ON THE FOURIER COEFFICIENTS OF WORD MAPS ON UNITARY GROUPS

NIR AVNI AND ITAY GLAZER

In memory of Steve Zelditch

Abstract. Given a word $w(x_1,\ldots,x_r)$, i.e., an element in the free group on r elements, and an integer $d\geq 1$, we study the characteristic polynomial of the random matrix $w(X_1,\ldots,X_r)$, where X_i are Haar-random independent $d\times d$ unitary matrices. If $c_m(X)$ denotes the m-th coefficient of the characteristic polynomial of X, our main theorem implies that there is a positive constant $\epsilon(w)$, depending only on w, such that

$$\left|\mathbb{E}\left(c_m\left(w(X_1,\ldots,X_r)\right)\right)\right|\leq\binom{d}{m}^{1-\epsilon(w)},$$

for every d and every $1\leq m\leq d$.

Our main computational tool is the Weingarten Calculus, which allows us to express integrals on unitary groups such as the expectation above, as certain sums on symmetric groups. We exploit a hidden symmetry to find cancellations in the sum expressing $\mathbb{E}\left(c_m(w)\right)$. These cancellations, coming from averaging a Weingarten function over cosets, follow from Schur's orthogonality relations.

1. Introduction

Let w be a word on r letters, i.e., an element in the free group on the letters $x_1,\ldots,x_r$. Let $X_1,\ldots,X_r$ be random $d\times d$ unitary matrices, chosen independently at random according to the Haar probability measure, and consider the random matrix $w(X_1,\ldots,X_r)$, obtained by substituting X_i for x_i in w. For example, if $w=x_1x_2x_1^{-1}x_2^{-1}$, then $w(X_1,X_2)=X_1X_2X_1^{-1}X_2^{-1}$. In this paper, we study the distribution of the characteristic polynomial of $w(X_1,\ldots,X_r)$.

To set notation, given a $d\times d$-matrix A and $1\leq m\leq d$, let $c_m(A)$ be the coefficient of t^{d-m} in the characteristic polynomial $\det(t\cdot\mathrm{Id}-A)$ of A. Note that $c_m(A)=(-1)^m\mathrm{tr}\left(\bigwedge^m A\right)$, where $\bigwedge^m A:\bigwedge^m\mathbb{C}^d\to\bigwedge^m\mathbb{C}^d$ is the m-th exterior power of A. If A is unitary, all eigenvalues have absolute value 1, so we get the trivial bound $|c_m(A)|\leq\binom{d}{m}$.

Our main theorem is the following:

Theorem 1.1. *For every non-trivial word $w\in F_r$, there exists a constant $\epsilon(w)>0$ such that*

$$\mathbb{E}\left(\left|c_m\left(w(X_1,\ldots,X_r)\right)\right|^2\right)\leq\binom{d}{m}^{2(1-\epsilon(w))},$$

for every d and every $1\leq m\leq d$. In particular, we have

$$\mathbb{E}\left(\left|c_m\left(w(X_1,\ldots,X_r)\right)\right|\right)\leq\binom{d}{m}^{1-\epsilon(w)}.$$

Remark 1.2.

(1) In the proof of Theorem 1.1, we show that, if the length of w is ℓ and $d\geq(25\ell)^{7\ell}$, then one can take $\epsilon(w)=\frac{1}{72}(25\ell)^{-2\ell}$. We believe $\epsilon(w)^{-1}$ can be taken to be a polynomial in ℓ, for $d\gg_\ell 1$.

2020 *Mathematics Subject Classification.* Primary 60B15, 60B20; Secondary 43A75, 20P05, 20B30.
Key words and phrases. Word maps, word measures, unitary groups, Weingarten Calculus, Fourier coefficients, random matrices, characteristic polynomial.

(2) On the other hand, it follows from [ET15, Theorem 5.2] that, for a fixed d, one has to take $\epsilon(w) \lesssim e^{-\sqrt{\ell}}$, for some arbitrarily long words, even for $m = 1$.

Theorem 1.1 relies on the following:

Theorem 1.3. *For every $m, \ell \in \mathbb{N}$, every $d \geq m\ell$, and every word $w \in F_r$ of length ℓ, one has:*

$$\mathbb{E}\left(|c_m(w(X_1, \dots, X_r))|^2\right) \leq (22\ell)^{m\ell}. \tag{1.1}$$

In particular, if $d \geq (22\ell)^{\ell} m$, we have

$$\mathbb{E}\left(|c_m(w(X_1, \dots, X_r))|^2\right) \leq \binom{d}{m}.$$

In addition, we show that similar bounds hold for symmetric powers:

Theorem 1.4. *For every $\ell \in \mathbb{N}$, every $d \geq m\ell$, and every word $w \in F_r$ of length ℓ, one has:*

$$\mathbb{E}\left(|\mathrm{tr}(\mathrm{Sym}^m w(X_1, \dots, X_r))|^2\right) \leq (16\ell)^{m\ell}.$$

In particular, if $d \geq (16\ell)^{\ell} m$, we have

$$\mathbb{E}\left(|\mathrm{tr}(\mathrm{Sym}^m w(X_1, \dots, X_r))|^2\right) \leq \binom{d+m-1}{m} = \dim \mathrm{Sym}^m \mathbb{C}^d,$$

and by the Cauchy–Schwarz inequality,

$$|\mathbb{E}(\mathrm{tr}(\mathrm{Sym}^m w(X_1, \dots, X_r)))| \leq \left(\dim \mathrm{Sym}^m \mathbb{C}^d\right)^{\frac{1}{2}}.$$

Remark 1.5. Theorem 1.4 is an analogue of Theorem 1.3. It is also an analogue of Theorem 1.1 for m at most linear in d. Contrary to exterior powers, the methods of this paper are insufficient for finding bounds similar to Theorem 1.1 for $|\mathbb{E}(\mathrm{tr}(\mathrm{Sym}^m w(X_1, \dots, X_r)))|$, in the regime where m is superlinear in d.

1.1. Related work.

Word maps on unitary groups and their eigenvalues have been extensively studied in the past few decades.

The case $w = x$, namely, the study of a Haar-random unitary matrix X, also known as the Circular Unitary Ensemble (CUE), is an important object of study in random matrix theory (see e.g. [AGZ10, Mec19] and the references within). The joint density of the eigenvalues of X is given by the Weyl Integration Formula [Wey39]. Schur's orthogonality relations immediately imply that $\mathbb{E}\left(|c_m(X)|^2\right) = 1$ for all $1 \leq m \leq d$. Various other properties of the characteristic polynomial of a random unitary matrix X have been extensively studied (see e.g. [KS00, HKO01, CFK+03, DG06, BG06, BHNY08, ABB17, CMN18, PZ18]).

Diaconis and Shahshahani [DS94] have shown that, for a fixed $m \in \mathbb{N}$, the sequence of random variables $\mathrm{tr}(X), \mathrm{tr}(X^2), \dots, \mathrm{tr}(X^m)$ converges in distribution, as $d \to \infty$, to a sequence of independent complex normal random variables. For the proof, which relies on the moment method, they computed the joint moments of those random variables and showed that

$$\mathbb{E}\left(\prod_{j=1}^{m} \mathrm{tr}(X^j)^{a_j} \mathrm{tr}(\overline{X}^j)^{b_j}\right) = \delta_{a,b} \prod_{j=1}^{m} j^{a_j} a_j!, \tag{1.2}$$

for $d \geq \sum_{j=1}^{m}(a_j + b_j)j$. The rate of convergence was later shown to be super-exponential by Johansson [Joh97].

When $w = x^\ell$, (1.2) gives a formula for the moments of traces, and one can use Newton's identities relating elementary symmetric polynomials and power sums, to deduce that

$$\mathbb{E}\left(\left|c_m(X^\ell)\right|^2\right) = \mathbb{E}\left(\left|\mathrm{tr}\left(\mathrm{Sym}^m w\right)\right|^2\right) = \binom{\ell+m-1}{m},$$

for $d \geq 2m\ell$ (see Appendix A). In [Rai97, Rai03], Rains partially extended (1.2) for small d and gave an explicit formula for the joint density of the eigenvalues of X^ℓ (see [Rai03, Theorem 1.3]).

We now move to general words $w \in F_r$. The case $m = 1$, namely, the asymptotics as $d \to \infty$ of the distribution of the random variable $\mathrm{tr}\left(w(X_1, \dots, X_r)\right)$, was studied in the context of Voiculescu's free probability (see e.g. [VDN92, MS17]). In particular, in [Voi91, R0̆6, MSS07] it was shown that, for a fixed $w \in F_r$, the sequence of random variables $\mathrm{tr}\left(w(X_1, \dots, X_r)\right)$, for $d = 1, 2, \dots$, converges in distribution, as $d \to \infty$, to a complex normal random variable (with suitable normalization). As a direct consequence, for a fixed $m \in \mathbb{N}$, the random variables $c_m\left(w(X_1, \dots, X_r)\right)$ converge, as $d \to \infty$, to a certain explicit polynomial of Gaussian random variables. This is done in Appendix A, Corollary A.4, following [DG06].

In [MP19], Magee and Puder have shown that $\mathbb{E}\left(\mathrm{tr}\left(w(X_1, \dots, X_r)\right)\right)$ coincides with a rational function of d, if d is sufficiently large, and bounded its degree in terms of the commutator length of w. They also found a geometric interpretation for the coefficients of the expansion of that rational function as a power series in d^{-1}, see [MP19, Corollaries 1.8 and 1.11]. See [Bro24] for additional work in this direction.

1.2. **Ideas of proofs.** With a few exceptions, the results stated in §1.1 are asymptotic in d, but not uniform in both m and d. We will try to explain some of the challenges in proving results that are uniform in m, while explaining the idea of the proof of Theorem 1.1.

Our main tool (which is also used in the papers [R0̆6, MSS07, MP19] above) to study integrals on unitary groups is the *Weingarten Calculus* ([Wei78, Col03, CS06]). Roughly speaking, the Weingarten Calculus utilizes the Schur–Weyl Duality to express integrals on unitary groups as sums of so called Weingarten functions over symmetric groups. In our case, in order to prove Theorem 1.1, we need to estimate the integral

$$\mathbb{E}\left(|c_m(w)|^2\right) = \int_{\mathrm{U}_d^r} \left|\mathrm{tr}\left(\bigwedge\nolimits^m w(X_1, \dots, X_r)\right)\right|^2 dX_1 \dots dX_r. \tag{1.3}$$

Using Weingarten calculus (Theorem 2.12), we express (1.3) as a finite sum

$$\sum_{(\pi_1,\dots,\pi_{2r}) \in \prod_{i=1}^{2r} S_{m\ell_i}} F(\pi_1, \dots, \pi_{2r}) \prod_{i=1}^{r} \mathrm{Wg}_d^{(i)}(\pi_i \pi_{i+r}^{-1}), \tag{1.4}$$

where $\ell_1, \dots, \ell_{2r} \in \mathbb{N}$ and $F : \prod_{i=1}^{2r} S_{m\ell_i} \to \mathbb{Z}$ are related to combinatorial properties of w, and each $\mathrm{Wg}_d^{(i)} : S_{m\ell_i} \to \mathbb{R}$ is a *Weingarten function* (see Definition 2.10). There are two main difficulties when dealing with sums such as (1.4) in the region when m is unbounded:

(1) While the asymptotics of Weingarten functions $\mathrm{Wg}_d : S_m \to \mathbb{R}$ are well understood when $d \gg m$ (see [Col03, Section 2.2] and [CM17, Therem 1.1]), much less is known in the regime where m is comparable with d.
(2) Even if we have a good understanding of a single Weingarten function, the number of summands in (1.4) is large and it is not enough to bound each individual Weingarten function.

Luckily, there are plenty of cancellations in the sum (1.4). To understand these cancellations, we identify a symmetry of (1.4). More precisely, we find a group H acting on $\prod_{i=1}^{2r} S_{m\ell_i}$ such that F is equivariant with respect to H, and such that the contribution of any H-orbit to the sum (1.4) is a product of terms, each of which has the form

$$\frac{1}{m!^{2\ell_i}} \sum_{h,h' \in S_m^{\ell_i}} \operatorname{sgn}(hh') \operatorname{Wg}_d^{(i)}\left(h'\pi_i h \pi_{i+r}^{-1}\right), \tag{1.5}$$

where $\operatorname{sgn}(x)$ is the sign of x and the sum is over the Young subgroup $S_m^{\ell_i} \subseteq S_{m\ell_i}$, see Corollary 5.3. Weingarten functions are class functions, so they are linear combinations of irreducible characters of $S_{m\ell_i}$. Explicitly, we have (see [CS06, Eq. (13)]):

$$\operatorname{Wg}_d^{(i)}(\sigma) = \frac{1}{(m\ell_i)!^2} \sum_{\lambda \vdash m\ell_i, \ell(\lambda) \leq d} \frac{\chi_\lambda(1)^2}{\rho_\lambda(1)} \chi_\lambda(\sigma), \qquad \sigma \in S_{m\ell_i}, \tag{1.6}$$

where each λ is a partition of $m\ell_i$ with at most d parts and χ_λ and ρ_λ are the corresponding irreducible characters of $S_{m\ell_i}$ and U_d, respectively. The cancellations that we get in the sum (1.5) come from averaging irreducible characters of $S_{m\ell_i}$ over $S_m^{\ell_i}$-cosets. $S_m^{\ell_i}$ is a large subgroup of $S_{m\ell_i}$, so these cancellations will be significant as well. For example, all terms in (1.6) for which λ has more than ℓ_i columns vanish. See Lemmas 2.7 and 2.8 for the precise bounds.

After we bound the average contribution of each H-orbit in the sum (1.4) by a function $C(m,d,w)$, we bound (1.4) by $|Z| \cdot C(m,d,w)$ for some finite set Z. This becomes a counting problem, which we solve in §6, see Proposition 6.1.

The proof of Theorem 1.1 occupies Sections 4, 5, 6 and 7. Since the combinatorics of general words is a bit complicated, we prove a simplified version of Theorem 1.3 for the special case of the Engel word $[[x,y],y]$ in §3. The proof for this special case contains the main ideas of the paper, while being easier to understand.

1.3. **Further discussion and some open questions.** The results of this paper fit in the larger framework of the study of word measures and their Fourier coefficients.

Let G be a compact group, and let μ_G be the Haar probability measure on G. To each word $w(x_1,\ldots,x_r) \in F_r$ we associate the corresponding *word map* $w_G : G^r \to G$, defined by $(g_1,\ldots,g_r) \mapsto w(g_1,\ldots,g_r)$. The pushforward measure $(w_G)_*(\mu_G^r)$ is called the *word measure* $\tau_{w,G}$ associated with w and G. Let $\operatorname{Irr}(G)$ be the set of irreducible characters of G. The *Fourier coefficient* of $\tau_{w,G}$ at $\rho \in \operatorname{Irr}(G)$ is

$$a_{w,G,\rho} := \int_{G^r} \rho(w(x_1,\ldots,x_r))\mu_G^r = \int_G \rho(y)\tau_{w,G}. \tag{1.7}$$

If $w \neq 1$ and G is a compact semisimple Lie group, then by Borel's theorem [Bor83], the map $w_G : G^r \to G$ is a submersion outside a proper subvariety in G^r. It follows that $\tau_{w,G}$ is absolutely continuous with respect to μ_G and, therefore, $\tau_{w,G} = f_{w,G} \cdot \mu_G$, where $f_{w,G} \in L^1(G)$ is the Radon–Nikodym density. In this case, $f_{w,G} = \sum_{\rho \in \operatorname{Irr}(G)} \overline{a_{w,G,\rho}} \cdot \rho$.

In [LST19, Theorem 4], Larsen, Shalev, and Tiep proved uniform L^∞-mixing time for convolutions of word measures on sufficiently large finite simple groups. From this, the following can be deduced:

Theorem 1.6. *For every $w \in F_r$, there exists $N(w) \in \mathbb{N}$ such that if G is a finite simple group with at least $N(w)$ elements, then*

$$|a_{w,G,\rho}| \leq (\dim \rho)^{1-\epsilon(w)}, \tag{1.8}$$

for $\epsilon(w) = C \cdot \ell(w)^{-4}$ and some absolute constant C.

The proof of Theorem 1.6 is given at the end of §7.

We believe that a similar statement should be true for compact semisimple Lie groups.

Conjecture 1.7. *For every $1 \neq w \in F_r$, there exists $\epsilon(w) > 0$ such that, for every compact connected semisimple Lie group G and every $\rho \in \mathrm{Irr}(G)$,*

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

It is natural to estimate $\epsilon(w)$ in terms of the length $\ell(w)$ of the word w. For simple groups of bounded rank, Item (2) of Remark 1.2 (i.e. [ET15, Theorem 5.2]) shows that there are arbitrarily long words w for which $\epsilon(w)$ cannot be larger than $e^{-\sqrt{\ell(w)}}$. However, we believe that better Fourier decay can be achieved for the high rank case.

Question 1.8. *Can one take $\epsilon(w)$ to be a polynomial in $\ell(w)$, if $\mathrm{rk}(G) \gg_{\ell(w)} 1$?*

Theorem 1.1 gives evidence to Conjecture 1.7 for $G = \mathrm{SU}_d$ and the collection of *fundamental representations* $\left\{\bigwedge^m \mathbb{C}^d\right\}_{m=1}^d$. Indeed, for every $\rho \in \mathrm{Irr}(\mathrm{U}_d)$, since $|\rho(\lambda A)| = |\rho(A)|$ for $A \in \mathrm{SU}_d$ and $\lambda \in \mathrm{U}_1$, and since μ_{U_d} is the pushforward of $\mu_{\mathrm{U}_1} \times \mu_{\mathrm{SU}_d}$ by the multiplication map $(\lambda, A) \mapsto \lambda A$, we have,

$$\begin{aligned}
(1.9) \qquad |a_{w,\mathrm{SU}_d,\rho}|^2 &\leq \mathbb{E}_{X_1,\ldots,X_r \in \mathrm{SU}_d} \left(|\rho\left(w(X_1,\ldots,X_r)\right)|^2 \right) \\
&= \mathbb{E}_{(\lambda_1,X_1),\ldots,(\lambda_r,X_r) \in \mathrm{SU}_d \times \mathrm{U}_1} \left(|\rho\left(w(\lambda_1,\ldots,\lambda_r) w(X_1,\ldots,X_r)\right)|^2 \right) \\
&= \mathbb{E}_{(\lambda_1,X_1),\ldots,(\lambda_r,X_r) \in \mathrm{SU}_d \times \mathrm{U}_1} \left(|\rho\left(w(\lambda_1 X_1,\ldots,\lambda_r X_r)\right)|^2 \right) = \mathbb{E}_{\mathrm{U}_d} \left(|\rho\left(w(X_1,\ldots,X_r)\right)|^2 \right),
\end{aligned}$$

Theorem 1.4 deals with another family of irreducible representations $\left\{\mathrm{Sym}^m \mathbb{C}^d\right\}_{m=1}^{\lfloor d/(16\ell)^\ell \rfloor}$, giving further evidence for Conjecture 1.7.

Verifying Conjecture 1.7 will imply that, for every word w, the random walks induced by the collection of measures $\{\tau_{w,G}\}_G$, where G runs over all compact connected simple Lie groups, admit a uniform L^∞-mixing time. Namely, using [GLM12, Theorem 1], it will show the existence of $t(w) \in \mathbb{N}$ such that

$$(1.10) \qquad \left\| \frac{\tau_{w,G}^{*t(w)}}{\mu_G} - 1 \right\|_\infty < 1/2,$$

for every compact connected simple Lie group G. By the above discussion, $t(w)$ grows at least exponentially with $\sqrt{\ell(w)}$ under no restriction on the rank. If the condition (1.10) is replaced by the condition that $\tau_{w,G}^{*t(w)}$ has bounded density, one might hope for polynomial bounds.

