

THE CYLINDRICAL WIDTH OF TRANSITIVE SETS

ASHWIN SAH, MEHTAAB SAWHNEY, AND YUFEI ZHAO

ABSTRACT. We show that for every $1 \leq k \leq d/(\log d)^C$, every finite transitive set of unit vectors in \mathbb{R}^d lies within distance $O(1/\sqrt{\log(d/k)})$ of some codimension k subspace, and this distance bound is best possible. This extends a result of Ben Green, who proved it for $k = 1$.

1. INTRODUCTION

The following counterintuitive fact was conjectured by the third author and proved by Green [4]. It says that every finite transitive subset of a high dimensional sphere is close to some hyperplane. Here a subset X of a sphere in \mathbb{R}^d is *transitive* if for every $x, x' \in X$, there is some $g \in \mathcal{O}(\mathbb{R}^d)$ so that $gX = X$ and $gx = x'$. We say that X has *width* at most $2r$ if it lies within distance r of some hyperplane. The finiteness assumption is important since otherwise the whole sphere is a counterexample.

Theorem 1.1 (Green [4]). *Let X be a finite transitive subset of the unit sphere in \mathbb{R}^d . Then the width of X is at most $O(1/\sqrt{\log d})$. Furthermore, this upper bound is best possible up to a constant factor.*

The bound in the theorem is tight since the set X obtained by taking all permutations and coordinate-wise \pm signings of the unit vector $(1, 1/\sqrt{2}, \dots, 1/\sqrt{d})/\sqrt{H_d}$, where $H_d = 1 + 1/2 + \dots + 1/d \sim \log d$, has width on the order of $1/\sqrt{\log d}$.

Green's proof uses a clever induction scheme along with sophisticated group theoretic arguments, including an application of the classification of finite simple groups.

We generalize Green's result by showing that a finite transitive set lies not only near some hyperplane, but in fact it lies near a subspace of codimension k , as long as k is not too large.

We say that $X \subset \mathbb{R}^d$ has *k -cylindrical width* at most $2r$ if X lies within distance r of some affine codimension k subspace. The case $k = 1$ corresponds to the usual notion of width. Our main result below implies that every finite transitive subset of the unit sphere in \mathbb{R}^d has k -cylindrical width $O(1/\sqrt{\log(d/k)})$ as long as k is not too large.

Theorem 1.2. *There is an absolute constant $C > 0$ so that the following holds. Let $1 \leq k \leq d/(\log(3d))^C$. Let X be a finite transitive subset of the unit sphere in \mathbb{R}^d . Then there is a real k -dimensional subspace W such that*

$$\sup_{\mathbf{x} \in X} \|\text{proj}_W \mathbf{x}\|_2 \lesssim \frac{1}{\sqrt{\log(d/k)}}.$$

Here and throughout $a \lesssim b$ means that $a \leq C'b$ for some absolute constant C' . We write $\|\mathbf{x}\|_2$ for the usual Euclidean norm of a vector \mathbf{x} . Also proj_W is the orthogonal projection onto W .

We deduce the above theorem from a complex version using a theorem on restricted invertibility (see Section 6). A transitive subset of the complex unit sphere is defined to be the orbit of a point under the action of some subgroup of the unitary group.

Sah and Sawhney were supported by NSF Graduate Research Fellowship Program DGE-1745302. Zhao was supported by NSF Award DMS-1764176, a Sloan Research Fellowship, and the MIT Solomon Buchsbaum Fund.

Theorem 1.3. *There is an absolute constant $C > 0$ so that the following holds. Let $1 \leq k \leq d/(\log(3d))^C$. Let X be a finite transitive subset of the unit sphere in \mathbb{C}^d . Then there is a complex k -dimensional subspace W such that*

$$\sup_{\mathbf{x} \in X} \|\text{proj}_W \mathbf{x}\|_2 \lesssim \frac{1}{\sqrt{\log(d/k)}}.$$

We suspect that the $1 \leq k \leq d/(\log(3d))^C$ hypothesis is unnecessary in both [Theorems 1.2](#) and [1.3](#).

Conjecture 1.4. *Let $1 \leq k \leq d$. Let X be a finite transitive subset of the unit sphere in \mathbb{C}^d . Then there is a complex k -dimensional subspace W such that*

$$\sup_{\mathbf{x} \in X} \|\text{proj}_W \mathbf{x}\|_2 \lesssim \frac{1}{\sqrt{\log(2d/k)}}.$$

One particularly intriguing special case of [Conjecture 1.4](#) is that every finite transitive set of unit vectors in \mathbb{R}^d has k -cylindrical width $o(1)$ for all $k = o(d)$.

We prove a matching lower bound on the cylindrical radius (See [Section 7](#) for proof.)

Theorem 1.5. *Let $1 \leq k \leq d$. There exists a transitive set X in \mathbb{R}^d such that for any (real or complex) k -dimensional subspace W we have*

$$\sup_{\mathbf{x} \in X} \|\text{proj}_W \mathbf{x}\|_2 \gtrsim \frac{1}{\sqrt{\log(2d/k)}}.$$

We propose another closely related conjecture: every finite transitive set in \mathbb{R}^d lies inside a small cube.

Conjecture 1.6. *Let X be a finite transitive subset of the unit sphere in \mathbb{R}^d (or \mathbb{C}^d). Then there is a unitary basis L such that*

$$\sup_{\mathbf{x} \in X, \mathbf{v} \in L} |\langle \mathbf{v}, \mathbf{x} \rangle| \lesssim \frac{1}{\sqrt{\log d}}. \quad (1.1)$$

Establishing an upper bound that decays to zero as $d \rightarrow \infty$ would already be interesting. Note that [Theorem 1.3](#) implies the existence of a set L of orthonormal vectors with $|L| \geq d^{0.99}$ so that (1.1) holds (and likely extendable to $|L| \geq d/(\log d)^C$ via our techniques). Proving either conjecture in full appears to require additional ideas.

Remark. Green's proof [\[4\]](#) of [Theorems 1.2](#) and [1.3](#) in the case $k = 1$ contains two errors. The first error is due to a missing supremum inside the integral in the first and second lines of the last display equation in proof of Proposition 2.1 on page 560. The second error occurs at the final equality step of the top display equation on page 569, after right after (4.4); here an orthogonality relation was incorrectly applied as it requires an unjustified exchange of the integral and supremum. Our proof here corrects these errors. Green has also updated the arXiv version of his paper [\[4\]](#) incorporating these corrections.

2. PROOF STRATEGY

The subspace W in [Theorem 1.3](#) must vary according to the transitive set X . On other hand, the strategy is to construct a single probability distribution μ (depending only on the symmetry group $G \leq \text{U}(\mathbb{C}^d)$ but not on X) on the set $\text{Gr}_{\mathbb{C}}(k, d)$ of k -dimensional subspaces of \mathbb{C}^d . This is an important idea introduced by Green (for $k = 1$).

Definition 2.1. Let $1 \leq k \leq d$. Let $f_k(d)$ be the smallest value so that for every finite $G \leq \text{U}(\mathbb{C}^d)$, there is a probability measure μ on $\text{Gr}_{\mathbb{C}}(k, d)$ such that for all $\mathbf{v} \in \mathbb{S}(\mathbb{C}^d)$,

$$\int \sup_{g \in G} \|\text{proj}_W(g\mathbf{v})\|_2^2 d\mu(W) \leq f_k(d)^2$$

The values $f_k(d)$ are well defined since the space of probability measures μ in question is closed under weak limits.

Our main result about $f_k(d)$ is stated below.

Theorem 2.2. *If $k \leq d/(\log d)^{20}$, then*

$$f_k(d) \lesssim \frac{1}{\sqrt{\log(d/k)}}.$$

Proof of Theorem 1.3 given Theorem 2.2. Let our transitive set X be the orbit of $\mathbf{v} \in \mathbb{S}(\mathbb{C}^d)$ under the action of the finite subgroup $G \leq \mathbf{U}(\mathbb{C}^d)$. By Theorem 2.2 and Definition 2.1, there is a measure μ on $\text{Gr}_{\mathbb{C}}(k, d)$ such that

$$\int \sup_{g \in G} \|\text{proj}_W(g\mathbf{v})\|_2^2 d\mu(W) \leq f_k(d)^2.$$

Therefore there is some k -dimensional subspace W with

$$\sup_{g \in G} \|\text{proj}_W(g\mathbf{v})\|_2 \leq f_k(d) \lesssim \frac{1}{\sqrt{\log(d/k)}}. \quad \square$$

To prove Theorem 2.2, we will decompose G to “smaller”, more restricted cases, namely irreducible and primitive representations. We will also need to consider permutation groups (for both the reduction step as well as the primitive case).

2.1. Preliminaries.

Definition 2.3. We say that $G \leq \mathbf{U}(\mathbb{C}^d)$ is *imprimitive* if there is a *system of imprimitivity*: a decomposition

$$\mathbb{C}^d = \bigoplus_{i=1}^{\ell} V_i$$

with $0 < \dim V_i < d$ such that for every $g \in G$ and $i \in [\ell]$ one has $gV_i = V_j$ for some $j \in [\ell]$. (The subspaces V_i need not be orthogonal.) Otherwise we say that G is *primitive*.

