, where each $V_i$ is an irreducible representation of $\mathsf{S}_{n/m}$, and the $m$-tuples $(V_1, \ldots, V_m)$ appearing in tensor decompositions of the different irreducible factors of $V|_K$ are the same up to permutation.

There exists a unique partition $\pi = (a_1, a_2, \ldots, a_r) \vdash m$, and an irreducible representation $W_i$ of $\mathsf{S}_{n/m}$ for each $i$ (so that $W_1, \ldots, W_r$ are pairwise non-isomorphic) such that, after permuting tensor factors, $V_1 \boxtimes \cdots \boxtimes V_m$ can be rewritten

$$W_\pi = W_1^{\boxtimes a_1} \boxtimes \cdots \boxtimes W_r^{\boxtimes a_r}.$$

This representation has inertia group $H_\pi = \mathsf{S}_{n/m} \wr \mathsf{S}_\pi$ in $H$, where $\mathsf{S}_{a_1} \times \cdots \times \mathsf{S}_{a_r} \subset \mathsf{S}_m$, $W_\pi$ extends to a representation $\tilde{W}_\pi$ of $H_\pi$, and

$$V = \mathrm{Ind}_{H_\pi}^H(\tilde{W}_\pi \otimes U)$$

for a suitable irreducible representation $U$ of $\mathsf{S}_\pi$ (inflated to $H_\pi$).

To calculate the trace $\chi(h)$ of $h \in H$ acting on $V$, we first consider the image $\overline{h}$ of $h$ in $\mathsf{S}_m$. In general, $\overline{h}$ permutes the $\mathsf{S}_\pi$-cosets and therefore the summands of $V|_{H_\pi}$. Only the summands which are stabilized by $\overline{h}$ contribute to $\chi(h)$, and the number of those summands is certainly bounded by $[\mathsf{S}_m : \mathsf{S}_\pi]$. Also, the absolute value of the trace of $h$ acting on $U$ is at most $|\mathsf{S}_\pi|$. Hence, it suffices to prove that assuming $\overline{h} \in \mathsf{S}_\pi$, we have

(i) If $\mathsf{fix}(h) \leq (n/m)^{1-\epsilon}$, then $\mathrm{tr}(h|\tilde{W}_\pi) \leq \dim \tilde{W}_\pi^{1-\epsilon/3}$.
(ii) If $\mathsf{fix}(h) = k$, then $\mathrm{tr}(h|\tilde{W}_\pi) \leq (n/m)^{k/4} \dim \tilde{W}_\pi^{1/2+\epsilon}$.

Writing $h = (h_1, \ldots, h_r)$ where $h_i \in \mathsf{S}_{n/m} \wr S_{a_i}$ acts on the extension $\widetilde{W_i^{\boxtimes a_i}}$ of $W_i^{\boxtimes a_i}$ to $\mathsf{S}_{n/m} \wr S_{a_i}$, it suffices to prove

(i) If $\mathsf{fix}(h_i) \leq (n/m)^{1-\epsilon}$, then $\mathrm{tr}(h_i|\widetilde{W_i^{\boxtimes a_i}}) \leq \dim \widetilde{W_i^{\boxtimes a_i}}^{1-\epsilon/3}$.
(ii) If $\mathsf{fix}(h_i) = k$, then $\mathrm{tr}(h_i|\widetilde{W_i^{\boxtimes a_i}}) \leq (n/m)^{k/4} \dim \widetilde{W_i^{\boxtimes a_i}}^{1/2+\epsilon}$.

Thus, we can reduce to the case $r = 1$. Decomposing $\overline{h} \in \mathsf{S}_{a_1}$ into a product of disjoint cycles, we further reduce to the case that $\overline{h}$ is an $a_1$-cycle $\sigma$, say $(1, 2, \ldots, a_1)$. Writing $h = ((t_1, ..., t_{a_1}), \sigma)$ with $t_i \in \mathsf{S}_{n/m}$, we then obtain

$$|\mathrm{tr}(h \mid W_1^{\boxtimes a_1})| = |\mathrm{tr}(t_1 \cdots t_{a_1})|;$$

in particular, it is at most $\dim W_1 \leq (\dim W_1^{\boxtimes a_1})^{1/2}$ if $a_1 \geq 2$, implying both (i) and (ii). If $a_1 = 1$, then $\mathsf{fix}(t_1 \cdots t_{a_1}) = \mathsf{fix}(h)$, and we are again done by Proposition 6.1(ii) applied to $\mathsf{S}_{n/m}$. □

Let $w \in F_d$ be a non-trivial word. We write it in reduced form: $w = y_l \cdots y_2 y_1$, where each $y_i$ can be regarded as a function $G^d \to G$ which is either projection on the $k^{\text{th}}$ factor for some $k \in [1, d]$ or projection composed with the inverse map.

**Lemma 6.3.** *For integers $1 \le a < b \le l$, $m \ge 1$, and $n > 2m(b-a)$, and $G \in \{\mathsf{A}_n, \mathsf{S}_n\}$, we define*

$$X_{m,n}(w,a,b) \subset G^d \times [1,n]^{m(b-a)}$$

*to be the set of tuples*

$$(g_1, \dots, g_d, r_{1,a}, \dots, r_{1,b-1}, \dots, r_{m,a}, \dots, r_{m,b-1})$$

*satisfying:*

(6.3.1) *all the $r_{i,j}$, $1 \le i \le m$, $a \le j \le b-1$, are pairwise distinct;*
(6.3.2) *for all $1 \le i \le m$ and $a \le j \le b-2$, $y_j(g_1, \dots, g_d)(r_{i,j}) = r_{i,j+1}$; and*
(6.3.3) *for all $1 \le i \le m$, $y_{b-1}(g_1, \dots, g_d)(r_{i,b-1}) = r_{i,a}$.*

*Then the projection $p_1$ of $X_{m,n}(w,a,b)$ onto $G^d$ has cardinality less than or equal to*

$$\frac{e^{2m^2(b-a)^2/n}}{m!}|G|^d.$$

*Proof.* If $y_a = y_{b-1}^{-1}$, then (6.3.2) for $j = a$ together with (6.3.3) implies $r_{i,a+1} = r_{i,b-1}$ for all $i$, contrary to (6.3.1), so it follows that $X_{m,n}(w,a,b)$ is empty. We therefore assume $y_a \ne y_{b-1}^{-1}$.

For each choice of $r_{i,j}$, $1 \le i \le m$, $a \le j \le b-1$ satisfying (6.3.1), the conditions (6.3.2) and (6.3.3) impose a total of $m(b-a)$ conditions on the $(g_1, \dots, g_d)$, where each condition is of the form $g_h u = v$ or $g_h^{-1} u = v$, for some $h \in [1,d]$ and $u, v \in [1,n]$. These conditions are independent because the $r_{i,j}$ are all distinct from one another, and $y_a \ne y_{b-1}^{-1}$. Let $c_h$ for $1 \le h \le d$ denote the number of conditions on $g_h$. As $c_1 + \cdots + c_d = m(b-a) \le n-2$, each $c_h$ is less than $n-1$, and the number of elements $g_h \in G$ satisfying the $c_h$ conditions is

$$\frac{|G|}{n(n-1)\cdots(n-c_h+1)} \ge \frac{|G|}{(n-c_h)^{c_h}}.$$

Overall,

$$|X_{m,n}(w,a,b)| \le n^{m(b-a)} \prod_{h=1}^{d} \frac{|G|}{(n-c_h)^{c_h}} \le \frac{n^{m(b-a)}}{(n-m(b-a))^{m(b-a)}}|G|^d.$$

The projection of $X_{m,n}(w,a,b)$ onto $G^d$ is at least $m!$ to 1 since $\mathsf{S}_m$ acts faithfully on $X_{m,n}(w,a,b)$ through its action on the $i$-coordinate of $r_{i,j}$. Thus, the cardinality of the projection is bounded above by

$$\frac{n^{m(b-a)}}{m!(n-m(b-a))^{m(b-a)}}|G|^d.$$

Setting $M = m(b-a) < n/2$, we have

$$\frac{n^M}{(n-M)^M} = \exp\Big(M \log\Big(1 + \frac{M}{n-M}\Big)\Big) < \exp\Big(\frac{M^2}{n-M}\Big) < \exp\Big(\frac{2M^2}{n}\Big),$$

which implies the lemma. □

For $G \in \{\mathsf{A}_n, \mathsf{S}_n\}$ and $w$ a word of length $l$ in $F_d$, let $W_n$ be the corresponding random variable on $G$ with distribution $p_{w,G}$.