Question 1.9. *Let $1 \neq w \in F_r$. Can one find $t(w) \in \mathbb{N}$ such that for every compact connected semisimple Lie group G, $\tau_{w,G}^{*t(w)}$ has bounded density with respect to μ_G? can $t(w)$ be chosen to have polynomial dependence on $\ell(w)$?*

Question 1.9 can be seen as an analytic specialization of a geometric phenomenon. Let $\varphi : X \to Y$ be a polynomial map between smooth $\mathbb{Q}$-varieties. We say that φ is *(FRS)* if it is flat and its fibers all have rational singularities. In [AA16, Theorem 3.4], Aizenbud and the first author showed that if φ is (FRS), then for every non-Archimedean local field F and every smooth, compactly supported measure μ on $X(F)$, the pushforward $\varphi_*\mu$ has bounded density. This result was extended in [Rei] to the Archimedean case, $F = \mathbb{R}$ or $\mathbb{C}$, and, moreover, if one runs over a large enough family of local fields, the condition of (FRS) is in fact necessary as well for the densities of pushforwards to be bounded (see [AA16, Theorem 3.4] and [GHS, Corollary 6.2]).

To rephrase Question 1.9 in geometric term, we further need the following notion from [GH19, GH21].

Definition 1.10 ([GH19, Definition 1.1]). Let $\varphi : X \to G$ and $\psi : Y \to G$ be morphisms from algebraic varieties X, Y to an algebraic group G. We define their *convolution* by

$$\varphi * \psi : X \times Y \to G,\ (x, y) \mapsto \varphi(x) \cdot \psi(y).$$

We denote by $\varphi^{*k} : X^k \to G$ the *k-fold convolution* of φ with itself.

Based on the above discussion, a positive answer to the following question will answer Question 1.9 positively.

Question 1.11 ([GH24, Question 1.15]). *Can we find $\alpha, C > 0$ such that, for every $w \in F_r$ of length ℓ and every simple algebraic group G, the word map $w_G^{*C\ell^\alpha}$ is (FRS)?*

In [GH19, GH21], Yotam Hendel and the second author and have shown that any dominant map $\varphi : X \to G$ from a smooth variety to a connected algebraic group becomes (FRS) after sufficiently many self-convolutions. Concrete bounds were given in [GHS, Corollary 1.9]. Based on these results, we prove Conjecture 1.7 and answer Question 1.9 for the bounded rank case (see Proposition 7.2).

To conclude the discussion, we remark that a positive answer for Question 1.11 will answer Question 1.9 for compact semisimple p-adic groups as well. A significant progress in this direction was done in the work [GH24], by Yotam Hendel and the second author, where singularities of word maps on semisimple Lie algebras and algebraic groups were studied.

1.4. **Conventions and notations.**

(1) We denote the set $\{1, \ldots, N\}$ by $[N]$.
(2) For a finite set X, we denote the symmetric group on X by $\mathrm{Sym}(X)$ and the space of functions $f : X \to \mathbb{C}$ by $\mathbb{C}[X]$.
(3) We write $(-1)^\sigma$ for the sign of a permutation σ.
(4) For a group G, a representation is a pair (π, V), with $\pi : G \to \mathrm{GL}(V)$ a homomorphism. We denote the character of (π, V) by χ_π and denote its dual by $(\pi^\vee, V^\vee)$.

Acknowledgement. We thank Rami Aizenbud, Yotam Hendel, Michael Larsen, Michael Magee, Doron Puder, Yotam Shomroni, Ofer Zeitouni and Steve Zelditch for useful conversations. We thank the referees for their useful comments and for improving the readability of the paper. NA was supported by NSF grant DMS–1902041, IG was supported by AMS–Simons travel grant, and both of us were supported by BSF grant 2018201.

2. Preliminaries

2.1. **Some facts in representation theory.** For a compact group G, we denote the set of irreducible complex characters of G by $\mathrm{Irr}(G)$. Given a subgroup $H \leq G$ and a character $\chi \in \mathrm{Irr}(H)$, we denote the induction of χ to G by $\mathrm{Ind}_H^G \chi$. We normalize the Haar measure to be a probability measure and denote the expectation with respect to the Haar measure by $\mathbb{E}$. The standard inner product on functions on G is $\langle f_1, f_2 \rangle_G = \mathbb{E} f_1 \overline{f_2}$.

2.1.1. *Representation theory of the symmetric group.* Given $m \in \mathbb{N}$, a *partition* of m is a non-increasing sequence $\lambda = (\lambda_1, .., \lambda_k)$ of non-negative integers that sum to m. In this case, we write $\lambda \vdash m$. Two partitions are equivalent if they differ only by a string of 0's at the end. A partition $\lambda = (\lambda_1, .., \lambda_k)$, with $\lambda_k > 0$, is graphically encoded by a *Young diagram*, which is a finite collection of boxes (or cells) arranged in k left-justified rows, where the j-th row has λ_j boxes. The *length* $\ell(\lambda)$

of a partition $\lambda \vdash m$ is the number of non-zero parts λ_i or equivalently the number of rows in the corresponding Young diagram.

The irreducible representations of S_m are in bijection with partitions $\lambda \vdash m$. We write $\chi_\lambda \in \mathrm{Irr}(S_m)$ for the corresponding character. For each cell (i,j) in the Young diagram of λ, the *hook length* $h_\lambda(i,j)$ is the number of cells (a,b) in the Young diagram of λ such that either $a=i$ and $b \geq j$, or $a \geq i$ and $b=j$. The hook-length formula states that

$$\chi_\lambda(1) = \frac{m!}{\prod_{(i,j)\in\lambda} h_\lambda(i,j)}. \tag{2.1}$$

Definition 2.1.

(1) Fix a Young Diagram λ and let $n \in \mathbb{N}$. An *n-expansion* of λ is any Young diagram obtained by adding n boxes to λ in such a way that no two boxes are added in the same column.

(2) Given a partition $\lambda = (\lambda_1, \dots, \lambda_{l_1}) \vdash k$ and a partition $\mu = (\mu_1, \dots, \mu_{l_2}) \vdash l$, a μ-expansion of λ is defined to be a μ_{l_2}-expansion of a μ_{l_2-1}-expansion of a $\cdots$ of a μ_1-expansion of the Young diagram of λ. For a μ-expansion of λ, we label the boxes added in the μ_{l_j}-expansion by the number j and order the boxes lexicographically by their position, first from top to bottom and then from right to left. We say that a μ-expansion of λ is *strict* if, for every $p \in \{1, \dots, l_2-1\}$ and every box t, the number of boxes coming before t that are labeled p is greater than or equal to the number of boxes coming before t that are labeled $(p+1)$.

Theorem 2.2 (Littlewood–Richardson rule, [Mac95, I.9])**.** *Let $\lambda \vdash k$ and $\mu \vdash m$. Then,*

$$\mathrm{Ind}_{\mathrm{S}_k \times S_m}^{S_{k+m}}(\chi_\lambda \otimes \chi_\mu) = \bigoplus_{\nu \vdash k+m} N_{\lambda\mu\nu}\chi_\nu,$$

where $N_{\lambda\mu\nu}$ is the number of strict μ-expansions of λ that are a Young diagram of the partition ν.

We will need the following consequence of Theorem 2.2.

Lemma 2.3. *Let $l \in \mathbb{Z}_{\geq 2}$ and identify S_m^l with its image in the standard embedding $S_m^l \hookrightarrow S_{ml}$. Then, each $\chi_\nu \in \mathrm{Irr}(S_{ml})$ appearing in $\mathrm{Ind}_{\mathrm{S}_m^l}^{S_{ml}}(1)$ (resp., $\mathrm{Ind}_{\mathrm{S}_m^l}^{S_{ml}}(\mathrm{sgn})$) corresponds to a partition $\nu \vdash ml$ with at most l rows (resp., l columns).*

Proof. We prove the statement for the trivial representation 1 by induction on l. The proof for sgn is similar. The character 1 of S_m corresponds to the partition λ consisting of one row of length m. By the induction hypothesis, we may assume that $\mathrm{Ind}_{\mathrm{S}_m^j}^{S_{mj}}(1) = \bigoplus_{\mu \vdash mj} m_\mu \chi_\mu$, with $m_\mu > 0$ only if μ has at most j rows, for all $j < l$. Hence we can write

$$\mathrm{Ind}_{\mathrm{S}_m^l}^{S_{ml}}(1) = \mathrm{Ind}_{\mathrm{S}_{m(l-1)} \times S_m}^{S_{ml}}(\mathrm{Ind}_{\mathrm{S}_m^{l-1}}^{\mathrm{S}_{m(l-1)}}(1) \otimes 1) = \bigoplus_{\mu \vdash m(l-1)} m_\mu \mathrm{Ind}_{\mathrm{S}_{m(l-1)} \times S_m}^{S_{ml}}(\chi_\mu \otimes 1). \tag{2.2}$$

By Theorem 2.2 and since a strict λ-expansion of μ increases the number of rows by at most one, the lemma follows. $\square$

2.1.2. *Representation theory of the unitary group.* The irreducible representations of U_d can be identified with the irreducible rational representations of $\mathrm{GL}_d(\mathbb{C})$. More precisely, the restriction map $\rho \mapsto \rho|_{\mathrm{U}_d}$ induces a bijection $\mathrm{Irr}(\mathrm{GL}_d(\mathbb{C})) \to \mathrm{Irr}(\mathrm{U}_d)$. Moreover, the set $\mathrm{Irr}(\mathrm{U}_d)$ is in bijection with the set Λ of dominant weights,

$$\Lambda := \{(\lambda_1, \dots, \lambda_d) : \lambda_1 \geq \ldots \geq \lambda_d,\ \lambda_i \in \mathbb{Z}\}.$$

We denote the representation corresponding to $\lambda \in \Lambda$ by $(\rho_\lambda, V_\lambda)$. The irreducible representations

$$\mathbb{C}^d, \bigwedge^2 \mathbb{C}^d, \dots, \bigwedge^d \mathbb{C}^d,$$

are called *the fundamental representations*, and we have $\bigwedge^m \mathbb{C}^d \simeq V_{(1,\dots,1,0,\dots,0)}$, with 1 appearing m times. In particular, the standard representation $\mathbb{C}^d$ is $V_{(1,0,\dots,0)}$. Note that $\bigwedge^d \mathbb{C}^d$ is the determinant representation $\chi_{\det}$. We identify a weight $\lambda \in \Lambda$ such that $\lambda_d \geq 0$ with a partition $(\lambda_1, \dots, \lambda_d)$.

Remark 2.4 ([FH91, I.6, Exc. 6.4]). For each $\lambda = (\lambda_1, \dots, \lambda_d) \vdash m$,

$$\rho_\lambda(1) = \frac{\chi_\lambda(1) \cdot \prod_{(i,j)\in\lambda}(d+j-i)}{m!}, \tag{2.3}$$

where (i,j) are the coordinates of the cells in the Young diagram with shape λ.

Given $\lambda, \mu \in \Lambda$, the irreducible subrepresentations of $\rho_\lambda \otimes \rho_\mu$ are determined by the Littlewood–Richardson rule as follows.

Theorem 2.5 (Littlewood–Richardson rule, see e.g. [FH91, I.6, Eq. (6.7)]). *Let $\lambda, \mu \in \Lambda$ and suppose that $\lambda_d, \mu_d \geq 0$. Let $N_{\lambda\mu\nu}$ be the coefficients from Theorem 2.2. Then:*

$$\rho_\lambda \otimes \rho_\mu = \bigoplus_{\nu:\nu_d\geq 0} N_{\lambda\mu\nu}\rho_\nu.$$

Remark 2.6. For $\lambda, \mu \in \Lambda$, set $\widetilde{\lambda} := \lambda - (\lambda_d, \dots, \lambda_d)$ and $\widetilde{\mu} := \mu - (\mu_d, \dots, \mu_d)$. Then $\rho_\lambda = \chi_{\det}^{\lambda_d} \cdot \rho_{\widetilde{\lambda}}$ and $\rho_\mu = \chi_{\det}^{\mu_d} \cdot \rho_{\widetilde{\mu}}$, and hence by Theorem 2.5, one has:

$$\rho_\lambda \otimes \rho_\mu = \chi_{\det}^{\lambda_d+\mu_d} \rho_{\widetilde{\lambda}} \otimes \rho_{\widetilde{\mu}} = \chi_{\det}^{\lambda_d+\mu_d} \bigoplus_{\nu} N_{\widetilde{\lambda}\widetilde{\mu}\nu}\rho_\nu. \tag{2.4}$$

2.1.3. *Averaging characters over cosets.*

Lemma 2.7. *Let G be a finite group, let (π, V) be an irreducible representation of G, let $H \leq G$ be a subgroup, and let λ be any one-dimensional character of H. Then, for every $g \in G$,*

$$\left| \frac{1}{|H|} \sum_{h\in H} \lambda^{-1}(h)\chi_\pi(gh) \right| \leq \langle \chi_\pi|_H, \lambda \rangle_H .$$

In particular, if $\langle \chi_\pi|_H, \lambda \rangle_H = 0$, then $\sum_{h\in H} \lambda^{-1}(h)\chi_\pi(gh) = 0$.

Proof. Write $\pi|_H = \bigoplus_{i=1}^{\widetilde{N}} \pi_i$ with each (π_i, V_i) an irreducible representation of H. For each i and $h' \in H$,

$$\begin{aligned}\left(\sum_{h\in H} \lambda^{-1}(h)\pi_i(h)\right) \pi_i(h') &= \sum_{h\in H} \lambda^{-1}(h)\pi_i(hh') = \sum_{h\in H} \lambda^{-1}(hh'^{-1})\pi_i(h) \\ &= \sum_{h\in H} \lambda^{-1}(h'^{-1}h)\pi_i(h) = \sum_{h\in H} \lambda^{-1}(h)\pi_i(h'h) = \pi_i(h') \left(\sum_{h\in H} \lambda^{-1}(h)\pi_i(h)\right).\end{aligned}$$

By Schur's lemma, $\sum_{h\in H} \lambda^{-1}(h)\pi_i(h)$ is a scalar matrix $\alpha \cdot I_{V_i}$, for some $\alpha \in \mathbb{C}$. Hence,

$$\alpha \cdot \chi_{\pi_i}(1) = \operatorname{tr}\left(\sum_{h\in H} \lambda^{-1}(h)\pi_i(h)\right) = \sum_{h} \lambda^{-1}(h)\chi_{\pi_i}(h) = \begin{cases} |H| & \text{if } \chi_{\pi_i} = \lambda \\ 0 & \text{else} \end{cases}. \tag{2.5}$$

Let $L := \{v \in V : \pi(h)v = \lambda(h) \cdot v, \ \forall h \in H\}$ be the subspace of (H, λ)-equivariant vectors in V and let $L^\perp$ be an H-invariant subspace of V with $V = L \oplus L^\perp$. By (2.5), the map $A := \sum_{h\in H} \lambda^{-1}(h)\pi(h) \in \operatorname{End}(V)$ satisfies $A|_{L^\perp} = 0$ and $A|_L = |H| \cdot I_L$. Take an orthonormal basis $v_1, \dots, v_N$ for V with $L = \langle v_1, \dots, v_M \rangle$, $L^\perp = \langle v_{M+1}, \dots, v_N \rangle$. Then:

$$\left| \sum_{h\in H} \lambda^{-1}(h)\chi_\pi(gh) \right| = \left| \sum_{i=1}^{N} \langle \pi(g) \left(\sum_{h\in H} \lambda^{-1}(h)\pi(h)\right) v_i, v_i \rangle \right| = |H| \left| \sum_{i=1}^{M} \langle \pi(g)v_i, v_i \rangle \right| \leq M\,|H|,$$

and the lemma follows. □

The following lemma gives a different estimate on the average of a character over a coset, and this estimate is sharper when the double coset HgH is large. We will not need these alternative estimates, but we thought it could be useful to state them.

Lemma 2.8. *Let G be a finite group, and let $H \leq G$ be a subgroup. Then, for each $\chi \in \mathrm{Irr}(G)$ and each $g \in G$,*

$$\left| \frac{1}{|H|} \sum_{h \in H} \chi(hg) \right| \leq \frac{\langle \chi, 1 \rangle_H^{1/2} \cdot |G|^{1/2}}{|HgH|^{1/2}\, \chi(1)^{1/2}}.$$

Proof. Let G be a finite group. For each $\chi \in \mathrm{Irr}(G)$, we denote by (π_χ, V_χ) the representation corresponding to χ. The non-commutative Fourier transform (see e.g. [App14, Section 2.3]) is the map $\mathcal{F} : \mathbb{C}[G] \to \bigoplus_{\chi \in \mathrm{Irr}(G)} \mathrm{End}(V_\chi)$ defined by $f \mapsto \widehat{f} := \left(\widehat{f}(\chi)\right)_{\chi \in \mathrm{Irr}(G)}$, where $\widehat{f}(\chi) = \frac{1}{|G|} \sum_{g' \in G} f(g') \pi_\chi(g'^{-1})$. We denote by $\|f\|_2 := \left(\frac{1}{|G|} \sum_{g' \in G} |f(g')|^2\right)^{\frac{1}{2}}$. Similarly, for a collection of endomorphisms $(A_\chi)_{\chi \in \mathrm{Irr}(G)} \in \bigoplus_{\chi \in \mathrm{Irr}(G)} \mathrm{End}(V_\chi)$, with $A_\chi \in \mathrm{End}(V_\chi)$, we define

$$\left\| (A_\chi)_{\chi \in \mathrm{Irr}(G)} \right\|_2 := \left(\sum_{\chi \in \mathrm{Irr}(G)} \chi(1) \cdot \|A_\chi\|_{\mathrm{HS}}^2 \right)^{\frac{1}{2}},$$

where $\|A_\chi\|_{\mathrm{HS}} := \mathrm{tr}(A_\chi \cdot A_\chi^*)^{\frac{1}{2}}$ is the Hilbert–Schmidt norm on $\mathrm{End}(V_\chi)$. The Plancherel Theorem (see e.g. [App14, Theorem 2.3.1(2)]), states that

$$\|f\|_2 = \left\| \widehat{f} \right\|_2. \tag{2.6}$$

Let $\psi_{HgH} := \frac{1}{|HgH|} 1_{HgH}$. For each $\chi \in \mathrm{Irr}(G)$, one has

$$\widehat{\psi_{HgH}}(\chi) = \frac{1}{|G|} \sum_{g' \in G} \psi_{HgH}(g') \pi_\chi(g'^{-1}) = \frac{1}{|HgH|\,|G|} \sum_{g' \in HgH} \pi_\chi(g'^{-1}).$$

The square of the L^2-norm of ψ_{HgH} is given by:

$$\|\psi_{HgH}\|_2^2 = \frac{1}{|G|} \sum_{g' \in G} \left(\psi_{HgH}(g')\right)^2 = \frac{1}{|G|} \sum_{g' \in HgH} \frac{1}{|HgH|^2} = \frac{1}{|HgH|\,|G|}. \tag{2.7}$$

Let $v_1, \dots, v_M$ be an orthonormal basis of $V_\chi^H := \{v \in V_\chi : \pi_\chi(h).v = v,\ \forall h \in H\}$ with respect to some G-invariant inner product $\langle\, , \rangle$ on V_χ, with $M = \langle \chi, 1 \rangle_H$. Let $\left(V_\chi^H\right)^\perp$ be the orthogonal complement to V_χ^H in V_χ. In the proof of Lemma 2.7, in the case that $\lambda = 1$, we have seen that

$$\sum_{h \in H} \pi_\chi(h).v = \begin{cases} 0 & \text{if } v \in \left(V_\chi^H\right)^\perp \\ |H| \cdot v & \text{if } v \in V_\chi^H. \end{cases} \tag{2.8}$$