Remark. Both primitivity and irreducibility are properties of a representation, rather than intrinsic to a group. We identify $G \leq \mathbf{U}(\mathbb{C}^d)$ with its natural representation on \mathbb{C}^d .

It follows from Maschke’s theorem that primitive group representations are irreducible.

Definition 2.4. Given $\mathbf{v} = (v_1, \dots, v_d) \in \mathbb{C}^d$, let

$$\mathbf{v}^{\succ} = (|v_{\sigma(1)}|, \dots, |v_{\sigma(d)}|) \in \mathbb{R}^d$$

where σ is a permutation of $[d]$ so that

$$|v_{\sigma(1)}| \geq \dots \geq |v_{\sigma(d)}|.$$

We write v_i^{\succ} for the i -th coordinate of \mathbf{v}^{\succ} . Let

$$\text{Dom}(\mathbf{v}) = \{\mathbf{w} \in \mathbb{C}^d : w_i^{\succ} \leq v_i^{\succ} \text{ for all } i \in [d]\}.$$

Let (here \mathfrak{S}_d denotes the symmetric group)

$$\Gamma_d := \mathfrak{S}_d \ltimes (\mathbb{S}^1)^d \leq \mathbf{U}(\mathbb{C}^d)$$

be the group that acts on \mathbb{C}^d by permuting its coordinates and multiplying individual coordinates by unit complex numbers. Then $\text{Dom}(\mathbf{v})$ is the convex hull of the Γ_d -orbit of \mathbf{v} .

We define some variants of $f_k(d)$ when the group G is restricted to special types.

Definition 2.5. Given $k \in [d]$, let $f_k^{\text{irred}}(d)$ (resp. $f_k^{\text{prim}}(d)$) be the smallest value so that for every finite $G \leq \text{U}(\mathbb{C}^d)$ which is irreducible (resp. primitive), there is a probability measure μ on $\text{Gr}_{\mathbb{C}}(k, d)$ such that for every $\mathbf{v} \in \mathbb{S}(\mathbb{C}^d)$,

$$\int \sup_{g \in G} \|\text{proj}_W(g\mathbf{v})\|_2^2 d\mu(W) \leq f_k^{\text{irred}}(d)^2 \quad (\text{resp. } f_k^{\text{prim}}(d)^2).$$

The permutation action on \mathbb{C}^d deserves special attention.

Definition 2.6. Let $f_k^{\text{sym}}(d)$ be the smallest value so that there is a probability measure μ on $\text{Gr}_{\mathbb{C}}(k, d)$ such that for every $\mathbf{v} \in \mathbb{S}(\mathbb{C}^d)$,

$$\int \sup_{\mathbf{u} \in \text{Dom}(\mathbf{v})} \|\text{proj}_W(\mathbf{u})\|_2^2 d\mu(W) \leq f_k^{\text{sym}}(d)^2.$$

Define $f_k^{\text{alt}}(d)$ to be the same with the additional constraint that μ is supported on the set of k -dimensional subspaces of the hyperplane $x_1 + \dots + x_d = 0$.

We will often equivalently consider, instead of μ on $\text{Gr}_{\mathbb{C}}(k, d)$, the corresponding measure μ^* on the complex Stiefel manifold $V_k(\mathbb{C}^d)$, that is, μ^* is derived from μ by first sampling a μ -random k -dimensional subspace W of \mathbb{C}^d , and then outputting a uniformly sampled a unitary basis $(\mathbf{w}_1, \dots, \mathbf{w}_k)$ of W . We have $\|\text{proj}_W(\mathbf{u})\|_2^2 = \sum_{\ell=1}^k |\langle g\mathbf{v}_1, \mathbf{w}_\ell \rangle|^2$.

2.2. Reductions. We first reduce the general problem to the irreducible case.

Proposition 2.7. *If $1 \leq k < \ell \leq d$ then*

$$f_k(d) \leq \max\left\{\sqrt{k/\ell}, \sup_{d' \geq d/(2\ell)} f_{\lceil 2kd'/d \rceil}^{\text{irred}}(d')\right\}.$$

We then reduce the irreducible case to the primitive case and the alternating case.

Proposition 2.8. *If $k \leq d/2$, then*

$$f_k^{\text{irred}}(d) \leq \max_{d_1 d_2 = d} \left(\min\left\{f_{\lceil k/d_1 \rceil}^{\text{prim}}(d_2), f_k^{\text{alt}}(d_1) + \mathbb{1}_{k \geq d_1}\right\} \right).$$

The symmetric and alternating cases can be handled explicitly, yielding the following.

Proposition 2.9. *If $k \leq d/(\log d)^5$, then*

$$f_k^{\text{sym}}(d) \leq f_k^{\text{alt}}(d) \lesssim 1/\sqrt{\log(d/k)}.$$

This leaves the primitive case, which we prove by invoking a group theoretic result proved by Green [4, Proposition 4.2] that allows us to once again reduce to the alternating case once again.

Proposition 2.10. *There is an absolute constant $c > 0$ such that for $k \leq cd/(\log d)^4$ we have*

$$f_k^{\text{prim}}(d) \lesssim \sup_{d' \geq cd/(\log d)^4} f_k^{\text{alt}}(d').$$

2.3. Putting everything together. We are now in position to derive [Theorem 2.2](#) using the preceding statements.

Proposition 2.11. *If $k \leq 2d/(\log d)^{10}$ then $f_k^{\text{prim}}(d) \lesssim 1/\sqrt{\log(d/k)}$.*

Proof. Combine [Propositions 2.9](#) and [2.10](#). □

Proposition 2.12. *If $k \leq d/(\log d)^{10}$ then $f_k^{\text{irred}}(d) \lesssim 1/\sqrt{\log(d/k)}$.*

Proof. By [Proposition 2.8](#), we have

$$f_k^{\text{irred}}(d) \leq \max_{d_1 d_2 = d} (\min(f_{\lceil k/d_1 \rceil}^{\text{prim}}(d_2), f_k^{\text{alt}}(d_1) + \mathbb{1}_{k \geq d_1})).$$

First consider the case $d_1 \leq k$. We have

$$\lceil k/d_1 \rceil \leq \frac{2d}{d_1 (\log d)^{10}} \leq \frac{2d_2}{(\log d_2)^{10}}.$$

By [Proposition 2.11](#), we have

$$f_{\lceil k/d_1 \rceil}^{\text{prim}}(d_2) \lesssim \frac{1}{\sqrt{\log(d_2/\lceil k/d_1 \rceil)}} \leq \frac{1}{\sqrt{\log(d/(2k))}}.$$

Now consider the case $d_1 > k$. Since $d_2(d_1/k) = d/k$, we have $\max\{d_1, d_2/k\} \geq \sqrt{d/k}$. If $d_2 \geq \sqrt{d/k}$, then

$$f_{\lceil k/d_1 \rceil}^{\text{prim}}(d_2) = f_1^{\text{prim}}(d_2) \lesssim \frac{1}{\sqrt{\log d_2}} \lesssim \frac{1}{\sqrt{\log(d/k)}}.$$

On the other hand, if $d_1/k \geq \sqrt{d/k}$, then $d_1/k \geq (\log d)^5$ so

$$k \leq \frac{d_1}{(\log d)^5} \leq \frac{d_1}{(\log d_1)^5}.$$

Hence [Proposition 2.9](#) yields

$$f_k^{\text{alt}}(d_1) \lesssim \frac{1}{\sqrt{\log(d_1/k)}} \lesssim \frac{1}{\sqrt{\log(d/k)}}.$$

Thus it follows that, for all $d_1 d_2 = d$,

$$\min(f_{\lceil k/d_1 \rceil}^{\text{prim}}(d_2), f_k^{\text{alt}}(d_1) + \mathbb{1}_{k \geq d_1}) \lesssim \frac{1}{\sqrt{\log(d/k)}},$$

and the result follows. \square

Now we show the main result assuming the above statements.

Proof of Theorem 2.2. Let $\ell = \lceil \sqrt{dk} \rceil \geq 2k$. We have

$$k/\ell \lesssim \sqrt{k/d} \lesssim \frac{1}{\sqrt{\log(d/k)}}.$$

Also, if $d' \geq d/(2\ell)$ then

$$\left\lceil \frac{2kd'}{d} \right\rceil \leq \frac{d'}{d/(2\ell)} \leq \frac{d'}{(\log d)^{10}} \leq \frac{d'}{(\log d')^{10}}.$$

By [Proposition 2.12](#), we have

$$f_{\lceil 2kd'/d \rceil}^{\text{irred}}(d') \lesssim \frac{1}{\sqrt{\log(d'/\lceil 2kd'/d \rceil)}} \lesssim \frac{1}{\sqrt{\log(d/(2\ell))}} \lesssim \frac{1}{\sqrt{\log(d/k)}}.$$

Applying [Proposition 2.7](#) to k and $\ell = \lceil \sqrt{dk} \rceil$, we find

$$f_k(d) \leq \max(\sqrt{k/\ell}, \sup_{d' \geq d/(2\ell)} f_{\lceil 2kd'/d \rceil}^{\text{irred}}(d')) \lesssim \frac{1}{\sqrt{\log(d/k)}}. \quad \square$$

2.4. Paper outline. In [Section 3](#), we prove the two key reductions, [Propositions 2.7](#) and [2.8](#). In [Section 4](#), we prove the key estimate for the symmetric and alternating cases, [Proposition 2.9](#). In [Section 5](#), we prove the primitive case, [Proposition 2.10](#). Finally, in [Section 6](#) we deduce a real version from the complex version, proving [Theorem 1.2](#). In [Section 7](#) we demonstrate optimality of our results by exhibiting the matching lower bound [Theorem 1.5](#).