**Proposition 6.4.** *If $k \geq (el)^{36}$, then for all positive integers $n$,*

$$\mathbf{P}[\mathsf{fix}(W_n) \geq k] \leq k^{-\frac{k}{3l^4}}.$$

*Proof.* Let $X_i$, $i = 1, \ldots, d$, be independent uniform random variables on $G$. We write $w = y_l \cdots y_2 y_1$ in reduced form, and let $z_i = y_i \cdots y_2 y_1$. Let $Y_i$ denote the random variable $y_i(X_1, \ldots, X_d)$.

Let us first assume that $l^4$ divides $k$ and if $l = 1$ we assume also $k \leq n/2$. Since there are less than $l^2$ pairs of integers $(a, b)$ with $0 \leq a < b \leq l$, if $\mathsf{fix}(w(g_1, \ldots, g_d)) \geq k$, there exist $a$ and $b$ and at least $k/l^2$ integers $r \in [1, n]$ such that the following two conditions hold:

(6.4.1) the terms of the sequence

$$z_a(g_1, \ldots, g_d)r, z_{a+1}(g_1, \ldots, g_d)r, \ldots, z_{b-1}(g_1, \ldots, g_d)r$$

are pairwise distinct, and

(6.4.2) $z_a(g_1, \ldots, g_d)r = z_b(g_1, \ldots, g_d)r$.

For $a \leq i, j < b$ and any given $r \in [1, n]$, there is at most one element $s \in \mathsf{fix}(w(g_1, \ldots, g_d))$ such that

$$z_i(g_1, \ldots, g_d)r = z_j(g_1, \ldots, g_d)s.$$

Thus, if there exist $k/l^2$ elements $r$ satisfying (6.4.1) and (6.4.2), there exists a subset of $m = k/l^4$ elements $\{r_1, \ldots, r_m\}$ for which the sets

$$\{\{z_a(g_1, \ldots, g_d)r_j, z_{a+1}(g_1, \ldots, g_d)r_j, \ldots, z_{b-1}(g_1, \ldots, g_d)r_j\} \mid 1 \leq j \leq m\}$$

are pairwise disjoint. Setting

$$r_{j,i} = z_i(g_1, \ldots, g_d)(r_j),$$

we see that the tuple

$$(g_1, \ldots, g_d, r_{1,a}, \ldots, r_{1,b-1}, \ldots, r_{m,b-1})$$

satisfies conditions (6.3.1)–(6.3.3). Thus, the set

$$\{(g_1, \ldots, g_d) \mid \mathsf{fix}(w(g_1, \ldots, g_d)) \geq k\}$$

is contained in

$$\bigcup_{0 \leq a < b \leq l} p_1(X_{m,n}(w, a, b)).$$

As $2ml = 2k/l^3 \leq n$, Lemma 6.3 applies, so

$$\begin{aligned}\mathbf{P}[\mathsf{fix}(W_n) \geq k] &= \mathbf{P}[\mathsf{fix}(Y_l \cdots Y_1) \geq k] \\ &\leq \frac{l^2 \max_{a,b} |p_1(X_{m,n}(w, a, b))|}{|G|^d} \leq \frac{l^2 e^{2k^2/l^6 n}}{(k/l^4)!}.\end{aligned}$$

We now consider the general case $k > (el)^{36}$. If $k_1$ denotes the largest multiple of $l^4$ not exceeding $k$ (or $n/2$ if $l = 1$), we have $k_1 \geq k/2$, and

$$(k_1/l^4)! \geq (k_1/el^4)^{k_1/l^4} \geq (k/2el^4)^{k/2l^4}.$$

By (i), we now have

$$\begin{aligned} -\log \mathbf{P}[\mathsf{fix}(W_n) \geq k] &\geq -\log \mathbf{P}[\mathsf{fix}(W_n) \geq k_1] \\ &\geq -2\log l - \frac{2k_1^2}{l^6 n} + \frac{k}{2l^4}\log\frac{k}{2el^4} \\ &\geq -2\log l - \frac{2k}{l^6} - \frac{k}{2l^4}\log 2el^4 + \frac{k\log k}{2l^4}. \end{aligned}$$

As $2\log l$, $\frac{2k}{l^6}$, $\frac{k}{2l^4}\log 2el^4$ all do not exceed $\frac{k\log k}{18l^4}$, we conclude that

$$-\log \mathbf{P}[\mathsf{fix}(W_n) \geq k] \geq \frac{k\log k}{3l^4},$$

as claimed. □

The following variant of Proposition 6.4, where $H$ is allowed to be any permutation group and $W_n$ is the random variable on $H$ corresponding to $w$, is needed in §7. We say $H$ is *$\epsilon$-roughly transitive* if for all $1 \leq i \leq t = n^{1-\epsilon}$, the size of every $H$-orbit of ordered $i$-tuples of pairwise distinct integers in $[1, n]$ is at least $t^i$.

**Proposition 6.5.** *Let $l$ be a positive integer and $0 < \epsilon < 1/4l$. If $n$ is sufficiently large in terms of $l$, $2el^4 n^{1/2} < k < n^{1-\epsilon}$, and $H < \mathsf{S}_n$ is $\epsilon$-roughly transitive, then*

$$\mathbf{P}[\mathsf{fix}(W_n) \geq k] \leq n^{-k/8l^4}.$$

*Proof.* The number of elements $g \in H$ satisfying $c$ conditions of the form $gu = v$ or $g^{-1}u = v$ is at most $|H|$ divided by the cardinality of the smallest $H$-orbit of a $c$-tuple $(s_1, \ldots, s_c)$ of pairwise distinct integers in $[1, n]$, which is bounded above by $t^{-c}|H|$. Thus, following the notation of Lemma 6.3,

$$|X_{m,n}(w, a, b)| \leq n^{m(b-a)}\frac{|H|^d}{t^{m(b-a)}} \leq (n/t)^{ml}|H|^d \leq n^{\epsilon ml}|H|^d.$$

Let $k_1$ be the largest multiple of $l^4$ which is bounded above by $k$. As in Proposition 6.4, $\mathsf{fix}(w(g_1, \ldots, g_d)) \geq k_1$ implies there exist at least $k_1/l^2$ elements satisfying (6.4.1) and (6.4.2) and therefore $(g_1, \ldots, g_d)$ lies in the projection of $X_{k_1/l^4,n}(w, a, b)$ for some $a, b$ with $1 \leq a < b \leq l$.

As $\epsilon < 1/4l$ and $k/2el^4 > \sqrt{n}$, if $n$ is sufficiently large,

$$\mathbf{P}[\mathsf{fix}(W_n) \geq k_1] \leq \frac{l^2 n^{\epsilon k_1/l^3}}{(k_1/l^4)!} \leq \frac{l^2 n^{\epsilon k/2l^3}}{(k_1/el^4)^{k_1/l^4}} \leq \frac{l^2 n^{\epsilon k/2l^3}}{(k/2el^4)^{k/2l^4}} \leq \frac{n^{k/8l^4}}{n^{k/4l^4}},$$

which gives the proposition. □

**Definition 6.6.** (i) Recall from the introduction that a word $w$ in the free group $F_d$ is *even* if $w \in \langle [F_d, F_d], x_1^2, \dots, x_d^2 \rangle$. Otherwise we say that $w$ is *odd*. We define $\gamma(w) = 1$ (resp. $\gamma(w) = 0$) when $w$ is even (resp. odd).