In particular, we have

$$\left\langle \sum_{g' \in HgH} \pi_\chi(g'^{-1}).v, v \right\rangle = \frac{|HgH|}{|H|^2} \left\langle \sum_{h', h \in H} \pi_\chi(h' g^{-1} h).v, v \right\rangle = \frac{|HgH|}{|H|^2} \left\langle (\sum_{h' \in H} \pi_\chi(h')).(\sum_{h \in H} \pi_\chi(g^{-1}h).v), v \right\rangle$$

$$= \frac{|HgH|}{|H|^2} \left\langle \sum_{h \in H} \pi_\chi(g^{-1}h).v, \sum_{h' \in H} \pi_\chi(h')v \right\rangle = \begin{cases} 0 & \text{if } v \in \left(V_\chi^H\right)^\perp \\ |HgH| \left\langle \pi_\chi(g^{-1}).v, v \right\rangle & \text{if } v \in V_\chi^H. \end{cases}$$

Hence,

$$\left\|\widehat{\psi_{HgH}}\right\|_2^2 = \sum_{\chi\in\mathrm{Irr}(G)} \chi(1) \left\| \frac{1}{|HgH|\,|G|} \sum_{g'\in HgH} \pi_\chi(g'^{-1}) \right\|_{\mathrm{HS}}^2 = \sum_{\chi\in\mathrm{Irr}(G)} \frac{\chi(1)}{|G|^2} \sum_{i,j=1}^{M} \left|\langle \pi_\chi(g^{-1}).v_i, v_j\rangle\right|^2, \tag{2.9}$$

By (2.6), (2.7) is equal to (2.9), hence,

$$\begin{aligned}\left|\frac{1}{|H|}\sum_{h\in H}\chi(hg^{-1})\right|^2 &= \left|\sum_{i=1}^{M}\langle \pi_\chi(g^{-1}).v_i, v_i\rangle\right|^2 \le M \sum_{i=1}^{M} \left|\langle \pi_\chi(g^{-1}).v_i, v_i\rangle\right|^2 \\ &\le M \sum_{i,j=1}^{M} \left|\langle \pi_\chi(g^{-1}).v_i, v_j\rangle\right|^2 \le \frac{M\,|G|}{\chi(1)\,|HgH|},\end{aligned}$$

where the first equality follows from (2.8), and the first inequality follows from Cauchy–Schwarz inequality. ☐

2.2. **Weingarten calculus.** In §2.1.1, 2.1.2 we stated that each partition $\lambda \vdash m$ with $\ell(\lambda) \le d$ induces two different representations, $\rho_\lambda \in \mathrm{Irr}(\mathrm{U}_d)$ and $\chi_\lambda \in \mathrm{Irr}(S_m)$. There is a deeper connection between ρ_λ and χ_λ coming from the Schur–Weyl duality: the space $\left(\mathbb{C}^d\right)^{\otimes m}$ carries a natural action of $\mathrm{U}_d \times S_m$, where $A \in \mathrm{U}_d$ acts diagonally $A.\,(v_1 \otimes \cdots \otimes v_m) = Av_1 \otimes \cdots \otimes Av_m$, and $\sigma \in S_m$ acts by $\sigma.\,(v_1 \otimes \cdots \otimes v_m) = v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(m)}$. The Schur–Weyl duality can be phrased as follows.

Theorem 2.9 (Schur–Weyl duality, [Wey39])**.** *The space $\left(\mathbb{C}^d\right)^{\otimes m}$ is a multiplicity-free representation of $\mathrm{U}_d \times S_m$. The decomposition of $\left(\mathbb{C}^d\right)^{\otimes m}$ into irreducible components is given by*

$$\left(\mathbb{C}^d\right)^{\otimes m} = \bigoplus_{\lambda\vdash m, \ell(\lambda)\le d} \rho_\lambda \otimes \chi_\lambda. \tag{2.10}$$

There are two special functions on S_m which come from (2.10). Firstly, writing $\ell(\sigma)$ for the number of disjoint cycles in $\sigma \in S_m$, the character of $\left(\mathbb{C}^d\right)^{\otimes m}$ as a representation of S_m is the function $\sigma \mapsto d^{\ell(\sigma)}$. Recall we have an isomorphism of algebras $\mathbb{C}[S_m] \simeq \bigoplus_{\lambda\vdash m} \mathrm{End}(V_{\chi_\lambda})$, where the multiplication in $\mathbb{C}[S_m]$ is the convolution operation $f_1 * f_2(y) := \sum_{x\in S_m} f(x)g(x^{-1}y)$. We denote by $\mathbb{C}_d[S_m]$ the subalgebra corresponding to $\bigoplus_{\lambda\vdash m,\ell(\lambda)\le d} \mathrm{End}(V_{\chi_\lambda})$.

Definition 2.10 ([CS06, Proposition 2.3])**.** Let $d \in \mathbb{N}$. The *Weingarten function* $\mathrm{Wg}_d : S_m \to \mathbb{C}$ is the inverse of the function $d^{\ell(\sigma)}$ in the ring $\mathbb{C}_d[S_m]$. It has the following Fourier expansion:

$$\mathrm{Wg}_d(\sigma) = \frac{1}{m!^2} \sum_{\lambda\vdash m,\ell(\lambda)\le d} \frac{\chi_\lambda(1)^2}{\rho_\lambda(1)} \chi_\lambda(\sigma). \tag{2.11}$$

Remark 2.11. Since in this paper we only consider $\mathrm{Wg}_{d'}(\sigma)$ for $d' = d$, we write Wg instead of Wg_d.

The *Weingarten Calculus*, developed in [Wei78, Col03, CS06] utilizes the Schur–Weyl duality to express integrals on unitary groups as finite sums of Weingarten functions on symmetric groups. One formulation is the following theorem by Collins and Śniady:

Theorem 2.12 ([CS06, Corollary 2.4])**.** *Let $(i_1,\dots,i_m)$, $(j_1,\dots,j_m)$, $(i'_1,\dots,i'_m)$, and $(j'_1,\dots,j'_m)$ be tuples of integers in $[d]$. Then:*

$$\begin{aligned}&\mathbb{E}_{X\in\mathrm{U}_d}\left(X_{i_1,j_1}\cdots X_{i_m,j_m}\cdot \overline{X_{i'_1,j'_1}\cdots X_{i'_m,j'_m}}\right) \\ &= \sum_{\sigma,\tau\in S_m} \delta_{i_1,i'_{\sigma(1)}}\cdots\delta_{i_m,i'_{\sigma(m)}}\delta_{j_1,j'_{\tau(1)}}\cdots\delta_{j_m,j'_{\tau(m)}}\cdot \mathrm{Wg}_d(\sigma^{-1}\tau).\end{aligned} \tag{2.12}$$

We will use a coordinate-free version of Theorem 2.12 which we proceed to state.

Definition 2.13. Let Ω be a set.

(1) A *symmetric partition* Φ of Ω is a partition $\Omega = \bigsqcup_{i=1}^r A_i \sqcup \bigsqcup_{i=1}^r B_i$, where $|A_i| = |B_i|$.
(2) Given a symmetric partition $\Phi = (A_1, \dots, A_r, B_1, \dots, B_r)$, let

$$S_\Phi = \{\Sigma \in \mathrm{Sym}(\Omega) : \Sigma(A_i) = B_i, \Sigma(B_i) = A_i\}.$$

(3) If $\Sigma \in S_\Phi$, then $\Sigma^2(A_i) = A_i$ and we define $\widetilde{\mathrm{Wg}}(\Sigma^2) = \prod_{i=1}^r \mathrm{Wg}(\Sigma^2|_{A_i})$.

Proposition 2.14. *Let $\Phi = (A, B)$ be a symmetric partition of Ω and let $F, H : \Omega \to [d]$. Then*

$$\mathbb{E}_{X \in \mathrm{U}_d}\left(\prod_{x \in A} X_{F(x),H(x)} \prod_{y \in B} X^{-1}_{F(y),H(y)}\right) = \mathbb{E}\left(\prod_{x \in A} X_{F(x),H(x)} \prod_{y \in B} \overline{X_{H(y),F(y)}}\right) = \sum_{\Sigma \in S_\Phi \colon H = F \circ \Sigma} \widetilde{\mathrm{Wg}}(\Sigma^2).$$

Proof. Identify $A \cong \{1, \dots, m\}$ and $B \cong \{-1, \dots, -m\}$ and let $\vec{i}, \vec{j}, \vec{i}', \vec{j}' \in [d]^m$ be

$$i_k = F(k) \quad j_k = H(k) \quad i'_k = H(-k) \quad j'_k = F(-k).$$

Then, by Theorem 2.12,

$$\mathbb{E}_{X \in \mathrm{U}_d}\left(\prod_{x \in A} X_{F(x),H(x)} \prod_{y \in B} X^{-1}_{F(y),H(y)}\right) = \mathbb{E}_{X \in \mathrm{U}_d}\left(X_{i_1,j_1} \cdots X_{i_m,j_m} \overline{X_{i'_1,j'_1} \cdots X_{i'_m,j'_m}}\right)$$
$$= \sum_{\sigma,\tau \in S_m} \delta_{i_1,i'_{\sigma(1)}} \cdots \delta_{i_m,i'_{\sigma(m)}} \cdot \delta_{j_1,j'_{\tau(1)}} \cdots \delta_{j_m,j'_{\tau(m)}} \cdot \mathrm{Wg}(\sigma^{-1}\tau).$$

For $\sigma, \tau \in S_m$, let $\Sigma_{(\sigma,\tau)} \in \mathrm{Sym}(A \sqcup B) \cong \mathrm{Sym}(\{-m, \dots, -1, 1, \dots, m\})$ be the permutation

$$\Sigma_{(\sigma,\tau)}(x) = \begin{cases} -\tau(x) & x \in \{1, \dots, m\} \\ \sigma^{-1}(-x) & x \in \{-1, \dots, -m\}. \end{cases}$$

The map $(\sigma, \tau) \mapsto \Sigma_{(\sigma,\tau)}$ is a bijection $S_m^2 \cong S_\Phi$ and the condition $\delta_{i_1,i'_{\sigma(1)}} \cdots \delta_{i_m,i'_{\sigma(m)}} \cdot \delta_{j_1,j'_{\tau(1)}} \cdots \delta_{j_m,j'_{\tau(m)}} = 1$ is equivalent to $H = F \circ \Sigma_{(\sigma,\tau)}$. Finally, the permutation $\left(\Sigma_{(\sigma,\tau)}\right)^2$ acts on A as $\sigma^{-1}\tau$, and the result follows. ☐

Corollary 2.15. *Let $\Phi = (A_1, \dots, A_r, B_1, \dots, B_r)$ be a symmetric partition of Ω and let $F, H : \Omega \to [d]$. Then*

$$\mathbb{E}\left(\prod_{i=1}^r \left(\prod_{x \in A_i} (X_i)_{F(x),H(x)} \prod_{y \in B_i} (X_i^{-1})_{F(y),H(y)}\right)\right) = \sum_{\Sigma \in S_\Phi \colon H = F \circ \Sigma} \widetilde{\mathrm{Wg}}(\Sigma^2).$$

3. The Engel word as a model case

"Those who run to long words are mainly the unskillful and tasteless; they confuse pomposity with dignity, flaccidity with ease, and bulk with force."[1]

In this section we prove the following simplified version of Theorem 1.3 for the Engel word. We chose the Engel word since it is short enough to make the proof easier to digest, while at same time complicated enough so that the proof contains most of the key ideas in the paper.

Theorem 3.1. *Let X, Y be independent random variables with respect to the normalized Haar measure on U_d. For every $d \geq 2m$, one has:*

$$\mathbb{E}\left(c_m\left([[X,Y],Y]\right)\right) < 2^{17m}.$$

[1] H.W Fowler, *A Dictionary of Modern English Usage*, 1965.

Let $w = [[x,y],y] = xyx^{-1}yxy^{-1}x^{-1}y^{-1}$ be the Engel word. We would like to compute $\mathbb{E}\left(\mathrm{tr} \bigwedge^m w(X,Y)\right)$. Denote $\mathcal{I}_{m,d} := \{a_1 < \ldots < a_m : a_i \in [d]\}$, and note that

$$\mathrm{tr}\left(\bigwedge^m w(X,Y)\right) = \sum_{\vec{a} \in \mathcal{I}_{m,d}} \sum_{\pi \in S_m} (-1)^\pi w(X,Y)_{a_1 a_{\pi(1)}} \cdots w(X,Y)_{a_m a_{\pi(m)}}. \tag{3.1}$$

We have

$$\begin{aligned} w(X,Y)_{a_i a_{\pi(i)}} &= \sum_{b_i,c_i,d_i,A_i,B_i,C_i,D_i \in [d]} X_{a_i,D_i} Y_{D_i,c_i} X^{-1}_{c_i,A_i} Y_{A_i,b_i} X_{b_i,C_i} Y^{-1}_{C_i,d_i} X^{-1}_{d_i,B_i} Y^{-1}_{B_i,a_{\pi(i)}} \\ &= \sum_{b_i,c_i,d_i,A_i,B_i,C_i,D_i \in [d]} X_{a_i,D_i} X_{b_i,C_i} \overline{X_{A_i,c_i} X_{B_i,d_i}} Y_{A_i,b_i} Y_{D_i,c_i} \overline{Y_{a_{\pi(i)},B_i} Y_{d_i,C_i}}. \end{aligned} \tag{3.2}$$

The group S_m acts on $[d]^m$ by $\sigma(\vec{v})_i = \vec{v}_{\sigma^{-1}(i)}$ for any $\sigma \in S_m$ and $\vec{v} \in [d]^m$. Similarly, given $\vec{v}, \vec{w} \in [d]^m$ and $\tau \in S_{2m}$, we denote by $(\vec{v}, \vec{w})$ the element in $[d]^{2m}$ given by $(\vec{v}, \vec{w})_i = \begin{cases} \vec{v}_i & \text{if } i \le m \\ \vec{w}_{i-m} & \text{if } m < i \le 2m \end{cases}$, and denote by $\tau(\vec{v}, \vec{w})_i = (\vec{v}, \vec{w})_{\tau^{-1}(i)}$. In particular, writing $X_{\vec{v},\vec{u}} := \prod_{i=1}^m X_{v_i,u_i}$ for $\vec{v}, \vec{u} \in [d]^m$, we have:

$$\mathrm{tr}\left(\bigwedge^m w(X,Y)\right) = \sum_{\vec{a} \in \mathcal{I}_{m,d}} \sum_{\vec{b},\ldots,\vec{D} \in [d]^m} \sum_{\pi \in S_m} (-1)^\pi \left(X_{\vec{a},\vec{D}} X_{\vec{b},\vec{C}} \overline{X_{\vec{A},\vec{c}} X_{\vec{B},\vec{d}}}\right) \left(Y_{\vec{A},\vec{b}} Y_{\vec{D},\vec{c}} \overline{Y_{\pi^{-1}(\vec{a}),\vec{B}} Y_{\vec{d},\vec{C}}}\right). \tag{3.3}$$

We now rewrite the expected value of (3.3) using Weingarten calculus. For this, define:

$$S(\vec{a}, \ldots, \vec{D}) := \left\{ (\sigma_1, \sigma_2, \tau_1, \tau_2) \in S_{2m}^4 : \begin{array}{l} (\vec{A}, \vec{B}) = \sigma_1(\vec{a}, \vec{b}),\ (\vec{c}, \vec{d}) = \tau_1(\vec{D}, \vec{C}) \\ (\vec{a}, \vec{d}) = \sigma_2(\vec{A}, \vec{D}),\ (\vec{B}, \vec{C}) = \tau_2(\vec{b}, \vec{c}) \end{array} \right\},$$

and

$$Z := \left\{ (\vec{a}, \ldots, \vec{D}, \sigma_1, \sigma_2, \tau_1, \tau_2) \in \mathcal{I}_{m,d} \times [d]^{7m} \times S_{2m}^4 : (\sigma_1, \sigma_2, \tau_1, \tau_2) \in S(\vec{a}, \ldots, \vec{D}) \right\}. \tag{3.4}$$

Lemma 3.2. *We have:*

$$\mathbb{E}\left(\mathrm{tr} \bigwedge^m w(X,Y)\right) = \sum_{(\vec{a},\ldots,\vec{D},\sigma_1,\sigma_2,\tau_1,\tau_2) \in Z} \sum_{\pi \in S_m} (-1)^\pi \mathrm{Wg}(\sigma_1^{-1}\tau_1) \mathrm{Wg}(\sigma_2^{-1}(\pi \times \mathrm{Id})\tau_2). \tag{3.5}$$

Proof. Using Weingarten calculus, i.e., Theorem 2.12, and (3.3),

$$\begin{aligned} \mathbb{E}\left(\mathrm{tr} \bigwedge^m w(X,Y)\right) = \sum_{\vec{a} \in \mathcal{I}_{m,d}} \sum_{\vec{b},\ldots,\vec{D} \in [d]^m} \sum_{\pi \in S_m} (-1)^\pi \sum_{\sigma_1,\widetilde{\sigma}_2,\tau_1,\tau_2 \in S_{2m}} \delta_{(\vec{a},\vec{b}),\sigma_1^{-1}(\vec{A},\vec{B})} \cdot \delta_{(\vec{D},\vec{C}),\tau_1^{-1}(\vec{c},\vec{d})} \mathrm{Wg}(\sigma_1^{-1}\tau_1) \\ \cdot \delta_{(\vec{A},\vec{D}),\widetilde{\sigma}_2^{-1}(\pi^{-1}(\vec{a}),\vec{d})} \cdot \delta_{(\vec{b},\vec{c}),\tau_2^{-1}(\vec{B},\vec{C})} \mathrm{Wg}(\widetilde{\sigma}_2^{-1}\tau_2). \end{aligned} \tag{3.6}$$

Applying the change of coordinate $\sigma_2 := (\pi \times \mathrm{Id}) \circ \widetilde{\sigma}_2$, and observing that $\widetilde{\sigma}_2^{-1}(\pi^{-1}(\vec{a}), \vec{d}) = \sigma_2^{-1}(\vec{a}, \vec{d})$, (3.6) becomes:

$$\mathbb{E}\left(\mathrm{tr} \bigwedge^m w(X,Y)\right) = \sum_{(\vec{a},\ldots,\vec{D},\sigma_1,\sigma_2,\tau_1,\tau_2) \in Z} \sum_{\pi \in S_m} (-1)^\pi \mathrm{Wg}(\sigma_1^{-1}\tau_1) \cdot \mathrm{Wg}(\sigma_2^{-1}(\pi \times \mathrm{Id})\tau_2). \qquad \square$$

In order to bound (3.5), we consider a natural action of S_m^7 on Z, and find a suitable change of coordinates such that the average of the product of the Weingarten functions in (3.5) over any S_m^7-orbit is equal to a product of averages of individual Weingarten functions over cosets (see (3.8)). We then use Lemma 2.7 to estimate the contribution in (3.5) of each S_m^7-orbit. To conclude the estimates of (3.5), we will further provide estimates for $|Z|$.