3. REDUCTION TO PRIMITIVE REPRESENTATIONS

We first reduce the general case to the alternating and irreducible cases.

Proof of Proposition 2.7. Consider $G \leq \mathbf{U}(\mathbb{C}^d)$. By Maschke's theorem, we can decompose into irreducible representations of G :

$$\mathbb{C}^d = \bigoplus_{j=1}^m V_j.$$

Let $d_j = \dim V_j$. Let

$$J = \{j \in [m] : d_j \geq d/(2\ell)\}.$$

First suppose $\sum_{j \in J} d_j \geq d/2$. Then in each such V_j , we consider the probability measure μ_j that witnesses $f_{[2kd_j/d]}^{\text{irred}}(d_j)$ for the irreducible representation of G on V_j . That is, μ_j samples a $[2kd_j/d]$ -dimensional subspace of V_j and satisfies

$$\int \sup_{g \in G} \|\text{proj}_W(g\mathbf{v})\|_2^2 d\mu_j(W) \leq f_{[2kd_j/d]}^{\text{irred}}(d_j)^2 \|\mathbf{v}\|_2^2$$

for each $\mathbf{v} \in V_j$. We define μ to be a uniformly random k -dimensional subspace of $\bigoplus_{j \in J} W_j$, where each W_j is an independent μ_j -random $[2kd_j/d]$ -dimensional subspace of V_j . (Note the W_j 's are orthogonal as the V_j 's are.) The total dimension of this direct sum is at least k , so μ is well-defined.

Given $\mathbf{v} \in \mathbb{C}^d$, write $\mathbf{v} = \sum_{j=1}^m \mathbf{v}_j$ with $\mathbf{v}_j \in V_j$. We have

$$\begin{aligned} \int \sup_{g \in G} \|\text{proj}_W(g\mathbf{v})\|_2^2 d\mu(W) &\leq \int \sup_{g \in G} \|\text{proj}_{\bigoplus_{j \in J} W_j}(g\mathbf{v})\|_2^2 \prod_{j \in J} d\mu_j(W_j) \\ &\leq \sum_{j \in J} \int \sup_{g \in G} \|\text{proj}_{W_j}(g\mathbf{v})\|_2^2 d\mu_j(W_j) \\ &\leq \sum_{j \in J} f_{[2kd_j/d]}^{\text{irred}}(d_j)^2 \|\mathbf{v}_j\|_2^2 \\ &\leq \sup_{d' \geq d/(2\ell)} f_{[2kd'/d]}^{\text{irred}}(d')^2 \|\mathbf{v}\|_2^2 \end{aligned}$$

by orthogonality of the V_j .

Next suppose $\sum_{j \in J} d_j < d/2$. Then $|[m] \setminus J| \geq \ell$. Let I be an ℓ -element subset of $[m] \setminus J$. Choose arbitrary $\mathbf{w}_j \in \mathbb{S}(V_j) \subseteq \mathbb{C}^d$ for $j \in I$, which are clearly orthogonal. Let μ be the probability measure on k -dimensional subspaces of \mathbb{C}^d obtained by taking the span of k uniform random elements in $\{\mathbf{w}_1, \dots, \mathbf{w}_\ell\}$.

For each $g \in G$, write

$$\mathbf{u}_g = (\langle g\mathbf{v}, \mathbf{w}_1 \rangle, \dots, \langle g\mathbf{v}, \mathbf{w}_\ell \rangle)$$

and

$$\mathbf{v}' = (\|\text{proj}_{V_1} \mathbf{v}\|_2, \dots, \|\text{proj}_{V_\ell} \mathbf{v}\|_2).$$

Given $S \subseteq [\ell]$, let proj_S take the projection of an ℓ -dimensional vector down to that subset of coordinates. We have

$$\begin{aligned} \int \sup_{g \in G} \|\text{proj}_W(g\mathbf{v})\|_2^2 d\mu(W) &= \frac{1}{\binom{\ell}{k}} \sum_{S \in \binom{[\ell]}{k}} \sup_{g \in G} \|\text{proj}_S(\mathbf{u}_g)\|_2^2 \\ &\leq \frac{1}{\binom{\ell}{k}} \sum_{S \in \binom{[\ell]}{k}} \sum_{j \in S} (v'_j)^2 \\ &= \frac{k}{\ell} \sum_{j=1}^{\ell} (v'_j)^2 \leq \frac{k}{\ell} \|\mathbf{v}\|_2^2. \end{aligned}$$

The first equality follows by the definition of μ , the subsequent inequality follows by $|\langle g\mathbf{v}, \mathbf{w}_j \rangle| \leq v'_j$, and the last line is by direct computation and orthogonality of the V_j . \square

We next reduce the irreducible case to the primitive case. We first collect a few facts proved in [4] regarding systems of imprimitivity.

Lemma 3.1 ([4, Section 2]). *Let $G \leq \text{U}(\mathbb{C}^d)$ be irreducible but imprimitive. Consider a system of imprimitivity*

$$\mathbb{C}^d = \bigoplus_{j=1}^{d_1} V_j$$

with d_1 maximal over all such systems of primitivity. Let $H = \{g \in G : gV_1 = V_1\}$ and choose $\gamma_1, \dots, \gamma_{d_1}$ such that $\gamma_j V_1 = V_j$. Then the following hold:

1. The V_j are orthogonal and have the same dimension, and G acts transitively on them.
2. H has primitive action on V_1 (i.e. the representation of H on V_1 is primitive).
3. $\gamma_1, \dots, \gamma_{d_1}$ form a complete set of left coset representatives for H in G .
4. For each $g \in G$ there is $\sigma_g \in \mathfrak{S}_{d_1}$ so that $\gamma_{\sigma_g(j)}^{-1} g \gamma_j \in H$ for all $j \in [d_1]$ (i.e., σ_g records how g permutes $\{V_1, \dots, V_{d_1}\}$).

Now we are ready to prove Proposition 2.8, which recall says that for all $k \leq d/2$,

$$f_k^{\text{irred}}(d) \leq \max_{d_1 d_2 = d} (\min\{f_{\lceil k/d_1 \rceil}^{\text{prim}}(d_2), f_k^{\text{alt}}(d_1) + \mathbb{1}_{k \geq d_1}\}).$$

Proof of Proposition 2.8. Let $G \leq \text{U}(\mathbb{C}^d)$ be irreducible but imprimitive. Consider a system of imprimitivity

$$\mathbb{C}^d = \bigoplus_{j=1}^{d_1} V_j$$

with d_1 maximal among all systems of imprimitivity. By Lemma 3.1, the spaces V_j are orthogonal and all the $\dim V_j$ are equal. Let $d_2 = \dim V_1$, so that $d_1 d_2 = d$. Furthermore, $H = \{g \in G : gV_1 = V_1\}$ acts primitively on V_1 , that G acts transitively on the V_j , and that there are $\gamma_1, \dots, \gamma_{d_1}$ so that $\gamma_j V_1 = V_j$ which form a complete set of left coset representatives for H in G . For each $g \in G$ we have some $\sigma_g \in \mathfrak{S}_{d_1}$ so that $\gamma_{\sigma_g(j)}^{-1} g \gamma_j \in H$ for all $j \in [d_1]$. Define $h(g, j) = \gamma_{\sigma_g(j)}^{-1} g \gamma_j$.

Let $\mathbf{v} \in \mathbb{C}^d$. There is a unique orthogonal decomposition

$$\mathbf{v} = \sum_{j=1}^{d_1} \gamma_j \mathbf{v}_j$$

where $\mathbf{v}_j \in V_1$ for all $j \in [d_1]$. We have

$$g\mathbf{v} = \sum_{j=1}^{d_1} g\gamma_j \mathbf{v}_j = \sum_{j=1}^{d_1} \gamma_j h(g, \sigma_g^{-1}(j)) \mathbf{v}_{\sigma_g^{-1}(j)}.$$

Finally, if

$$\mathbf{w} = \sum_{j=1}^{d_1} \lambda_j \gamma_j \mathbf{x}$$

for some $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_{d_1}) \in \mathbb{C}^{d_1}$ and $\mathbf{x} \in V_1$ then we see from the above and orthogonality that

$$\langle g\mathbf{v}, \mathbf{w} \rangle = \sum_{j=1}^{d_1} \lambda_j \langle h(g, \sigma_g^{-1}(j)) \mathbf{v}_{\sigma_g^{-1}(j)}, \mathbf{x} \rangle.$$

Now we return to the situation at hand: we need to choose a k -dimensional space with a good projection for our transitive set. Consider the map $\psi : V_1 \times \mathbb{C}^{d_1} \rightarrow \mathbb{C}^d$ given by

$$\psi(\mathbf{x}, \boldsymbol{\lambda}) = \sum_{j=1}^{d_1} \lambda_j \gamma_j \mathbf{x}.$$

It clearly maps the pair of unit spheres into the unit sphere. Given probability measures μ_1 on $\text{Gr}_{\mathbb{C}}(k_1, V_1)$ and μ_2 on $\text{Gr}_{\mathbb{C}}(k_2, \mathbb{C}^{d_1})$, we define the pushforward measure μ on $\text{Gr}_{\mathbb{C}}(k_1 k_2, d)$ by taking the image of these two subspaces under ψ . Equivalently, suppose μ_1^* samples a unitary basis $\mathbf{x}_1, \dots, \mathbf{x}_{k_1}$ of a subspace of V_1 and μ_2^* samples a unitary basis $\boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_{k_2}$ of a subspace of \mathbb{C}^{d_1} , then μ samples the subspace of \mathbb{C}^d with basis $\{\psi(\mathbf{x}_i, \boldsymbol{\lambda}_j) : i \in [k_1], j \in [k_2]\}$. It is easy to check this basis is in fact unitary.

