Low-degree Boolean functions on S_n, with an application to isoperimetry

David Ellis,[*] Yuval Filmus,[†] and Ehud Friedgut[‡]

November 2015

Abstract

We prove that Boolean functions on S_n, whose Fourier transform is highly concentrated on irreducible representations indexed by partitions of n whose largest part has size at least $n-t$, are close to being unions of cosets of stabilizers of t-tuples. We also obtain an edge-isoperimetric inequality for the transposition graph on S_n which is asymptotically sharp for subsets of S_n of size $n!/\text{poly}(n)$, using eigenvalue techniques. We then combine these two results to obtain a sharp edge-isoperimetric inequality for subsets of S_n of size $(n-t)!$, where n is large compared to t, confirming a conjecture of Ben Efraim in these cases.

1 Introduction

This paper (together with [13] and [14]) is part of a trilogy dealing with stability and 'quasi-stability' results concerning Boolean functions on the symmetric group, which are (in a sense) of 'low complexity'.

Let us begin with some notation and definitions which will enable us to present the Fourier-theoretic context of our results.

If X is a set and $S \subset X$, we denote by 1_S the *characteristic function* (or *indicator function*) of S, i.e. $1_S : X \to \{0,1\}$ with $1_S(x) = 1$ if and only if $x \in S$. If $f, g : X \to \{0,1\}$ are Boolean functions, we say that f and g are *ϵ-close* if $|\{x \in X : f(x) \neq g(x)\}| \leq \epsilon|X|$.

Let S_n denote the symmetric group of order n, the group of all permutations of $[n] := \{1, 2, \ldots, n\}$. For $i, j \in [n]$, we let

$$T_{ij} = \{\sigma \in S_n : \sigma(i) = j\}.$$

We call the T_{ij}'s the *1-cosets* of S_n, since they are the cosets of stabilizers of points.

[*]School of Mathematical Sciences, Queen Mary, University of London, 327 Mile End Road, London, E1 4NS, UK.

[†]Computer Science Department, Technion – Israel Institute of Technology, Technion City, Haifa 3200003, Israel. This research was supported in part by the Canadian Friends of the Hebrew University / University of Toronto Permanent Endowment.

[‡]Faculty of Mathematics and Computer Science, Weizmann Institute of Science, Rehovot 76100, Israel. This research was supported in part by I.S.F. grant 0398246, and B.S.F. grant 2010247.

Similarly, for $t > 1$ and for two ordered t-tuples of distinct elements of $[n]$, $I = (i_1, \ldots, i_t)$ and $J = (j_1, \ldots, j_t)$, we let

$$T_{IJ} = \{\sigma \in S_n \colon \sigma(I) = J\} = \{\sigma \in S_n \colon \sigma(i_k) = j_k \ \forall k \in [t]\};$$

we call these the t-*cosets* of S_n. Abusing notation, we will sometimes use T_{ij} and T_{IJ} to denote their own characteristic functions.

For any non-negative integer t, we let U_t be the vector space of real-valued functions on S_n whose Fourier transform is supported on irreducible representations of S_n which are indexed by partitions of n whose largest part has size at least $n - t$. In [15] it is proved that U_t is the space spanned by the t-cosets:

Theorem 1. *If n and t are integers with $0 \leq t \leq n$ then*

$$U_t = \operatorname{Span}\{T_{IJ} : I, J \text{ are ordered t-tuples of distinct elements of } [n]\}.$$

If t is fixed and n is large, the irreducible representations of S_n which are indexed by partitions of n with largest part of size t have dimension $\Theta(n^t)$. The functions in U_t therefore have Fourier transform supported on irreducible representations of dimension $O(n^t)$, and can be viewed as having low complexity. If f is a real-valued function on S_n, we define the *degree* of f to be the minimum t such that $f \in U_t$. This can be seen as a measure of the complexity of f, analogous to the degree of a Boolean function on $\{0, 1\}^n$. Indeed, by Theorem 1, it is precisely the minimum possible total degree of a polynomial in the T_{ij}'s which is equal to f.

The following theorem from [15] characterizes the Boolean functions in U_1.

Theorem 2 (Ellis, Friedgut, Pilpel). *Let $f \colon S_n \to \{0, 1\}$ be in U_1. Then f is the characteristic function of a disjoint union of 1-cosets.*

Note that a disjoint union of 1-cosets is necessarily of the form

$$\bigcup_{j \in S} T_{ij}$$

for some $i \in [n]$ and some $S \subset [n]$, or of the form

$$\bigcup_{i \in S} T_{ij}$$

for some $j \in [n]$ and some $S \subset [n]$, so is determined by the image or the pre-image of a fixed element. We call such a family a *dictatorship*.

In [13], we proved that a Boolean function of expectation $O(1/n)$, whose Fourier transform is highly concentrated on irreducible representations corresponding to the partitions (n) and $(n-1, 1)$, is close in structure to a union of 1-cosets. Put another way, a Boolean function of expectation $O(1/n)$, which is close to U_1 (in Euclidean distance), has small symmetric difference with a union of 1-cosets. This is not true stability, as the Boolean function corresponding to $T_{11} \cup T_{22}$ is $O(1/n^2)$-close to U_1, but is not $1/(2n)$-close to any dictatorship, whereas a Boolean function *in* U_1 must be a dictatorship. We call it a 'quasi-stability' result. (In [14], on the other hand, we

2

prove that a Boolean function of expectation bounded away from 0 and 1, which is close to U_1, must be close in structure to a dictatorship; this is 'genuine' stability.)

Our aim in this paper is to prove an analogue of our quasi-stability result in [13], for Boolean functions of degree at most t. Namely, we show that a Boolean function on S_n with expectation $O(n^{-t})$, whose Fourier transform is highly concentrated on irreducible representations indexed by partitions of n with first row of length at least $n - t$, is close in structure to a union of cosets of stabilizers of t-tuples. Put another way, a Boolean function of expectation $O(n^{-t})$, which is close to U_t (in Euclidean distance), has small symmetric difference with a union of t-cosets. Here is the precise statement, our main theorem.

Theorem 3. *For each $t \in \mathbb{N}$, there exists $C_t > 0$ such that the following holds. Let $n \in \mathbb{N}$ with $n \geq t$, and let $\mathcal{A} \subset S_n$ with $|\mathcal{A}| = c(n-t)!$, where $c \geq 0$. Let $f = 1_{\mathcal{A}} \colon S_n \to \{0, 1\}$ denote the characteristic function of $\mathcal{A}$, so that $\mathbb{E}[f] = c/(n)_t$. Let f_t denote the orthogonal projection of f onto U_t. Let $0 \leq \epsilon \leq 1$. If $\mathbb{E}[(f - f_t)^2] \leq \epsilon c/(n)_t$, then there exists $\mathcal{C} \subset S_n$ such that $\mathcal{C}$ is a union of $\mathrm{round}(c)$ t-cosets of S_n, and*

$$|\mathcal{A} \triangle \mathcal{C}| \leq C_t(\epsilon^{1/2} + c/\sqrt{n})|\mathcal{A}|. \tag{1}$$

Moreover, we have $|c - \mathrm{round}(c)| \leq C_t(\epsilon^{1/2} + c/\sqrt{n})c$.

Here, $(n)_t := n(n-1)\ldots(n-t+1)$ denotes the falling factorial, and if $c \geq 0$, $\mathrm{round}(c)$ denotes the closest integer to c, rounding up if $c + \frac{1}{2} \in \mathbb{N}$. The term 'orthogonal projection onto U_t' will be made precise in section 2, where we define the standard inner product on the space of complex-valued functions on S_n.

Theorem 3 can be viewed as a non-Abelian analogue of the theorems in [5] and [25], which concern Boolean functions on $\{0, 1\}^n$ whose Fourier transform is highly concentrated on small sets. In [25], for example, it is shown that a Boolean function on $\{0, 1\}^n$ whose Fourier transform is highly concentrated on sets of size at most t, must be close in structure to a junta depending upon at most $j(t)$ coordinates, for some function $j : \mathbb{N} \to \mathbb{N}$.

We remark that if $t \geq 2$, then a Boolean function in U_t is not necessarily the characteristic function of a union of t-cosets. Theorem 27 in [15] states that a Boolean function in U_t is the characteristic function of a disjoint union of t-cosets, but this is false; a counterexample, and the error in the proof, is pointed out by the second author in [19]. A counterexample when $t = 2$ is as follows. Let $n \geq 8$. For any permutation $\sigma \in S_n$, define $x = x(\sigma) \in \{0, 1\}^4$ by $x_i = 1\{\sigma(i) \in [4]\}$, and consider the function

$$f : S_n \to \{0, 1\}; \quad \sigma \mapsto 1\{x_1 \geq x_2 \geq x_3 \geq x_4 \text{ or } x_1 \leq x_2 \leq x_3 \leq x_4\}. \tag{2}$$

It can be checked that $f \in U_2$, but the value of f clearly cannot be determined by fixing the images of at most two numbers, so neither f nor $1 - f$ is a union of 2-cosets. It is easy to use f to construct a counterexample for each $t \geq 3$, by considering a product of f with the characteristic function of the pointwise stabilizer of a $(t-2)$-set. We note that the main application of Theorem 27 in [15] was to characterize (for large n) the t-intersecting families in S_n of maximum size (i.e., to characterize the cases of equality in the Deza-Frankl conjecture); fortunately, this characterization follows

3

immediately e.g. from the Hilton-Milner type result of the first author in [12], where the proof does not depend on Theorem 27 in [15] (and indeed predates the latter).

The example (2) above shows that for $t \geq 2$, the right-hand side of (1) must necessarily contain a term which does not tend to zero as $\epsilon \to 0$. (This contrasts with the situation when $t = 1$; as mentioned above, any Boolean function in U_1 is a characteristic function of a disjoint union of 1-cosets.) We believe that the $\sqrt{n}$ factor in Theorem 3 is not sharp, but it suffices for our applications, and for the sake of brevity we have not attempted to improve it.