We first describe the action of S_m^7. The element $(\pi_b, \pi_c, \dots, \pi_D) \in S_m^7$ acts on $(\overrightarrow{a}, \dots, \overrightarrow{D})$ by $(\overrightarrow{a}, \pi_b(\overrightarrow{b}), \pi_c(\overrightarrow{c}), \dots, \pi_D(\overrightarrow{D}))$ and it acts on $(\sigma_1, \sigma_2, \tau_1, \tau_2)$ by:

$$\begin{aligned}\sigma_1 &\mapsto (\pi_A \times \pi_B) \circ \sigma_1 \circ (\mathrm{Id} \times \pi_b^{-1})\\ \tau_1 &\mapsto (\pi_c \times \pi_d) \circ \tau_1 \circ (\pi_D^{-1} \times \pi_C^{-1})\\ \sigma_2 &\mapsto (\mathrm{Id} \times \pi_d) \circ \sigma_2 \circ (\pi_A^{-1} \times \pi_D^{-1})\\ \tau_2 &\mapsto (\pi_B \times \pi_C) \circ \tau_2 \circ (\pi_b^{-1} \times \pi_c^{-1}).\end{aligned}$$

This gives rise to an action of S_m^7 on Z. The action on the input of the Weingarten functions becomes

$$\mathrm{Wg}((\pi_D^{-1} \times \pi_C^{-1}\pi_b)\sigma_1^{-1}(\pi_A^{-1}\pi_c \times \pi_B^{-1}\pi_d)\tau_1) \text{ and } \mathrm{Wg}(\pi_b^{-1}\pi_A \times \pi_c^{-1}\pi_D)\sigma_2^{-1}(\pi\pi_B \times \pi_d^{-1}\pi_C)\tau_2), \tag{3.7}$$

where we used the conjugacy invariance of Wg to move permutations from right to left. Consider the bijection $\psi : S_m^8 \to S_m^8$, defined by $(x_1, \dots, x_8) \mapsto (x_1, x_1x_2, \dots, x_1x_2\cdots x_8)$. Under the change of coordinates $(\theta_D, \theta_c, \theta_A, \theta_b, \theta_C, \theta_d, \theta_B, \theta) := \psi^{-1}(\pi_D, \pi_c, \pi_A, \pi_b, \pi_C, \pi_d, \pi_B, \pi^{-1})$, the summation of (3.5) over an S_m^7-orbit splits into a product of two separate sums:

$$\begin{aligned}&\sum_{(\pi_D,\dots,\pi)\in S_m^8} (-1)^{\pi}\mathrm{Wg}((\pi_D^{-1} \times \pi_C^{-1}\pi_b)\sigma_1^{-1}(\pi_A^{-1}\pi_c \times \pi_B^{-1}\pi_d)\tau_1)\mathrm{Wg}(\pi_b^{-1}\pi_A \times \pi_c^{-1}\pi_D)\sigma_2^{-1}(\pi\pi_B \times \pi_d^{-1}\pi_C)\tau_2)\\ &= \sum_{(\theta_D,\dots,\theta)\in S_m^8} (-1)^{\theta_D\cdots\theta}\mathrm{Wg}((\theta_D^{-1} \times \theta_C^{-1})\sigma_1^{-1}(\theta_A^{-1} \times \theta_B^{-1})\tau_1)\mathrm{Wg}(\theta_b^{-1} \times \theta_c^{-1})\sigma_2^{-1}(\theta^{-1} \times \theta_d^{-1})\tau_2)\\ &= \sum_{\eta_1,\eta_1'\in S_m^2} (-1)^{\eta_1\eta_1'}\mathrm{Wg}(\eta_1\sigma_1^{-1}\eta_1'\tau_1) \sum_{\eta_2,\eta_2'\in S_m^2} (-1)^{\eta_2\eta_2'}\mathrm{Wg}(\eta_2\sigma_2^{-1}\eta_2'\tau_2).\end{aligned} \tag{3.8}$$

We can now use the Fourier expansion of Wg (2.11) and the estimates in §2.1.3 to bound the contribution of an S_m^7-orbit in Z to (3.5):

Proposition 3.3. *Let $\widetilde{v} := (\widetilde{\overrightarrow{a}}, \dots, \widetilde{\overrightarrow{D}}, \widetilde{\sigma}_1, \widetilde{\sigma}_2, \widetilde{\tau}_1, \widetilde{\tau}_2) \in Z$ and let $\mathcal{O}_{\widetilde{v}} := S_m^7\widetilde{v}$ be its S_m^7-orbit. Then,*

$$\left|\frac{1}{|\mathcal{O}_{\widetilde{v}}|}\sum_{(\overrightarrow{a},\dots,\tau_2)\in\mathcal{O}_{\widetilde{v}}}\sum_{\pi\in S_m} (-1)^{\pi}\mathrm{Wg}(\sigma_1^{-1}\tau_1)\mathrm{Wg}(\sigma_2^{-1}(\pi \times \mathrm{Id})\tau_2)\right| \le \frac{1}{(2m)!^2m!^3}\binom{d}{2m}^{-2}. \tag{3.9}$$

Proof. By the Orbit-Stabilizer Theorem, the LHS of (3.9) is the same as summing over all $(\pi_D, \dots, \pi_B) \in S_m^7$ and dividing by $m!^7$. By (3.8), the LHS of (3.9) is equal to

$$\frac{1}{m!^7}\left|\sum_{\eta_1,\eta_1'\in S_m^2} (-1)^{\eta_1\eta_1'}\mathrm{Wg}(\eta_1\widetilde{\sigma}_1^{-1}\eta_1'\widetilde{\tau}_1) \sum_{\eta_2,\eta_2'\in S_m^2} (-1)^{\eta_2\eta_2'}\mathrm{Wg}(\eta_2\widetilde{\sigma}_2^{-1}\eta_2'\widetilde{\tau}_2)\right|.$$

Note that $(S_{2m}, S_m \times S_m)$ is a sgn-twisted Gelfand pair, that is, the representation $\mathrm{Ind}_{S_m^2}^{S_{2m}}\mathrm{sgn}$ is multiplicity-free. By Frobenius reciprocity, each irreducible subrepresentation (V_λ, π_λ) of $\mathrm{Ind}_{S_m^2}^{S_{2m}}\mathrm{sgn}$ has a unique (S_m^2, sgn)-invariant unit vector, so $\langle \chi_\lambda|_{S_m^2}, \mathrm{sgn}\rangle_{S_m^2} = 1$. By Lemma 2.7, for each $\sigma \in S_{2m}$, we have

$$\left|\sum_{h\in S_m^2} (-1)^h\chi_\lambda(h\sigma)\right| \le m!^2 \langle \chi_\lambda|_{S_m^2}, \mathrm{sgn}\rangle_{S_m^2} = \begin{cases} m!^2 & \text{if } \pi_\lambda \hookrightarrow \mathrm{Ind}_{S_m^2}^{S_{2m}}\mathrm{sgn} \\ 0 & \text{otherwise.}\end{cases} \tag{3.10}$$

By Lemma 2.3 it follows that $\pi_\lambda \hookrightarrow \mathrm{Ind}_{S_m^2}^{S_{2m}}\mathrm{sgn}$ if and only if the young diagram of $\lambda \vdash 2m$ has at most two columns. Combining with (3.10), we have:

$$\left|\sum_{\eta_1,\eta_1'\in S_m^2}(-1)^{\eta_1\eta_1'}\chi_\lambda(\eta_1\widetilde{\sigma}_1^{-1}\eta_1'\widetilde{\tau}_1)\right| = \left|\sum_{\eta_1'\in S_m^2}(-1)^{\eta_1'}\sum_{\eta_1\in S_m^2}(-1)^{\eta_1}\chi_\lambda(\eta_1\widetilde{\sigma}_1^{-1}\eta_1'\widetilde{\tau}_1)\right| \tag{3.11}$$
$$\leq \sum_{\eta_1'\in S_m^2}\left|\sum_{\eta_1\in S_m^2}(-1)^{\eta_1}\chi_\lambda(\eta_1\widetilde{\sigma}_1^{-1}\eta_1'\widetilde{\tau}_1)\right| \leq \begin{cases} m!^4 & \text{if } \lambda\vdash 2m,\ \lambda \text{ has} \leq 2 \text{ columns,}\\ 0 & \text{otherwise.}\end{cases}$$

By (2.11), (3.11), (2.3), and by our assumption that $d \geq 2m$, we have

$$\left|\sum_{\eta_1,\eta_1'\in S_m^2}(-1)^{\eta_1\eta_1'}\mathrm{Wg}(\eta_1\widetilde{\sigma}_1^{-1}\eta_1'\widetilde{\tau}_1)\right| = \frac{1}{(2m)!^2}\left|\sum_{\lambda\vdash 2m}\frac{\chi_\lambda(1)^2}{\rho_\lambda(1)}\sum_{\eta_1,\eta_1'\in S_m^2}(-1)^{\eta_1\eta_1'}\chi_\lambda(\eta_1\widetilde{\sigma}_1^{-1}\eta_1'\widetilde{\tau}_1)\right|$$
$$\leq \frac{m!^4}{(2m)!^2}\sum_{\lambda\vdash 2m,\ \lambda \text{ has } \leq 2 \text{ columns}}\frac{\chi_\lambda(1)^2}{\rho_\lambda(1)} = \frac{m!^4}{(2m)!}\sum_{\lambda\vdash 2m,\ \lambda \text{ has } \leq 2 \text{ columns}}\frac{\chi_\lambda(1)}{\prod_{(i,j)\in\lambda}(d+j-i)}$$
$$\leq \frac{m!^4}{(2m)!}\cdot\frac{1}{d\cdot\dots\cdot(d-2m+1)}\sum_{\lambda\vdash 2m,\ \lambda \text{ has } \leq 2 \text{ columns}}\chi_\lambda(1) = \frac{m!^4}{(2m)!}\frac{\dim \mathrm{Ind}_{S_m^2}^{S_{2m}}\mathrm{sgn}}{d\cdot\dots\cdot(d-2m+1)} = \frac{m!^2}{(2m)!}\binom{d}{2m}^{-1}.$$

This concludes the proposition. ☐

We now turn to the last ingredient in the proof of Theorem 3.1.

Definition 3.4. Let $f : S \to [d]$ be a function on a set S. We define the *shape* $\nu_f : [d] \to \mathbb{N}$ of f as

$$\nu_f = (\nu_{f,1}, \dots, \nu_{f,d}) := (\left|f^{-1}(1)\right|, \dots, \left|f^{-1}(d)\right|),$$

and denote $\nu_f! := \prod_{u=1}^d \nu_{f,u}$.

Proposition 3.5. *Let Z be as in (3.4). Then:*

$$|Z| \leq m!^7\binom{2m}{m}^4\binom{d}{m}\binom{d+m-1}{m}^3.$$

Proof. We need to count all the possible tuples $(\overrightarrow{a}, \dots, \overrightarrow{D}, \sigma_1, \sigma_2, \tau_1, \tau_2)$ in Z. Suppose we have already fixed $\overrightarrow{a}$ and the shapes $\nu_{\overrightarrow{b}}, \nu_{\overrightarrow{c}}$ and $\nu_{\overrightarrow{d}}$ of $\overrightarrow{b}$, $\overrightarrow{c}$ and $\overrightarrow{d}$, where $\overrightarrow{b}$, $\overrightarrow{c}$, $\overrightarrow{d}$ are considered as a functions $[m] \to [d]$. Given this data:

1. There are $\frac{m!^3}{\nu_{\overrightarrow{b}}!\nu_{\overrightarrow{c}}!\nu_{\overrightarrow{d}}!}$ options for $\overrightarrow{b}$, $\overrightarrow{c}$, $\overrightarrow{d}$ with the above shapes.
2. There are $(2m)!^2$ options for σ_2 and τ_2.
3. There are at most $\binom{2m}{m}^2$ options for choosing $\tau_1^{-1}([m])$ and $\sigma_1([m])$, as subsets of $[2m]$. Note that we count both valid and invalid options.
4. After fixing the subsets $\tau_1^{-1}([m])$ and $\sigma_1([m])$, there are at most $\nu_{\overrightarrow{c}}!\nu_{\overrightarrow{d}}!$ options for τ_1 and $\nu_{\overrightarrow{b}}!$ options for σ_1.

Summarizing the above items, we get there are at most $\frac{m!^3(2m)!^2\nu_{\overrightarrow{b}}!\nu_{\overrightarrow{c}}!\nu_{\overrightarrow{d}}!}{\nu_{\overrightarrow{b}}!\nu_{\overrightarrow{c}}!\nu_{\overrightarrow{d}}!}\binom{2m}{m}^2 = m!^7\binom{2m}{m}^4$ options for $(\overrightarrow{a}, \dots, \overrightarrow{D}, \sigma_1, \sigma_2, \tau_1, \tau_2) \in Z$ with the initial data $\overrightarrow{a}, \nu_{\overrightarrow{b}}, \nu_{\overrightarrow{c}}, \nu_{\overrightarrow{d}}$. Note that there are $\binom{d}{m}$ possible options for $\overrightarrow{a}$, and $\binom{d+m-1}{m}^3$ options for $\nu_{\overrightarrow{b}}, \nu_{\overrightarrow{c}}, \nu_{\overrightarrow{d}}$. This gives the desired upper bound. ☐

We can now finish the proof of Theorem 3.1.

Proof of Theorem 3.1. Note that for every $k \geq 1$ and $n \geq k$ we have

$$\left(\frac{n}{k}\right)^k \leq \prod_{j=0}^{k-1}\left(\frac{n-j}{k-j}\right) = \binom{n}{k} \leq \frac{n^k}{k!} \leq \left(\frac{n}{k}\right)^k e^k, \tag{3.12}$$

where the rightmost inequality follows from Stirling's approximation. By Lemma 3.2 and by Propositions 3.3 and 3.5,

$$\begin{aligned}\left|\mathbb{E}\left(\operatorname{tr}\bigwedge^m w(X,Y)\right)\right| &= |Z| \cdot \left|\frac{1}{|Z|}\sum_{(\vec{a},\ldots,\vec{D},\sigma_1,\sigma_2,\tau_1,\tau_2)\in Z}\ \sum_{\pi\in S_m}(-1)^\pi \mathrm{Wg}(\sigma_1^{-1}\tau_1)\cdot \mathrm{Wg}(\sigma_2^{-1}(\pi\times \mathrm{Id})\tau_2)\right| \\ &\leq |Z|\cdot\frac{1}{(2m)!^2 m!^3}\binom{d}{2m}^{-2} \leq \binom{2m}{m}^2\binom{d}{m}\binom{d+m}{m}^3\binom{d}{2m}^{-2}. \end{aligned}\tag{3.13}$$

By (3.13), (3.12), by the inequality $\binom{2m}{m} \leq 2^{2m}$, and by our assumption that $d \geq 2m$,

$$\left|\mathbb{E}\left(\operatorname{tr}\bigwedge^m w(X,Y)\right)\right| \leq \frac{2^{4m}e^{4m}\left(\frac{d}{m}\right)^m\left(\frac{d+m}{m}\right)^{3m}}{\left(\frac{d}{2m}\right)^{4m}} \leq \frac{2^{7m}e^{4m}\left(\frac{d}{m}\right)^{4m}}{\left(\frac{d}{2m}\right)^{4m}} \leq 2^{11m}e^{4m} \leq 2^{17m}. \qquad \square$$

Remark 3.6. The current proof of Proposition 3.5 depends on the special structure of the Engel word. One can give a slightly more complicated argument, which can be easily generalized for every word w (this is done in §6). Here are the main ideas of this alternative argument.

We encode the expression

$$X_{a,D}Y_{D,c}X^{-1}_{c,A}Y_{A,b}X_{b,C}Y^{-1}_{C,d}X^{-1}_{d,B}Y^{-1}_{B,a} \tag{3.14}$$

from (3.2), graphically, by the 4×4 matrix

$$\begin{pmatrix} \cdot & C & D & \cdot \\ c & \cdot & \cdot & b \\ d & \cdot & \cdot & a \\ \cdot & B & A & \cdot \end{pmatrix}, \tag{3.15}$$

which is constructed as follows. The rows and columns are indexed by x, y, x^{-1}, y^{-1}. We order the rows by $x < y < y^{-1} < x^{-1}$ and order the columns by $x^{-1} < y^{-1} < y < x$. To find the (x, y^{-1})-entry of this matrix (i.e. the $(1,2)$-entry), we look for the subword XY^{-1} in (3.14) and record the letter of the common index, which is C. All other entries are determined in similar fashion. Note that we do not have elements in the main diagonal since w is cyclically reduced.

We denote $\eta_1 = \tau_1$, $\eta_2 = \tau_2$, $\eta_3 = \sigma_2^{-1}$, $\eta_4 = \sigma_1^{-1}$. Note that η_i sends the ith row of (3.15) into a permuted copy of its ith column. The alternative counting argument in Proposition 3.5 goes as follows. We fix the upper triangular part, i.e., $\vec{C}, \vec{D}, \vec{a}, \vec{b}$ (instead of $\vec{a}, \vec{b}, \vec{c}, \vec{d}$ in the proof above). We then choose η_1 (with $2m!$ options), which gives us $\vec{c}, \vec{d}$ and, in particular, reveals the second row. Next, we choose all possible $\eta_2 : (\vec{b}, \vec{c}) \to (\vec{B}, \vec{C})$, taking into consideration the fact that $\vec{C}$ is already known. We then proceed to the next row and guess η_3, taking into consideration that we already know $\vec{D}$. At this point, the vectors $\vec{a}, \vec{b}, \ldots, \vec{C}, \vec{D}$ and the permutations η_1, η_2, η_3 are known, and the number of options for η_4 is determined by the shapes of $\vec{a}, \vec{b}$. This argument will be generalized in §6 for arbitrary words, where, instead of a 4×4 matrix, we will have a $2r \times 2r$ matrix and, each time we choose $\eta_1, \ldots, \eta_k$, the $k+1$-st row is revealed, allowing us to proceed by induction.

4. Rewriting Theorem 1.3 using Weingarten calculus

In this section, we rewrite the expression $\mathbb{E}\left(|\mathrm{tr}\bigwedge^m w(X_1,\dots,X_r)|^2\right)$ of Theorem 1.3 as a finite sum of Weingarten functions.