Next, we choose μ_1 and μ_2 based on the sizes of d_1 and d_2 .

First let $k_1 = \lceil k/d_1 \rceil \leq d_2$ (as $k \leq d/2$) and $k_2 = d_1$. We let μ_1 be the measure guaranteed by [Definition 2.1](#) so that

$$\int \sup_{h \in H} \|\text{proj}_W(h\mathbf{u})\|_2^2 d\mu_1(W) \leq f_{k_1}^{\text{prim}}(d_2)^2 \|\mathbf{u}\|_2^2$$

for all $\mathbf{u} \in V_1$ and let μ_2 be the atom on the space \mathbb{C}^{d_1} in $\text{Gr}_{\mathbb{C}}(d_1, d_1)$. Let μ be the ψ -pushforward of (μ_1, μ_2) as described earlier. We find

$$\begin{aligned} \int \sup_{g \in G} \|\text{proj}_W(g\mathbf{v})\|_2^2 d\mu(W) &= \int \sup_{g \in G} \sum_{\ell=1}^{k_1} \sum_{j=1}^{d_1} |\langle g\mathbf{v}, \psi(\mathbf{x}_\ell, \mathbf{e}_j) \rangle|^2 d\mu_1^*(\mathbf{x}_1, \dots, \mathbf{x}_{k_1}) \\ &= \int \sup_{g \in G} \sum_{\ell=1}^{k_1} \sum_{j=1}^{d_1} |\langle h(g, \sigma_g^{-1}(j)) \mathbf{v}_{\sigma_g^{-1}(j)}, \mathbf{x}_\ell \rangle|^2 d\mu_1^*(\mathbf{x}_1, \dots, \mathbf{x}_{k_1}) \\ &= \int \sup_{g \in G} \sum_{\ell=1}^{k_1} \sum_{j=1}^{d_1} |\langle h(g, j) \mathbf{v}_j, \mathbf{x}_\ell \rangle|^2 d\mu_1^*(\mathbf{x}_1, \dots, \mathbf{x}_{k_1}) \\ &\leq \sum_{j=1}^{d_1} \int \sup_{g \in G} \sum_{\ell=1}^{k_1} |\langle h(g, j) \mathbf{v}_j, \mathbf{x}_\ell \rangle|^2 d\mu_1^*(\mathbf{x}_1, \dots, \mathbf{x}_{k_1}) \\ &\leq \sum_{j=1}^{d_1} \int \sup_{h \in H} \sum_{\ell=1}^{k_1} |\langle h \mathbf{v}_j, \mathbf{x}_\ell \rangle|^2 d\mu_1^*(\mathbf{x}_1, \dots, \mathbf{x}_{k_1}) \\ &\leq \sum_{j=1}^{d_1} f_{k_1}^{\text{prim}}(d_2)^2 \|\mathbf{v}_j\|_2^2 = f_{k_1}^{\text{prim}}(d_2)^2 \|\mathbf{v}\|_2^2. \end{aligned}$$

The last equality is by orthogonality of V_1, \dots, V_{d_1} and unitarity of γ_j for $j \in [d_1]$.

Now suppose that $k < d_1$. Let $k_1 = 1$ and $k_2 = k$. Choose an arbitrary unit vector $\mathbf{x} \in V_1$ and μ_1 be an atom on $\text{Gr}_{\mathbb{C}}(1, V_1)$ supported on the line $\mathbb{C}\mathbf{x}$. Let μ_2 be guaranteed by [Definition 2.6](#) so that

$$\int \sup_{\mathbf{u} \in \text{Dom}(\mathbf{w})} \sum_{\ell=1}^k |\langle \mathbf{u}, \boldsymbol{\lambda}_\ell \rangle|^2 d\mu_2^*(\boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_k) \leq f_k^{\text{alt}}(d_1)^2 \|\mathbf{w}\|_2^2$$

for all $\mathbf{w} \in V_1$. Let μ be the ψ -pushforward of (μ_1, μ_2) as described earlier. We find

$$\begin{aligned} \int \sup_{g \in G} \|\text{proj}_W(g\mathbf{v})\|_2^2 d\mu(W) &= \int \sup_{g \in G} \sum_{\ell=1}^k |\langle g\mathbf{v}, \mathbf{w}_\ell \rangle|^2 d\mu^*(\mathbf{w}_1, \dots, \mathbf{w}_k) \\ &= \int \sup_{g \in G} \sum_{\ell=1}^k |\langle g\mathbf{v}, \psi(\mathbf{x}, \boldsymbol{\lambda}_\ell) \rangle|^2 d\mu_2^*(\boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_k) \\ &= \int \sup_{g \in G} \sum_{\ell=1}^k \left| \sum_{j=1}^{d_1} \lambda_{\ell,j} \langle h(g, \sigma_g^{-1}(j)) \mathbf{v}_{\sigma_g^{-1}(j)}, \mathbf{x} \rangle \right|^2 d\mu_2^*(\boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_k) \\ &\leq \int \sup_{\mathbf{u} \in \text{Dom}(\mathbf{y})} \sum_{\ell=1}^k |\langle \mathbf{u}, \boldsymbol{\lambda}_\ell \rangle|^2 d\mu_2^*(\boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_k), \end{aligned}$$

where \mathbf{y} has coordinates $y_j = \sup_{h \in H} |\langle h\mathbf{v}_j, \mathbf{x} \rangle|$ for $j \in [d_1]$. We immediately deduce

$$\begin{aligned} \int \sup_{g \in G} \|\text{proj}_W(g\mathbf{v})\|_2^2 d\mu(W) &\leq \int \sup_{\mathbf{u} \in \text{Dom}(\mathbf{y})} \sum_{\ell=1}^k |\langle \mathbf{u}, \boldsymbol{\lambda}_\ell \rangle|^2 d\mu_2^*(\boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_k) \\ &\leq f_k^{\text{alt}}(d_1)^2 \|\mathbf{y}\|_2^2 \leq f_k^{\text{alt}}(d_1)^2 \sum_{j=1}^{d_1} \|\mathbf{v}_j\|_2^2 = f_k^{\text{alt}}(d_1)^2 \|\mathbf{v}\|_2^2. \end{aligned}$$

Note that the above constructed measures in both cases are independent of \mathbf{v} . The second construction is only valid when $k < d_1$. Therefore since the f values are clearly bounded by 1, we have an upper bound of

$$f_k^{\text{prim}}(d) \leq \max_{d_1 d_2 = d} (\min(f_{\lceil k/d_1 \rceil}^{\text{prim}}(d_2), f_k^{\text{alt}}(d_1) + \mathbb{1}_{k \geq d_1})),$$

as claimed. \square

4. PERMUTATION GROUPS

In this section, we establish upper bounds for $f_k^{\text{sym}}(d)$ and $f_k^{\text{alt}}(d)$, extending the previous construction [\[4, Section 3\]](#) for $k = 1$.

A useful high dimensional intuition is that, for small k , a random k -dimensional subspace of \mathbb{R}^d has the property that *all* its unit vectors have distribution of coordinate magnitudes similar to that of a random Gaussian vector.

We first need the existence of a large dimension subspace of \mathbb{R}^d with certain delocalization properties. We encode this through the following norm.

Definition 4.1. Given $\mathbf{v} \in \mathbb{R}^d$, let

$$\|\mathbf{v}\|_T^2 = \sup_{\emptyset \subsetneq S \subseteq [d]} \log^4(2d/|S|) \sum_{j \in S} v_j^2$$

and let

$$T^* = \{\mathbf{t} \in \mathbb{R}^d : |\langle \mathbf{t}, \mathbf{w} \rangle| \leq 1 \text{ whenever } \|\mathbf{w}\|_T \leq 1\}.$$

Remark. Note that $\|\cdot\|_T$ is a norm as it can be represented as a supremum of seminorms. Hence

$$\|\mathbf{w}\|_T = \sup_{\mathbf{t} \in T^*} |\langle \mathbf{t}, \mathbf{w} \rangle|.$$

We next recall a classical lemma regarding the concentration of norms on Gaussian space (see e.g. [5]); we provide a short proof for convenience.