Our proof of Theorem 3 is similar in some respects to the proof of the 'quasi-stability' theorem in [13], but the representation-theoretic tools are significantly more involved. It also turns out to be easier to deal with fourth moments (rather than third moments, as in [13]); these have the advantage of always being non-negative, although we pay the price of having to substitute approximations for many of the exact expressions in [13].

We use one of our representation-theoretic lemmas to obtain an edge-isoperimetric inequality for the *transposition graph* on S_n (i.e. the Cayley graph on S_n generated by all transpositions). This inequality, Theorem 22, is a little technical, so we defer its statement until section 5. It is asymptotically sharp for subsets of S_n of size $n!/\mathrm{poly}(n)$. We combine it with Theorem 3 and some additional combinatorial arguments to prove the following.

Theorem 4. *Let $\mathcal{A} \subset S_n$ with $|\mathcal{A}| = (n - t)!$. If n is sufficiently large depending on t, then*
$$|\partial \mathcal{A}| \geq |\partial T_{(1,2,\ldots,t)(1,2,\ldots,t)}|.$$
Equality holds if and only if $\mathcal{A}$ is a t-coset of S_n.

Here, $\partial \mathcal{A}$ denotes the *edge-boundary* of $\mathcal{A}$ in the transposition graph, i.e. the set of all edges of the transposition graph which go between $\mathcal{A}$ and $S_n \setminus \mathcal{A}$. Theorem 4 confirms a conjecture of Ben Efraim [3] in the relevant cases.

We note that some of the methods in this paper bear a high level resemblance to methods used more recently by Dinur, Khot, Kindler, Minzer and Safra in [9] to obtain structural results on subsets of the Grassmann graph with small edge expansion.

The remainder of this paper is structured as follows. In section 2, we give some background on general representation theory, on the representation theory of S_n, and on normal Cayley graphs. In section 3, we prove some preliminary representation-theoretic results we will need. In section 4, we prove our 'quasi-stability' result, Theorem 3. In section 5, we prove our isoperimetric inequality for the transposition graph on S_n, Theorem 4. In section 6, we outline how our results can be used to give an alternative proof of a stability result for t-intersecting families of permutations, obtained by the first author in [12]. Finally, in section 7, we conclude with some open problems.

4

2 Background

Background on general representation theory

In this section, we recall the basic notions and results we need from general representation theory. For more background, the reader may consult [29].

Let G be a finite group. A *representation of G over $\mathbb{C}$* is a pair (ρ, V), where V is a finite-dimensional complex vector space, and $\rho\colon G \to GL(V)$ is a group homomorphism from G to the group of all invertible linear endomorphisms of V. The vector space V, together with the linear action of G defined by $gv = \rho(g)(v)$, is sometimes called a $\mathbb{C}G$-*module*. A *homomorphism* between two representations (ρ, V) and (ρ', V') is a linear map $\phi\colon V \to V'$ such that such that $\phi(\rho(g)(v)) = \rho'(g)(\phi(v))$ for all $g \in G$ and $v \in V$. If ϕ is a linear isomorphism, the two representations are said to be *equivalent* or *isomorphic*, and we write $(\rho, V) \cong (\rho', V')$. If $\dim(V) = n$, we say that ρ *has dimension n*, and we write $\dim(\rho) = n$.

The representation (ρ, V) is said to be *irreducible* if it has no proper subrepresentation, i.e. there is no proper subspace of V which is $\rho(g)$-invariant for all $g \in G$.

It turns out that for any finite group G there are only finitely many equivalence classes of irreducible representations of G, and *any* representation of G is isomorphic to a direct sum of irreducible representations of G. Hence, we may choose a set of representatives $\mathcal{R}$ for the equivalence classes of irreducible representations of G.

If (ρ, V) is a representation of G, the *character χ_ρ* of ρ is the map defined by

$$
\begin{aligned}
\chi_\rho\colon G &\to \mathbb{C}; \\
\chi_\rho(g) &= \mathrm{Tr}(\rho(g)),
\end{aligned}
$$

where $\mathrm{Tr}(\alpha)$ denotes the trace of the linear map α (i.e. the trace of any matrix of α). Note that $\chi_\rho(\mathrm{Id}) = \dim(\rho)$, and that χ_ρ is a *class function* on G (meaning that it is constant on each conjugacy-class of G.)

The usefulness of characters lies in the following.

Fact. *Two representations of G are isomorphic if and only if they have the same character. The irreducible characters form an orthonormal basis for the vector space of complex-valued class functions on G.*

We now define the Fourier transform of a function on an arbitrary finite group.

Definition. *Let $\mathcal{R}$ be a complete set of non-isomorphic, irreducible representations of G, i.e. containing one representative from each isomorphism class of irreducible representations of G. Let $f\colon G \to \mathbb{C}$ be a complex-valued function on G. We define the* Fourier transform *of f at an irreducible representation $\rho \in \mathcal{R}$ as*

$$
\hat{f}(\rho) = \frac{1}{|G|} \sum_{\sigma \in G} f(\sigma)\rho(\sigma); \tag{3}
$$

Note that $\hat{f}(\rho)$ is a linear endomorphism of V.

Let G be a finite group. Let $\mathbb{C}[G]$ denote the vector space of all complex-valued functions on G. Let $\mathbb{P}$ denote the uniform probability measure on G,

$$\mathbb{P}(\mathcal{A}) = |\mathcal{A}|/|G| \quad (\mathcal{A} \subset G),$$

let

$$\langle f, g \rangle = \frac{1}{|G|} \sum_{\sigma \in G} f(\sigma)\overline{g(\sigma)}$$

denote the corresponding inner product, and let

$$\|f\|_2 = \sqrt{\mathbb{E}[f^2]} = \sqrt{\frac{1}{|G|} \sum_{\sigma \in G} |f(\sigma)|^2}$$

denote the induced Euclidean norm.

For each irreducible representation $\rho \in \mathcal{R}$, let

$$U_\rho := \{f \in \mathbb{C}[G] : \hat{f}(\pi) = 0 \text{ for all } \pi \in \mathcal{R} \setminus \{\rho\}\}.$$

We refer to this as the subspace of functions whose Fourier transform is supported on the irreducible representation ρ, and we refer to the U_ρ as the *isotypical subspaces*. It turns out that the U_ρ are pairwise orthogonal, and that

$$\mathbb{C}[G] = \bigoplus_{\rho \in \mathcal{R}} U_\rho. \tag{4}$$

For each $\rho \in \mathcal{R}$, let f_ρ denote the orthogonal projection of f onto the subspace U_ρ. It follows from the above that

$$\|f\|_2^2 = \sum_{\rho \in \mathcal{R}} \|f_\rho\|_2^2. \tag{5}$$

The *group algebra* $\mathbb{C}G$ denotes the complex vector-space with basis G and multiplication defined by extending the group multiplication linearly. In other words,

$$\mathbb{C}G = \left\{ \sum_{g \in G} x_g g : x_g \in \mathbb{C} \ \forall g \in G \right\},$$

and

$$\left(\sum_{g \in G} x_g g \right) \left(\sum_{h \in G} y_h h \right) = \sum_{g,h \in G} x_g y_h (gh).$$

As a vector space, $\mathbb{C}G$ may be identified with $\mathbb{C}[G]$, by identifying $\sum_{g \in G} x_g g$ with the function $g \mapsto x_g$.

Background on the representation theory of S_n

Our treatment below follows [26] and [30].

Definition. *A partition of n is a non-increasing sequence of positive integers summing to n, i.e. a sequence $\lambda = (\lambda_1, \ldots, \lambda_k)$ with $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k \geq 1$ and $\sum_{i=1}^{k} \lambda_i = n$; we write $\lambda \vdash n$. For example, $(3,2,2) \vdash 7$.*

If a part a appears t times in a partition then we write a^t instead of $\overbrace{a, \ldots, a}^{t \text{ times}}$. For example, $(3, 2^2) = (3,2,2)$.

The following two orders on partitions of n will be useful.

Definition. *(Dominance order) Let $\lambda = (\lambda_1, \ldots, \lambda_k)$ and $\mu = (\mu_1, \ldots, \mu_l)$ be partitions of n. We say that $\lambda \trianglerighteq \mu$ (λ dominates μ) if $\sum_{j=1}^{i} \lambda_i \geq \sum_{j=1}^{i} \mu_i \ \forall i$ (where we define $\lambda_i = 0 \ \forall i > k$, $\mu_i = 0 \ \forall i > l$).*

It is easy to see that this is a partial order.

Definition. *(Lexicographic order) Let $\lambda = (\lambda_1, \ldots, \lambda_r)$ and $\mu = (\mu_1, \ldots, \mu_s)$ be partitions of n. We say that $\lambda > \mu$ if $\lambda_j > \mu_j$, where $j = \min\{i \in [n] \colon \lambda_i \neq \mu_i\}$.*

It is easy to see that this is a total order which extends the dominance order.

It is well-known that there is an explicit 1-1 correspondence between irreducible representations of S_n (up to isomorphism) and partitions of n. The reader may refer to [30] for a full description of this correspondence, or to the paper [15] for a shorter description.

For each partition α of n, we write $[\alpha]$ for the corresponding isomorphism class of irreducible representations of S_n, and we write $U_\alpha = U_{[\alpha]}$ for the vector space of complex-valued functions on S_n whose Fourier transform is supported on $[\alpha]$. Similarly, if $f \in \mathbb{C}[S_n]$, we write f_α for the orthogonal projection of f onto U_α. By (4), we have

$$\mathbb{C}[S_n] = \bigoplus_{\alpha \vdash n} U_\alpha,$$

and the U_α are pairwise orthogonal. For any $f \in \mathbb{C}[S_n]$, we have

$$f = \sum_{\alpha \vdash n} f_\alpha,$$

and

$$\|f\|_2^2 = \sum_{\alpha \vdash n} \|f_\alpha\|^2.$$

We write χ_α for the character of the irreducible representation corresponding to α, and we write $\dim[\alpha]$ for the dimension of $[\alpha]$.