Let ℓ, m, d, w be as in Theorem 1.3. We may assume that w is *cyclically reduced*, i.e., it does not contain a subword of the form $x_j x_j^{-1}$ and the first and last letters of w are not inverse of each other. For $u \in [\ell]$, let

$$w(u) = \begin{cases} a & \text{if the } u^{\text{th}} \text{ letter of } w \text{ is } x_a \\ -a & \text{if the } u^{\text{th}} \text{ letter of } w \text{ is } x_a^{-1} \end{cases}.$$

If we denote $x_{-a} = x_a^{-1}$, then $w = \prod_u x_{w(u)}$. We write w^{-1} for the inverse word,

$$w^{-1} := x_{-w(\ell)} x_{-w(\ell-1)} \cdots x_{-1}. \tag{4.1}$$

We start by noting that

$$\begin{aligned} \mathbb{E}\left(\left|\mathrm{tr}\left(\bigwedge^m w(X_1,\dots,X_r)\right)\right|^2\right) &= \mathbb{E}\left(\mathrm{tr}\left(\bigwedge^m w(X_1,\dots,X_r)\right)\cdot\overline{\mathrm{tr}\left(\bigwedge^m w(X_1,\dots,X_r)\right)}\right) \\ &= \mathbb{E}\left(\mathrm{tr}\left(\bigwedge^m w(X_1,\dots,X_r)\right)\cdot\mathrm{tr}\left(\bigwedge^m w^{-1}(X_1,\dots,X_r)\right)\right). \end{aligned} \tag{4.2}$$

Define $\widetilde{T} \in \mathrm{Sym}([\ell]\times[m])$ by

$$\widetilde{T}(u,k) = \begin{cases} (u+1,k) & u \neq \ell \\ (1,k) & u = \ell \end{cases}. \tag{4.3}$$

Recall that $\mathcal{I}_{m,d} = \{a_1 < \ldots < a_m : a_i \in [d]\}$. We have

$$\begin{aligned} \mathrm{tr}\left(\bigwedge^m w(X_1,\dots,X_r)\right) &= \sum_{\vec{a}\in\mathcal{I}_{m,d}} \sum_{\pi\in S_m} (-1)^\pi \prod_{k=1}^m w(X_1,\dots,X_r)_{a_k,a_{\pi(k)}} \\ &= \sum_{\vec{a}\in\mathcal{I}_{m,d}} \sum_{\pi\in S_m} (-1)^\pi \prod_{k=1}^m \sum_{\substack{f_k:[\ell+1]\to[d] \\ f_k(1)=a_k, f_k(\ell+1)=a_{\pi(k)}}} \prod_{u=1}^{\ell} \left(X_{w(u)}\right)_{f_k(u),f_k(u+1)} \\ &= \sum_{\vec{a}\in\mathcal{I}_{m,d}} \sum_{\pi\in S_m} (-1)^\pi \sum_{\substack{f:[\ell+1]\times[m]\to[d] \\ f(1,k')=a_{k'}, f(\ell+1,k')=a_{\pi(k')},\forall k'}} \prod_{(u,k)\in[\ell]\times[m]} \left(X_{w(u)}\right)_{f(u,k),f(u+1,k)} \\ &= \sum_{\pi\in S_m} (-1)^\pi \sum_{\substack{f:[\ell+1]\times[m]\to[d] \\ f(\ell+1,k')=f(1,\pi(k')),\forall k' \\ f(1,-)\text{ increasing}}} \prod_{(u,k)\in[\ell]\times[m]} \left(X_{w(u)}\right)_{f(u,k),f(u+1,k)} \\ &= \sum_{\pi\in\mathrm{Sym}(\{\ell\}\times[m])} (-1)^\pi \sum_{\substack{F:[\ell]\times[m]\to[d] \\ F(1,-)\text{ increasing}}} \prod_{(u,k)\in[\ell]\times[m]} \left(X_{w(u)}\right)_{F(u,k),F(\widetilde{T}\pi(u,k))}, \end{aligned} \tag{4.4}$$

where in the last equality we use the natural embedding $\mathrm{Sym}(\{\ell\}\times[m]) \hookrightarrow \mathrm{Sym}([\ell]\times[m])$ obtained by acting trivially on $[\ell-1]\times[m]$. Applying this to w^{-1}, we get

$$\overline{\mathrm{tr}\bigwedge^m w(X_1,\dots,X_r)} = \sum_{\pi'\in\mathrm{Sym}(\{\ell\}\times[m])} (-1)^{\pi'} \sum_{\substack{F':[\ell]\times[m]\to[d] \\ F'(1,-)\text{ increasing}}} \prod_{(u,k)\in[\ell]\times[m]} \left(X_{w^{-1}(u)}\right)_{F'(u,k),F'(\widetilde{T}\pi'(u,k))}. \tag{4.5}$$

Set $\Omega = [2] \times [\ell] \times [m]$, $\Omega_{s,u} = \{s\} \times \{u\} \times [m]$, and for $\gamma \in \Omega$, define

$$\widetilde{w}(\gamma) = \begin{cases} w(u) & \gamma = (1,u,k) \\ w^{-1}(u) & \gamma = (2,u,k) \end{cases}.$$

Define $T \in \mathrm{Sym}(\Omega)$ by

$$T(s,u,k) := (s, \widetilde{T}(u,k)). \tag{4.6}$$

By combining (4.4) and (4.5), we get

$$\left|\mathrm{tr} \bigwedge^m w(X_1,\dots,X_r)\right|^2 = \sum_{(\pi,\pi')\in\prod_{s=1}^2 \mathrm{Sym}(\Omega_{s,\ell})} (-1)^{\pi\pi'} \sum_{\substack{F:\Omega\to[d] \\ F(1,1,-)\text{ increasing} \\ F(2,1,-)\text{ increasing}}} \prod_{\gamma\in\Omega} \left(X_{\widetilde{w}(\gamma)}\right)_{F(\gamma),F(T\pi\pi'(\gamma))}. \tag{4.7}$$

The map $\pi \mapsto T\pi T^{-1}$ is an isomorphism $\mathrm{Sym}(\Omega_{s,\ell}) \xrightarrow{\cong} \mathrm{Sym}(\Omega_{s,1})$, for $s \in [2]$. Hence,

$$\begin{aligned}
\left|\mathrm{tr} \bigwedge^m w(X_1,\dots,X_r)\right|^2 &= \sum_{(\pi,\pi')\in \mathrm{Sym}(\Omega_{1,1})\times \mathrm{Sym}(\Omega_{2,1})} (-1)^{\pi\pi'} \sum_{\substack{F:\Omega\to[d] \\ F(1,1,-)\text{ increasing} \\ F(2,1,-)\text{ increasing}}} \prod_{\gamma\in\Omega} \left(X_{\widetilde{w}(\gamma)}\right)_{F(\gamma),F(\pi\pi' T(\gamma))} \\
&= \sum_{(\pi,\pi')\in \mathrm{Sym}(\Omega_{1,1})\times \mathrm{Sym}(\Omega_{2,1})} (-1)^{\pi\pi'} \sum_{\substack{F:\Omega\to[d] \\ F(1,1,-)\text{ increasing} \\ F(2,1,-)\text{ increasing}}} \prod_{\gamma\in\Omega} \left(X_{\widetilde{w}(\gamma)}\right)_{F(\gamma),F((\pi\pi')^{-1} T(\gamma))} \\
&= \sum_{(\pi,\pi')\in \mathrm{Sym}(\Omega_{1,1})\times \mathrm{Sym}(\Omega_{2,1})} (-1)^{\pi\pi'} \sum_{\substack{F:\Omega\to[d] \\ F\circ\pi(1,1,-)\text{ increasing} \\ F\circ\pi'(2,1,-)\text{ increasing}}} \prod_{\gamma\in\Omega} \left(X_{\widetilde{w}(\gamma)}\right)_{F(\pi\pi'\gamma),F(T(\gamma))},
\end{aligned} \tag{4.8}$$

where, in the last equality, we replaced F by $F \circ (\pi'\pi)^{-1}$.

Let $\Phi = (A_1,\dots,A_r,B_1,\dots,B_r)$ be the partition given by

$$A_i = \{(s,u,k) \mid \widetilde{w}(s,u,k) = i\} \quad B_i = \{(s,u,k) \mid \widetilde{w}(s,u,k) = -i\}. \tag{4.9}$$

For each $(\pi,\pi') \in \mathrm{Sym}(\Omega_{1,1}) \times \mathrm{Sym}(\Omega_{2,1})$, set

$$Z_{\pi,\pi'} := \left\{ (F,\Sigma) : {\scriptstyle \substack{F:\Omega\to[d],\Sigma\in S_\Phi \\ F\circ\pi(1,1,-),F\circ\pi'(2,1,-)\text{ increasing} \\ F\circ T = F\circ\pi\pi'\circ\Sigma}} \right\}. \tag{4.10}$$

The sets $Z_{\pi,\pi'}$ are disjoint. We denote

$$Z := \bigcup_{\pi,\pi'} Z_{\pi,\pi'}. \tag{4.11}$$

Remark 4.1. Note that we have a map $Z \to \mathrm{Sym}(\Omega_{1,1}) \times \mathrm{Sym}(\Omega_{2,1})$ sending (F,Σ) to the unique pair (π_F,π'_F) such that $(F,\Sigma) \in Z_{\pi_F,\pi'_F}$.

Rewriting (4.8) using Weingarten calculus (Corollary 2.15), we have:

Proposition 4.2. *Let $w \in F_r$ be a cyclically reduced word. Then:*

$$\mathbb{E}\left(\left|\mathrm{tr}\left(\bigwedge^m w(X_1,\dots,X_r)\right)\right|^2\right) = \sum_{(F,\Sigma)\in Z} (-1)^{\pi_F\pi'_F} \widetilde{\mathrm{Wg}}(\Sigma^2). \tag{4.12}$$

5. Estimating the contribution of a single orbit in Z

In this section we introduce an action of $H := \prod_{(s,u)\in[2]\times[\ell]} \mathrm{Sym}(\Omega_{s,u})$ on Z, and estimate (4.12) restricted to each H-orbit. The action can be described as follows.

For every $(s,u)\in[2]\times[\ell]$, the group $\mathrm{Sym}(\Omega_{s,u})$ acts on Z in the following way: if $u\neq 1$, the action of $\pi_{s,u}\in\mathrm{Sym}(\Omega_{s,u})$ is

$$\pi_{s,u}.(F,\Sigma)=(F\circ\pi_{s,u}^{-1},\pi_{s,u}\circ\Sigma\circ T^{-1}\pi_{s,u}^{-1}T). \tag{5.1}$$

If $s\in[2]$ and $\pi_{s,1}\in\mathrm{Sym}(\Omega_{s,1})$, then

$$\pi_{s,1}.(F,\Sigma)=(F\circ\pi_{s,1}^{-1},\Sigma\circ T^{-1}\pi_{s,1}^{-1}T). \tag{5.2}$$

The above group actions commute, which gives rise to an action of H. Note that $(\pi_{1,1},\pi_{2,1}).Z_{\pi,\pi'}=Z_{\pi_{1,1}\pi,\pi_{2,1}\pi'}$. If $u\neq 1$, then $\pi_{s,u}.(Z_{\pi,\pi'})=Z_{\pi,\pi'}$.

Definition 5.1. For each $u,v\in[\ell]$, we define $*$:$\mathrm{Sym}(\Omega_{s,u})\times\mathrm{Sym}(\Omega_{s,v})\to\mathrm{Sym}(\Omega_{s,v})$ by

$$\pi_{s,u}*\pi_{s,v}:=T^{v-u}\pi_{s,u}T^{u-v}\pi_{s,v}. \tag{5.3}$$

Note that $*$ is associative.

Let $h:=\prod_{(s,u)}\pi_{s,u}\in H$ and denote $\overline{h}:=\prod_{(s,u)\neq(1,1),(2,1)}\pi_{s,u}$. Then $h.\Sigma=\overline{h}\circ\Sigma\circ T^{-1}h^{-1}T$. Since $\widetilde{\mathrm{Wg}}$ is invariant under conjugation in H,

$$\widetilde{\mathrm{Wg}}\left((h.(\Sigma))^2\right)=\widetilde{\mathrm{Wg}}\left(\Psi_h\circ\Sigma\circ\Psi_h\circ\Sigma\right),$$

where $\Psi_h=T^{-1}h^{-1}T\overline{h}\in H$. On each $\Omega_{s,u}$, Ψ_h has the following form:

Lemma 5.2. *We have*

$$\Psi_h|_{\Omega_{s,u}}=\begin{cases}T^{-1}\pi_{s,2}^{-1}T & \text{if } u=1\\ \pi_{s,u+1}^{-1}*\pi_{s,u} & \text{if } u\neq 1,\ell\\ \pi_{s,1}^{-1}*\pi_{s,\ell} & \text{if } u=\ell.\end{cases}$$

Corollary 5.3. *Let $(\widehat{F},\widehat{\Sigma})$ be a representative of an H-orbit $\mathcal{O}_{(\widehat{F},\widehat{\Sigma})}$, with $(\pi_{\widehat{F}},\pi'_{\widehat{F}})=(\mathrm{Id},\mathrm{Id})$. Then:*

$$\frac{1}{\left|\mathcal{O}_{(\widehat{F},\widehat{\Sigma})}\right|}\sum_{(F,\Sigma)\in\mathcal{O}_{(\widehat{F},\widehat{\Sigma})}}(-1)^{\pi_F\pi'_F}\widetilde{\mathrm{Wg}}(\Sigma^2)=\frac{1}{m!^{2\ell}}\prod_{i=1}^{r}\sum_{\substack{h_i\in\prod_{\widetilde{w}=i}\mathrm{Sym}(\Omega_{s,u})\\ h'_i\in\prod_{\widetilde{w}=-i}\mathrm{Sym}(\Omega_{s,u})}}(-1)^{h_ih'_i}\mathrm{Wg}\left(h_i\widehat{\Sigma}|_{B_i}h'_i\widehat{\Sigma}|_{A_i}\right). \tag{5.4}$$

Proof. For each $h=\prod_{(s,u)}\pi_{s,u}\in H$, write $\nu(h):=(-1)^{\pi_{1,1}\pi_{2,1}}$. Consider the bijection $\psi:H\to H$, $\psi\left(\prod_{(s,u)}\pi_{s,u}\right)=\prod_{(s,u)}\theta_{s,u}$, where, for $s=1,2$,

$$(\theta_{s,2},\dots,\theta_{s,\ell},\theta_{s,1})=(\pi_{s,2},\pi_{s,2}*\pi_{s,3},\dots,\pi_{s,2}*\cdots*\pi_{s,\ell},\pi_{s,2}*\cdots*\pi_{s,\ell}*\pi_{s,1}),$$

and observe that $\nu(\psi(h))=(-1)^h=(-1)^{T^{-1}h^{-1}T}$. Further note that

$$(\pi_{s,2}*\cdots*\pi_{s,u+1})^{-1}*\pi_{s,2}*\cdots*\pi_{s,u}=T^{-1}\pi_{s,u+1}^{-1}T,$$

and hence $\Psi_{\psi(h)}=\prod_{(s,u)}T^{-1}\pi_{s,u}^{-1}T$. Changing variables using ψ, the left hand side of (5.4) is:

$$\begin{aligned}\frac{1}{|H|}\sum_{h\in H}\nu(h)\widetilde{\mathrm{Wg}}\left(\Psi_h\circ\widehat{\Sigma}\circ\Psi_h\circ\widehat{\Sigma}\right)&=\frac{1}{m!^{2\ell}}\sum_{h\in H}\nu(\psi(h))\widetilde{\mathrm{Wg}}\left(\prod_{(s,u)}T^{-1}\pi_{s,u}^{-1}T\circ\widehat{\Sigma}\circ\prod_{(s,u)}T^{-1}\pi_{s,u}^{-1}T\circ\widehat{\Sigma}\right)\\&=\frac{1}{m!^{2\ell}}\sum_{h\in H}(-1)^h\widetilde{\mathrm{Wg}}\left(\prod_{(s,u)}\pi_{s,u}\circ\widehat{\Sigma}\circ\prod_{(s,u)}\pi_{s,u}\circ\widehat{\Sigma}\right)\\&=\frac{1}{m!^{2\ell}}\sum_{h\in H}(-1)^h\prod_{i=1}^{r}\mathrm{Wg}\left(\prod_{(s,u):\widetilde{w}=i}\pi_{s,u}\widehat{\Sigma}|_{B_i}\prod_{(s,u):\widetilde{w}=-i}\pi_{s,u}\widehat{\Sigma}|_{A_i}\right),\end{aligned}$$

where in each line above, $h=\prod_{(s,u)}\pi_{s,u}$. ☐

Corollary 5.4. *Set $\ell_i := \frac{|A_i|}{m}$ for each $i\in[r]$. Then the following holds:*

$$\mathbb{E}\left(\left|\operatorname{tr}\left(\bigwedge^m w(X_1,\dots,X_r)\right)\right|^2\right)\le\frac{|Z|}{m!^{\ell}}\prod_{i=1}^{r}\frac{1}{d\cdots(d-m\ell_i+1)}. \tag{5.5}$$

Proof. By Proposition 4.2, Corollary 5.3, Equation (2.11), Lemma 2.7, and by Equation (2.3),

$$\begin{aligned}\mathbb{E}\left(\left|\operatorname{tr}\left(\bigwedge^m w(X_1,\dots,X_r)\right)\right|^2\right) &= \sum_{(\widehat{F},\widehat{\Sigma})\in Z}(-1)^{\pi_{\widehat{F}}\pi'_{\widehat{F}}}\widetilde{\mathrm{Wg}}(\widehat{\Sigma}^2)\\
&=\sum_{(\widehat{F},\widehat{\Sigma})\in Z}\frac{1}{\left|\mathcal{O}_{(\widehat{F},\widehat{\Sigma})}\right|}\sum_{(F,\Sigma)\in\mathcal{O}_{(\widehat{F},\widehat{\Sigma})}}(-1)^{\pi_F\pi'_F}\widetilde{\mathrm{Wg}}(\Sigma^2)\\
&\le\sum_{(\widehat{F},\widehat{\Sigma})\in Z}\frac{1}{m!^{2\ell}}\prod_{i=1}^{r}\left|\sum_{\substack{h_i\in\prod_{\widetilde{w}=i}\mathrm{Sym}(\Omega_{s,u})\\ h'_i\in\prod_{\widetilde{w}=-i}\mathrm{Sym}(\Omega_{s,u})}}(-1)^{h_ih'_i}\mathrm{Wg}\left(h_i\widehat{\Sigma}|_{B_i}h'_i\widehat{\Sigma}|_{A_i}\right)\right|\\
&\le\sum_{(\widehat{F},\widehat{\Sigma})\in Z}\frac{1}{m!^{2\ell}}\prod_{i=1}^{r}\frac{m!^{2\ell_i}}{(m\ell_i)!^2}\sum_{\lambda\vdash m\ell_i:\chi_\lambda\subseteq\mathrm{Ind}_{\mathrm{S}_m^{\ell_i}}^{S_{m\ell_i}}(\mathrm{sgn})}\langle\chi_\lambda,\mathrm{sgn}\rangle_{\mathrm{S}_m^{\ell_i}}\frac{\chi_\lambda(1)^2}{\rho_\lambda(1)}\\
&=\frac{|Z|}{m!^{\ell}}\prod_{i=1}^{r}\frac{m!^{\ell_i}}{(m\ell_i)!}\sum_{\lambda\vdash m\ell_i:\chi_\lambda\subseteq\mathrm{Ind}_{\mathrm{S}_m^{\ell_i}}^{S_{m\ell_i}}(\mathrm{sgn})}\frac{\langle\chi_\lambda,\mathrm{sgn}\rangle_{\mathrm{S}_m^{\ell_i}}\chi_\lambda(1)}{\prod_{(a,b)\in\lambda}(d+b-a)}.\end{aligned} \tag{5.6}$$

Note that the irreducible characters χ_λ in $\mathrm{Ind}_{\mathrm{S}_m^{\ell_i}}^{S_{m\ell_i}}(\mathrm{sgn})$ correspond to Young diagrams $\lambda\vdash m\ell_i$ with at most ℓ_i columns. If the columns of λ are of lengths $j_1\ge\ldots\ge j_{\ell_i}$ then

$$\prod_{(a,b)\in\lambda}(d+b-a)\ge d\cdots(d-j_1+1)\cdot d\cdots(d-j_2+1)\cdot\ldots\cdot d\cdots(d-j_{\ell_i}+1)\ge d\cdots(d-m\ell_i+1). \tag{5.7}$$

Combining (5.6) with (5.7) implies the corollary. ☐

6. Estimates on $|Z|$

In this section we give upper bounds on $|Z|$, defined in (4.11). We first set some notation. For each $0\neq i,j\in[-r,r]$, set

$$R_i:=\left\{\gamma\in\Omega:\widetilde{w}\circ T^{-1}(\gamma)=i\right\}=\begin{cases}T(A_i) & i>0\\ T(B_{-i}) & i<0\end{cases},$$

$$C_j:=\{\gamma\in\Omega:\widetilde{w}(\gamma)=-j\}=\begin{cases}B_j & j>0\\ A_{-j} & j<0\end{cases},$$

$$V_{ij}:=\left\{\gamma\in\Omega:\widetilde{w}\circ T^{-1}(\gamma)=i,\widetilde{w}(\gamma)=-j\right\}=R_i\cap C_j.$$

Following Remark 3.6, it is helpful to picture a $2r\times 2r$ matrix, whose (i,j)-th entry is the set V_{ij}, with $R_{-r},\dots,R_r$ correspond to rows, and $C_{-r},\dots,C_r$ correspond to columns. Denote

$$\ell_{i,j}:=\frac{|V_{ij}|}{m}\text{ and }\ell_i:=\frac{|R_i|}{m}=\frac{|C_i|}{m}. \tag{6.1}$$

Observe that $\ell_i = \ell_{-i}$, $\ell_{i,j} = \ell_{j,i}$ and note that $\ell_i = \frac{|A_i|}{m}$ if $i > 0$, so that (6.1) extends the definition of ℓ_i in Corollary 5.4. For each $0 \neq j \in [-r, r]$ set

$$C_j^+ := \bigcup_{i<j} V_{ij},\ C_+ := \bigcup_j C_j^+.$$

For each $i \in [r]$ and each $\Sigma \in S_\Phi$, denote $\eta_i := T \circ (\Sigma^{-1})\,|_{B_i}$ and $\eta_{-i} := T \circ (\Sigma^{-1})\,|_{A_i}$. Notice that $\eta_i(C_i) = R_i$ for all i. Define the following sets:

$$W' := \{(F : \Omega \to [d], \Sigma \in S_\Phi) : F \circ T = F \circ \Sigma\}, \tag{6.2}$$

and

$$W := \left\{(F, \Sigma) \in W' : F(s, 1, -) \text{ is one-to-one } \forall s \in [2]\right\}.$$

Proposition 6.1. *We have*

$$|Z| = |W| \leq |W'| \leq \binom{d+m\ell}{m\ell}(m\ell)! \prod_{0\neq k=-r}^{r} \frac{(m\ell_k)!}{\left(\sum_{i<k} m\ell_{i,k}\right)!}.$$

Proof. The map $(F, \Sigma) \mapsto \left(F \circ \pi_F \pi_F', \Sigma \circ T^{-1}\pi_F\pi_F' T\right)$ is a bijection between Z and W, giving the first equality. Clearly, $|W| \leq |W'|$.