Lemma 4.2. *There is an absolute constant $C > 0$ so that for all $p \geq 1$, a Gaussian random vector $\mathbf{w} \sim \mathcal{N}(0, I_d)$ satisfies*

$$(\mathbb{E}_{w_1+\dots+w_d=0} \|\mathbf{w}\|_T^p)^{1/p} \leq (\mathbb{E} \|\mathbf{w}\|_T^p)^{1/p} \leq \mathbb{E} \|\mathbf{w}\|_T + C\sqrt{p} \sup_{\mathbf{t} \in T^*} \|\mathbf{t}\|_2.$$

Proof. For first inequality note that $\mathbf{w} \sim \mathcal{N}(0, I_d)$ can be written as $\mathbf{w}' + G\mathbf{1}$ where \mathbf{w}' is drawn from $\mathcal{N}(0, I_d)$ conditioned on having coordinate sum zero and $G \in \mathcal{N}(0, 1)$ is independent of \mathbf{w}' . Then by convexity note that

$$(\mathbb{E} \|\mathbf{w}\|_T^p)^{1/p} = (\mathbb{E} \|\mathbf{w}' + G\mathbf{1}\|_T^p)^{1/p} \geq (\mathbb{E}_{\mathbf{w}'} \|\mathbb{E}[\mathbf{w}' + \mathbf{v}' | \mathbf{w}']\|_T^p)^{1/p} = (\mathbb{E}_{w_1+\dots+w_d=0} \|\mathbf{w}\|_T^p)^{1/p}.$$

To prove the second inequality first note that

$$\|\mathbf{w}\|_T - \|\mathbf{v}\|_T \leq \|\mathbf{w} - \mathbf{v}\|_T = \sup_{\mathbf{t} \in T^*} |\langle \mathbf{t}, \mathbf{w} - \mathbf{v} \rangle| \leq \|\mathbf{w} - \mathbf{v}\|_2 \sup_{\mathbf{t} \in T^*} \|\mathbf{t}\|_2.$$

Therefore if $L = \sup_{\mathbf{t} \in T^*} \|\mathbf{t}\|_2$ then $\mathbf{w} \mapsto \|\mathbf{w}\|_T$ is an L -Lipschitz function with respect to Euclidean distance. Therefore by Gaussian concentration for Lipschitz functions (see e.g. [1, p. 125]) we have that

$$\mathbb{P}[|\|\mathbf{w}\|_T - \mathbb{E}[\|\mathbf{w}\|_T]| \geq t] \leq 2 \exp(-ct^2/L^2)$$

where c is an absolute constant. Using standard moment bounds for sub-Gaussian random variables (see e.g. [8, Proposition 2.5.2]), we find that

$$(\mathbb{E} |\|\mathbf{w}\|_T - \mathbb{E} \|\mathbf{w}\|_T|^p)^{1/p} \leq C\sqrt{p} \sup_{\mathbf{t} \in T^*} \|\mathbf{t}\|_2$$

for an absolute constant $C > 0$. Finally, Minkowski's inequality implies that

$$(\mathbb{E} \|\mathbf{w}\|_T^p)^{1/p} \leq \mathbb{E} \|\mathbf{w}\|_T + (\mathbb{E} |\|\mathbf{w}\|_T - \mathbb{E} \|\mathbf{w}\|_T|^p)^{1/p}$$

and therefore the result follows. \square

We now prove an upper bound for $\mathbb{E}[\|\mathbf{w}\|_T]$.

Lemma 4.3. *A Gaussian random vector $\mathbf{w} \sim \mathcal{N}(0, I_d)$ satisfies $\mathbb{E} \|\mathbf{w}\|_T \lesssim \sqrt{d}$.*

Proof. Recall w_i^\rceil from Definition 2.4. We have

$$\begin{aligned} \mathbb{E}(w_i^\rceil)^2 &= \int_0^\infty \mathbb{P}[w_i^\rceil \geq \sqrt{t}] dt \leq \int_0^\infty \min\left(1, \binom{d}{i} (2e^{-t/2})^i\right) dt \\ &\leq \int_0^\infty \min(1, (2de^{1-t/2}/i)^i) dt \lesssim \log(2d/i). \end{aligned}$$

Therefore

$$\begin{aligned} (\mathbb{E} \|\mathbf{w}\|_T)^2 &\leq \mathbb{E} \|\mathbf{w}\|_T^2 \leq \sum_{i=1}^d \log^4(2d/i) (w_i^\rceil)^2 \lesssim \sum_{i=1}^d \log^5(2d/i) \\ &\leq d \int_0^1 \log(2/x)^5 dx = d \int_0^\infty (y + \log 2)^5 e^{-y} dy \lesssim d. \end{aligned} \quad \square$$

We are in position to derive a high-probability version.

Lemma 4.4. *With probability at least $1 - \exp(-2d/(\log d)^4)$, a standard Gaussian vector $\mathbf{w} \sim \mathcal{N}(0, I_d)$ satisfies $\|\mathbf{w}\|_T \lesssim \sqrt{d}$. In fact, the same is true after conditioning \mathbf{w} to have coordinate sum 0.*

Proof. Note that if $\mathbf{t} \in T^*$, then

$$\|\mathbf{t}\|_2^2 = \|\mathbf{t}\|_T \left| \left\langle \mathbf{t}, \frac{\mathbf{t}}{\|\mathbf{t}\|_T} \right\rangle \right| \leq \|\mathbf{t}\|_T \leq \log^2(2d) \|\mathbf{t}\|_2.$$

Hence

$$\sup_{\mathbf{t} \in T^*} \|\mathbf{t}\|_2 \leq \log^2(2d).$$

To deduce the claimed bound, note that

$$\begin{aligned} \mathbb{P}[\|\mathbf{w}\|_T \geq K\sqrt{d}] &\leq (K\sqrt{d})^{-p} \mathbb{E}[\|\mathbf{w}\|_T^p] \\ &\leq (K\sqrt{d})^{-p} (\mathbb{E}\|\mathbf{w}\|_T + C\sqrt{p} \sup_{\mathbf{t} \in T^*} \|\mathbf{t}\|_2)^p \\ &\leq (K\sqrt{d})^{-p} (C'\sqrt{d} + C'\sqrt{p} \log^2(2d))^p \end{aligned}$$

for appropriate absolute constants $C, C' > 0$, using Lemmas 4.2 and 4.3 and the above inequality. Letting $p = d/(\log d)^4$ and $K > 0$ be a sufficiently large absolute constant yields

$$\mathbb{P}[\|\mathbf{w}\|_T \geq K\sqrt{d}] \leq \exp(-2p),$$

as desired. The same holds if we condition on sum 0, using the moment bound for the conditional variable derived in Lemma 4.2 instead. \square

Lemma 4.5. *There is a $\lceil d/(\log d)^4 \rceil$ -dimensional subspace of the hyperplane $\mathbf{1}^\perp$ in \mathbb{R}^d such that each of its unit vectors \mathbf{v} satisfies*

$$\|\mathbf{v}\|_T \lesssim 1.$$

Proof. We can assume d is sufficiently large. Let $k = \lceil d/(\log d)^4 \rceil$, and consider a uniformly random k -dimensional subspace W of $\mathbf{1}^\perp$. Let U be a $d \times k$ matrix whose columns form an orthonormal basis of W , chosen uniformly at random.

By a standard volume packing argument (e.g., see [7, Lemma 4.3]), there exists $\mathcal{N} \subset \mathbb{S}(\mathbb{R}^k)$ with $|\mathcal{N}| \leq 6^k$ such that for every $\mathbf{v} \in \mathbb{S}(\mathbb{R}^k)$ there is $\mathbf{v}' \in \mathcal{N}$ so that $\|\mathbf{v} - \mathbf{v}'\|_2 \leq 1/2$. Thus if \mathbf{u} is a unit vector in the direction of $\mathbf{v} - \mathbf{v}'$, we have

$$\|U\mathbf{v}\|_T \leq \|U\mathbf{v}'\|_T + \|U(\mathbf{v} - \mathbf{v}')\|_T \leq \|U\mathbf{v}'\|_T + \frac{1}{2}\|U\mathbf{u}\|_T.$$

We deduce

$$\sup_{\mathbf{v} \in \mathbb{S}(\mathbb{R}^k)} \|U\mathbf{v}\|_T \leq \sup_{\mathbf{v}' \in \mathcal{N}} \|U\mathbf{v}'\|_T + \frac{1}{2} \sup_{\mathbf{u} \in \mathbb{S}(\mathbb{R}^k)} \|U\mathbf{u}\|_T$$

and thus

$$\sup_{\mathbf{v} \in \mathbb{S}(\mathbb{R}^k)} \|U\mathbf{v}\|_T \leq 2 \sup_{\mathbf{v}' \in \mathcal{N}} \|U\mathbf{v}'\|_T.$$

Now fix some $\mathbf{v} \in \mathcal{N}$. Note the distribution of $U\mathbf{v}$ is uniform among unit vectors in $\mathbf{1}^\perp$ since W was chosen uniformly. Now note that for any constant C we have that

$$\mathbb{P}[\|U\mathbf{v}\|_T \geq C] = \mathbb{P}[\|\mathbf{G}/\|\mathbf{G}\|_2\|_T \geq C]$$

where $\mathbf{G} \sim N(0, I_d - (\mathbf{1}^T \mathbf{1})/d)$. Now since $\mathbf{G}/\|\mathbf{G}\|_2$ and $\|\mathbf{G}\|_2$ are independent we have that

$$\begin{aligned} \mathbb{P}[\|\mathbf{G}/\|\mathbf{G}\|_2\|_T \geq C] &= \mathbb{P}[\|\mathbf{G}\|_2 \leq 2\sqrt{d}]^{-1} \mathbb{P}[\|\mathbf{G}/\|\mathbf{G}\|_2\|_T \geq C \text{ and } \|\mathbf{G}\|_2 \leq 2\sqrt{d}] \\ &\leq 2\mathbb{P}[\|\mathbf{G}/\|\mathbf{G}\|_2\|_T \geq C \text{ and } \|\mathbf{G}\|_2 \leq 2\sqrt{d}] \\ &\leq 2\mathbb{P}[\|\mathbf{G}/\|\mathbf{G}\|_2\|_T \geq 2C\sqrt{d}]. \end{aligned}$$

By Lemma 4.4, the last expression is at most $2 \exp(-2d/(\log d)^4)$. The result follows upon taking the union bound over at most 6^k vectors in \mathcal{N} , since $6 < e^2$. \square

Finally, we will need a form of Selberg's inequality (see [3, Chapter 27, Theorem 1]).