Definition. *If $\lambda = (\lambda_1, \ldots, \lambda_l)$ is a partition of n, the Young diagram of shape λ is an array of left-justified cells with λ_i cells in row i, for each $i \in [l]$.*

For example, the Young diagram of the partition $(3, 2^2)$ is:

Definition. *A λ-tableau is a Young diagram of shape λ, each of whose cells contains a number between 1 and n. If μ is a partition of n, a Young tableau is said to have content μ if it contains μ_i i's for each $i \in \mathbb{N}$.*

Definition. *A Young tableau is said to be* standard *if it has content $(1, 1, \ldots, 1)$ and the numbers are strictly increasing down each row and along each column.*

Definition. *A Young tableau is said to be* semistandard *if the numbers are non-decreasing along each row and strictly increasing down each column.*

The relevance of standard Young tableaux stems from the following

Theorem 5. *If α is a partition of n, then $\dim[\alpha]$ is the number of standard α-tableaux.*

We say that two λ-tableaux of content $(1, 1, \ldots, 1)$ are *row-equivalent* if they contain the same set of numbers in each row. A row-equivalence-class of λ-tableaux with content $(1, 1, \ldots, 1)$ is called a λ-*tabloid*. Consider the natural left action of S_n on the set of λ-tabloids, i.e. if a λ-tabloid $\mathcal{T}$ has ith row $R_i \in [n]^{(\lambda_i)}$ for each i, then $\sigma(\mathcal{T})$ has ith row $\sigma(R_i)$ for each i. Let M^λ denote the induced permutation representation. We write ξ_λ for the character of M^λ; the ξ_λ are called the *permutation characters* of S_n.

Young's theorem gives the decomposition of each permutation representation into irreducible representations of S_n, in terms of the *Kostka numbers*.

Definition. *Let λ and μ be partitions of n. The* Kostka number *$K_{\lambda,\mu}$ is the number of semistandard λ-tableaux of content μ.*

Theorem 6 (Young's theorem). *If μ is a partition of n, then*

$$M^\mu \cong \bigoplus_{\substack{\lambda \vdash n : \\ \lambda \geq \mu}} K_{\lambda,\mu}[\lambda].$$

It follows that for each partition μ of n, we have

$$\xi_\mu = \sum_{\substack{\lambda \vdash n : \\ \lambda \geq \mu}} K_{\lambda,\mu} \chi_\lambda.$$

On the other hand, we can express the irreducible characters in terms of the permutation characters using the *determinantal formula*: for any partition α of n,

$$\chi_\alpha = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \xi_{\alpha - \operatorname{id} + \pi}. \tag{6}$$

Here, if $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_l)$, $\alpha - \mathrm{id} + \pi$ is defined to be the sequence

$$(\alpha_1 - 1 + \pi(1), \alpha_2 - 2 + \pi(2), \ldots, \alpha_l - l + \pi(l)).$$

If this sequence has all its entries non-negative, we let $\overline{\alpha - \mathrm{id} + \pi}$ be the partition of n obtained by reordering its entries, and we define $\xi_{\alpha - \mathrm{id} + \pi} = \xi_{\overline{\alpha - \mathrm{id} + \pi}}$. If the sequence has a negative entry, we define $\xi_{\alpha - \mathrm{id} + \pi} = 0$. Note that if ξ_β appears on the right-hand side of (6), then $\beta \geq \alpha$, so the determinantal formula expresses χ_α in terms of $\{\xi_\beta : \beta \geq \alpha\}$. (See [26] for a proof of (6).)

For each $t \in \mathbb{N}$, we define

$$U_t = \bigoplus_{\substack{\alpha \vdash n: \\ \alpha_1 \geq n - t}} U_\alpha,$$

i.e. U_t is the subspace of functions on S_n whose Fourier transform is concentrated on irreducible representations corresponding to partitions with first row of length at least $n - t$. As mentioned in the introduction, it was proved in [15] that U_t is the linear span of the characteristic functions of the t-cosets of S_n:

$$U_t = \mathrm{Span}\{T_{IJ} : I, J \text{ are ordered } t\text{-tuples of distinct elements of } [n]\}.$$

For brevity, we shall sometimes write $\mathcal{C}(n, t)$ for the set of t-cosets of S_n.

If $f\colon S_n \to \mathbb{C}$, we write f_t for the orthogonal projection of f onto U_t, for each $t \in [n]$; equivalently,

$$f_t = \sum_{\substack{\alpha \vdash n: \\ \alpha_1 \geq n - t}} f_\alpha.$$

For each $t \in \mathbb{N}$, we write

$$V_t = \bigoplus_{\substack{\alpha \vdash n: \\ \alpha_1 = n - t}} U_\alpha.$$

Note that the V_t are pairwise orthogonal, and that

$$U_t = U_{t-1} \oplus V_t$$

for each $t \leq n$.

Background on normal Cayley graphs

Several of our results will involve the analysis of *normal Cayley graphs* on S_n. We recall the definition of a Cayley graph on a finite group.

Definition. *Let G be a finite group, and let $S \subset G \setminus \{Id\}$ be inverse-closed (meaning that $S^{-1} = S$). The (left) Cayley graph on G with generating set S is the graph with vertex-set G, where we join g to sg for every $g \in G$ and $s \in S$; we denote it by $\mathrm{Cay}(G, S)$. Formally,*

$$V(\mathrm{Cay}(G, S)) = G, \quad E(\mathrm{Cay}(G, S)) = \{\{g, sg\}\colon g \in G, \ s \in S\}.$$

Note that the Cayley graph $\mathrm{Cay}(G, S)$ is $|S|$-regular. If the generating set S is conjugation-invariant, i.e. is a union of conjugacy classes of G, then the Cayley graph $\mathrm{Cay}(G, S)$ is said to be a normal *Cayley graph.*

The connection between normal Cayley graphs and representation theory arises from the following fundamental theorem, which states that for any normal Cayley graph, the eigenspaces of its adjacency matrix are in 1-1 correspondence with the isomorphism classes of irreducible representations of the group.

Theorem 7 (Frobenius / Schur / Diaconis-Shahshahani). *Let G be a finite group, let $S \subset G$ be an inverse-closed, conjugation-invariant subset of G, let $\Gamma = \mathrm{Cay}(G,S)$ be the Cayley graph on G with generating set S, and let A be the adjacency matrix of Γ. Let $\mathcal{R}$ be a complete set of non-isomorphic complex irreducible representations of G. Then we have*

$$\mathbb{C}[G] = \bigoplus_{\rho \in \mathcal{R}} U_\rho,$$

and each U_ρ is an eigenspace of A with dimension $\dim(\rho)^2$ and eigenvalue

$$\lambda_\rho = \frac{1}{\dim(\rho)} \sum_{g \in S} \chi_\rho(g). \tag{7}$$

Note that the adjacency matrix A in the above theorem is given by

$$A_{g,h} = 1_S(gh^{-1}) \quad (g, h \in G).$$

We will need a slight generalization of Theorem 7. Recall that if G is a finite group, a *class function* on G is a function on G which is constant on each conjugacy-class of G. If S is conjugation-invariant, then 1_S is clearly a class function. Observe that if A is any matrix in $\mathbb{C}[G \times G]$ defined by

$$A_{g,h} = w(gh^{-1}),$$

where w is a class function on G, then A is a linear combination of the adjacency matrices of normal Cayley graphs. Therefore, the isotypical subspaces U_ρ are again eigenspaces of A, and the corresponding eigenvalues are given by

$$\lambda_\rho = \frac{1}{\dim(\rho)} \sum_{g \in S} w(g)\chi_\rho(g). \tag{8}$$

Applying Theorem 7 to S_n, we see that if $\Gamma = \mathrm{Cay}(S_n, X)$ is a normal Cayley graph on S_n and A is the adjacency matrix of Γ then we have

$$\mathbb{C}[S_n] = \bigoplus_{\alpha \vdash n} U_\alpha,$$

and each U_α is an eigenspace of A with dimension $\dim[\alpha]$ and eigenvalue

$$\lambda_\alpha = \frac{1}{\dim[\alpha]} \sum_{\sigma \in X} \chi_\alpha(\sigma). \tag{9}$$

3 Preliminary results

In this section, we prove some preliminary representation-theoretic lemmas.

Preliminary results on the characters of S_n

In this section, we prove some easy results on the characters of S_n. We will use the following standard notation. For $n, r \in \mathbb{N}$, we will write

$$(n)_r := n(n-1)\ldots(n-r+1)$$

for the rth falling factorial moment of n, and we will write $([n])_r$ for the set of all r-tuples of distinct elements of $[n]$.

If $f = f(n,t)$, $g = g(n,t) : \mathbb{N} \times \mathbb{N} \to \mathbb{R}_{\geq 0}$, we will write $f = O_t(g)$ if for every $t \in \mathbb{N}$, there exists $C_t > 0$ such that for all $n \in \mathbb{N}$, $f(n,t) \leq C_t g(n,t)$. Similarly, we write $f = \Omega_t(g)$ if for every $t \in \mathbb{N}$, there exists $c_t > 0$ such that for all $n \in \mathbb{N}$, $f(n,t) \geq c_t g(n,t)$. We write $f = \Theta_t(g)$ if $f = O_t(g)$ and $f = \Omega_t(g)$.

We begin with a crude upper bound on the dimensions of the irreducible characters χ_α.