In order to prove the last inequality, we use the map $\Phi_+ : W' \to \{f : C_+ \to [d]\}$, sending $(F, \Sigma) \in W'$ to $F|_{C_+}$. We estimate $|W'|$ by analyzing the fibers of Φ_+. Let $f \in \Phi_+(W')$ and suppose it has a shape ν_+ (see Definition 3.4). We write $\nu_{j,+}$ for the shapes of $f|_{C_j^+}$. We reveal $(F, \Sigma) \in \Phi_+^{-1}(f)$ row by row, starting with the $-r$-th row R_{-r} and making sure that, in each step, $F \circ T|_{T^{-1}(R_k)} = F \circ \Sigma|_{T^{-1}(R_k)}$, or, equivalently, $F|_{R_k} = F \circ \eta_k^{-1}|_{R_k}$.

1. There are at most $(m\ell_{-r})!$ options for η_{-r}. Note that $R_{-r} \subseteq C_+$ and hence by (6.2), the choice of η_{-r} determines $F|_{C_{-r}}$. At this point, $F|_{R_{-r+1}}$ is determined as well.
2. Note that $C_{-r+1}^+ = V_{-r,-r+1}$. There are at most:
   (a) $\binom{m\ell_{-r+1}}{m\ell_{-r,-r+1}}$ options for the sets $\eta_{-r+1}(C_{-r+1}^+)$ and $\eta_{-r+1}(C_{-r+1}\backslash C_{-r+1}^+)$.
   (b) $(m(\ell_{-r+1} - \ell_{-r,-r+1}))!$ options for $\eta_{-r+1}|_{C_{-r+1}\backslash C_{-r+1}^+} : C_{-r+1}\backslash C_{-r+1}^+ \to \eta_{-r+1}(C_{-r+1}\backslash C_{-r+1}^+)$.
   (c) $(\nu_{-r+1,+})!$ options for $\eta_{-r+1} : C_{-r+1}^+ \to \eta_{-r+1}(C_{-r+1}^+)$.
3. More generally, assume, by induction, that we have fixed $(\eta_i)_{i<k}$, and, thus, we have already determined $F|_{R_i}$ for $i \leq k$, $F|_{C_i}$ for $i < k$, and $F|_{C_+}$. Then there are at most:
   (a) $\binom{m\ell_k}{\sum_{i<k} m\ell_{i,k}}$ options for the sets $\eta_k(C_k^+)$ and $\eta_k(C_k\backslash C_k^+)$.
   (b) $\left(m(\ell_k - \sum_{i<k} \ell_{i,k})\right)!$ options for $\eta_k|_{C_k\backslash C_k^+} : C_k\backslash C_k^+ \to \eta_k(C_k\backslash C_k^+)$.
   (c) $(\nu_{k,+})!$ options for $\eta_k|_{C_k^+} : C_k^+ \to \eta_k(C_k^+)$.

After choosing $\eta_{-r}, \dots, \eta_r$, we have determined F. Furthermore, since $\sum_{0\neq k=-r}^r \nu_{k,+} = \nu_+$, we have $\prod_{0\neq k=-r}^r (\nu_{k,+})! \leq \nu_+!$. Hence,

$$\begin{aligned}\left|\Phi_+^{-1}(f)\right| &\leq \prod_{0\neq k=-r}^{r} \left(\binom{m\ell_k}{\sum_{i<k} m\ell_{i,k}} \left(m(\ell_k - \sum_{i<k} \ell_{i,k})\right)!\,(\nu_{k,+})!\right) \\ &= \prod_{0\neq k=-r}^{r} (\nu_{k,+})! \frac{(m\ell_k)!}{\left(\sum_{i<k} m\ell_{i,k}\right)!} \leq \nu_+! \prod_{0\neq k=-r}^{r} \frac{(m\ell_k)!}{\left(\sum_{i<k} m\ell_{i,k}\right)!}.\end{aligned} \tag{6.3}$$

Since $|C_+| = m\ell$, we have

$$\left|\left\{f \in \Phi_+(W') : f \text{ is of shape } \nu_+\right\}\right| \leq \frac{(m\ell)!}{\nu_+!}, \tag{6.4}$$

and there are at most $\binom{d+m\ell}{m\ell}$ possible shapes ν_+. Combining (6.3) and (6.4) we conclude:

$$\left|W'\right| \leq \sum_{\nu_+} \left|\left\{f \in \Phi_+(W') : f \text{ is of shape } \nu_+\right\}\right| \cdot \left|\Phi_+^{-1}(f)\right|$$

$$\leq \sum_{\nu_+} \frac{(m\ell)!}{\nu_+!} \cdot \nu_+! \prod_{0\neq k=-r}^{r} \frac{(m\ell_k)!}{\left(\sum_{i<k} m\ell_{i,k}\right)!} \leq \binom{d+m\ell}{m\ell}(m\ell)! \prod_{0\neq k=-r}^{r} \frac{(m\ell_k)!}{\left(\sum_{i<k} m\ell_{i,k}\right)!}. \qquad \square$$

7. Proof of Theorems 1.1 and 1.3

In this section we use the results of Sections 4, 5 and 6 to prove Theorems 1.1 and 1.3. We end the section with the proof of Theorem 1.6.

Proof of Theorem 1.3. Assume that $d = am$ for $a \geq \ell \geq 2$. By (3.12), we have:

$$\binom{d}{m\ell} = \binom{am}{m\ell} \geq \frac{a^{m\ell}}{\ell^{m\ell}}, \tag{7.1}$$

$$\binom{d+m\ell}{m\ell} \leq \left(\frac{a+\ell}{\ell}\right)^{m\ell} e^{m\ell} \leq \frac{a^{m\ell}(2e)^{m\ell}}{\ell^{m\ell}}. \tag{7.2}$$

We remind the reader the definition of ℓ_i and $\ell_{i,j}$ in (6.1). Concretely, for each $0 \neq i,j \in [-r,r]$, ℓ_i is the combined number of appearances of the letter x_i (with the convention that $x_{-i} = x_i^{-1}$) in w and w^{-1}, and $\ell_{i,j}$ is the combined number of appearances of the string "$x_i x_j^{-1}$" in w and in w^{-1}. In particular, we have $\sum_{i=1}^{r} \ell_i = \ell$, $\ell_{i,i} = 0$ and $\sum_{0\neq i\in[-r,r]} \ell_{i,k} = \ell_k$ and therefore:

$$\prod_{i=1}^{r} d\cdots(d-m\ell_i+1) \geq d\cdots(d-m\ell+1) \quad \text{and} \quad \frac{(m\ell_k)!}{\left(\sum_{i<k} m\ell_{i,k}\right)!\left(\sum_{i>k} m\ell_{i,k}\right)!} = \binom{m\ell_k}{\sum_{i>k} m\ell_{i,k}} \leq 2^{m\ell_k}. \tag{7.3}$$

By Corollary 5.4, Proposition 6.1 and by (7.3), (7.1) and (7.2), we obtain:

$$\mathbb{E}\left(\left|\operatorname{tr}\left(\bigwedge^m w(X_1,\dots,X_r)\right)\right|^2\right) \leq \frac{|Z|}{m!^\ell} \prod_{i=1}^{r} \frac{1}{d\cdots(d-m\ell_i+1)}$$

$$\leq \binom{d+m\ell}{m\ell} \cdot \frac{(m\ell)!}{\prod_{i=1}^{r} d\cdots(d-m\ell_i+1)} \cdot \frac{1}{m!^\ell} \cdot \prod_{0\neq k=-r}^{r} \frac{(m\ell_k)!}{\left(\sum_{i<k} m\ell_{i,k}\right)!}$$

$$\leq \binom{d+m\ell}{m\ell} \cdot \frac{(m\ell)!}{d\cdots(d-m\ell+1)} \cdot \frac{\prod_{0\neq k=-r}^{r}\left(\sum_{i>k} m\ell_{i,k}\right)!}{m!^\ell} \cdot \prod_{0\neq k=-r}^{r} \frac{(m\ell_k)!}{\left(\sum_{i<k} m\ell_{i,k}\right)!\left(\sum_{i>k} m\ell_{i,k}\right)!}$$

$$\leq \binom{d+m\ell}{m\ell} \cdot \binom{d}{m\ell}^{-1} \cdot \frac{(m\ell)!}{m!^\ell} \cdot \prod_{0\neq k=-r}^{r} 2^{m\ell_k} \leq (2e)^{m\ell}\ell^{m\ell} \cdot 2^{2m\ell} \leq (8e\ell)^{m\ell} \leq (22\ell)^{m\ell}.$$

Finally, note that if $d \geq (22\ell)^\ell m$ then $(22\ell)^{m\ell} \leq (\frac{d}{m})^m \leq \binom{d}{m}$. $\square$

We now turn to the proof of Theorem 1.1. We first deal with the case when the rank is bounded (and prove Conjecture 1.7 in this case) and then prove Theorem 1.1 in the unbounded case.

Definition 7.1. Given $w_1 \in F_{r_1}$ and $w_2 \in F_{r_2}$, we denote by $w_1 * w_2 \in F_{r_1+r_2}$ their *concatenation. For example, if $w = [x,y]$, then $w * w = [x,y] \cdot [z,w]$.*

We remind the reader the for a compact group G, and a word $w \in F_r$, we denote by $\tau_{w,G} := (w_G)_*(\mu_G^r)$ the word measure associated to w and G, and the Fourier coefficient of $\tau_{w,G}$ at $\rho \in \operatorname{Irr}(G)$ is $a_{w,G,\rho} := \int_{G^r} \rho(w(x_1,\dots,x_r))\mu_G^r = \int_G \rho(y)\tau_{w,G}$. If G is a compact connected semisimple Lie group, by [Bor83], the map $w_G : G^r \to G$ is a submersion outside a proper subvariety in G^r. It follows that in this case, or e.g. when G is a finite group, $\tau_{w,G}$ is absolutely continuous with respect to μ_G, and we can write $\tau_{w,G} = f_{w,G} \cdot \mu_G$, with $f_{w,G} \in L^1(G)$. Since $\tau_{w,G}$ is conjugate invariant, $f_{w,G}$ is a class function, and it can be written as a linear combination of characters $f_{w,G} = \sum_{\rho \in \mathrm{Irr}(G)} \overline{a_{w,G,\rho}} \cdot \rho$.

By Definition 7.1, we see that $\tau_{w_1 * w_2,G} = \tau_{w_1,G} * \tau_{w_2,G}$ for every $w_1 \in F_{r_1}$ and $w_2 \in F_{r_2}$. Since $\rho_1 * \rho_2 = \frac{\delta_{\rho_1,\rho_2}}{\rho_1(1)} \cdot \rho_1$ for every $\rho_1, \rho_2 \in \mathrm{Irr}(G)$, we have:

$$a_{w_1 * w_2,G,\rho} = \int_G \rho(g) \tau_{w_1 * w_2,G}(g) = \int_G \rho(g) \tau_{w_1,G} * \tau_{w_2,G}(g) = \frac{a_{w_1,G,\rho} \cdot a_{w_2,G,\rho}}{\rho(1)}. \tag{7.4}$$

Proposition 7.2. *For every $1 \neq w \in F_r$ and $d \in \mathbb{N}$, there exists $\epsilon(d, w) > 0$ such that:*

1. *For every compact connected semisimple Lie group G of rank d and every $\rho \in \mathrm{Irr}(G)$, we have $|a_{w,G,\rho}| \leq \rho(1)^{1-\epsilon(d,w)}$.*
2. *In particular, for every $1 \leq m \leq d$,*

$$\mathbb{E}_{\mathrm{U}_d}\left(\left| \mathrm{tr}\left(\bigwedge^m w(X_1, \ldots, X_r) \right) \right|^2 \right) = \mathbb{E}_{\mathrm{SU}_d}\left(\left| \mathrm{tr}\left(\bigwedge^m w(X_1, \ldots, X_r) \right) \right|^2 \right) \leq \binom{d}{m}^{2(1-\epsilon(d,w))}.$$

Proof. We first prove Item (1). Fix $w \in F_r$ and a compact connected semisimple Lie group G. Let $\tau_{w,G} = f_{w,G} \mu_G$ be the word measure. By (7.4), and since $a_{w^{-1},G,\rho} = \overline{a_{w,G,\rho}}$ for each $\rho \in \mathrm{Irr}(G)$, we have

$$a_{w * w^{-1},G,\rho} = \frac{|a_{w,G,\rho}|^2}{\rho(1)}. \tag{7.5}$$

Replacing w by $w * w^{-1}$, we may assume that all Fourier coefficients $a_{w,G,\rho}$ are in $\mathbb{R}_{\geq 0}$.

It follows from [GHS, Theorem 1.1] that $f_{w,G} \in L^{1+\epsilon'}(G)$ for some $\epsilon' = \epsilon'(G, w) > 0$. By Young's convolution inequality, it follows that $f_{w,G}^{*t} \in L^\infty(G)$ for all $t \geq t_0(G, w) := \left\lceil \frac{1+\epsilon'(G,w)}{\epsilon'(G,w)} \right\rceil$ (see e.g. [GHS, Section 1.1, end of p.3]). In particular, by (7.4), we deduce that:

$$f_{w,G}^{*t_0}(1) = \sum_{\rho \in \mathrm{Irr}(G)} \rho(1)^{2-t_0} a_{w,G,\rho}^{t_0} < \infty.$$

Since $a_{w,G,\rho} \geq 0$, we deduce that $a_{w,G,\rho} < \rho(1)^{1-\frac{2}{t_0(G,w)}}$ for all but finitely many $\rho \in \mathrm{Irr}(G)$. To deal with the remaining finitely many (non-trivial) representations of G, we simply use the bound $a_{w,G,\rho} < \rho(1)$, which follows e.g. by the Itô–Kawada equidistribution theorem [KI40] (see also [App14, Theorem 4.6.3]), since $\mathrm{Supp}(\tau_{w,G})$ generates G. Since there are only finitely many compact semisimple connected Lie groups of rank d, this implies Item (1).

Note that the character $\rho_{(1^m)} \otimes \rho_{(1^m)}^\vee$ of the representation $\bigwedge^m \mathbb{C}^d \otimes \left(\bigwedge^m \mathbb{C}^d\right)^\vee$ of SU_d is given by $|\mathrm{tr}\left(\bigwedge^m(A)\right)|^2$. Since $\rho_{(1^m)} \otimes \rho_{(1^m)}^\vee$ is a sum of irreducible characters, by applying the Itô–Kawada equidistribution theorem to each irreducible character, for each $1 \leq m \leq d$, we have

$$\mathbb{E}_{\mathrm{SU}_d}\left(\left| \mathrm{tr}\left(\bigwedge^m (w(X_1, \ldots, X_r)) \right) \right|^2 \right) = \mathbb{E}_{\mathrm{SU}_d}\left(\rho_{(1^m)} \otimes \rho_{(1^m)}^\vee (w(X_1, \ldots, X_r)) \right) < \rho_{(1^m)} \otimes \rho_{(1^m)}^\vee(1) = \binom{d}{m}^2.$$

Since there are only finitely many such m's, this implies Item (2). ∎

Theorem 1.1 now follows from Proposition 7.2 and the following Theorem.

Theorem 7.3. *For every $\ell \in \mathbb{N}$, there exist $\epsilon(\ell), C(\ell) > 0$ such that, for every $d \geq C(\ell)$, every $1 \leq m \leq d$, and every word $w \in F_r$ of length ℓ, one has:*

$$\mathbb{E}\left(\left| \mathrm{tr}\left(\bigwedge^m w(X_1, \ldots, X_r) \right) \right|^2 \right) \leq \binom{d}{m}^{2(1-\epsilon(\ell))}.$$

In order to prove Theorem 7.3, we need the following technical lemma.