Lemma 4.6. *For $\mathbf{v}_1, \dots, \mathbf{v}_m \in \mathbb{C}^d$ we have that*

$$\sup_{\mathbf{w} \in \mathbb{S}(\mathbb{C}^d)} \sum_{i=1}^m |\langle \mathbf{w}, \mathbf{v}_i \rangle|^2 \leq \sup_{i \in [m]} \sum_{j=1}^m |\langle \mathbf{v}_i, \mathbf{v}_j \rangle|.$$

Now we prove Proposition 2.9, which recall says that for $k \leq d/(\log d)^5$, one has

$$f_k^{\text{sym}}(d) \leq f_k^{\text{alt}}(d) \lesssim 1/\sqrt{\log(d/k)}.$$

The first inequality is immediate as the set of allowable μ 's in the definition of f_k^{alt} is a subset of those of f_k^{sym} . So we just need to prove the second inequality.

Proof of Proposition 2.9. Let \mathbf{e}_i be the i -th coordinate vector. For each j with $k \leq 2^j/(\log 2^j)^4 \leq d$, we apply Lemma 4.5 to the space $V_j = \text{span}_{\mathbb{R}}\{\mathbf{e}_1, \dots, \mathbf{e}_{2^j}\}$. Here the T -norm is defined with respect to this 2^j -dimensional space. In particular, there exists a k -dimensional (real) subspace of the orthogonal complement of $\mathbf{e}_1 + \dots + \mathbf{e}_{2^j}$ within V_j , call it W_j , so that every unit vector $\mathbf{u} \in W_j$ satisfies

$$\sum_{i \in S} u_i^2 \lesssim \frac{1}{\log^4(2^{j+1}/|S|)}$$

for every nonempty $S \subseteq [2^j]$. Let $V'_j = \text{span}_{\mathbb{C}} V_j$ and $W'_j = \text{span}_{\mathbb{C}} W_j$. We immediately deduce that every unit vector $\mathbf{u} \in W'_j$ satisfies

$$\sum_{i \in S} |u_i|^2 \lesssim \frac{1}{\log^4(2^{j+1}/|S|)} \quad (4.1)$$

since we can write it as $\mathbf{u} = \alpha \mathbf{u}_r + \beta \sqrt{-1} \mathbf{u}_c$ where $\mathbf{u}_r, \mathbf{u}_c \in W_j$ are real unit vectors and $\alpha, \beta \in \mathbb{R}$ satisfy $\alpha^2 + \beta^2 = 1$.

Now we construct our random subspace as follows: let $W = W_j$ where j is a random integer uniformly chosen from

$$J = \{\lceil \log_2(2k \log^4 d) \rceil, \dots, \lfloor \log_2 d \rfloor\}.$$

Let μ be the probability measure on $\text{Gr}_{\mathbb{C}}(k, n)$ that gives W .

For every $\mathbf{v} \in \mathbb{S}(\mathbb{C}^d)$, we have

$$\sup_{\gamma \in \Gamma_d} \|\text{proj}_W(\gamma \mathbf{v})\|_2 = \sup_{\gamma \in \Gamma_d} \sup_{\mathbf{w} \in \mathbb{S}(W)} |\langle \gamma \mathbf{v}, \mathbf{w} \rangle| = \sup_{\mathbf{w} \in \mathbb{S}(W)} \langle \mathbf{v}^{\rceil}, \mathbf{w}^{\rceil} \rangle.$$

Therefore

$$\int \sup_{\gamma \in \Gamma_d} \|\text{proj}_W(\gamma \mathbf{v})\|_2^2 d\mu(W) = \frac{1}{|J|} \sum_{j \in J} \sup_{\gamma \in \Gamma_d} \|\text{proj}_{W_j}(\gamma \mathbf{v})\|_2^2 = \frac{1}{|J|} \sum_{j \in J} \sup_{\mathbf{w} \in \mathbb{S}(W_j)} \langle \mathbf{v}^{\rceil}, \mathbf{w}^{\rceil} \rangle^2.$$

Let $\mathbf{w}'_j \in \mathbb{S}(W_j)$ be such that

$$\sup_{\mathbf{w} \in \mathbb{S}(W_j)} \langle \mathbf{v}^{\rceil}, \mathbf{w}^{\rceil} \rangle^2 = \langle \mathbf{v}^{\rceil}, (\mathbf{w}'_j)^{\rceil} \rangle^2,$$

which exists by compactness. For $i, j \in J$ with $i \geq j$, we have

$$|\langle (\mathbf{w}'_i)^{\rceil}, (\mathbf{w}'_j)^{\rceil} \rangle| \leq \|\text{proj}_{V_i}((\mathbf{w}'_j)^{\rceil})\|_2 \lesssim \frac{1}{\log^2(2^{i+1}/2^j)}.$$

The first inequality follows from $\mathbf{w}'_j \in V_j$, which implies $(\mathbf{w}'_j)^{\rceil} \in V_j$. The second follows from (4.1) applied to \mathbf{w}'_j and S a subset of $[2^i]$ composed of the 2^j largest magnitude coordinates of \mathbf{w}'_j .

Applying Lemma 4.6, we deduce

$$\begin{aligned} \int \sup_{\gamma \in \Gamma_d} \|\text{proj}_W(\gamma \mathbf{v})\|_2^2 d\mu(W) &= \frac{1}{|J|} \sum_{j \in J} \langle \mathbf{v}^\gamma, (\mathbf{w}'_j)^\gamma \rangle^2 \leq \sup_{i \in J} \frac{1}{|J|} \sum_{j \in J} |\langle (\mathbf{w}'_i)^\gamma, (\mathbf{w}'_j)^\gamma \rangle| \\ &\lesssim \frac{1}{|J|} \left(\sum_{j \in J, j \geq i} \frac{1}{\log^2(2^{j+1}/2^i)} + \sum_{j \in J, j < i} \frac{1}{\log^2(2^{i+1}/2^j)} \right) \lesssim \frac{1}{|J|}. \end{aligned}$$

This μ thus shows that $f_k^{\text{alt}}(d) \lesssim 1/\sqrt{\log(d/k)}$. \square

5. PRIMITIVE REPRESENTATIONS

We now turn to the case of bounding $f_k^{\text{prim}}(d)$. First, we show that if the group $G \leq \text{U}(\mathbb{R}^d)$ is sufficiently small, then a random basis achieves the necessary bound for $f_k^{\text{prim}}(d)$. This is a minor modification of [4, Proposition 4.1].

Proposition 5.1. *Let $G \leq \text{U}(\mathbb{C}^d)$. Suppose that $[G : Z_d \cap G] \leq e^{d/\log d}$, where $Z_d := \{\lambda I_d : |\lambda| = 1\}$. Then for $k \in [d]$ there exists a probability measure μ on $\text{Gr}_{\mathbb{C}}(k, d)$ such that*

$$\int \sup_{g \in G} \|\text{proj}_W(g\mathbf{v})\|_2^2 d\mu(W) \lesssim \frac{1}{\log(2d/k)} \|\mathbf{v}\|_2^2$$

for all $\mathbf{v} \in \mathbb{C}^d$.

Proof. We let μ be the uniform measure on $\text{Gr}_{\mathbb{C}}(k, d)$. By scaling, we may assume that \mathbf{v} is a unit vector. Furthermore let W' be the subspace generated by the first k coordinate vectors $\mathbf{e}_1, \dots, \mathbf{e}_k$. Note that

$$\begin{aligned} \mathbb{P}_W \left[\sup_{g \in G} \|\text{proj}_W(g\mathbf{v})\|_2 \geq t \right] &\leq e^{d/\log d} \mathbb{P}_W [\|\text{proj}_W(\mathbf{v})\|_2 \geq t] \\ &\leq e^{d/\log d} \mathbb{P}_{\mathbf{v}' \in \mathbb{S}(\mathbb{C}^d)} [\|\text{proj}_{W'}(\mathbf{v}')\|_2 \geq t] \end{aligned}$$

using a union bound and then orthogonal invariance. Now note that

$$\mathbb{E}[\|\text{proj}_{W'}(\mathbf{v}')\|_2]^2 \leq \mathbb{E}[\|\text{proj}_{W'}(\mathbf{v}')\|_2^2] = k/d$$

and that $\|\text{proj}_{W'}(\mathbf{v}')\|_2$ is a 1-Lipschitz function of \mathbf{v}' . Therefore by Lévy concentration on the sphere we have that

$$\mathbb{P}_{\mathbf{v}' \in \mathbb{S}(\mathbb{C}^d)} [\|\text{proj}_{W'}(\mathbf{v}')\|_2 \geq \sqrt{k/d} + C/\sqrt{\log d}] \leq e^{-2d/\log d}$$

for a suitably large absolute constant C . Finally, using $\sqrt{k/d} \lesssim 1/\sqrt{\log(2d/k)}$ and using the bound $\|\text{proj}_{W'}(\mathbf{v}')\|_2 \leq 1$, the desired result follows immediately. \square

We need the following key group theoretic result from Green [4], which in turn builds on ideas from Collins' work on optimal bounds for Jordan's theorem [2]. Roughly, it says that if $[G : Z_d \cap G]$ is large then G has a large normal alternating subgroup. The first part of the following theorem is [4, Proposition 4.2], while the rest is implicit in the proof of [4, Proposition 1.11].