Lemma 8. *Let α be a partition of n with $\alpha_1 = n - s$. Then*

$$\dim[\alpha] \leq (n)_s.$$

Proof. Since $\alpha_1 = n - s$, we have $\alpha \geq (n - s, 1^s)$. Recall (or observe from Theorem 6) that for each partition $\beta \geq (n - s, 1^s)$, $[\beta]$ is an irreducible constituent of $M^{(n-s,1^s)}$, which has dimension $(n)_s$. It follows that $\dim[\alpha] \leq (n)_s$, as required. $\square$

We also need crude bounds on the Kostka numbers $K_{\alpha,(n-s,1^s)}$.

Lemma 9. *Let α be a partition of n with $\alpha_1 = n - s$. Then*

$$\binom{n-s}{s} K_{\alpha,(n-s,1^s)} \leq \dim[\alpha] \leq \binom{n}{s} K_{\alpha,(n-s,1^s)}.$$

Proof. Let us write $\alpha = (n - s, \gamma)$, where γ is a partition of s. Recall from Theorem 5 that $\dim[\alpha]$ is the number of standard α-tableaux. Observe that we may construct $\binom{n-s}{s} \dim[\gamma]$ distinct standard α-tableaux as follows:

1. Place i in cell $(1, i)$ for $i = 1, 2, \ldots, s$.

2. Choose any s-set $\{i_1, \ldots, i_s\}$ from $\{s + 1, \ldots, n\}$ ($\binom{n-s}{s}$ choices).

3. Place $\{i_1, \ldots, i_s\}$ in the cells below row 1 so that they are strictly increasing along each row and down each column. The number of ways of doing this is precisely the number of standard γ-tableaux, which is $\dim[\gamma]$.

4. Place the other numbers in the remaining cells of row 1, in increasing order from left to right.

Hence, there are at least $\binom{n-s}{s} \dim[\gamma]$ standard α-tableaux. On the other hand, removing the first row from any standard α-tableau produces a standard γ-tabeau (filled

with some s of the numbers between 1 and n, rather than with $\{1, 2, \ldots, s\}$). Hence, the number of standard α-tableaux is at most $\binom{n}{s} \dim[\gamma]$. Therefore,

$$\binom{n-s}{s} \dim[\gamma] \leq \dim[\alpha] \leq \binom{n}{s} \dim[\gamma].$$

Recall from Theorem 6 that $K_{\alpha,(n-s,1^s)}$ is the number of semistandard α-tableaux of content $(n-s, 1^s)$. A semistandard α-tableau of content $(n-s, 1^s)$ must have all $n-s$ of its 1's in the first row; after deleting the first row, what is left is precisely a standard γ-tableau, filled with the numbers $\{2, 3, \ldots, s+1\}$, rather than with $\{1, 2, \ldots, s\}$. Hence, the number of semistandard α-tableaux of content $(n-s, 1^s)$ is precisely the number of standard γ-tableaux, so $K_{\alpha,(n-s,1^s)} = \dim[\gamma]$. Therefore,

$$\binom{n-s}{s} K_{\alpha,(n-s,1^s)} \leq \dim[\alpha] \leq \binom{n}{s} K_{\alpha,(n-s,1^s)},$$

as required. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Finally, we need a crude lower bound on the L^1-norm of the characters of the symmetric group.

Lemma 10. *For each $s \in \mathbb{N}$, there exists $K_s > 0$ such that for any $n \in \mathbb{N}$ and any partition α of n with $\alpha_1 \geq n - s$, we have*

$$\frac{1}{n!} \sum_{\sigma \in S_n} |\chi_\alpha(\sigma)| \geq K_s.$$

To assist with the proof, we introduce the following notation/definition.

Definition. *Let $s \in \mathbb{N}$, and let $\beta = (\beta_1, \ldots, \beta_l)$ be a partition of n with $\beta_1 \geq n - s$. We define $X_{\beta,s}$ to be the set of all permutations with cycle-type in the set*

$$\{(\lambda_1, \ldots, \lambda_k, \beta_2, \ldots, \beta_l) : (\lambda_1, \ldots, \lambda_k) \vdash \beta_1, \ |\lambda_i| > s \ \forall i \in [k]\}.$$

Definition. *For $m, s \in \mathbb{N}$ with $m > s$, we let $N_{m,s}$ denote the number of permutations in S_m whose cycles all have lengths greater than s.*

We also need three preparatory claims.

Claim 1. *For any $m, s \in \mathbb{N}$ with $m > s$, we have $N_{m,s} \geq m!/(2s+1)$.*

Proof of Claim. By induction on m, for each fixed $s \in \mathbb{N}$. If $s + 1 \leq m \leq 2s + 1$, then $N_{m,s}$ is precisely the number of m-cycles in S_m, which is $(m-1)! \geq m!/(2s+1)$. We now derive a recurrence relation for $N_{m,s}$, for all $m \geq 2s + 2$. Let $\sigma \in S_m$ be a permutation with all cycles of length greater than s. Let $i = \sigma(m)$; then $i < m$ and we may write $\sigma = (i \ m)\rho$, where $\rho \in S_{[m-1]}$ is a permutation with all its cycles of length greater than s, except possibly an s-cycle containing i. Conversely, given such a pair (ρ, i), $(i \ m)\rho$ has all its cycles of length greater than s. It follows that

$$N_{m,s} = (m-1)(N_{m-1,s} + (m-2)(m-3)\ldots(m-s)N_{m-s-1,s}) \quad \forall m > s,$$

12

where we define $N_{0,s} := 1$. We now perform the inductive step. Let $m \geq 2s + 2$, and assume that the claim holds when m is replaced by any $m' < m$. Since $m - 1 - s > s$, by the inductive hypothesis we have

$$
\begin{aligned}
N_{m,s} &= (m-1)(N_{m-1,s} + (m-2)(m-3)\ldots(m-s)N_{m-s-1,s}) \\
&\geq (m-1)((m-1)! + (m-2)(m-3)\ldots(m-s)(m-s-1)!)/(2s+1) \\
&= m!/(2s+1),
\end{aligned}
$$

completing the inductive step and proving the claim. $\qquad\square$

Claim 2. *Let $\beta = (\beta_1, \ldots, \beta_l)$ be a partition of n with $\beta_1 \geq n - s$, where $n > 2s$. Then*
$$
|X_{\beta,s}| \geq L_s n!,
$$
where $L_s > 0$ depends upon s alone.

Proof of Claim. Let $\beta_1 = n - r$, where $r \leq s$. Observe that

$$
|X_{\beta,s}| \geq \binom{n}{r} N_{n-r,s},
$$

since there are $\binom{n}{r}$ choices for the r numbers to go in the cycles of lengths $\beta_2, \ldots, \beta_l$, and then $N_{n-r,s}$ choices for placing the other $n - r$ numbers in the cycles of lengths greater than s. It follows, using Claim 1, that

$$
|X_{\beta,s}| \geq \binom{n}{r}(n-r)!/(2s+1) = \frac{n!}{r!(2s+1)} \geq \frac{n!}{s!(2s+1)} n! = L_s n!,
$$

where $L_s := 1/(s!(2s+1))$, proving the claim. $\qquad\square$

Claim 3. *If α, β are partitions of n with $\alpha_1, \beta_1 \geq n - s$, where $n > 2s$, then χ_α is constant on $X_{\beta,s}$.*

Proof of Claim. Recall from (6) that χ_α can be expressed as a linear combination of the permutation characters $\{\xi_\gamma : \gamma \geq \alpha\}$. Recall also that $\xi_\gamma(\sigma)$ is simply the number of γ-tabloids fixed by σ, which is the number of γ-tabloids that can be produced by taking each row to be a union of cycles of σ. This is clearly the same for all $\sigma \in X_{\beta,s}$, since in order to produce a γ-tabloid from $\sigma \in X_{\beta,s}$ as above, the top row must contain the union of all the cycles of length greater than s.

It follows that $\chi_\alpha(X_{\beta,s}) = \chi_\alpha(\sigma_\beta)$, where σ_β is a permutation of cycle-type β — so $\chi_\alpha(X_{\beta,s})$ is simply the (α, β)-entry of the character table of S_n. $\qquad\square$

We can now prove the lemma.

Proof of Lemma 10. Let α be a partition of n with $\alpha_1 \geq n - s$. Recall that for any partition α of n, the top-left minor of the character table of S_n indexed by partitions $\geq \alpha$, is invertible (see, for example, [15, Theorem 20]). It follows that it cannot have

a row of zeros, so there exists $\beta \geq \alpha$ such that $\chi_\alpha(\sigma_\beta) \neq 0$. Since the irreducible characters of S_n are all integer-valued, we have $|\chi_\alpha(\sigma_\beta)| \geq 1$. Hence, we have

$$\frac{1}{n!} \sum_{\sigma \in S_n} |\chi_\alpha(\sigma)| \geq \frac{1}{n!} |X_{\beta,s}||\chi_\alpha(\sigma_\beta)| \geq L_s.$$

This proves the lemma when $n > 2s$. Since no character vanishes on all of S_n, by taking K_s to be sufficiently small, we see that the lemma holds for all n. $\qquad \square$

A preliminary result on the projections of Boolean functions

The following result will be useful both in the proof of our quasi-stability theorem, and in that of our approximate isoperimetric inequality for the transposition graph.

Lemma 11. *For each $t \in \mathbb{N}$ there exists $C'_t > 0$ such that the following holds. If $\mathcal{A} \subset S_n$ and $f = 1_\mathcal{A}$ denotes the characteristic function of $\mathcal{A}$, then*

$$\|f_{t-1}\|_2^2 \leq C'_t n^{t-1}(|\mathcal{A}|/n!)^2.$$

Proof. Let α be a partition of n with $\alpha_1 = n - s$. Our aim is to prove that

$$\|f_\alpha\|_2^2 \leq C''_s n^s (|\mathcal{A}|/n!)^2,$$

where $C''_s > 0$ depends upon s alone. If $s = 0$ then $\alpha = (n)$, and we have

$$\|f_{(n)}\|_2^2 = (|\mathcal{A}|/n!)^2,$$

so we may take $C''_0 = 1$.