Lemma 7.4. *Let $H(x) = -x\log(x) - (1-x)\log(1-x)$ be the binary entropy function. Then:*

(1) *For every $d \in \mathbb{N}$ and every $0 < x < 1$ such that $dx \in \mathbb{N}$, we have $\frac{2^{dH(x)}}{\sqrt{8dx(1-x)}} \le \binom{d}{xd} \le \frac{2^{dH(x)}}{\sqrt{\pi dx(1-x)}} \le 2^{dH(x)}$.*

(2) *Let $0 < \delta \le \frac{1}{2}$. Then for every $b \in [\delta, \frac{1}{2}]$, $a \in [\delta, b]$, and $d > \frac{1}{\delta^4}$ such that bd, ad, d are integers, one has:*

$$\binom{d}{(b-a)d} \le \binom{d}{bd}^{1-\delta^2}.$$

Proof. Item (1) follows e.g. from [CT06, Lemma 17.5.1]. The Taylor series of $H(x)$ around $1/2$ is

$$H(x) = 1 - \frac{1}{2\ln 2}\sum_{n=1}^{\infty}\frac{(1-2x)^{2n}}{n(2n-1)}. \tag{7.6}$$

Since $H'(x) = \log(\frac{1-x}{x})$, $H(x)$ is monotone increasing in $(0, 1/2)$, and therefore,

$$\begin{aligned} H(b) - H(b-a) \ge H(b) - H(b-\delta) &= \frac{1}{2\ln 2}\left(\sum_{n=1}^{\infty}\frac{(1-2b+2\delta)^{2n} - (1-2b)^{2n}}{n(2n-1)}\right) \\ &\ge \frac{1}{2\ln 2}\left((1-2b+2\delta)^2 - (1-2b)^2\right) = \frac{1}{2\ln 2}\left(4\delta^2 + 4\delta(1-2b)\right) \ge 2\delta^2. \end{aligned}$$

Since $d > \frac{1}{\delta^4} \ge 16$, we have $\frac{\log(d)}{d} \le \frac{1}{\sqrt{d}} \le \delta^2$. Combining with Item (1), we have:

$$\begin{aligned} \binom{d}{(b-a)d} &\le 2^{dH(b-a)} \le \sqrt{8db(1-b)}2^{d(H(b-a)-H(b))}\binom{d}{bd} \le 2^{-2d\delta^2+\log(d)}\binom{d}{bd} \\ &\le 2^{-d\delta^2}\binom{d}{bd} = \left(2^{-dH(b)}\right)^{\frac{\delta^2}{H(b)}}\binom{d}{bd} \le \binom{d}{bd}^{1-\frac{\delta^2}{H(b)}} \le \binom{d}{bd}^{1-\delta^2}. \end{aligned}$$ ∎

Proof of Theorem 7.3. Since $\bigwedge^m V \simeq \left(\bigwedge^{d-m} V\right)^\vee \otimes \chi_{\det}$, we may assume that $2m \le d$. Let $\delta(\ell) := (25\ell)^{-\ell}$, let $C(\ell) = \delta(\ell)^{-7}$, and suppose that $d \ge C(\ell)$. By Theorem 1.3, we may assume that $d \le \delta(\ell)^{-1}m$, and, in particular, $m \ge \delta(\ell)^{-6}$. As in the proof of Proposition 7.2, by replacing w by $w * w^{-1}$, we may assume that $a_{w,\mathrm{U}_d,\rho} \in \mathbb{R}_{\ge 0}$ for all $\rho \in \mathrm{Irr}(\mathrm{U}_d)$. By Theorem 2.5, we have for all $c \le \frac{d}{2}$:

$$\bigwedge^c V \otimes \bigwedge^c V^\vee \simeq \left(\bigwedge^c V \otimes \bigwedge^{d-c} V\right) \otimes \chi_{\det}^{-1} \simeq \bigoplus_{j=0}^{c} V_{\lambda_{(j)}},$$

where $\lambda_{(j)} = (1, \dots, 1, 0, \dots, 0, -1, \dots, -1)$, with -1 and 1 appearing j times. Moreover, $V_{\lambda_{(c)}}$ is the largest irreducible subrepresentation of $\bigwedge^c V \otimes \bigwedge^c V^\vee$, and we have $\rho_{\lambda_{(c)}}(1) \ge \frac{1}{c+1}\binom{d}{c}^2 \ge \binom{d}{c}^{3/2}$. By Theorem 1.3, and since all $a_{w,\mathrm{U}_d,\rho}$ are non-negative, if $c \le \lceil \delta(\ell)d \rceil \le (22\ell)^{-\ell}d$, then

$$\mathbb{E}\left(\rho_{\lambda_{(c)}} \circ w\right) \le \sum_{j=0}^{c}\mathbb{E}\left(\rho_{\lambda_{(j)}} \circ w\right) = \mathbb{E}\left(\rho_{\bigwedge^c V \otimes \bigwedge^c V^\vee} \circ w\right) = \mathbb{E}\left(\left|(\rho_{\bigwedge^c V} \circ w)\right|^2\right) \le \binom{d}{c} \le \rho_{\lambda_{(c)}}(1)^{2/3}.$$

Applying the last inequality for w^{*9}, recalling that $a_{w^{*t},\mathrm{U}_d,\rho} = \frac{a^t_{w,\mathrm{U}_d,\rho}}{\rho(1)^{t-1}}$ for all $\rho \in \mathrm{Irr}(\mathrm{U}_d)$, we get

$$\mathbb{E}\left(\left|\mathrm{tr}\left(\bigwedge^{\lceil\delta(\ell)d\rceil} w^{*9}(X_1, \dots, X_r)\right)\right|^2\right) = \sum_{j=0}^{\lceil\delta(\ell)d\rceil}\mathbb{E}\left(\rho_{\lambda_{(j)}} \circ w^{*9}\right) \le \sum_{j=0}^{\lceil\delta(\ell)d\rceil}\rho_{\lambda_{(j)}}(1)^{-2} \le \sum_{j=1}^{\infty}\frac{1}{j^2} < 2. \tag{7.7}$$

Note that, for each $\delta(\ell)d \le m \le \frac{d}{2}$, $\bigwedge^m V$ is a subrepresentation of $\bigwedge^{\lceil\delta(\ell)d\rceil} V \otimes \bigwedge^{m-\lceil\delta(\ell)d\rceil} V$, so

$$\bigwedge^m V \otimes \left(\bigwedge^m V\right)^\vee \hookrightarrow \left(\bigwedge^{\lceil\delta(\ell)d\rceil} V \otimes \left(\bigwedge^{\lceil\delta(\ell)d\rceil} V\right)^\vee\right) \otimes \left(\bigwedge^{m-\lceil\delta(\ell)d\rceil} V \otimes \left(\bigwedge^{m-\lceil\delta(\ell)d\rceil} V\right)^\vee\right).$$

Finally, by the positivity of the Fourier coefficients of w, by (7.7), by Lemma 7.4 (note that $m \geq \lceil \delta(\ell) d \rceil$) and by (3.12) (note that $\delta(\ell)^2 m \geq 1$),

$$\begin{aligned}\mathbb{E}\left(\left|\operatorname{tr}\left(\bigwedge^m w^{*9}(X_1,\ldots,X_r)\right)\right|^2\right) &\leq \mathbb{E}\left(\left|\operatorname{tr}\left(\bigwedge^{\lceil\delta(\ell)d\rceil} w^{*9}(X_1,\ldots,X_r)\right)\right|^2 \left|\operatorname{tr}\left(\bigwedge^{m-\lceil\delta(\ell)d\rceil} w^{*9}(X_1,\ldots,X_r)\right)\right|^2\right) \\ &\leq \mathbb{E}\left(\left|\operatorname{tr}\left(\bigwedge^{\lceil\delta(\ell)d\rceil} w^{*9}(X_1,\ldots,X_r)\right)\right|^2\right)\cdot \binom{d}{m-\lceil\delta(\ell)d\rceil}^2 \\ &\leq 2\binom{d}{m-\lceil\delta(\ell)d\rceil}^2 \leq \frac{d}{m}\binom{d}{m}^{2-2\delta(\ell)^2} \leq \binom{d}{m}^{2-\delta(\ell)^2}.\end{aligned}$$

By (3.12), $m+1 \leq 2^{2\sqrt{m}} \leq \binom{d}{m}^{2/\sqrt{m}}$ for each $m \leq \frac{d}{2}$. Hence,

$$\rho_{\lambda_{(m)}}(1) \geq \frac{1}{m+1}\binom{d}{m}^2 \geq \binom{d}{m}^{2(1-\frac{1}{\sqrt{m}})} \geq \binom{d}{m}^{2-2\delta(\ell)^3}. \tag{7.8}$$

Consequently, we get

$$\begin{aligned}\left(\mathbb{E}\left(\rho_{\lambda_{(m)}}\circ w\right)\right)^9 &= \mathbb{E}\left(\rho_{\lambda_{(m)}}\circ w^{*9}\right)\rho_{\lambda_{(m)}}(1)^8 \leq \mathbb{E}\left(\left|\operatorname{tr}\left(\bigwedge^m w^{*9}(X_1,\ldots,X_r)\right)\right|^2\right)\rho_{\lambda_{(m)}}(1)^8 \\ &\leq \binom{d}{m}^{2-\delta(\ell)^2}\rho_{\lambda_{(m)}}(1)^8 \leq \rho_{\lambda_{(m)}}(1)^{9-\frac{\delta(\ell)^2}{4}},\end{aligned}$$

and thus $\mathbb{E}\left(\rho_{\lambda_{(m)}}\circ w\right) \leq \rho_{\lambda_{(m)}}(1)^{1-\frac{\delta(\ell)^2}{36}}$. Taking $\epsilon(\ell) := \frac{\delta(\ell)^2}{72}$, and using $m+1 \leq \binom{d}{m}^{2\delta(\ell)^3}$, we get

$$\mathbb{E}\left(\left|\operatorname{tr}\left(\bigwedge^m w(X_1,\ldots,X_r)\right)\right|^2\right) = \sum_{j=0}^m \mathbb{E}\left(\rho_{\lambda_{(j)}}\circ w\right) \leq (m+1)\binom{d}{m}^{2-4\epsilon(\ell)} \leq \binom{d}{m}^{2(1-\epsilon(\ell))}. \qquad \square$$

We end the section with a proof of Theorem 1.6.

Proof of Theorem 1.6. Let $w \in F_r$. Denote $\widetilde{w} := w * w^{-1}$. Recall that $\tau_{\widetilde{w},G} = f_{\widetilde{w},G}\mu_G$ and note that for every $t \in \mathbb{N}$,

$$f_{\widetilde{w}^{*t},G} = f_{\widetilde{w},G}^{*t}.$$

Applying [LST19, Theorem 4], there are $C', M(w) \in \mathbb{N}$ such that, for $N(w) := C'\ell(w)^4$ and for every finite simple group G of size $> M(w)$, one has

$$\left|\sum_{1\neq\rho\in\operatorname{Irr}(G)} \frac{a_{\widetilde{w},G,\rho}^{N(w)}}{\rho(1)^{N(w)-1}}\rho(1)\right| = \left|\sum_{1\neq\rho\in\operatorname{Irr}(G)} a_{\widetilde{w}^{*N(w)},G,\rho}\rho(1)\right| = \left|f_{\widetilde{w}^{*N(w)},G}(1)-1\right| = \left|f_{\widetilde{w},G}^{*N(w)}(1)-1\right| < 1,$$

where the first equality follows from (7.4). Since $a_{\widetilde{w},G,\rho} = \frac{|a_{w,G,\rho}|^2}{\rho(1)} \geq 0$, we deduce that for each $1 \neq \rho \in \operatorname{Irr}(G)$

$$\frac{|a_{w,G,\rho}|^{2N(w)}}{\rho(1)^{2N(w)-2}} = \frac{|a_{w,G,\rho}|^{2N(w)}}{\rho(1)^{2N(w)-1}}\rho(1) = \frac{a_{\widetilde{w},G,\rho}^{N(w)}}{\rho(1)^{N(w)-1}}\rho(1) < 1,$$

from which the theorem follows for $\epsilon = \frac{1}{N(w)} = \frac{1}{C'\ell(w)^4}$. $\square$

8. Fourier coefficients of symmetric powers

In this section, we prove Theorem 1.4. Denote $\mathcal{J}_{m,d} = \{c_1 \leq \ldots \leq c_m : c_i \in [d]\}$. We first claim that, for each $A \in \operatorname{End}(\mathbb{C}^d)$ and $m \geq 1$,

$$\operatorname{tr}\left(\operatorname{Sym}^m A\right) = \frac{1}{m!}\sum_{\overrightarrow{a}\in[d]^m}\sum_{\pi\in S_m} A_{a_1 a_{\pi(1)}}\cdots A_{a_m a_{\pi(m)}}.$$

Indeed, for each $\overrightarrow{c}\in\mathcal{J}_{m,d}$, let $\nu_{\overrightarrow{c}}$ be the shape of $\overrightarrow{c}$ (see Definition 3.4) and set

$$v_{\overrightarrow{c}}:=\sqrt{\frac{1}{m!\cdot\nu_{\overrightarrow{c}}!}}\sum_{\pi\in S_m}e_{c_{\pi(1)}}\otimes\ldots\otimes e_{c_{\pi(m)}}.$$

Then $\{v_{\overrightarrow{c}}\}_{\overrightarrow{c}\in\mathcal{J}_{m,d}}$ is an orthonormal basis for $\mathrm{Sym}^m(\mathbb{C}^d)$. Given $A\in\mathrm{End}(\mathbb{C}^d)$, we have:

$$\begin{aligned}\mathrm{tr}\,(\mathrm{Sym}^m A)&=\sum_{\overrightarrow{c}\in\mathcal{J}_{m,d}}\langle A.v_{\overrightarrow{c}},v_{\overrightarrow{c}}\rangle=\sum_{\overrightarrow{c}\in\mathcal{J}_{m,d}}\frac{1}{m!\cdot\nu_{\overrightarrow{c}}!}\sum_{\pi,\pi'\in S_m}\langle Ae_{c_{\pi(1)}}\otimes\ldots\otimes Ae_{c_{\pi(m)}},e_{c_{\pi'(1)}}\otimes\ldots\otimes e_{c_{\pi'(m)}}\rangle\\&=\sum_{\overrightarrow{c}\in\mathcal{J}_{m,d}}\frac{1}{\nu_{\overrightarrow{c}}!}\sum_{\pi\in S_m}\langle Ae_{c_1}\otimes\ldots\otimes Ae_{c_m},e_{c_{\pi(1)}}\otimes\ldots\otimes e_{c_{\pi(m)}}\rangle\\&=\sum_{\overrightarrow{c}\in\mathcal{J}_{m,d}}\frac{1}{\nu_{\overrightarrow{c}}!}\sum_{\pi\in S_m}A_{c_1c_{\pi(1)}}\cdots A_{c_mc_{\pi(m)}}=\frac{1}{m!}\sum_{\overrightarrow{a}\in[d]^m}\sum_{\pi\in S_m}A_{a_1a_{\pi(1)}}\cdots A_{a_ma_{\pi(m)}},\end{aligned}$$

where the last equality follows since $\sum_{\pi\in S_m}A_{c_1c_{\pi(1)}}\cdots A_{c_mc_{\pi(m)}}$ is invariant under permuting $c_1,\ldots,c_m$, and since there are $\frac{m!}{\nu_{\overrightarrow{c}}!}$ vectors $\overrightarrow{a}\in[d]^m$ of a shape $\nu_{\overrightarrow{c}}$. In particular, for any word w,

$$\mathrm{tr}\,(\mathrm{Sym}^m w(X_1,\ldots,X_r))=\frac{1}{m!}\sum_{\overrightarrow{a}\in[d]^m}\sum_{\pi\in S_m}w(X_1,\ldots,X_r)_{a_1a_{\pi(1)}}\cdots w(X_1,\ldots,X_r)_{a_ma_{\pi(m)}}.\tag{8.1}$$

Proposition 8.1. *Let $w\in F_r$ be a cyclically reduced word. With $\Phi,T,\Omega,\Omega_{s,u}$ as in §4, we have:*

$$\mathbb{E}\left(\left|\mathrm{tr}\,(\mathrm{Sym}^m w(X_1,\ldots,X_r))\right|^2\right)=\frac{1}{m!^2}\sum_{(\pi,\pi',F,\Sigma)\in\widetilde{Z}}\widetilde{\mathrm{Wg}}(\Sigma^2),\tag{8.2}$$

where

$$\widetilde{Z}:=\left\{(\pi,\pi',F,\Sigma):\substack{F:\Omega\to[d],\Sigma\in S_\Phi\\ \pi,\pi'\in\mathrm{Sym}(\Omega_{1,1})\times\mathrm{Sym}(\Omega_{2,1})\\ F\circ T=F\circ\pi\pi'\circ\Sigma}\right\}.$$

Proof. Similarly to (4.4), we have

$$\begin{aligned}\mathrm{tr}\,(\mathrm{Sym}^m w(X_1,\ldots,X_r))&=\frac{1}{m!}\sum_{\overrightarrow{a}\in[d]^m}\sum_{\pi\in S_m}\sum_{\substack{f:[\ell+1]\times[m]\to[d]\\ f(1,k)=a_k,f(\ell+1,k)=a_{\pi(k)}}}\prod_{(u,k)\in[\ell]\times[m]}\left(X_{w(u)}\right)_{f(u,k),f(u+1,k)}\\&=\sum_{\pi\in\mathrm{Sym}(\{\ell\}\times[m])}\sum_{F:[\ell]\times[m]\to[d]}\prod_{(u,k)\in[\ell]\times[m]}\left(X_{w(u)}\right)_{F(u,k),F(\widetilde{T}\pi(u,k))}.\end{aligned}$$

Consequently, as in (4.8), we have:

$$\mathbb{E}\left(\left|\mathrm{tr}\,(\mathrm{Sym}^m w(X_1,\ldots,X_r))\right|^2\right)=\frac{1}{m!^2}\sum_{(\pi,\pi')\in\mathrm{Sym}(\Omega_{1,1})\times\mathrm{Sym}(\Omega_{2,1})}\sum_{F:\Omega\to[d]}\prod_{\gamma\in\Omega}\left(X_{\widetilde{w}(\gamma)}\right)_{F(\pi\pi'\gamma),F(T(\gamma))}.$$

The Proposition now follows from Corollary 2.15. □

We next define an action of $H:=\prod_{(s,u)\in[2]\times[\ell]}\mathrm{Sym}(\Omega_{s,u})$ on $\widetilde{Z}$ in the same way as in §5. For $(s,u)\in[2]\times([\ell]\backslash\{1\})$ and $\pi_{s,u}\in\mathrm{Sym}\,(\Omega_{s,u})$,

$$\pi_{s,u}.\left(\pi,\pi',F,\Sigma\right):=\left(\pi,\pi',F\circ\pi_{s,u}^{-1},\pi_{s,u}\circ\Sigma\circ T^{-1}\pi_{s,u}^{-1}T\right),$$

and if $(\pi_{1,1},\pi_{2,1})\in\mathrm{Sym}(\Omega_{1,1})\times\mathrm{Sym}(\Omega_{2,1})$,

$$(\pi_{1,1},\pi_{2,1}).\left(\pi,\pi',F,\Sigma\right):=\left(\pi_{1,1}\pi,\pi_{2,1}\pi',F\circ\pi_{1,1}^{-1}\pi_{2,1}^{-1},\Sigma\circ T^{-1}\pi_{1,1}^{-1}\pi_{2,1}^{-1}T\right).$$

Proof of Theorem 1.4. The proof is similar to the proof of Theorem 1.3. The only difference is that now, summing over the H-orbit kills all representations that do not appear in $\mathrm{Ind}_{S_m^{\ell_i}}^{S_{m\ell_i}}(1)$, rather than the representations not in $\mathrm{Ind}_{S_m^{\ell_i}}^{S_{m\ell_i}}(\mathrm{sgn})$. By Lemma 2.3, the irreducible subrepresentations

χ_λ of $\mathrm{Ind}_{S_m^{\ell_i}}^{S_{m\ell_i}}(1)$ correspond to partitions $\lambda = (\lambda_1, \dots, \lambda_{\ell_i})$ with at most ℓ_i rows, and, therefore, $\prod_{(a,b)\in\lambda}(d+b-a) \geq (d-\ell)^{m\ell_i}$. As in Corollary 5.3 and (5.6), the average of $\widetilde{\mathrm{Wg}}(\Sigma^2)$ over an H-orbit $H.(\widehat{\pi}, \widehat{\pi}', \widehat{F}, \widehat{\Sigma})$ is bounded by

$$\frac{1}{m!^{2\ell}} \prod_{i=1}^{r} \left| \sum_{\substack{h_i \in \prod_{\widetilde{w}=i} \mathrm{Sym}(\Omega_{s,u}) \\ h_i' \in \prod_{\widetilde{w}=-i} \mathrm{Sym}(\Omega_{s,u})}} \mathrm{Wg}\left(h_i \widehat{\Sigma}|_{B_i} h_i' \widehat{\Sigma}|_{A_i}\right) \right|$$

$$\leq \frac{1}{m!^{\ell}} \prod_{i=1}^{r} \frac{m!^{\ell_i}}{(m\ell_i)!} \sum_{\lambda \vdash m\ell_i : \chi_\lambda \subseteq \mathrm{Ind}_{S_m^{\ell_i}}^{S_{m\ell_i}}(1)} \frac{\chi_\lambda(1)\langle \chi_\lambda, 1\rangle_{S_m^{\ell_i}}}{\prod_{(a,b)\in\lambda}(d+b-a)} \leq \frac{1}{m!^{\ell}} \frac{1}{(d-\ell)^{m\ell}}. \tag{8.3}$$

Denote $\widetilde{Z}_{\pi,\pi'} := \left\{(F,\Sigma) : (\pi,\pi',F,\Sigma) \in \widetilde{Z}\right\}$. Since $\widetilde{Z}_{\mathrm{Id},\mathrm{Id}} = W'$, Proposition 6.1 implies that

$$\left|\widetilde{Z}\right| = m!^2 \left|\widetilde{Z}_{\mathrm{Id},\mathrm{Id}}\right| = m!^2 \left|W'\right| \leq m!^2 \binom{d+m\ell}{m\ell}(m\ell)! \prod_{0\neq k=-r}^{r} \frac{(m\ell_k)!}{(\sum_{i<k} m\ell_{i,k})!}. \tag{8.4}$$

As in the proof of Theorem 1.3, if $d \geq m\ell$, then

$$\begin{aligned}
\mathbb{E}\left(\left|\mathrm{tr}\left(\mathrm{Sym}^m w(X_1,\dots,X_r)\right)\right|^2\right) &= \frac{1}{m!^2} \sum_{(\pi,\pi',F,\Sigma)\in\widetilde{Z}} \widetilde{\mathrm{Wg}}(\Sigma^2) \leq \left|\widetilde{Z}\right| \frac{1}{m!^{\ell+2}} \frac{1}{(d-\ell)^{m\ell}} \\
&\leq \frac{(d+m\ell)\cdots(d+1)}{(d-\ell)^{m\ell} m!^{\ell}} \prod_{0\neq k=-r}^{r} \frac{(m\ell_k)!}{(\sum_{i<k} m\ell_{i,k})!} \\
&\leq 4^{m\ell}\ell^{m\ell} \prod_{0\neq k=-r}^{r} \binom{m\ell_k}{m\ell_k/2} 4^{m\ell}\ell^{m\ell}2^{2m\ell} = (16\ell)^{m\ell}. \qquad \square
\end{aligned}$$

Appendix A. Fourier coefficients of the power word and a Diaconis–Shahshahani type result

In this Appendix, we formulate two results. The first is a computation of the Fourier coefficients of the power word $w = x^l$ for representations $\rho_\lambda \in \mathrm{Irr}(\mathrm{U}_d)$, where $\widetilde{\lambda}$ (see Remark 2.6) has at most $\frac{d}{2l}$ boxes. The second is a Diaconis–Shahshahani type result for the m-th coefficient of the characteristic polynomial of a word w in random unitary matrices. Both statements are consequences of known results.