(ii) For $G = \mathsf{S}_n$, let $U_G^0 = U_G$, the uniform distribution on $G$, and let $U^1(g) = 2/|G|$ for $g \in \mathsf{A}_n$ and $U^1(g) = 0$ for $g \in G \smallsetminus \mathsf{A}_n$.

Note that if $w = w_1 w_2 \cdots w_N$ is a product of pairwise disjoint words, then

$$\gamma(w) = \gamma(w_1)\gamma(w_2)\cdots\gamma(w_N). \tag{6.3}$$

The relevance of Definition 6.6 follows from the following statement, where $\mathsf{sgn}$ denotes the sign character of $\mathsf{S}_n$:

**Lemma 6.7.** *Let $G = \mathsf{S}_n$, $w \in F_d$, and let $X$ be the random variable on $G$ with distribution $p_{w,G}$. Then*

$$\sum_C \mathbf{P}(X \in C)\mathsf{sgn}(C) = \gamma(w),$$

*where the summation runs over conjugacy classes in $G$.*

*Proof.* Note that the sum in question is $\Sigma = \mathbf{P}(X \in \mathsf{A}_n) - \mathbf{P}(X \notin \mathsf{A}_n)$. If $w$ is even, then $X$ is always in $\mathsf{A}_n$, whence $\Sigma = 1 = \gamma(w)$. If $w$ is odd, then half of the time $X$ belongs to $\mathsf{A}_n$ and half of the time it does not, whence $\Sigma = 0 = \gamma(w)$. □

**Proposition 6.8.** *Let $l, N \geq 1$ be integers such that $N > 8l^4 + 41$. If $w_1, w_2, \dots, w_N$ is a sequence of non-trivial words of length at most $l$, then we have*

$$\lim_{n\to\infty} \| \underbrace{p_{w_1,\mathsf{A}_n} * \cdots * p_{w_N,\mathsf{A}_n}}_{N} - U_{\mathsf{A}_n} \|_{L^\infty} = 0.$$

*Furthermore, for $\gamma = \gamma(w_1 w_2 \cdots w_N)$, we have*

$$\lim_{n\to\infty} \| \underbrace{p_{w_1,\mathsf{S}_n} * \cdots * p_{w_N,\mathsf{S}_n}}_{N} - U_{\mathsf{S}_n}^\gamma \|_{L^\infty} = 0.$$

*Proof.* For $G \in \{\mathsf{A}_n, \mathsf{S}_n\}$, let $X_{1,n}, \dots, X_{N,n}$ be independent random variables on $G$, with distribution $p_{w_1,G}, \dots, p_{w_N,G}$ which are invariant under conjugation in $G$. Let $X_n^N = X_{1,n} \cdots X_{N,n}$. For $g \in G$,

$$\mathbf{P}[X_n^N = g] = |G|^{-1} \sum_{C_1,\dots,C_N} \left( \prod_{i=1}^N \mathbf{P}[X_{i,n} \in C_i] \right) \sum_\chi \frac{\chi(C_1)\cdots\chi(C_N)\overline{\chi}(g)}{\chi(1)^{N-1}}, \tag{6.4}$$

where each $C_i$ in the first summation ranges over the conjugacy classes of $G$, and the second summation runs over all irreducible characters $\chi \in \mathrm{Irr}(G)$. In the second summation, the contribution of the $\chi = 1_G$ term is 1.

Suppose $G = \mathsf{S}_n$. Then, by Lemma 6.7 and (6.3), the contribution of $\chi = \mathsf{sgn}$ to (6.4) is

$$\begin{aligned} &\mathsf{sgn}(g)|G|^{-1} \sum_{C_1,\dots,C_N} \Big(\prod_{i=1}^N \mathbf{P}[X_{i,n} \in C_i]\mathsf{sgn}(C_i)\Big) = \\ &\mathsf{sgn}(g)|G|^{-1}\Big(\prod_{i=1}^N \sum_{C_i} \mathbf{P}[X_{i,n} \in C_i]\mathsf{sgn}(C_i)\Big) = \gamma \cdot \mathsf{sgn}(g)|G|^{-1}. \end{aligned} \tag{6.5}$$

Our goal is to show that the *error term* in (6.4), i.e., the contribution of the characters $\chi$ with $\chi(1) > 1$ to the sum, is $o(|G|^{-1})$. Indeed, in the case $G = \mathsf{A}_n$ we then have for all $g$:

$$|G| \cdot \big|\mathbf{P}[X_n^N = g] - U(\{g\})\big| = o(1).$$

Suppose $G = \mathsf{S}_n$. Then $U^\gamma(g) = (1 + \gamma \cdot \mathsf{sgn}(g))/|G|$, and so (6.5) yields

$$|G| \cdot \big|\mathbf{P}[X_n^N = g] - U^\gamma(\{g\})\big| = o(1).$$

We choose $\epsilon = 1/4$ in Proposition 6.1 and assume $n > A(\epsilon)$. We also assume that $n$ is large enough that $n^{3/4} > (el)^{36}$. For any conjugacy class $C_i$ in $G$ and any irreducible character $\chi$ of $G$, either $\mathsf{fix}(C_i) \le n^{3/4}$, in which case

$$\mathbf{P}[X_{i,n} \in C_i]|\chi(C_i)| \le \chi(1)^{11/12} \tag{6.6}$$

by Proposition 6.1(i), or $\mathsf{fix}(C_i) \ge n^{3/4} > (el)^{36}$, in which case

$$\mathbf{P}[X_{i,n} \in C_i]|\chi(C_i)| < n^{\mathsf{fix}(C_i)/4}\chi(1)^{3/4} \tag{6.7}$$

by Proposition 6.1(ii), and

$$\mathbf{P}[X_{i,n} \in C_i]|\chi(C_i)| < \mathsf{fix}(C_i)^{-\mathsf{fix}(C_i)/3l^4}\chi(1) < n^{-\mathsf{fix}(C_i)/4l^4}\chi(1) \tag{6.8}$$

by Proposition 6.4. We combine these inequalities by putting the right hand side of (6.7) to the $\frac{1}{l^4+2}$ power and the right hand side of (6.8) to the $\frac{l^4+1}{l^4+2}$ power and multiplying:

$$\mathbf{P}[X_{i,n} \in C_i]|\chi(C_i)| < n^{-\frac{\mathsf{fix}(C_i)}{4l^4+8}}\chi(1)^{\frac{4l^4+7}{4l^4+8}} \le n^{-\frac{n^{3/4}}{4l^4+8}}\chi(1)^{\frac{4l^4+7}{4l^4+8}}. \tag{6.9}$$

The error term

$$|G|^{-1} \sum_{C_1,\dots,C_N} \Big(\prod_{i=1}^N \mathbf{P}[X_{i,n} \in C_i]\Big) \sum_{\chi(1)>1} \frac{|\chi(C_1)\cdots\chi(C_N)\overline{\chi}(g)|}{\chi(1)^{N-1}}$$

can be rewritten as

$$|G|^{-1} \sum_{\chi(1)>1} |\overline{\chi}(g)|\chi(1) \sum_{C_1,\dots,C_N} \Big(\prod_{i=1}^N \frac{\mathbf{P}[X_{i,n} \in C_i]|\chi(C_i)|}{\chi(1)}\Big). \tag{6.10}$$

If $\mathsf{fix}(C_i) \geq n^{3/4}$ for at least 25 different values $i \in [1, N]$, then by (6.6),

$$\prod_{i=1}^{N} \frac{\mathbf{P}[X_{i,n} \in C_i]|\chi(C_i)|}{\chi(1)} \leq \chi(1)^{-25/12} \prod_{i=1}^{n} \mathbf{P}[X_{i,n} \in C_i].$$