Theorem 5.2 ([4, Section 4]). *Let $G \leq \text{U}(\mathbb{C}^d)$ be primitive and suppose that $[G : Z_d \cap G] \geq e^{d/\log d}$. If d is sufficiently large then all of the following hold.*

- (1) G has a normal subgroup isomorphic to the alternating group A_n for some $n \gtrsim d/(\log d)^4$.
- (2) G has a subgroup of index at most 2 of the form $A_n \times H$, with the same n .
- (3) The resulting representation $\rho : A_n \times H \hookrightarrow G \hookrightarrow \text{U}(\mathbb{C}^d)$ decomposes into irreducible representations, at least one of which (call it ρ_1) is of the form $\rho_1 \simeq \psi \otimes \psi'$, where ψ' is an irreducible representation of H and ψ is the representation of A_n acting via permutation of coordinates on $\{\mathbf{z} \in \mathbb{C}^n : z_1 + \dots + z_n = 0\}$.

We are now in position to prove [Proposition 2.10](#), which recall says that there is an absolute constant $c > 0$ such that for every $k \leq cd/(\log d)^4$ we have

$$f_k^{\text{prim}}(d) \lesssim \sup_{d' \geq cd/(\log d)^4} f_k^{\text{alt}}(d').$$

The proof mirrors that of [\[4, Proposition 1.11\]](#), but we correct an error of Green ([\[4, p. 20\]](#)) involving an incorrect orthogonality identity. This erroneous deduction is replaced by an argument which still allows one to reduce the primitive case to the alternating case.

Proof of [Proposition 2.10](#). We may assume d is sufficiently large. If $[G : Z_d \cap G] \leq e^{d/\log d}$, then the result follows by [Proposition 5.1](#). So we can assume $[G : Z_d \cap G] \geq e^{d/\log d}$, and thus by [Theorem 5.2](#), G has a normal subgroup isomorphic to A_n for some $n \gtrsim d/(\log d)^4$ and that G has a subgroup of index at most 2 which is of the form $A_n \times H$. If the index is 2, let τ be the nontrivial right coset representative of $A_n \times H$ in G (otherwise just let τ be the identity). Note that

$$\sup_{g \in G} \|\text{proj}_W(g\mathbf{v})\|_2^2 \leq \sup_{g \in A_n \times H} \|\text{proj}_W(g\mathbf{v})\|_2^2 + \sup_{g \in A_n \times H} \|\text{proj}_W(g\tau\mathbf{v})\|_2^2,$$

so it is easy to see that, up to losing a constant factor, we may reduce to studying groups of the form $G = A_n \times H$ where $n \gtrsim d/(\log d)^4$ (but note that the representation may no longer be primitive, or even irreducible).

Now [Theorem 5.2](#) shows that the representation $\rho : A_n \times H \rightarrow \text{U}(\mathbb{C}^d)$ coming from this setup has an irreducible component of the form $\rho_1 \simeq \psi \otimes \psi'$, where ψ' is an irreducible representation of H and ψ is the representation of A_n acting via permutation of coordinates on $\{\mathbf{z} \in \mathbb{C}^n : z_1 + \dots + z_n = 0\}$.

Note that $\dim \rho_1 \geq \dim \psi = n - 1 \gtrsim d/(\log d)^4$, so $\dim \rho_1 \geq k$ provided that $c > 0$ is sufficiently small. We will choose a k -dimensional subspace of the irreducible component ρ_1 .

We explicitly present this situation as follows. Let V' be the space acted on by ψ' (unitarily). Consider $V = \mathbf{1}^\perp \subseteq \mathbb{C}^n$, and consider the spaces $V \otimes V' \subseteq \mathbb{C}^n \otimes V'$, which has a natural unitary structure given by the tensor product. Note ψ acts on V by permutation of coordinates when represented in \mathbb{C}^n . Every vector in $V \otimes V'$ is spanned by pure tensors $\mathbf{v} \otimes \mathbf{v}'$ where \mathbf{v} has zero coordinate sum, and $\rho_1((a, h))$ acts by $\psi(a) \otimes \psi'(h)$ on pure tensors. In fact, we can extend this action to all of $\mathbb{C}^n \otimes V'$ in the natural way (and the resulting representation is isomorphic to a direct sum of ρ_1 and $\text{triv}_{A_n} \otimes \psi'$). At this point, the analysis will be similar to that in the proof of [Proposition 2.8](#).

Let ν be the measure on $\text{Gr}_{\mathbb{C}}(k, n)$ which is guaranteed by [Definition 2.6](#) (so is supported on subspaces of $V \subseteq \mathbb{C}^n$) and consider the measure which is supported on a single atom in $\text{Gr}_{\mathbb{C}}(1, V')$ in the direction of a fixed unit vector \mathbf{x} . Let μ be the tensor of these two measures, i.e., if ν^* samples k orthonormal (sum zero) vectors $\mathbf{u}_1, \dots, \mathbf{u}_k$ then we choose the subspace with basis $\mathbf{u}_1 \otimes \mathbf{x}, \dots, \mathbf{u}_k \otimes \mathbf{x}$.

Now consider some \mathbf{v} in the space $V \otimes V' \subseteq \mathbb{C}^n \otimes V'$, and write it as

$$\mathbf{v} = \sum_{j=1}^n \mathbf{e}_j \otimes \mathbf{v}'_j$$

where the \mathbf{e}_j is the j -th coordinate vector of \mathbb{C}^n . In fact, the \mathbf{v}'_j must add up to $\mathbf{0} \in V'$. We see that

$$\|\mathbf{v}\|_2^2 = \sum_{j=1}^n \|\mathbf{v}'_j\|_2^2.$$

We have

$$\begin{aligned}
& \int \sup_{g \in A_n \times H} \|\text{proj}_W(g\mathbf{v})\|_2^2 d\mu(W) \\
&= \int \sup_{a \in A_n, h \in H} \sum_{\ell=1}^k \left| \left\langle \sum_{j=1}^n \psi(a)\mathbf{e}_j \otimes \psi'(h)\mathbf{v}'_j, \mathbf{w}_\ell \right\rangle \right|^2 d\mu^*(\mathbf{w}_1, \dots, \mathbf{w}_k) \\
&= \int \sup_{a \in A_n, h \in H} \sum_{\ell=1}^k \left| \sum_{j=1}^n \langle \psi(a)\mathbf{e}_j, \mathbf{u}_\ell \rangle \langle \psi'(h)\mathbf{v}'_j, \mathbf{x} \rangle \right|^2 d\nu^*(\mathbf{u}_1, \dots, \mathbf{u}_k) \\
&\leq \int \sup_{\mathbf{w} \in \text{Dom}(\mathbf{y})} \sum_{\ell=1}^k |\langle \mathbf{w}, \mathbf{u}_\ell \rangle|^2 d\nu^*(\mathbf{u}_1, \dots, \mathbf{u}_k) \\
&\leq f_k^{\text{alt}}(n)^2 \|\mathbf{y}\|_2^2,
\end{aligned}$$

where $\mathbf{y} \in \mathbb{C}^n$ satisfies $y_j = \sup_{h \in H} |\langle \psi'(h)\mathbf{v}'_j, \mathbf{x} \rangle|$. The first inequality follows by noting that $\langle \psi(a)\mathbf{e}_j, \mathbf{u}_\ell \rangle$ as j varies simply records the coordinates of \mathbf{u}_ℓ in some permutation, and by considering $\mathbf{w} = (w_1, \dots, w_n)$ defined via $w_j = \langle \psi'(h)\mathbf{v}'_j, \mathbf{x} \rangle$, which is clearly on $\text{Dom}(\mathbf{y})$. Now we see

$$\int \sup_{g \in A_n \times H} \|\text{proj}_W(g\mathbf{v})\|_2^2 d\mu(W) \leq f_k^{\text{alt}}(n)^2 \|\mathbf{y}\|_2^2 \leq f_k^{\text{alt}}(n)^2 \sum_{j=1}^n \|\mathbf{v}'_j\|_2^2 = f_k^{\text{alt}}(n)^2 \|\mathbf{v}\|_2^2. \quad \square$$

This completes all the components of the proof of [Theorem 1.3](#).

6. REAL SUBSPACES

We already proved [Theorem 1.3](#), which finds a complex subspace. Now we use it to deduce [Theorem 1.2](#), which gives a real subspace. We will apply the following version of the restricted invertibility theorem, which is a special case of [6, Theorem 6]. We write $s_1(M) \geq s_2(M) \geq \dots$ for the singular values of a matrix M .