Proposition A.1. *Let $w = x^l$ be the l-th power word. Then, for every $m \in \mathbb{N}$ and every $d \geq 2ml$,*

(1) *We have*

$$\mathbb{E}\left(\left|\rho_\lambda \circ w\right|^2\right) = \frac{1}{m!} \sum_{\sigma\in S_m} l^{\ell(\sigma)} \left|\chi_\lambda(\sigma)\right|^2,$$

for all $\lambda \vdash m$. In particular, $\mathbb{E}\left(\left|\rho_\lambda \circ w\right|^2\right) \leq l^m$.

(2) *We have*

$$\mathbb{E}\left(\left|\mathrm{tr}\left(\bigwedge^m w\right)\right|^2\right) = \mathbb{E}\left(\left|\mathrm{tr}\left(\mathrm{Sym}^m w\right)\right|^2\right) = \binom{l+m-1}{m}.$$

Proof. For every matrix $A \in \mathrm{U}_d$ and every $\mu \vdash m$, set

$$\mathrm{tr}_\mu(A) := \prod_{j=1}^{m} \mathrm{tr}(A^j)^{a_j}, \tag{A.1}$$

where $\mu = (1^{a_1} \cdots m^{a_m})$ is the partition $m = \underbrace{(1 + \ldots + 1)}_{a_1 \text{ times}} + \ldots + \underbrace{(m + \ldots + m)}_{a_m \text{ times}}$. The functions tr_μ correspond to the power-sums symmetric functions p_μ. Given $\lambda \vdash m$, the character $\rho_\lambda(A)$ is a Schur polynomial in the eigenvalues of A, and, hence, it can be expressed in terms of $\mathrm{tr}_\mu(A)$ via the following formula (see e.g. [Mac95, I.7, page 114]),

$$\rho_\lambda(A) = \sum_{\mu \vdash m} \frac{\chi_\lambda(\mu)}{\prod_{j=1}^m a_j! j^{a_j}} \cdot \mathrm{tr}_\mu(A), \tag{A.2}$$

where $\chi_\lambda(\mu)$ is the value of the character $\chi_\lambda \in \mathrm{Irr}(S_m)$ on the elements with cycle type μ. In addition, by (1.2), for every pair of partitions $\mu = (1^{a_1} \cdots m^{a_m})$ and $\mu' = (1^{b_1} \cdots m^{b_m})$ of m, we have:

$$\mathbb{E}\left(\mathrm{tr}_\mu(X^l)\mathrm{tr}_{\mu'}(\overline{X}^l)\right) = \mathbb{E}\left(\prod_{j=1}^m \mathrm{tr}(X^{jl})^{a_j} \mathrm{tr}(\overline{X}^{jl})^{b_j}\right) = \delta_{\mu,\mu'} \prod_{j=1}^m (jl)^{a_j} a_j!. \tag{A.3}$$

Combining (A.2) and (A.3), and using the fact that the number of permutations $\sigma \in S_m$ of cycle type $\mu = (1^{a_1} \cdots m^{a_m})$ is $\frac{m!}{\prod_{j=1}^m a_j! j^{a_j}}$, we obtain:

$$\begin{aligned}
\mathbb{E}\left(\left|\rho_\lambda\left(X^l\right)\right|^2\right) &= \sum_{\mu \vdash m} |\chi_\lambda(\mu)|^2 \frac{\mathbb{E}\left(\left|\mathrm{tr}_\mu\left(X^l\right)\right|^2\right)}{\left(\prod_{j=1}^m a_j! j^{a_j}\right)^2} = \sum_{\mu \vdash m} \frac{l^{\ell(\mu)} |\chi_\lambda(\mu)|^2}{\prod_{j=1}^m a_j! j^{a_j}} \\
&= \frac{1}{m!} \sum_{\mu \vdash m} \frac{m!}{\prod_{j=1}^m a_j! j^{a_j}} l^{\ell(\mu)} |\chi_\lambda(\mu)|^2 = \frac{1}{m!} \sum_{\sigma \in S_m} l^{\ell(\sigma)} |\chi_\lambda(\sigma)|^2.
\end{aligned} \tag{A.4}$$

The second claim of Item (1) follows from Schur orthogonality and the inequality $l^{\ell(\sigma)} \leq l^m$.
For Item (2), note that $\mathrm{tr}\left(\bigwedge^m w\right) = \rho_{(1^m)} \circ w$ and $\mathrm{tr}\left(\mathrm{Sym}^m w\right) = \rho_{(m^1)} \circ w$. The corresponding characters of S_m are the sign and the trivial characters. Thus, (A.4) becomes

$$\mathbb{E}\left(\left|\mathrm{tr}\left(\bigwedge^m w\right)\right|^2\right) = \mathbb{E}\left(\left|\mathrm{tr}\left(\mathrm{Sym}^m w\right)\right|^2\right) = \mathbb{E}_{S_m}\left(l^{\ell(\sigma)}\right) = \frac{1}{m!} \sum_{k=1}^m \begin{bmatrix} m \\ k \end{bmatrix} l^k = \binom{l+m-1}{m},$$

where $\begin{bmatrix} m \\ k \end{bmatrix}$ is the number of permutations of m elements with exactly k disjoint cycles, also known as the unsigned Stirling number of the first kind. The last equality follows e.g. from [GKP94, Equation (6.11)]. This concludes Item (2). □

We next prove a Diaconis–Shahshahani type result. We first recall the following proposition, which is a consequence of [MSS07, Theorem 2] and [Rŏ6, Theorem 4.1] (see also [MP19, Corollary 1.13]).

Proposition A.2. *Let $w \in F_r$, and let $\mu = (1^{a_1} \cdots m^{a_m})$, $\mu' = (1^{b_1} \cdots m^{b_m})$ be partitions of m. Let $p(w) \in \mathbb{N}$ be such that $w = u^{p(w)}$ with $u \in F_r$ a non-power. Then:*

$$\lim_{d \to \infty} \mathbb{E}_{\mathrm{U}_d}\left(\mathrm{tr}_\mu(w)\mathrm{tr}_{\mu'}(w^{-1})\right) = \lim_{d \to \infty} \mathbb{E}_{\mathrm{U}_d}\left(\prod_{j=1}^m \mathrm{tr}(w^j)^{a_j} \mathrm{tr}(w^{-j})^{b_j}\right) = \delta_{\mu,\mu'} \prod_{j=1}^m a_j!(jp(w))^{a_j}. \tag{A.5}$$

Since the joint moments of $\mathrm{tr}(w^1), \ldots, \mathrm{tr}(w^m)$ converge, as $d \to \infty$, to the joint moments of independent complex normal random variables, an application of the moment method (as was done in [DS94] for $w = x$, and later in [Rŏ6, MSS07] for a general word) implies

Corollary A.3 ([Rŏ6, Theorem 4.1], [MSS07, Theorem 2])**.** *The random variables $\mathrm{tr}(w^1), \ldots, \mathrm{tr}(w^m)$ converge in distribution to $\sqrt{p(w)}Z_1, \ldots, \sqrt{mp(w)}Z_m$, as $d \to \infty$, where $Z_1, \ldots, Z_m$ are independent complex normal variables.*

In [DG06], Diaconis and Gamburd combined Corollary A.3 for $w = x$ (namely [DS94]), together with Newton's identities relating elementary and power sum symmetric functions to give a formula for the limit behavior of the random variables $\mathrm{tr} \bigwedge^m X$ with X is a random unitary matrix in U_d. Repeating the argument for a general word w yields the following description of $\lim_{d\to\infty} \mathrm{tr}_{\mathrm{U}_d} \bigwedge^m w$.

Corollary A.4 (cf. [DG06, Proposition 4]). *Let $w \in F_r$ be a word and let $m \in \mathbb{N}$. Then the sequence of random variables $\mathrm{tr}_{\mathrm{U}_d} \bigwedge^m w$ converges in distribution, as $d \to \infty$, to the polynomial in the normal variables $Z_1, \dots, Z_m$ given by:*

$$\frac{1}{m!}\det\begin{pmatrix} \sqrt{p(w)}Z_1 & 1 & 0 & \dots & 0 \\ \sqrt{2p(w)}Z_2 & \sqrt{p(w)}Z_1 & 2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sqrt{(m-1)p(w)}Z_{m-1} & \sqrt{(m-2)p(w)}Z_{m-2} & \sqrt{(m-3)p(w)}Z_{m-3} & \dots & (m-1) \\ \sqrt{mp(w)}Z_m & \sqrt{(m-1)p(w)}Z_{m-1} & \sqrt{(m-2)p(w)}Z_{m-2} & \dots & \sqrt{p(w)}Z_1 \end{pmatrix}.$$

Example A.5. Let $m = 3$. Then for every Borel set $A \subseteq \mathbb{C}$,

$$\lim_{d\to\infty} \mathbb{P}\left(\mathrm{tr}_{\mathrm{U}_d} \bigwedge^3 w(X_1, ..., X_r) \in A\right) = \mathbb{P}\left(f(Z_1, Z_2, Z_3) \in A\right),$$

where Z_1, Z_2, Z_3 are i.i.d normal variables, and

$$f(Z_1, Z_2, Z_3) = \frac{p(w)^{3/2}}{6}Z_1^3 - \frac{p(w)}{\sqrt{2}}Z_1Z_2 + \frac{p(w)^{1/2}}{\sqrt{3}}Z_3.$$

References

[AA16] A. Aizenbud and N. Avni. Representation growth and rational singularities of the moduli space of local systems. *Invent. Math.*, 204(1):245–316, 2016.

[ABB17] L. Arguin, D. Belius, and P. Bourgade. Maximum of the characteristic polynomial of random unitary matrices. *Comm. Math. Phys.*, 349(2):703–751, 2017.

[AGZ10] G. W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

[App14] D. Applebaum. *Probability on compact Lie groups*, volume 70 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2014. With a foreword by Herbert Heyer.

[BG06] D. Bump and A. Gamburd. On the averages of characteristic polynomials from classical groups. *Comm. Math. Phys.*, 265(1):227–274, 2006.

[BHNY08] P. Bourgade, C. P. Hughes, A. Nikeghbali, and M. Yor. The characteristic polynomial of a random unitary matrix: a probabilistic approach. *Duke Math. J.*, 145(1):45–69, 2008.

[Bor83] A. Borel. On free subgroups of semisimple groups. *Enseign. Math. (2)*, 29(1-2):151–164, 1983.

[Bro24] Y. Brodsky. Word Measures on Unitary groups: Improved Bounds for Small Representations. *Int. Math. Res. Not.*, 2024(14):10790–10821, 05 2024.

[CFK+03] J. B. Conrey, D. W. Farmer, J. P. Keating, M. O. Rubinstein, and N. C. Snaith. Autocorrelation of random matrix polynomials. *Comm. Math. Phys.*, 237(3):365–395, 2003.

[CM17] B. Collins and S. Matsumoto. Weingarten calculus via orthogonality relations: new applications. *ALEA Lat. Am. J. Probab. Math. Stat.*, 14(1):631–656, 2017.

[CMN18] R. Chhaibi, T. Madaule, and J. Najnudel. On the maximum of the CβE field. *Duke Math. J.*, 167(12):2243–2345, 2018.

[Col03] B. Collins. Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral, and free probability. *Int. Math. Res. Not.*, (17):953–982, 2003.

[CS06] B. Collins and P. Śniady. Integration with respect to the Haar measure on unitary, orthogonal and symplectic group. *Comm. Math. Phys.*, 264(3):773–795, 2006.

[CT06] T. M. Cover and J. A. Thomas. *Elements of information theory*. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, second edition, 2006.

[DG06] P. Diaconis and A. Gamburd. Random matrices, magic squares and matching polynomials. *Electron. J. Combin.*, 11(2):Research Paper 2, 26, 2004/06.

[DS94] P. Diaconis and M. Shahshahani. On the eigenvalues of random matrices. volume 31A, pages 49–62. 1994. Studies in applied probability.

[ET15] A. Elkasapy and A. Thom. On the length of the shortest non-trivial element in the derived and the lower central series. *J. Group Theory*, 18(5):793–804, 2015.

[FH91] W. Fulton and J. Harris. *Representation theory*, volume 129 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991. A first course, Readings in Mathematics.

[GH19] I. Glazer and Y. I. Hendel. On singularity properties of convolutions of algebraic morphisms. *Selecta Math. (N.S.)*, 25(1):Art. 15, 41, 2019.

[GH21] I. Glazer and Y. I. Hendel. On singularity properties of convolutions of algebraic morphisms—the general case. *J. Lond. Math. Soc. (2)*, 103(4):1453–1479, 2021. With an appendix by Glazer, Hendel and Gady Kozma.

[GH24] I. Glazer and Y. I. Hendel. On singularity properties of word maps and applications to probabilistic Waring-type problems. *Mem. Amer. Math. Soc.*, 299(1497), 2024.

[GHS] I. Glazer, Y. I. Hendel, and S. Sodin. Integrability of pushforward measures by analytic maps. *To Appear in Algebraic Geometry*.

[GKP94] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete mathematics*. Addison-Wesley Publishing Company, Reading, MA, second edition, 1994. A foundation for computer science.

[GLM12] R. Guralnick, M. Larsen, and C. Manack. Low degree representations of simple Lie groups. *Proc. Amer. Math. Soc.*, 140(5):1823–1834, 2012.

[HKO01] C. P. Hughes, J. P. Keating, and N. O'Connell. On the characteristic polynomial of a random unitary matrix. *Comm. Math. Phys.*, 220(2):429–451, 2001.

[Joh97] K. Johansson. On random matrices from the compact classical groups. *Ann. of Math. (2)*, 145(3):519–545, 1997.

[KI40] Y. Kawada and K. Itô. On the probability distribution on a compact group. I. *Proc. Phys.-Math. Soc. Japan (3)*, 22:977–998, 1940.

[KS00] J. P. Keating and N. C. Snaith. Random matrix theory and $\zeta(1/2+it)$. *Comm. Math. Phys.*, 214(1):57–89, 2000.

[LST19] M. Larsen, A. Shalev, and P. H. Tiep. Probabilistic Waring problems for finite simple groups. *Ann. of Math. (2)*, 190(2):561–608, 2019.

[Mac95] I. G. Macdonald. *Symmetric functions and Hall polynomials*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, second edition, 1995. With contributions by A. Zelevinsky, Oxford Science Publications.

[Mec19] E. S. Meckes. *The random matrix theory of the classical compact groups*, volume 218 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2019.

[MP19] M. Magee and D. Puder. Matrix group integrals, surfaces, and mapping class groups I: $\mathcal{U}(n)$. *Invent. Math.*, 218(2):341–411, 2019.

[MS17] J. A. Mingo and R. Speicher. *Free probability and random matrices*, volume 35 of *Fields Institute Monographs*. Springer, New York; Fields Institute for Research in Mathematical Sciences, Toronto, ON, 2017.

[MSS07] J. A. Mingo, P. Śniady, and R. Speicher. Second order freeness and fluctuations of random matrices. II. Unitary random matrices. *Adv. Math.*, 209(1):212–240, 2007.

[PZ18] E. Paquette and O. Zeitouni. The maximum of the CUE field. *Int. Math. Res. Not. IMRN*, (16):5028–5119, 2018.

[Rai97] E. M. Rains. High powers of random elements of compact Lie groups. *Probab. Theory Related Fields*, 107(2):219–241, 1997.

[Rai03] E. M. Rains. Images of eigenvalue distributions under power maps. *Probab. Theory Related Fields*, 125(4):522–538, 2003.

[Rei] A. Reiser. Pushforwards of measures on real varieties under maps with rational singularities. *arXiv:1807.00079*.

[R0̆6] F. Rădulescu. Combinatorial aspects of Connes's embedding conjecture and asymptotic distribution of traces of products of unitaries. In *Operator theory 20*, volume 6 of *Theta Ser. Adv. Math.*, pages 197–205. Theta, Bucharest, 2006.

[VDN92] D. V. Voiculescu, K. J. Dykema, and A. Nica. *Free random variables*, volume 1 of *CRM Monograph Series*. American Mathematical Society, Providence, RI, 1992. A noncommutative probability approach to free products with applications to random matrices, operator algebras and harmonic analysis on free groups.

[Voi91] D. Voiculescu. Limit laws for random matrices and free products. *Invent. Math.*, 104(1):201–220, 1991.

[Wei78] D. Weingarten. Asymptotic behavior of group integrals in the limit of infinite rank. *J. Mathematical Phys.*, 19(5):999–1001, 1978.

[Wey39] H. Weyl. *The Classical Groups. Their Invariants and Representations*. Princeton University Press, Princeton, N.J., 1939.

Department of Mathematics, Northwestern University, 2033 Sheridan Road, Evanston, IL 60208, USA

Email address: avni.nir@gmail.com

URL: https://sites.math.northwestern.edu/~nir/

Department of Mathematics, Northwestern University, 2033 Sheridan Road, Evanston, IL 60208, USA

Email address: itayglazer@gmail.com

URL: https://sites.google.com/view/itay-glazer