Suppose now that $s \geq 1$. Let $\Gamma = \mathrm{Cay}(S_n, X)$ be any normal Cayley graph on S_n, and let A denote its adjacency matrix. If $\mathcal{B} \subset S_n$, let $e(\mathcal{B})$ denote the number of edges of Γ within $\mathcal{B}$, and let $(\lambda_\beta)_{\beta \vdash n}$ denote the eigenvalues of A. By Theorem 7, we have

$$Af = \sum_{\beta \vdash n} \lambda_\beta f_\beta,$$

and therefore

$$2e(\mathcal{A}) = n!\langle f, Af \rangle = n! \sum_{\beta \vdash n} \lambda_\beta \|f_\beta\|^2.$$

Trivially, we have $e(\mathcal{A}) \leq \binom{|\mathcal{A}|}{2}$, and therefore

$$|\mathcal{A}|^2 \geq |\mathcal{A}|(|\mathcal{A}| - 1) \geq 2e(\mathcal{A}) = n! \sum_{\beta \vdash n} \lambda_\beta \|f_\beta\|^2 \geq n!\lambda_\alpha \|f_\alpha\|_2^2.$$

It follows that

$$\|f_\alpha\|_2^2 \leq \frac{|\mathcal{A}|^2}{n!\lambda_\alpha}.$$

In order to minimize the right-hand side, we will choose Γ with λ_α as large as possible.

By (9), we have
$$\lambda_\alpha = \frac{1}{\dim[\alpha]} \sum_{\sigma \in X} \chi_\alpha(\sigma).$$

To maximize λ_α, we simply take $X = \{\sigma \in S_n : \chi_\alpha(\sigma) > 0\}$. The corresponding Cayley graph $\mathrm{Cay}(S_n, X)$ then has
$$\lambda_\alpha = \frac{1}{\dim[\alpha]} \sum_{\substack{\sigma \in S_n: \\ \chi_\alpha(\sigma) > 0}} \chi_\alpha(\sigma).$$

Since $\alpha \neq (n)$, by the orthogonality of the irreducible characters, we have
$$\langle \chi_\alpha, \chi_{(n)} \rangle = 0,$$

i.e.
$$\sum_{\sigma \in S_n} \chi_\alpha(\sigma) = 0,$$

and therefore
$$\sum_{\substack{\sigma \in S_n: \\ \chi_\alpha(\sigma) > 0}} \chi_\alpha(\sigma) = \tfrac{1}{2} \sum_{\sigma \in S_n} |\chi_\alpha(\sigma)|.$$

Hence,
$$\lambda_\alpha = \frac{1}{2\dim[\alpha]} \sum_{\sigma \in S_n} |\chi_\alpha(\sigma)|.$$

We therefore obtain
$$\|f_\alpha\|_2^2 \leq \frac{2\dim[\alpha]|\mathcal{A}|^2}{n! \sum_{\sigma \in S_n} |\chi_\alpha(\sigma)|}.$$

By Lemma 8, we have $\dim[\alpha] \leq (n)_s \leq n^s$, and by Lemma 10, we have
$$\frac{1}{n!} \sum_{\sigma \in S_n} |\chi_\alpha(\sigma)| \geq K_s.$$

Hence,
$$\|f_\alpha\|_2^2 \leq \frac{2n^s}{K_s}(|\mathcal{A}|/n!)^2 = C_s'' n^s (|\mathcal{A}|/n!)^2,$$

where $C_s'' = 2/K_s$.

Recall that
$$\|f_{t-1}\|_2^2 = \sum_{\substack{\alpha \vdash n: \\ \alpha_1 \geq n-t+1}} \|f_\alpha\|_2^2.$$

Note that for each $n \geq 2s$, the number of partitions α of n with $\alpha_1 = n - s$ is equal to the number of partitions of s. Hence, the number of terms in the above sum is bounded from above by a function of t alone, so there exists $C_t' > 0$ such that
$$\|f_{t-1}\|_2^2 \leq C_t' n^{t-1}(|\mathcal{A}|/n!)^2,$$

as required. $\qquad\square$

Observe that Lemma 11 implies that if $\mathcal{A} \subset S_n$ with $|\mathcal{A}| = o((n - t + 1)!)$, then $\|(1_\mathcal{A})_{t-1}\|_2^2 = o(|\mathcal{A}|/n!)$, i.e. the projection of $1_\mathcal{A}$ onto U_{t-1} has small L^2-norm. In other words, the Fourier transform of $1_\mathcal{A}$ has very little 'mass' on the irreducible representations $\{[\alpha] : \alpha_1 \geq n - t + 1\}$.

4 Proof of the quasi-stability result

Our aim in this section is to prove Theorem 3. We first give an overview of the proof.

Proof overview. The proof proceeds as follows. Instead of working with f_t, the projection of f onto U_t, it will be more convenient to work with two different functions, which are both close to f_t in L^2-norm. Namely, we will let

$$g_t = f_t - f_{t-1};$$

note that g_t is the orthogonal projection of f onto V_t, as V_t is the orthogonal complement of U_{t-1} in U_t. We will see that if $c = o(n)$, then $\|f_{t-1}\|_2^2 = o(\|f\|_2^2)$, i.e. the projection of f onto U_{t-1} is small, so g_t is close to f_t in L^2-norm.

For each t-coset T, we define

$$a_T = |\mathcal{A} \cap T|/(n - t)!;$$

note that $a_T = (n)_t\langle f, 1_T\rangle = \mathbb{E}[f|_T]$, the expectation of f restricted to T. Our aim is to show that there are approximately c t-cosets T such that a_T is close to 1. However, there are no simple relationships between the a_T's and the moments of f_t, so instead, we work with the quantities

$$b_T := (n)_t\langle g_t, 1_T\rangle = \mathbb{E}[g_t|_T].$$

We define

$$h_t := \sum_T b_T 1_T.$$

Because h_t is in V_t (as opposed to f_t, which is in U_t but not in V_t), there is a close relationship between $\mathbb{E}[h_t^2]$ and $\sum_T b_T^2$, and between $\mathbb{E}[h_t^4]$ and $\sum_T b_T^4$. Moreover, we will also see that h_t is close to g_t (and therefore to f) in L^2-norm, so $\mathbb{E}[h_t^2] \approx c/(n)_t$. This, together with the relationship between $\mathbb{E}[h_t^2]$ and $\sum_T b_T^2$, will imply that $\sum_T b_T^2 \approx c$.

Using the fact that $\mathbb{E}[(h_t - f)^2]$ is small, together with the fact that f is a Boolean function with expectation $c/(n)_t$, we will show that $\mathbb{E}[h_t^4] \approx c/(n)_t$. This, together with the relationship between $\sum_T b_T^4$ and $\mathbb{E}[h_t^4]$, will imply that $\sum_T b_T^4 \approx c$. As long as $c = o(n)$, we will have $|b_T - a_T| = o(1)$ for each T, and therefore $b_T \leq 1 + o(1)$. Hence, the only way we can have $\sum_T b_T^2 \approx c$ and $\sum_T b_T^4 \approx c$ is if approximately c of the b_T's are close to 1. This in turn implies that the corresponding a_T's are close to 1, completing the proof.

Note that this proof relies upon analysing fourth moments, as opposed to our proof of the $t = 1$ case in [13], which relied upon analysing the third moments of a

non-negative function. For $t > 1$, there is no simple relationship between h_t and the non-negative function

$$v_t = \sum_T a_T 1_T,$$

which is the natural analogue of non-negative function we used in the $t = 1$ case. The relationship between $\mathbb{E}[h_t^3]$ and $\mathbb{E}[v_t^3]$ seems too hard to analyse for general t, and we were unable to find another non-negative function whose third moment was simply related to that of h_t. This led us to consider $\mathbb{E}[h_t^4]$ instead; being non-negative, this is easier to bound from below than $\mathbb{E}[h_t^3]$. The price we pay is that the relationship between $\mathbb{E}[h_t^4]$ and $\sum_T b_T^4$ is considerably more complicated than that between $\mathbb{E}[h_t^3]$ and $\sum_T b_T^3$; we make do with an approximate relationship, as opposed to the exact one in [13]. $\qquad\square$

We now begin our formal proof.

Proof of Theorem 3

Let $\mathcal{A}, f, \epsilon$ satisfy the hypotheses of Theorem 3. If $\epsilon > 1/2$, then provided C_t is chosen to be sufficiently large, the conclusion of the theorem holds trivially, for any union $\mathcal{C}$ of round(c) t-cosets of S_n. Moreover, for any $c_t > 0$, if $c \geq c_t\sqrt{n}$, and C_t is chosen to be sufficiently large, the conclusion of the theorem holds trivially. Hence, we may assume throughout that $0 \leq \epsilon \leq 1/2$ and $c \leq c_t\sqrt{n}$, where $c_t > 0$ is to be chosen later.

For each t-coset T of S_n, define

$$a_T = \frac{|\mathcal{A} \cap T|}{(n-t)!}.$$

Our ultimate aim is to show that approximately c of the a_T's are close to 1. Observe that

$$a_T = (n)_t \langle f, 1_T \rangle = (n)_t \langle f_t, 1_T \rangle \quad \forall T \in \mathcal{C}(n, t),$$

since $1_T \in U_t$.

Define

$$g_t = f_t - f_{t-1};$$

note that g_t is the orthogonal projection of f onto V_t. Instead of working with the a_T's, for most of the proof we will work with the quantities

$$b_T := (n)_t \langle g_t, 1_T \rangle = \mathbb{E}[g_t|_T].$$

By Lemma 11, we have

$$\|f_{t-1}\|_2^2 \leq C_t' n^{t-1} (c/(n)_t)^2 = O_t(c/n)c/(n)_t = O_t(c/n)\mathbb{E}[f^2], \tag{10}$$

so if $c = o(n)$, f_t is close to $g_t := f_t - f_{t-1}$ in L^2-norm, as desired.