Now,

$$\sum_{\chi(1)>1} |\overline{\chi}(g)|\chi(1) \cdot \chi(1)^{-25/12} \sum_{C_1,\ldots,C_N} \prod_{i=1}^{n} \mathbf{P}[X_{i,n} \in C_i] = \zeta_G(1/12) - [G : \mathsf{A}_n]$$

goes to zero by Theorem 2.6 and Corollary 2.7 of [LiSh2]. Therefore, the contribution of $N$-tuples $(C_1, \ldots, C_N)$ of which at least 25 of the $C_i$ have fixity $\leq n^{3/4}$ to the the error term (6.10) is $o(|G|^{-1})$. That leaves the $N$-tuples for which at least $8l^4 + 16$ classes $C_i$ have fixity $\geq n^{3/4}$. For these, the bound (6.9) gives

$$\prod_{i=1}^{N} \frac{\mathbf{P}[X_{i,n} \in C_i]|\chi(C_i)|}{\chi(1)} \leq n^{-2n^{3/4}} \chi(1)^{-2}.$$

The total number of ordered $N$-tuples of conjugacy classes in $G$ is at most $(2p(n))^N \leq e^{cN\sqrt{n}}$, where $p(n)$ denotes the partition function. Thus,

$$\sum_{\chi(1)>1} |\overline{\chi}(g)|\chi(1) \sum_{C_1,\ldots,C_N} n^{-2n^{3/4}} \chi(1)^{-2} = o(1),$$

which implies the proposition. □

**Proposition 6.9.** *Let $l$ and $N$ be positive integers such that $N > (1.5) \cdot 10^{10} l^2$. If $w_1, w_2, \ldots, w_N$ is a sequence of non-trivial words of length at most $l$, then we have*

$$\lim_{|G|\to\infty} \| \underbrace{p_{w_1,G} * \cdots * p_{w_N,G}}_{N} - U_G \|_{L^\infty} = 0$$

*if the limit is taken over finite simple groups $G$ of Lie type of rank $r \geq 7 \cdot 10^8 l^2$.*

*Proof.* We follow the method of Proposition 6.8, and let $W_i$ denote a random variable with values in $G$ and distribution $p_{w_i,G}$. We write $G$ as the quotient of $H = \mathrm{Cl}(V)$ by its center, where $\mathbb{F} = \mathbb{F}_{q^f}$, $f \leq 2$, and $V = \mathbb{F}^n$. Let $\tilde{W}_i$ denote a random variable on $H$ with distribution $p_{w_i,H}$. Let $\delta_0 = 0.0011$ and $\delta = 1/7400$. Note that $\delta(2n+2)^2 < \delta_0 n^2$, since $n \geq r$. Hence, by [GLT3, Theorem 1.4], if $|\mathbf{C}_G(g)| \leq q^{2f\delta n^2}$ then $|\chi(g)| \leq \chi(1)^{1-.008}$.

By Lemma 5.3,

$$\mathbf{P}[|\mathbf{C}_G(W_i)| \geq q^{2f\delta n^2}] \leq \mathbf{P}[|\mathbf{C}_H(\tilde{W}_i)| \geq q^{2f\delta n^2}].$$

By Lemma 5.4 (and the discussion prior to (5.2)), $\mathbf{P}[|\mathbf{C}_H(\tilde{W}_i)| \geq q^{2f\delta n^2}]$ does not exceed the probability that there exists a non-constant polynomial

$Q(x) \in \mathbb{F}[x]$ such that $\dim_{\mathbb{F}} \ker Q(\tilde{W}_i) \geq \delta n \deg Q$. As $n \geq r > 7 \cdot 10^8 l$, $n\delta \geq 9 \cdot 10^4 l$, and so

$$k = \lfloor n\delta/2l \rfloor \geq n\delta/3l + 3.$$

From Proposition 5.5, it follows that

$$\mathbf{P}[|\mathbf{C}_G(W_i)| \geq q^{2f\delta n^2}] \leq q^{-fk((k-1)lD-2.5)} \leq q^{-\delta^2 f n^2/9l^2} \leq |G|^{-\delta^2/9l^2}.$$

Setting $\epsilon = \frac{1}{5\cdot 10^8 l^2} < .008$, for every $i$ and conjugacy class $C_i$, we have that either $|\chi(C_i)| \leq \chi(1)^{1-\epsilon}$ for all $\chi \in \mathrm{Irr}(G)$, or $\mathbf{P}[W_i \in C_i] \leq |G|^{-\epsilon}$.

By hypothesis, $G$ has rank $r \geq 11/(8\epsilon)$; in particular $1.1Nr \leq (4/5)Nr^2\epsilon$. Recall that by [FG], the number $k(G)$ of conjugacy classes in $G$ is less than $27.2q^r < q^{r+5} < q^{1.1r}$, while $|G| \geq q^{hr}/2(r+1)$, where $h \geq r+2$ is the Coxeter number of $G$, so $|G| \geq q^{r^2}$. As $N \geq 11/\epsilon$ and using the obvious estimate $|\chi(g)\chi(1)| \leq |G|$, we have that

$$\begin{aligned}|\overline{\chi}(g)|\chi(1) \sum_{C_1,\ldots,C_N} |G|^{-(9/10)Nr^2\epsilon} &\leq q^{1.1Nr}|G|^{1-(9/10)N\epsilon} \\ &\leq q^{1.1Nr-(4/5)Nr^2\epsilon} q^{r^2(1-N\epsilon/10)} \leq q^{-0.1r^2}.\end{aligned}$$

It follows that

$$\sum_{\chi} |\overline{\chi}(g)|\chi(1) \sum_{C_1,\ldots,C_N} \prod_{i=1}^{N} \left( \frac{\mathbf{P}[X_{i,n} \in C_i]|\chi(C_i)|}{\chi(1)} \right) \leq q^{-0.1r^2},$$

if the inner sum is taken only over $N$-tuples $(C_1, \ldots, C_N)$ for which at least $0.9N$ of the classes satisfy $\mathbf{P}[W_i \in C_i] \leq |G|^{-\epsilon}$. Certainly, $q^{-0.1r^2} \to 0$ when $|G| \to \infty$. On the other hand, by [LiSh3, Theorem 1.1],

$$\lim_{|G|\to\infty} \zeta_G(1) - 1 = 0.$$

As $N > 30/\epsilon$, we obtain that

$$\begin{aligned}\sum_{\chi \neq 1_G} |\overline{\chi}(g)|\chi(1) \sum_{C_1,\ldots,C_N} \prod_{i=1}^{N} \left( \frac{\mathbf{P}[X_{i,n} \in C_i]|\chi(C_i)|}{\chi(1)} \right) &\leq \zeta_G\left(\frac{N\epsilon}{10} - 2\right) - 1 \\ &\leq \zeta_G(1) - 1,\end{aligned}$$

if the inner sum is taken over $N$ tuples $(C_1, \ldots, C_N)$ for which at least $0.1N$ classes satisfy $|\chi(C_i)| \leq \chi(1)^{1-\epsilon}$ for all irreducible $\chi$. □

**Proposition 6.10.** *Let $r$ and $N$ be positive integers such that $N \geq (2r+1)^2$ and $w_1, w_2, \ldots, w_N$ be any sequence of non-trivial words. Then*

$$\lim_{|G|\to\infty} \| \underbrace{p_{w_1,G} * \cdots * p_{w_N,G}}_{N} - U_G \|_{L^\infty} = 0$$

*if the limit is taken over finite simple groups $G$ of Lie type of rank $r$.*

*Proof.* We follow the method of Proposition 6.8. Let $W_i$ denote a random variable with values in $G$ and distribution $p_{w_i,G}$. By [FG], the number $k(G)$ of conjugacy classes in $G$ is $< 28q^r$. (Note $q$ need not be in $\mathbb{Z}$ since $G$ may be of Suzuki or Ree type.) By classification of root systems, $\dim G \leq 2r^2 + r$, where by a slight abuse of notation we write $\dim G$ for the dimension of the simply connected algebraic group associated to $G$. Thus,

$$\chi(1) \leq |G|^{1/2} = O(q^{r^2+r/2}).$$

By Gluck's bound [Gl] for irreducible character values of groups of Lie type, $|\chi(g)| = O(q^{-1/2}\chi(1))$ for all $g \neq 1$. On the other hand $\mathbf{P}[W_i = 1] = \frac{|w^{-1}(1)|}{|G|^d}$. By [LS3, Proposition 3.4], this is $O(q^{-1})$. Thus, for every conjugacy class $C_i$, either $C_i = \{1\}$, in which case $\mathbf{P}[W_i \in C_i] = O(q^{-1})$, or $C_i \neq \{1\}$, in which case $|\chi(C_i)|/\chi(1) = O(q^{-1/2})$.