Theorem 6.1 ([6, Theorem 6]). *Let M be a real $2k \times 4k$ matrix of rank $2k$. There exists $S \subseteq [4k]$ with $|S| = k$ such that M_S , the restriction of M to the columns S , satisfies*

$$s_k(M_S) \gtrsim \sqrt{\frac{\sum_{j=3k/2}^{4k} s_j(M)^2}{k}}.$$

Proof of Theorem 1.2. Let $2k \leq d/(\log d)^C$, where C is as in [Theorem 1.3](#). By embedding X in $\mathbb{S}(\mathbb{C}^d)$ and using [Theorem 1.3](#) we can find a $2k$ -dimensional complex subspace W of \mathbb{C}^d such that

$$\sup_{\mathbf{x} \in X} \|\text{proj}_W \mathbf{x}\|_2 \lesssim 1/\sqrt{\log(d/k)}.$$

Let $\mathbf{v}_1, \dots, \mathbf{v}_{2k}$ be a unitary basis for the subspace W and let the matrix with these columns be denoted by B . Now consider the matrix M which has $4k$ columns which are $\text{Re } \mathbf{v}_1, \dots, \text{Re } \mathbf{v}_{2k}$ and $\text{Im } \mathbf{v}_1, \dots, \text{Im } \mathbf{v}_{2k}$. Note that M has $s_{2k}(M) \geq 1/\sqrt{2}$ as any vectors in \mathbb{C}^{4k} satisfying $iv_j = v_{j+2k}$ have $\|M\mathbf{v}\| = \|\mathbf{v}\|/\sqrt{2}$. Therefore by [Theorem 6.1](#) one can select k columns such that the matrix N with those k columns satisfies

$$s_k(N) \gtrsim 1.$$

Now consider any unit vector \mathbf{v} in the image of N . Such a vector can be represented as $\mathbf{v} = N\mathbf{w}$ where $\|\mathbf{w}\| \lesssim 1$. It therefore suffices to prove that

$$\sup_{\mathbf{x} \in X, \mathbf{w} \in \mathbb{S}(\mathbb{R}^k)} |\langle N\mathbf{w}, \mathbf{x} \rangle| \lesssim 1/\sqrt{\log(d/k)}.$$

To see this separate N into N_1 and N_2 where N_1 corresponds to columns chosen from the real parts of vectors \mathbf{v}_i and the columns are chosen from the complex parts of \mathbf{v}_i . Let these have ℓ and $k - \ell$ columns respectively. Then

$$\begin{aligned} \sup_{\mathbf{x} \in X, \mathbf{w} \in \mathbb{S}(\mathbb{R}^k)} |\langle N\mathbf{w}, \mathbf{x} \rangle| &\leq \sup_{\mathbf{x} \in X, \mathbf{w} \in \mathbb{S}(\mathbb{R}^\ell)} |\langle N_1\mathbf{w}, \mathbf{x} \rangle| + \sup_{\mathbf{x} \in X, \mathbf{w} \in \mathbb{S}(\mathbb{R}^{k-\ell})} |\langle N_2\mathbf{w}, \mathbf{x} \rangle| \\ &\leq 2 \sup_{\mathbf{x} \in X, \mathbf{w} \in \mathbb{S}(\mathbb{C}^k)} |\langle B\mathbf{w}, \mathbf{x} \rangle| \\ &\lesssim 1/\sqrt{\log(d/k)}. \end{aligned} \quad \square$$

7. LOWER BOUND

Finally, we show a lower bound of $\Omega(1/\sqrt{\log(2d/k)})$, which demonstrates optimality of our results.

Proof of Theorem 1.5. We prove the real case; an analogous proof works over \mathbb{C} by considering a suitably fine discretization of Γ_d , or we can repeat the proof in Section 6 to transfer a lower bound from real to complex.

The claim for $k = 1$ was already proved in [4, Sharpness after Theorem 1.3] (see the construction at the beginning of this article right after Theorem 1.1). The case $k = 1$ implies the result also for $k \leq d^{1-c}$ for any constant c , since we can project from W onto an arbitrary 1-dimensional subspace of W .

So from now on assume $k \geq d^{1/2}$. Consider the action of $G = \mathfrak{S}_d \ltimes (\mathbb{Z}/2\mathbb{Z})^d$ on \mathbb{R}^d by permutation and signing. Let

$$\mathbf{a} = \left(\frac{1}{\sqrt{\lfloor k/2 \rfloor + 1}}, \dots, \frac{1}{\sqrt{d}}, 0, \dots, 0 \right).$$

Let X be the G -orbit of $\mathbf{a}/\|\mathbf{a}\|_2$.

Let W be a k -dimensional subspace of \mathbb{R}^d . We wish to show $\sup_{\mathbf{x} \in X} \|\text{proj}_W \mathbf{x}\|_2 \gtrsim 1/\sqrt{\log(2d/k)}$.

Let $\mathbf{y} = (y_1, \dots, y_d)$ a uniform random vector in $\mathbb{S}(W)$. Let $\sigma_i = (\mathbb{E}y_i^2)^{1/2}$. We have

$$\sigma_1^2 + \dots + \sigma_d^2 = \mathbb{E}[y_1^2 + \dots + y_d^2] = 1 \quad (7.1)$$

and

$$\sigma_i^2 = \frac{1}{k} \|\text{proj}_W(\mathbf{e}_i)\|^2 \leq \frac{1}{k}. \quad (7.2)$$

Without loss of generality, assume that $1/\sqrt{k} \geq \sigma_1 \geq \dots \geq \sigma_d \geq 0$, so that $\sigma_i \leq 1/\sqrt{i}$ for each i . We claim that

$$a_i \geq \sqrt{\frac{2}{3}} \sigma_i \quad \text{for all } 1 \leq i \leq d - k/2.$$

Indeed, for $i \leq k$, we have $a_i \geq 1/\sqrt{3k/2} \geq \sqrt{3/2} \sigma_i$. For $k < i \leq d - \lfloor k/2 \rfloor$, we have $a_i = 1/\sqrt{\lfloor k/2 \rfloor + i} \geq \sigma_i \sqrt{i/(\lfloor k/2 \rfloor + i)} \geq \sqrt{2/3} \sigma_i$.

We have $\mathbb{E}|y_i| \gtrsim (\mathbb{E}y_i^2)^{1/2} = \sigma_i$ since y_i is distributed as the first coordinate of a random point on $\sigma_i \sqrt{k} \cdot \mathbb{S}(\mathbb{R}^k)$.

Putting everything together, we have

$$\begin{aligned} \|\mathbf{a}\|_2 \sup_{\mathbf{x} \in X} \|\text{proj}_W \mathbf{x}\|_2 &\geq \sup_{g \in G} \|\text{proj}_W g\mathbf{a}\| \geq \mathbb{E} \sup_{g \in G} \langle \mathbf{a}, g\mathbf{y} \rangle \\ &\geq \mathbb{E} \sum_{1 \leq i \leq d} a_i |y_i| \gtrsim \sum_{i=1}^d a_i \sigma_i \gtrsim \sum_{i=1}^{d-k/2} \sigma_i^2 \geq \frac{1}{2}, \end{aligned}$$

where the final step uses (7.1) and (7.2). Thus

$$\sup_{\mathbf{x} \in X} \|\text{proj}_W \mathbf{x}\|_2 \gtrsim \frac{1}{\|\mathbf{a}\|_2} \gtrsim 1/\sqrt{\log(2d/k)}. \quad \square$$

REFERENCES

- [1] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart, *Concentration inequalities*, Oxford University Press, Oxford, 2013, A nonasymptotic theory of independence, With a foreword by Michel Ledoux.
- [2] Michael J. Collins, *On Jordan's theorem for complex linear groups*, J. Group Theory **10** (2007), 411–423.
- [3] Harold Davenport, *Multiplicative number theory*, third ed., Graduate Texts in Mathematics, vol. 74, Springer-Verlag, New York, 2000, Revised and with a preface by Hugh L. Montgomery.
- [4] Ben Green, *On the width of transitive sets: Bounds on matrix coefficients of finite groups*, Duke Math. J. **169** (2020), 551–578.
- [5] Michel Ledoux and Michel Talagrand, *Probability in Banach spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 23, Springer-Verlag, Berlin, 1991, Isoperimetry and processes.
- [6] Assaf Naor and Pierre Youssef, *Restricted invertibility revisited*, A journey through discrete mathematics, Springer, Cham, 2017, pp. 657–691.
- [7] Mark Rudelson, *Recent developments in non-asymptotic theory of random matrices*, Modern aspects of random matrix theory, Proc. Sympos. Appl. Math., vol. 72, Amer. Math. Soc., Providence, RI, 2014, pp. 83–120.
- [8] Roman Vershynin, *High-dimensional probability*, Cambridge Series in Statistical and Probabilistic Mathematics, vol. 47, Cambridge University Press, Cambridge, 2018, An introduction with applications in data science, With a foreword by Sara van de Geer.

DEPARTMENT OF MATHEMATICS, MASSACHUSETTS INSTITUTE OF TECHNOLOGY, CAMBRIDGE, MA 02139, USA
 Email address: {asah,msawhney,yufeiz}@mit.edu