It follows that b_T is close to a_T, for each t-coset T. Indeed, we have

$$
\begin{aligned}
|a_T - b_T| &= (n)_t |\langle f_{t-1}, 1_T \rangle| \\
&\leq (n)_t \| f_{t-1} \|_2 \| 1_T \|_2 \\
&\leq (n)_t \sqrt{C'_t n^{t-1}} (c/(n)_t) \sqrt{1/(n)_t} \\
&\leq O_t(c/\sqrt{n}),
\end{aligned} \tag{11}
$$

using the Cauchy–Schwarz inequality and (10).

Since $a_T \in [0,1]$, it follows that

$$
-O_t(c/\sqrt{n}) \leq b_T \leq 1 + O_t(c/\sqrt{n}) \quad \forall T \in \mathcal{C}(n,t), \tag{12}
$$

so b_T cannot lie too far outside the interval $[0,1]$.

We now define the function

$$
h_t = \sum_{T \in \mathcal{C}(n,t)} b_T 1_T = (n)_t \sum_{T \in \mathcal{C}(n,t)} \langle g_t, 1_T \rangle 1_T;
$$

this function is the crucial one in our proof. Our next step is to show that h_t must be close to g_t. To this end, we define the linear operator

$$
\begin{aligned}
M \colon \mathbb{C}[S_n] &\to \mathbb{C}[S_n]; \\
g &\mapsto (n)_t \sum_{T \in \mathcal{C}(n,t)} \langle g, 1_T \rangle 1_T.
\end{aligned}
$$

(Note that $h_t = M g_t$.) We now prove the following.

Lemma 12. *The eigenspaces of M are precisely the U_α. The operator $M|_{V_t}$ is an invertible endomorphism of V_t, and all its eigenvalues are $1 + O_t(1/n)$.*

Proof. For each $\sigma \in S_n$, let us write e_σ for the function on S_n which is 1 at σ and 0 elsewhere. With slight abuse of notation, we use $(M_{\sigma,\pi})_{\sigma,\pi \in S_n}$ to denote the matrix of M with respect to the standard basis $\{e_\sigma \colon \sigma \in S_n\}$ of $\mathbb{C}[S_n]$. Observe that

$$
\begin{aligned}
M_{\sigma,\pi} &= (M e_\pi)(\sigma) \\
&= (n)_t \sum_{T \in \mathcal{C}(n,t)} \tfrac{1}{n!} 1\{\pi \in T\} 1\{\sigma \in T\} \\
&= \frac{(n)_t}{n!} |\{T \in \mathcal{C}(n,t) : \sigma, \pi \in T\}| \\
&= \frac{(n)_t}{n!} |\{x \in [n]^{(t)} : \sigma(i) = \pi(i) \; \forall i \in x\}| \\
&= \frac{(n)_t}{n! t!} |\{(i_1, \ldots, i_t) \in [n]^t : i_1, \ldots, i_t \text{ are all distinct}, \; \sigma(i_k) = \pi(i_k) \; \forall k \in [t]\}| \\
&= \tfrac{1}{n!} \binom{n}{t} \xi_{(n-t,1^t)}(\sigma \pi^{-1}).
\end{aligned}
$$

18

(Here and in the rest of the paper, $1\{B\}$ is equal to 1 if B is true, and to 0 if B is false.)

Define

$$w_\sigma = \tfrac{1}{n!}\binom{n}{t}\xi_{(n-t,1^t)}(\sigma);$$

observe that $\sigma \mapsto w_\sigma$ is a class function, and that $M_{\sigma,\pi} = w(\sigma\pi^{-1})$ for all $\sigma, \pi \in S_n$. By (8), the eigenvalues of M are given by

$$
\begin{aligned}
\lambda_\alpha &= \frac{\binom{n}{t}}{n!f^\alpha}\sum_{\sigma\in S_n}\xi_{(n-t,1^t)}(\sigma)\chi_\alpha(\sigma) \\
&= \frac{\binom{n}{t}}{f^\alpha}\langle \xi_{(n-t,1^t)},\chi_\alpha\rangle \\
&= \frac{\binom{n}{t}}{f^\alpha}K_{\alpha,(n-t,1^t)} \quad (\alpha \vdash n),
\end{aligned}
\tag{13}
$$

with the U_α being the corresponding eigenspaces.

In particular, it follows that if $g \in V_t = \bigoplus_{\alpha\vdash n:\,\alpha_1=n-t}U_\alpha$, then $Mg \in V_t$ also, so $M|_{V_t}$ is a linear endomorphism of V_t. Since $K_{\alpha,(n-t,1^t)} \geq 1$ for all $\alpha \geq (n-t,1^t)$, we have $\lambda_\alpha > 0$ for each $\alpha \vdash n$ with $\alpha_1 = n-t$, so $M|_{V_t}$ is invertible. Combining Lemma 9 and (13), we see that for each $\alpha \vdash n$ with $\alpha_1 = n-t$, we have $\lambda_\alpha = 1 + O_t(1/n)$, completing the proof. $\qquad\square$

Since $g_t \in V_t$, it follows that $h_t = Mg_t \in V_t$. Moreover,

$$\|h_t - g_t\|_2 = \|Mg_t - g_t\|_2 \leq O_t(1/n)\|g_t\|_2 < O_t(1/n)\sqrt{c/(n)_t}. \tag{14}$$

Combining (10) and (14), using the triangle inequality, yields

$$
\begin{aligned}
\|h_t - f\|_2 &\leq \|h_t - g_t\|_2 + \|g_t - f_t\|_2 + \|f_t - f\|_2 \\
&\leq O_t(1/n)\sqrt{c/(n)_t} + O_t(\sqrt{c/n})\sqrt{c/(n)_t} + \sqrt{\epsilon}\sqrt{c/(n)_t} \\
&= (O_t(1/n) + O_t(\sqrt{c/n}) + \sqrt{\epsilon})\sqrt{c/(n)_t} \\
&= \psi\sqrt{c/(n)_t} \\
&= \psi\|f\|_2,
\end{aligned}
\tag{15}
$$

where $\psi := \sqrt{\epsilon} + O_t(1/n) + O_t(\sqrt{c/n})$.

Moreover, by (14), we have

$$|\|h_t\|_2 - \|g_t\|_2| \leq \|h_t - g_t\|_2 \leq O_t(1/n)\|g_t\|_2 < O_t(1/n)\sqrt{c/(n)_t}, \tag{16}$$

and therefore

$$\mathbb{E}[h_t^2] \leq (1 + O_t(1/n))c/(n)_t. \tag{17}$$

By (15), we have

$$\|h_t\|_2 \geq (1 - \psi)\|f\|_2,$$

and therefore
$$\mathbb{E}[h_t^2] \geq (1 - \psi)^2 c/(n)_t \geq (1 - 2\psi)c/(n)_t, \tag{18}$$
so $\mathbb{E}[h_t^2]$ is close to $c/(n)_t$.

We can simplify our expression for ψ by noting that c cannot be too small. Indeed, since we are assuming that $\epsilon \leq 1/2$, we have
$$\|f_t\|_2^2 \geq (\|f\|_2 - \|f - f_t\|_2)^2 \geq (1 - \sqrt{\epsilon})^2 c/(n)_t \geq (1 - \sqrt{1/2})^2 c/(n)_t,$$
whereas by Lemma 11, we have
$$\|f_t\|_2^2 \leq C'_{t+1} n^t (c/(n)_t)^2.$$
It follows that
$$c \geq (1 - \sqrt{1/2})^2/(C_{t+1}). \tag{19}$$
Therefore, we can absorb the $O_t(1/n)$ term in the $O_t(\sqrt{c/n})$ term in our expression for ψ, giving $\psi = \sqrt{\epsilon} + O_t(\sqrt{c/n})$. Combining this with (18) and (17) implies the following.

Proposition 13.
$$(1 - 2\sqrt{\epsilon} - O_t(\sqrt{c/n}))c/(n)_t \leq \mathbb{E}[h_t^2] \leq (1 + O_t(1/n))c/(n)_t.$$

$\square$

Our next step is to obtain a lower bound on $\mathbb{E}[h_t^4]$, using (15). To this end, we prove the following.

Lemma 14. *Let $C > 0$ and let $\eta, \theta \in [0, 1]$ such that $\eta \leq C\theta$. Let $F : [0, 1] \to \{0, 1\}$ be a Lebesgue measurable function with $\mathbb{E}[F] = \theta$, and let $h : [0, 1] \to \mathbb{R}$ be a Lebesgue measurable function such that $\mathbb{E}[(h - F)^2] \leq \eta$. Then*
$$\mathbb{E}[h^4] \geq \theta - 4(1 + C)\sqrt{\eta\theta}.$$

Proof. We will solve the following optimization problem.

Problem P. Among all measurable functions $h : [0, 1] \to \mathbb{R}$ such that $\mathbb{E}[(h - F)^2] \leq \eta$, find the minimum value of $\mathbb{E}[h^4]$.