Let $S$ denote any subset of $\{1, \ldots, N\}$, and let $C(S)$ denote $N$-tuples $(C_1, \ldots, C_N)$ such that $C_i = \{1\}$ if and only if $i \in S$. Then,

$$\sum_{(C_1,\ldots,C_N)\in C(S)} \prod_{i=1}^{n} \left( \frac{\mathbf{P}[W_i \in C_i]|\chi(C_i)|}{\chi(1)} \right) = O(q^{-|S|-\frac{N-|S|}{2}}) \leq O(q^{-N/2}).$$

As $|\mathrm{Irr}(G)| = k(G) < 28q^r$,

$$\sum_{\chi\neq 1_G} |\overline{\chi}(g)|\chi(1) \sum_{S} \sum_{(C_1,\ldots,C_N)\in C(S)} \prod_{i=1}^{n} \left( \frac{\mathbf{P}[W_i \in C_i]|\chi(C_i)|}{\chi(1)} \right) = O(q^{2r^2+2r-N/2})$$
$$\leq O(q^{-1/2}),$$

which proves the proposition, since $q \to \infty$ when $|G| \to \infty$. □

*Proof of Theorem 4.* By the classification of finite simple groups, it suffices to prove the result for alternating groups, groups of Lie type of rank greater than $7 \cdot 10^8 l^2$, and groups of Lie type of lower rank. These three cases are covered by Propositions 6.8, 6.9, and 6.10 respectively. □

## 7. Some applications

In this section we derive various applications, proving Theorem 5.

**Proposition 7.1.** *There is an absolute constant $0 < \epsilon < 1$ such that the following statement holds for any prime power $q$ and for any positive integer $n$. Let $G = \mathrm{Cl}(V)$ be a classical group in dimension $n$ (in the sense of §5), and let $P$ be a maximal subgroup of $G$ of order $|P| > |G|^{1-\epsilon}$ not containing $[G, G]$. Then there is a classical group $H$ in dimension $m$, with $m < n$ and a normal subgroup $K \lhd P$ with $|K| < [G : P]^3$ such that $P/K \cong H$.*

*Proof.* The smallest index of proper subgroups of a simple finite classical group is listed in [KlL, Table 5.2.A]. It follows for $n \leq 9$ that, if we take $0 < \epsilon < 1/10$ small enough, then any maximal subgroup $P$ of $G$ of order $|P| > |G|^{1-\epsilon}$ contains $[G, G]$. Hence in the rest of the proof we may assume

$n \geq 10$. By Theorems 1 and 2 of [Ka], there is some $\theta(n) \geq 1$ such that $P$ acts reducibly on $V$ whenever $[G : P] < q^{\theta(n)}$ and $P \not\geq [G, G]$; furthermore $\theta(n) > d(G)/5$ when $n$ is large enough, where $d(n)$ is the degree of $|[G, G]|$ as a polynomial of $q$. By taking $0 < \epsilon \leq 1/5$ small enough, we may therefore assume that $P$ stabilizes a subspace $U \subset V$ of dimension $0 < k < n$. The maximality of $P$ then implies that $P = \mathrm{Stab}_G(U)$, and furthermore, if $G$ respects a form $\langle \cdot, \cdot \rangle$, $U$ is either totally singular or non-degenerate with respect to $\langle \cdot, \cdot \rangle$.

First suppose that $\mathrm{SL}_n(\mathbb{F}_q) \leq G \leq \mathrm{GL}_n(\mathbb{F}_q)$. Then $[G : P] > q^{k(n-k)}$. Replacing the action of $P$ on $U$ by its action on $V/U$ if necessary, we may assume that $k \geq n/2$, and get a surjection from $P$ onto $H = \mathrm{GL}_k(\mathbb{F}_q)$, with kernel $K$ of order less than

$$q^{k(n-k)+(n-k)^2} = q^{(n-k)n} \leq q^{2k(n-k)} < [G : P]^2.$$

Next suppose that $\mathrm{SU}_n(\mathbb{F}_q) \leq G \leq \mathrm{U}_n(\mathbb{F}_q)$. If $U$ is non-degenerate, then replacing it by $U^\perp$ if necessary, we may assume that $k \geq n/2$. The action of $P$ on $U$ yields a surjection from $P \leq \mathrm{U}_k(\mathbb{F}_q) \times \mathrm{U}_{n-k}(\mathbb{F}_q)$ onto $H = \mathrm{U}_k(\mathbb{F}_q)$, with kernel $K \leq \mathrm{U}_{n-k}(q)$ of order less than

$$q^{(n-k)^2+1} \leq q^{2k(n-k)-3} < [G : P].$$

If $U$ is totally singular, then $k \leq n/2$ and

$$q^{2kn-3k^2-4} < [G : P] < q^{2kn-3k^2+1}.$$

By taking $0 < \epsilon \leq 1/5$ small enough, the condition $|P| > |G|^{1-\epsilon}$ implies that $n - 2k \geq 6$. Now the action of $P$ on $U^\perp/U$ yields a surjection from $P$ onto $H = \mathrm{U}_{n-2k}(q)$ with kernel $K$ of order less than

$$q^{k(2n-3k)+2k^2} \leq q^{3k(2n-3k)-12} < [G : P]^3.$$

The orthogonal and symplectic cases are handled in the same way. □

Recall that the group $\Gamma$ is defined in Theorem 5 as the group with generators $x_1, \ldots, x_d$ and a single relator $w = w_1 \cdots w_N$, where $w_i \in F_d$ are pairwise disjoint non-trivial words of length at most $l$.

**Proposition 7.2.** *If $N$ is sufficiently large in terms of $l$ then there exists $\epsilon > 0$ such that for all positive integers $n$, if $m < n^\epsilon/2$ divides $n$, then*

$$|\mathrm{Hom}(\Gamma, \mathsf{S}_{n/m} \wr \mathsf{S}_m)| = (1 + o(1))|\mathsf{S}_{n/m} \wr \mathsf{S}_m)|^{d-1}.$$

*Proof.* The proof is essentially that of Proposition 6.8, but it requires three estimates for $H = \mathsf{S}_{n/m} \wr \mathsf{S}_m$ analogous to those for $\mathsf{S}_n$ and $\mathsf{A}_n$ used in that proof: an upper bound on the probability that a random variable with distribution $w_* U_{H^d}$ takes values with more than $2el^4\sqrt{n}$ fixed points, an upper bound on values of irreducible characters $\chi$ of $H$ on elements with $\leq 2el^4\sqrt{n}$ fixed points, and an upper bound on $|\mathrm{Irr}(H)|$.

For the first, we fix $\epsilon < 1/4l$ and apply Proposition 6.5. Indeed, since $n/m \geq 2n^{1-\epsilon}$, the orbit $\mathcal{O}_{H_S}(u)$ of any element $u \in [1, n]$ under the pointwise stabilizer $H_S$ of any subset $S \subset [1, n] \setminus \{u\}$ with $|S| \leq n^{1-\epsilon}$, satisfies

$|\mathcal{O}_{H_S}(u)| \geq n^{1-\epsilon}$. The second is given by Lemma 6.2. The third follows from the classification of irreducible characters of $H$ used in the proof of Lemma 6.2; namely, each such character is determined by an ordered $m$-tuple of irreducible characters of $\mathsf{S}_{n/m}$ together with an irreducible character of $\mathsf{S}_m$. Thus, $|\mathrm{Irr}(H)| \leq p(n/m)^m p(m)$. As $m < n^\epsilon$, we conclude that

$$\log |\mathrm{Irr}(H)| = O(n^{\frac{1+\epsilon}{2}}).$$

□

**Lemma 7.3.** *Let $d \in \mathbb{Z}_{\geq 1}$ and let $\Delta$ be any $d$-generated group such that*

$$\lim_{|G|\to\infty} |G|^{1-d}|\mathrm{Hom}(\Delta, G)| = 1,$$

*where $G$ ranges over the finite simple groups. If $0 < \epsilon < 1/(d-2)$, then the probability that a homomorphism $\varphi\colon \Delta \to G$ chosen uniformly from $\mathrm{Hom}(\Delta, G)$ has the property that $\varphi(\Delta)$ is contained in a maximal subgroup $M$ of $G$ of index greater than $|G|^\epsilon$ goes to 0 as $|G| \to \infty$.*

*Proof.* For any finite simple group $G$, let $Q(G)$ denote the probability that a random homomorphism from $\Delta$ to $G$ is not an epimorphism. Then

$$Q(G) \leq \sum_{M \overset{\max}{<} G} |\mathrm{Hom}(\Delta, M)|/|\mathrm{Hom}(\Delta, G)|.$$

Since $\Delta$ is $d$-generated, we trivially have $|\mathrm{Hom}(\Delta, M)| \leq |M|^d$. Thus

$$Q(G) \leq (1 + o(1)) \sum_{M \overset{\max}{<} G} |M|^d/|G|^{d-1}.$$

By [LMS, Theorem 1.1],

$$\sum_{M \overset{\max}{<} G} [G : M]^{-2} \to 0 \tag{7.1}$$

as $G$ ranges over the finite simple groups. If $|M| \leq |G|^{1-1/(d-2)}$ then $|M|^d/|G|^{d-1} \leq [G : M]^{-2}$. This implies

$$\sum_{M \overset{\max}{<} G, |M| \leq |G|^{1-1/(d-2)}} |M|^d/|G|^{d-1} \to 0 \text{ as } |G| \to \infty.$$

□

**Proposition 7.4.** *Let $l$ be a positive integer and $w = w_1 \cdots w_N \in F_d$ be a product of pairwise disjoint non-trivial words $w_i$, each of length at most $l$. If $N$ is sufficiently large in terms of $l$, then there exists $0 < \epsilon < 1$ such that the following statement holds. For every finite classical group $G$ and every $P < G$ maximal among subgroups not containing $[G, G]$ with $|P| > |G|^{1-\epsilon}$, we have*

$$|w_P^{-1}(1)| \leq [G : P]^{-2}|G|^{d-1}.$$

*Proof.* By Proposition 7.1, we can find $0 < \epsilon < 1$ such that, given any $P$ as in the theorem, there exists a classical quotient group $H$ with

$$|P|/|H| < [G : P]^3, \tag{7.2}$$

and so

$$|w_P^{-1}(1)| \leq (|P|/|H|)^d \max_{h \in H} |w_H^{-1}(h)| < [G : P]^{3d} \max_{h \in H} |w_H^{-1}(h)|.$$

We claim that the right hand side is $O([G : P]^{3d+3}|H|^{d-1})$. It suffices to prove that the maximum of $|w_H^{-1}(h)|$ is $O(q^3|H|^{d-1})$. We follow the method of proof of Proposition 6.8. We start with the inequality

$$|w_H^{-1}(h)| \leq |H|^{d-1} \sum_{\chi \in \mathrm{Irr}(H)} \sum_{C_1,\ldots,C_N} \frac{|\chi(C_1) \cdots \chi(C_N)|}{\chi(1)^{N-2}} \prod_{i=1}^{N} \mathbf{P}[W_i \in C_i],$$

where the $W_i$ are independent random variables of $H$ with distribution $p_{w_i,H}$, and the $C_i$ are the conjugacy classes of $H$. We separate this sum into two pieces according to whether the restriction of $\chi$ to $[H, H]$ has a trivial constituent. The contribution of the characters whose restriction to $[H, H]$ has a trivial constituent is at most

$$\sum_{\chi \in \mathrm{Irr}(H/[H,H])} \chi(1)^2 = |H/[H, H]| < 2q.$$

Let $\mathrm{Irr}(H)^*$ denote the set of characters whose restriction to $[H, H]$ has no trivial constituent. If for some positive constant $\epsilon$ depending only on $l$ we have that for every conjugacy class $C_i$ of $H$ and every irreducible character $\chi$ of $H$, either $\mathbf{P}[W_i \in C_i] < |H|^{-\epsilon}$ or $|\chi(C_i)| \leq \chi(1)^{1-\epsilon}$, and, moreover,

$$\lim_{|H| \to \infty} \sum_{\chi \in \mathrm{Irr}(H)^*} \chi(1)^{-3} = 0, \tag{7.3}$$

then we can finish as in Proposition 6.9. The dichotomy for $C_i$ follows for general classical groups by the same argument as for finite simple classical groups since Proposition 5.5 and the character estimate [GLT3, Theorem 1.3] hold for classical groups in full generality.

To estimate the sum in (7.3), we choose a function $f : \mathrm{Irr}(H)^* \to \mathrm{Irr}([H, H])$ mapping each $\chi$ to a non-trivial irreducible character of $[H, H]$ which appears as a factor of the restriction of $\chi$ to $[H, H]$. Thus, $f(\chi)(1) \leq \chi(1)$, and $f$ is at most $|H/[H, H]| \leq 2q$ to 1. Thus, it suffices to prove

$$\lim_{|H| \to \infty} \sum_{1 \neq \chi \in \mathrm{Irr}([H,H])} 2q\chi(1)^{-3} \to 0,$$

and since the minimum degree of a non-trivial representation of $[H, H]$ is greater than $q/3$, this follows from [LiSh3, Theorem 1.1]. □

*Proof of Theorem 5.* In the proof of parts (i)–(iv) we take $N^*(l) = N(l)$ (defined in Theorem 4). In the proof of parts (v) and (vi) we have $N^*(l) \geq$

$N(l)$. Hence in any case we have $N^*(l) > 3$ for all $l$, which implies that $d > 3$.

(i) This follows immediately from Theorem 4 and Proposition 2.5.

(ii) As $\mathrm{Hom}(\Gamma, \underline{G})$ is the fiber of the word map $w_{\underline{G}}$ over the identity, this follows from (i) and the fact [EGA $\mathrm{IV}_2$, Corollaire 6.1.2] that all non-empty fibers of a flat morphism of affine varieties $\underline{X} \to \underline{Y}$ have dimension $\dim \underline{X} - \dim \underline{Y}$.