Let $A = F^{-1}(\{1\})$, and let $B = F^{-1}(\{0\})$; then A and B are measurable sets, and $\lambda(A) = \theta$. Observe that if $h : [0, 1] \to \mathbb{R}$ is feasible for P, then the function
$$\tilde{h}(x) = \begin{cases} \frac{1}{\theta} \int_A h(x) \, dx & \text{if } x \in A, \\ \frac{1}{1-\theta} \int_B h(x) \, dx & \text{if } x \in B, \end{cases}$$

20

obtained by averaging h first over A and then over B, is also feasible. Indeed,

$$\mathbb{E}[(h-F)^2] = \int_A (h(x)-1)^2 \, dx + \int_B h(x)^2 \, dx$$

$$\geq \frac{1}{\theta} \left(\int_A (h(x)-1) \, dx \right)^2 + \frac{1}{1-\theta} \left(\int_B h(x) \, dx \right)^2$$

$$= \theta \left(\frac{1}{\theta} \int_A h(x) \, dx - 1 \right)^2 + (1-\theta) \left(\frac{1}{1-\theta} \int_B h(x) \, dx \right)^2$$

$$= \mathbb{E}[(\tilde{h}-F)^2],$$

by the Cauchy–Schwarz inequality. Moreover, we have

$$\mathbb{E}[h^4] = \int_A h(x)^4 \, dx + \int_B h(x)^4 \, dx$$

$$= \theta \cdot \frac{1}{\theta} \int_A h(x)^4 \, dx + (1-\theta) \cdot \frac{1}{1-\theta} \int_B h(x)^4 \, dx$$

$$\geq \theta \left(\frac{1}{\theta} \int_A h(x) \, dx \right)^4 + (1-\theta) \left(\frac{1}{1-\theta} \int_B h(x) \, dx \right)^4$$

$$= \mathbb{E}[\tilde{h}^4],$$

by the convexity of $y \mapsto y^4$. Hence, replacing h with $\tilde{h}$ if necessary, we may assume that h is constant on A and on B. In other words, we may assume that h has the following form:

$$h(x) = \begin{cases} r & \text{if } x \in A, \\ s & \text{if } x \in B. \end{cases}$$

Therefore, P is equivalent to the following problem:

Problem Q:

$$\text{Minimize} \quad \theta r^4 + (1-\theta)s^4$$

$$\text{subject to} \quad \theta(r-1)^2 + (1-\theta)s^2 \leq \eta.$$

Clearly $|r-1| \leq \sqrt{\eta/\theta}$ and so the minimum is at least $\theta(1-\sqrt{\eta/\theta})^4$. Conversely, $r = 1 - \sqrt{\eta/\theta}$, $s = 0$ is a feasible solution to Q, and so the minimum is obtained at

$$\theta(1-\sqrt{\eta/\theta})^4 \geq \theta - 4\sqrt{\eta\theta} - 4\theta\sqrt{\eta/\theta} \cdot (\eta/\theta) \geq \theta - 4(1+C)\sqrt{\eta\theta}. \qquad \square$$

We apply Lemma 14 to functions on the discrete probability space $(S_n, \mathbb{P})$ by considering them as step functions on $[0,1]$. Applying it with $h = h_t$, $F = f$, $\theta = c/(n)_t$, $\eta = \psi^2 c/(n)_t$ and $C = \psi^2$, we obtain

$$\mathbb{E}[h_t^4] \geq (1 - 4(1+\psi^2)\psi)c/(n)_t. \tag{20}$$

The next lemma will be used to relate $\mathbb{E}[h_t^2]$ to $\sum_{T \in \mathcal{C}(n,t)} b_T^2$.

Lemma 15. *Suppose $g \in V_t$ is real-valued, let $b_T = (n)_t \langle g, 1_T \rangle$ for each $T \in \mathcal{C}(n,t)$, and define $h \in V_t$ by*

$$h = Mg = \sum_{T \in \mathcal{C}(n,t)} b_T 1_T.$$

Then

$$\mathbb{E}[h^2] = (1 + O_t(1/n)) \frac{1}{(n)_t} \sum_{T \in \mathcal{C}(n,t)} b_T^2.$$

Proof. Let us abbreviate $\sum_{T \in \mathcal{C}(n,t)}$ to $\sum_T$. We expand as follows:

$$\mathbb{E}[h^2] = \sum_T b_T^2 \mathbb{E}[1_T] + \sum_{(S,T):\, S \neq T} b_S b_T \mathbb{E}[1_S 1_T] = \frac{1}{(n)_t} \sum_T b_T^2 + \sum_{\substack{(S,T):\, S \neq T \\ S \cap T \neq \emptyset}} b_S b_T \mathbb{E}[1_{S \cap T}].$$

We must prove that

$$\sum_{\substack{(S,T):\, S \neq T \\ S \cap T \neq \emptyset}} b_S b_T \mathbb{E}[1_{S \cap T}] = O_t(1/(n)_{t+1}) \sum_T b_T^2. \tag{21}$$

We pause to define some more notation. We define

$$\mathcal{D}(n,t) := \{\{(x_1, y_1), \ldots, (x_t, y_t)\} : (x_1, \ldots, x_t) \in ([n])_t,\ (y_1, \ldots, y_t) \in ([n])_t\};$$

note that there is a natural one-to-one correspondence between $\mathcal{D}(n,t)$ and $\mathcal{C}(n,t)$, the set of t-cosets, given by identifying the t-coset

$$\{\sigma \in S_n : \sigma(x_i) = y_i\ \forall i \in [t]\} \in \mathcal{C}(n,t)$$

with the set of ordered pairs

$$\{(x_1, y_1), \ldots, (x_t, y_t)\} \in \mathcal{D}(n,t).$$

We denote this correspondence by $\leftrightarrow$. If $T \leftrightarrow I$, we define $b_I := b_T$.

We say that two sets $\{(u_1, v_1), \ldots, (u_s, v_s)\} \in \mathcal{D}(n,s)$, $\{(x_1, y_1), \ldots, (x_t, y_t)\} \in \mathcal{D}(n,t)$ are *compatible* if the corresponding cosets have nonempty intersection, i.e. if $u_i = x_j \Leftrightarrow v_i = y_j$. Otherwise, we say that they are *incompatible*.

We say that the two sets are *independent* if $u_i \neq x_j$ and $v_i \neq y_j$ for all $i \in [s]$, $j \in [t]$. Observe that if $A, B \in \mathcal{D}(n,t)$ are compatible, then we may express

$$A = I \cup J,\ B = I \cup K,$$

where $I \in \mathcal{D}(n,e)$ for some $e \leq t$, $J, K \in \mathcal{D}(n, t-e)$, and I, J and K are pairwise independent.

We pause to observe some crucial linear dependencies between the b_T's. Firstly, we claim that

$$\sum_{k \neq x_1, \ldots, x_{t-1}} b_{\{(x_1, y_1), \ldots, (x_{t-1}, y_{t-1}), (k, y_t)\}} = 0 \tag{22}$$

22

for any distinct $x_1, \ldots, x_{t-1} \in [n]$, and any distinct $y_1, \ldots, y_t \in [n]$. Indeed, the left-hand side is precisely $(n)_t \langle g, 1_S \rangle$, where

$$\mathcal{C}(n, t-1) \ni S = \{\sigma \in S_n : \sigma(x_i) = y_i \ \forall i \in [t-1]\} \leftrightarrow \{(x_1, y_1), \ldots, (x_{t-1}, y_{t-1})\}.$$

Since $g \in V_t \perp U_{t-1} = \mathrm{Span}\{1_S : S \in \mathcal{C}(n, t-1)\}$, the claim follows.

Similarly, we claim that

$$\sum_{k \neq y_1, \ldots, y_{t-1}} b_{\{(x_1, y_1), \ldots, (x_{t-1}, y_{t-1}), (x_t, k)\}} = 0, \tag{23}$$

for any distinct $x_1, \ldots, x_t \in [n]$, and any distinct $y_1, \ldots, y_{t-1} \in [n]$. Indeed, the left-hand side is precisely $(n)_t \langle g, 1_S \rangle$, where S is as above.

We split up the sum in (21) as follows:

$$\sum_{\substack{(S,T): \ S \neq T, \\ S \cap T \neq \emptyset}} b_S b_T \mathbb{E}[1_{S \cap T}] = \sum_{e=0}^{t-1} \frac{1}{(n)_{2t-e}} \sum_{\substack{I \in \mathcal{D}(n,e), \ J,K \in \mathcal{D}(n,t-e), \\ I,J,K \ \text{independent}}} b_{I \cup J} b_{I \cup K}.$$

We now use an argument similar to the indicator-function proof of inclusion-exclusion. Observe that for any $I \in \mathcal{D}(n, e)$, we have

$$\sum_{\substack{J,K \in \mathcal{D}(n,t-e): \\ I,J,K \ \text{indep.}}} b_{I \cup J} b_{I \cup K} = \sum_{\substack{J,K \in \mathcal{D}(n,t-e): \\ I,K \ \text{indep.}, \ I,J \ \text{indep.}}} b_{I \cup J} b_{I \cup K} \prod_{p,q \in [t-e], d \in [2]} (1 - 1\{j_p^{(d)} = k_q^{(d)}\}),$$

where

$$J = \{(j_1^{(1)}, j_1^{(2)}), (j_2^{(1)}, j_2^{(2)}), \ldots, (j_{t-e}^{(1)}, j_{t-e}^{(2)})\},$$
$$K = \{(k_1^{(1)}, k_1^{(2)}), (k_2^{(1)}, k_2^{(2)}), \ldots, (k_{t-e}^{(1)}, k_{t-e}^{(2)})\}.$$

Consider what happens when we multiply out the product on the right-hand side, *à la* inclusion-exclusion. We obtain a (± 1)-linear combination of products of indicators of the form

$$\prod_{(p,q,d) \in \mathcal{I}} 1\{j_p^{(d)} = k_q^{(d)}\}.$$

For each such (non-zero) product of indicators, let us replace $k_q^{(d)}$ by $j_p^{(d)}$ in the sum whenever the product contains the indicator $1\{j_p^{(d)} = k_q^{(d)}\}$. We obtain a (± 1)-linear combination of terms of the form

$$\sum_{\substack{J,K': \\ I,J \ \text{indep}, \ I,K' \ \text{indep.}}} b_{I \cup J} b_{I \cup K'},$$

where K' is obtained from K by replacing $k_q^{(d)}$ with $j_p^{(d)}$ for various (p, q, d). If K' contains any (unreplaced) $k_q^{(d)}$, then the corresponding term is zero, by applying (22)

23

or (23) with $k = k_q^{(d)}$. Hence, the only non-zero terms are where all the $k_q^{(d)}$'s have been replaced, i.e. they are of the form

$$\sum_{\substack{J \in \mathcal{D}(n,t-e): \\ I,J \text{ indep.}}} b_{I \cup J} b_{I \cup \{(j_1^{(1)}, j_{\sigma(1)}^{(2)}), \ldots, (j_{t-e}^{(1)}, j_{\sigma(t-e)}^{(2)})\}}, \tag{24}$$

where $\sigma \in S_{t-e}$. We now sum each such term over I, and apply the Cauchy–Schwarz inequality:

$$\left| \sum_{\substack{I \in \mathcal{D}(n,e), \ J \in \mathcal{D}(n,t-e): \\ I,J \text{ indep.}}} b_{I \cup J} b_{I \cup \{(j_1^{(1)}, j_{\sigma(1)}^{(2)}), \ldots, (j_{t-e}^{(1)}, j_{\sigma(t-e)}^{(2)})\}} \right| \leq \binom{t}{e} \sum_T b_T^2.$$

Here, the factor of $\binom{t}{e}$ comes from the fact that there are $\binom{t}{e}$ ways of producing an ordered partition (I, J) of a t-set into an e-set I and a $(t-e)$-set J.