(iii) The quasi-finite morphisms (i.e., morphisms with finite fibers) $\mathrm{SL}_n \times \mathrm{GL}_1 \to \mathrm{GL}_n$ and $\mathrm{GL}_n \to \mathrm{PGL}_n \times \mathrm{GL}_1$ give rise to quasi-finite morphisms

$$\mathrm{Hom}(\Gamma, \mathrm{SL}_n) \times \mathrm{Hom}(\Gamma, \mathrm{GL}_1) \to \mathrm{Hom}(\Gamma, \mathrm{GL}_n) \to \mathrm{Hom}(\Gamma, \mathrm{PGL}_n) \times \mathrm{Hom}(\Gamma, \mathrm{GL}_1).$$

It follows from part (ii) above that

$$\dim \mathrm{Hom}(\Gamma, \mathrm{SL}_n) = \dim \mathrm{Hom}(\Gamma, \mathrm{PGL}_n) = (d-1)(n^2-1).$$

We therefore have the inequalities

$$\begin{aligned}(d-1)(n^2-1) + \dim \mathrm{Hom}(\Gamma, \mathrm{GL}_1) &\leq \dim \mathrm{Hom}(\Gamma, \mathrm{GL}_n) \\ &\leq (d-1)(n^2-1) + \dim \mathrm{Hom}(\Gamma, \mathrm{GL}_1),\end{aligned}$$

and so

$$\dim \mathrm{Hom}(\Gamma, \mathrm{GL}_n) = (d-1)(n^2-1) + \dim \mathrm{Hom}(\Gamma, \mathrm{GL}_1).$$

Note that $\dim \mathrm{Hom}(\Gamma, \mathrm{GL}_1)$ is $d$ or $d-1$ depending on whether $w$ belongs to $[F_d, F_d]$ or not. Thus $\dim \mathrm{Hom}(\Gamma, \mathrm{GL}_n) = (d-1)n^2 + a$ where $a = 0$ if $w \notin [F_d, F_d]$ and $a = 1$ otherwise.

(iv) Note that for any group $\Gamma$ we have

$$a_n(\Gamma) = |\mathrm{Hom}_{\mathrm{trans}}(\Gamma, \mathsf{S}_n)|/(n-1)!,$$

where $\mathrm{Hom}_{\mathrm{trans}}(\Gamma, \mathsf{S}_n)$ is the set of homomorphisms from $\Gamma$ to $\mathsf{S}_n$ with transitive image. See for instance [LuSe, 1.1.1].

By Proposition 6.8, we have for $\gamma = \gamma(w)$ (and so $b = 1 + \gamma$) that

$$\lim_{n \to \infty} \| p_{w, \mathsf{S}_n} - U^{\gamma}_{\mathsf{S}_n} \|_{L^\infty} = 0.$$

It follows that $P_{w, \mathsf{S}_n}(1) \sim b/n!$, hence

$$|\mathrm{Hom}(\Gamma, \mathsf{S}_n)| \sim b \cdot n!^{d-1}.$$

Now, the probability $Q_n$ that $\phi \in \mathrm{Hom}(\Gamma, \mathsf{S}_n)$ is not in $\mathrm{Hom}_{\mathrm{trans}}(\Gamma, \mathsf{S}_n)$ satisfies

$$Q_n \leq \sum_{1 \leq k \leq n/2} \binom{n}{k} |\mathrm{Hom}(\Gamma, M_k)|/|\mathrm{Hom}(\Gamma, \mathsf{S}_n)|,$$

where $M_k = \mathsf{S}_k \times \mathsf{S}_{n-k}$, the stabilizer of $\{1, \ldots, k\}$ in $\mathsf{S}_n$. We have

$$\begin{aligned}|\mathrm{Hom}(\Gamma, M_k)| &= |\mathrm{Hom}(\Gamma, \mathsf{S}_k)| \cdot |\mathrm{Hom}(\Gamma, \mathsf{S}_{n-k})| \\ &\leq (b + o_k(1))k!^{d-1} \cdot (b + o_n(1))(n-k)!^{d-1}.\end{aligned}$$

Therefore

$$Q_n \le \sum_{1\le k\le n/2} \binom{n}{k} \frac{(b+o_k(1))(b+o_n(1))}{b+o_n(1)} \cdot \frac{k!^{d-1}(n-k)!^{d-1}}{n!^{d-1}},$$

and since $d \ge 3$ we see that

$$Q_n \le \sum_{1\le k\le n/2} (b+o_k(1)) \binom{n}{k}^{-(d-2)} = O\Big( \sum_{1\le k\le n/2} \binom{n}{k}^{-1} \Big) \to 0 \text{ as } n \to \infty.$$

Therefore, as $n \to \infty$, almost all homomorphisms from $\Gamma$ to $\mathsf{S}_n$ have transitive image. This yields

$$|\mathrm{Hom}_{\mathrm{trans}}(\Gamma, \mathsf{S}_n)| \sim |\mathrm{Hom}(\Gamma, \mathsf{S}_n)| \sim bn!^{d-1}.$$

Dividing both sides by $(n-1)!$ we obtain $a_n(\Gamma) \sim bn \cdot n!^{d-2}$, proving the main assertion of Theorem 5(iv). To prove the second assertion, note that $a_n(F_{d-1}) \sim n \cdot n!^{d-2}$ (see [LuSe, 2.1]), and this yields $a_n(\Gamma)/a_n(F_{d-1}) \to b$ as $n \to \infty$.

To prove the remaining statements in Theorem 5, we use another consequence of Theorem 4 that

$$\lim_{|G|\to\infty} |G|^{1-d} |\mathrm{Hom}(\Gamma, G)| = 1$$

when $G$ runs over the finite simple groups. Thus Lemma 7.3 applies to $\Delta = \Gamma$.

(v) Denote by $\mathrm{Hom}_{\mathrm{prim}}(\Gamma, \mathsf{S}_n)$ the set of homomorphisms from $\Gamma$ to $\mathsf{S}_n$ with primitive image. Then we have

$$m_n(\Gamma) = |\mathrm{Hom}_{\mathrm{prim}}(\Gamma, \mathsf{S}_n)|/(n-1)!.$$

The argument is now similar to the one given above in (iv), except that we also have to take the maximal subgroups $M \cap \mathsf{A}_n$ with $M = \mathsf{S}_{n/k} \wr \mathsf{S}_k$ into account. Applying Lemma 7.3 to $\Delta = \Gamma$, we need only consider $k$-values which are less than $n^\epsilon$. By Proposition 7.2, the set of homomorphisms $\varphi\colon \Gamma \to \mathsf{A}_n$ for which there exists a partition into $k$ sets of cardinality $n/k$ which $\varphi(\Gamma)$ respects has cardinality less than

$$[\mathsf{S}_n : M]|\mathrm{Hom}(\Gamma, M)| = (1+o(1))n!|M|^{d-2} = O(n^{2-d}|\mathrm{Hom}(\Gamma, \mathsf{A}_n)|).$$

As $d > 3$, the sum over all $k$ values is $o(|\mathrm{Hom}(\Gamma, \mathsf{A}_n)|)$.

(vi) We apply Lemma 7.3 to $\Delta = \Gamma$ to bound the probability $Q(G)$ defined as in the proof of the lemma. First assume that $G$ is of Lie type of bounded rank $r$. Choosing $N$ (hence $d$) large, as we may, and using for instance Tables 5.2.A and 5.3.A of [KlL], we have that all the maximal subgroups of $G$ satisfy $|M| < |G|^{1-1/(d-1)}$. So by Lemma 7.3 we obtain that $Q(G) \to 0$ as $|G| \to \infty$ for such simple groups $G$.

For alternating groups $G = \mathsf{A}_n$, by Bochert's theorem (see [DM, 3.3B]), for $n$ sufficiently large, we need only consider maximal subgroups of the types considered in (iv) and (v), and so we are done in this case.

Now let $G$ be a simple classical group of large rank. Choosing $N$ (hence $d$) large, we may assume that Proposition 7.4 holds for a fixed $0 < \epsilon < 1/(d-2)$. If $\mathrm{Hom}(\Gamma, G)$ fails to be surjective, its image is contained in a subgroup $P$ of $G$ maximal among all subgroups not containing $[G, G] = G$. By Lemma 7.3, the probability of this event, but under the condition that $|P| \leq |G|^{1-\epsilon}$ tends to 0 when $|G| \to \infty$. So it remains to bound this probability under the condition that $|P| > |G|^{1-\epsilon}$. By Proposition 7.4, for each such $P$, the probability of this is less than $[G : P]^{-2}$. Together with (7.1), this implies that $Q(G) \to 0$ as $|G| \to \infty$. □

## References


[AA] A. Aizenbud and N. Avni, Representation growth and rational singularities of the moduli space of local systems, *Invent. Math.* **204** (2016), 245–316.

[BC] V. V. Benyash-Krivets and V. I. Chernousov, Varieties of representations of fundamental groups of compact non-oriented surfaces. (Russian), *Mat. Sb.* **188**, 47–92 (1997); translation in *Sb. Math.* **188**, 997–1039 (1997).

[Bor] A. Borel, On free subgroups of semisimple groups, *Enseign. Math.* **29** (1983), 151–164.

[CH] W. Cocke and M.-C. Ho, The probability distribution of word maps on finite groups, arXiv:1807.07111.

[De] P. Deligne, La conjecture de Weil. II, *Inst. Hautes Études Sci. Publ. Math.* **52** (1980), 137–252.

[DM] J. D. Dixon and B. Mortimer, '*Permutation Groups*', Graduate Texts in Mathematics **163**, Springer Verlag, New York, 1996.

[ET] P. Erdős and P. Turán, On some problems of a statistical group theory. I, *Z. Wahrsch. Verw. Gabiete* **4** (1965), 175–186.

[Fu] K. Fujiwara, Rigid geometry, Lefschetz-Verdier trace formula and Deligne's conjecture, *Invent. Math.* **127** (1997), no. 3, 489–533.

[FG] J. Fulman and R. M. Guralnick, Bounds on the number and sizes of conjugacy classes in finite Chevalley groups with applications to derangements, *Trans. Amer. Math. Soc.* **364** (2012), 3023–3070.

[GLL] S. Garion, M. Larsen and A. Lubotzky, Beauville surfaces and finite simple groups, *J. Reine Angew. Math.* **666** (2012), 225–243.

[GS] S. Garion and A. Shalev, Commutator maps, measure preservation, and $T$-systems, *Trans. Amer. Math. Soc.* **361** (2009), 4631–4651.

[Gl] D. Gluck, Character value estimates for nonsemisimple elements, *J. Algebra* **155** (1993), no. 1, 221–237.

[Go] W. M. Goldman, Topological components of spaces of representations, *Invent. Math.* **93**, 557–607 (1988).

[EGA IV$_1$] A. Grothendieck, Éléments de géométrie algébrique. IV. Étude locale des schémas et des morphismes de schémas. I. Inst. Hautes Études Sci. Publ. Math. No. **20**, 1964.

[EGA IV$_2$] A. Grothendieck, Éléments de géométrie algébrique. IV. Étude locale des schémas et des morphismes de schémas. II. Inst. Hautes Études Sci. Publ. Math. No. **24**, 1965.

[EGA IV$_3$] A. Grothendieck, Éléments de géométrie algébrique. IV. Étude locale des schémas et des morphismes de schémas. III. Inst. Hautes Études Sci. Publ. Math. No. **28**, 1966.

[EGA IV$_4$] A. Grothendieck, Éléments de géométrie algébrique. IV. Étude locale des schémas et des morphismes de schémas. IV. Inst. Hautes Études Sci. Publ. Math. No. **32**, 1967.

[GLT1] R. M. Guralnick, M. Larsen, and Pham Huu Tiep, Representation growth in positive characteristic and conjugacy classes of maximal subgroups, *Duke Math. J.* **161** (2012), 107–137.

[GLT2] R. M. Guralnick, M. Larsen, and Pham Huu Tiep, Character levels and character bounds, arXiv:1708.03844.

[GLT3] R. M. Guralnick, M. Larsen, and Pham Huu Tiep, Character levels and character bounds. II, arXiv:1904.08070.

[GT] R. M. Guralnick and Pham Huu Tiep, Lifting in Frattini covers and a characterization of finite solvable groups, *J. Reine Angew. Math.* **708** (2015), 49–72.

[Hu] J. E. Humphreys, '*Conjugacy Classes in Semisimple Algebraic Groups*', Mathematical Surveys and Monographs, **43**. American Mathematical Society, Providence, RI, 1995.

[JK] G. James and A. Kerber, '*The Representation Theory of the Symmetric Group*', Encyclopedia of Mathematics and its Applications, vol. **16**, Addison-Wesley Publishing Co., Reading, Mass., 1981.

[Ka] W. M. Kantor, Permutation representations of the finite classical groups of small degree or rank, *J. Algebra* **60** (1979), 158–168.

[KlL] P. B. Kleidman and M. W. Liebeck, '*The Subgroup Structure of the Finite Classical Groups*', London Math. Soc. Lecture Note Ser. no. **129**, Cambridge University Press, 1990.

[LS1] M. Larsen and A. Shalev, Characters of symmetric groups: sharp bounds and applications, *Invent. Math.* **174** (2008), 645–687.

[LS2] M. Larsen and A. Shalev, Word maps and Waring type problems, *J. Amer. Math. Soc.* **22** (2009), 437–466.

[LS3] M. Larsen and A. Shalev, Fibers of word maps and some applications, *J. Algebra* **354** (2012), 36–48.

[LS4] M. Larsen and A. Shalev, On the distribution of values of certain word maps, *Trans. Amer. Math. Soc.* **368** (2016), 1647–1661.

[LST] M. Larsen, A. Shalev and Pham Huu Tiep, The Waring problem for finite simple groups, *Annals of Math.* **174** (2011), 1885–1950.

[LMS] M. W. Liebeck, B. M. Martin and A. Shalev, On conjugacy classes of maximal subgroups of finite simple groups, and a related zeta function, *Duke Math. J.* **128** (2005), 541–557.

[LiSh1] M. W. Liebeck and A. Shalev, The probability of generating a finite simple group, *Geom. Dedicata* **56** (1995), 103–113.

[LiSh2] M. W. Liebeck and A. Shalev, Fuchsian groups, coverings of Riemann surfaces, subgroup growth, random quotients and random walks, *J. Algebra* **276** (2004), 552–601.

[LiSh3] M. W. Liebeck and A. Shalev, Fuchsian groups, finite simple groups, and representation varieties, *Invent. Math.* **159** (2005), 317–367.

[LiSh4] M. W. Liebeck and A. Shalev, Character degrees and random walks in finite groups of Lie type, *Proc. London Math. Soc.* **90** (2005), 61–86.

[LuSe] A. Lubotzky and D. Segal, '*Subgroup Growth*', Progress in Math. **212**, Birkhäuser Verlag, Basel, 2003.

[MP] T. W. Müller and J.-C. Puchta, Character theory of symmetric groups and subgroup growth of surface groups, *J. London Math. Soc.* **66** (2002), 623–640.

[RBC] A. S. Rapinchuk, V. V. Benyash-Krivetz and V. I. Chernousov, Representation varieties of the fundamental groups of compact orientable surfaces, *Isr. J. Math.* **93**, 29–71 (1996).

[Seg] D. Segal, '*Words: Notes on Verbal Width in Groups*', London Math. Soc. Lecture Note Series **361**, Cambridge University Press, Cambridge, 2009.

[Se] J-P. Serre, Zeta and L functions, 1965 Arithmetical Algebraic Geometry (Proc. Conf. Purdue Univ., 1963), pp. 82–92, Harper & Row, New York.

[Sh1] A. Shalev, Word maps, conjugacy classes, and a non-commutative Waring-type theorem, *Annals of Math.* **170** (2009), 1383–1416.

[Sh2] A. Shalev, Some problems and results in the theory of word maps, *Erdős Centennial*, eds: Lovász et al., Bolyai Soc. Math. Studies **25** (2013), pp. 611–649.

*E-mail address*: mjlarsen@indiana.edu

Department of Mathematics, Indiana University, Bloomington, IN 47405, U.S.A.

*E-mail address*: shalev@math.huji.ac.il

Einstein Institute of Mathematics, Hebrew University, Givat Ram, Jerusalem 91904, Israel

*E-mail address*: pht19@math.rutgers.edu

Department of Mathematics, Rutgers University, Piscataway, NJ 08854-8019, U.S.A.