There are at most $O_t(1)$ terms of the form (24) for each $e \in [t-1]$, so we obtain

$$\left| \sum_{e=0}^{t-1} \frac{1}{(n)_{2t-e}} \sum_{\substack{I \in \mathcal{D}(n,e), \ J,K \in \mathcal{D}(n,t-e), \\ I,J,K \text{ independent}}} b_{I \cup J} b_{I \cup K} \right| \leq O_t(1) \frac{1}{(n)_{t+1}} \sum_{T \in \mathcal{C}(n,t)} b_T^2,$$

as required. $\qquad \square$

The next lemma will be used to relate $\mathbb{E}[h_t^4]$ to $\sum_{T \in \mathcal{C}(n,t)} b_T^4$.

Lemma 16. *Suppose $g \in V_t$ is real-valued, let $b_T = (n)_t \langle g, 1_T \rangle$ for each $T \in \mathcal{C}(n,t)$, and define $h \in V_t$ by*

$$h = Mg = \sum_{T \in \mathcal{C}(n,t)} b_T 1_T.$$

Then

$$\mathbb{E}[h^4] = \frac{1}{(n)_t} \left(\sum_{T \in \mathcal{C}(n,t)} b_T^4 + O_t(1/n) \left(\sum_{T \in \mathcal{C}(n,t)} b_T^2 \right)^2 \right).$$

Proof. We need one more piece of notation. If X is a finite set, and $Z \subset X^2$ is a set of ordered pairs of elements of X, we let $Z(d)$ denote the set of all elements of X which occur in the dth coordinate of an ordered pair in Z, for $d \in \{1, 2\}$.

We expand as follows:

$$\mathbb{E}[h^4] = \sum_T b_T^4 \mathbb{E}[1_T] + \sum_{(T_1,T_2,T_3,T_4) \text{ not all equal}} b_{T_1} b_{T_2} b_{T_3} b_{T_4} \mathbb{E}[1_{T_1} 1_{T_2} 1_{T_3} 1_{T_4}]$$

$$= \frac{1}{(n)_t} \sum_T b_T^4 + \sum_{\substack{(T_1,T_2,T_3,T_4) \text{ not all equal}: \\ T_1 \cap T_2 \cap T_3 \cap T_4 \neq \emptyset}} b_{T_1} b_{T_2} b_{T_3} b_{T_4} \mathbb{E}[1_{T_1 \cap T_2 \cap T_3 \cap T_4}].$$

24

We must prove that

$$\sum_{\substack{(T_1,T_2,T_3,T_4) \text{ not all equal:} \\ T_1 \cap T_2 \cap T_3 \cap T_4 \neq \emptyset}} b_{T_1} b_{T_2} b_{T_3} b_{T_4} \mathbb{E}[1_{T_1 \cap T_2 \cap T_3 \cap T_4}] = O_t(1/(n)_{t+1}) \left(\sum_T b_T^2 \right)^2.$$

We now split up the sum above, by partitioning the set

$$\{(T_1, T_2, T_3, T_4) \text{ not all equal} : T_1 \cap T_2 \cap T_3 \cap T_4 \neq \emptyset\}$$

into a bounded number of classes. We define an equivalence relation $\sim$ on this set by

$$(T_1', T_2', T_3', T_4') \sim (T_1, T_2, T_3, T_4) \quad \text{iff there exist } \sigma, \pi \in S_n : \sigma T_r' \pi = T_r \ \forall r \in [4].$$

Note that the number of equivalence classes is at most $\binom{4t}{t}^4 = O_t(1)$, since we can choose a representative from each equivalence class which is a 4-tuple of t-cosets, each of which is the pointwise stabiliser of some t-element subset of $[4t]$.

Note also that $\mathbb{E}[1_{T_1 \cap T_2 \cap T_3 \cap T_4}]$ depends only upon the equivalence class of (T_1, T_2, T_3, T_4), and that we always have $\mathbb{E}[1_{T_1 \cap T_2 \cap T_3 \cap T_4}] = 1/(n)_l$, where $t+1 \leq l \leq 4t$, and therefore we always have $\mathbb{E}[1_{T_1 \cap T_2 \cap T_3 \cap T_4}] \leq 1/(n)_{t+1}$.

Let $\mathcal{S}$ be an equivalence class; we wish to bound

$$\left| \sum_{(T_1,T_2,T_3,T_4) \in \mathcal{S}} b_{T_1} b_{T_2} b_{T_3} b_{T_4} \mathbb{E}[1_{T_1 \cap T_2 \cap T_3 \cap T_4}] \right|.$$

Since we always have $\mathbb{E}[1_{T_1 \cap T_2 \cap T_3 \cap T_4}] \leq 1/(n)_{t+1}$, we have

$$\left| \sum_{(T_1,T_2,T_3,T_4) \in \mathcal{S}} b_{T_1} b_{T_2} b_{T_3} b_{T_4} \mathbb{E}[1_{T_1 \cap T_2 \cap T_3 \cap T_4}] \right| \leq \frac{1}{(n)_{t+1}} \left| \sum_{(T_1,T_2,T_3,T_4) \in \mathcal{S}} b_{T_1} b_{T_2} b_{T_3} b_{T_4} \right|.$$
(25)

Take any $(T_1, T_2, T_3, T_4) \in \mathcal{S}$, and let $T_r \leftrightarrow Z_r \in \mathcal{D}(n, t)$ for each $r \in [4]$. Since $T_1 \cap T_2 \cap T_3 \cap T_4 \neq \emptyset$, Z_1, Z_2, Z_3, Z_4 are pairwise compatible, so for any $r, s \in [4]$, if $Z_r = \{(u_i, v_i)\}_{i \in [t]}$ and $Z_s = \{(x_i, y_i)\}_{i \in [t]}$, then $u_i = x_j \Leftrightarrow v_i = y_j$ for all $i, j \in [t]$. It follows that $|\cup_{r=1}^4 Z_r(1)| = |\cup_{r=1}^4 Z_r(2)|$; we denote this number by N. Note that $N \leq 4t$.

Observe that we may write

$$\sum_{(T_1,T_2,T_3,T_4) \in \mathcal{S}} b_{T_1} b_{T_2} b_{T_3} b_{T_4} = \sum_{\substack{(i_1,\ldots,i_N) \text{ distinct,} \\ (i_{N+1},\ldots,i_{2N}) \text{ distinct}}} \prod_{r=1}^4 b_{\{(i_p,i_q)\}_{(p,q) \in A_r}}$$
(26)

where $A_r \subset [N] \times \{N+1, \ldots, 2N\}$ with $|A_r| = |A_r(1)| = |A_r(2)| = t$, for each $r \in [4]$. (The conditions $|A_r(1)| = |A_r(2)| = t$ say that any $p \in [2N]$ occurs in at most one ordered pair in A_r.)

For each $p \in [2N]$, we shall say that p is *'good'* if it appears in at least two of the A_r's; otherwise, we say that p is *'bad'*. Our aim is to use indicator functions (*à*

25

la inclusion-exclusion), together with the linear dependence relations (22) and (23), to rewrite (26) as a (± 1)-linear combination of at most $O_t(1)$ sums containing no 'bad' indices. Without loss of generality, we may assume that there is a bad index in $\{N+1, \ldots, 2N\}$, say $2N$. (The argument for a bad index in $[N]$ is identical, except that (22) is used instead of (23).) Without loss of generality, we may assume that $2N$ appears in A_4 alone. Let

$$P = \{p \in \{N+1, \ldots, 2N\} : p \text{ appears in } A_4\}, \quad Q = \{N+1, \ldots, 2N\} \setminus P.$$

Then

$$\sum_{\substack{(i_1, \ldots, i_N) \text{ distinct,} \\ (i_{N+1}, \ldots, i_{2N}) \text{ distinct}}} \prod_{r=1}^{4} b_{\{(i_p, i_q)\}_{(p,q) \in A_r}}$$

$$= \sum_{\substack{(i_1, \ldots, i_N) \text{ distinct,} \\ (i_{N+1}, \ldots, i_{2N-1}) \text{ distinct,} \\ i_{2N} : i_{2N} \neq i_p \ \forall p \in P}} \prod_{r=1}^{4} b_{\{(i_p, i_q)\}_{(p,q) \in A_r}} \prod_{q \in Q} (1 - 1\{i_{2N} = i_q\}).$$

Consider what happens when we expand out the product over all $q \in Q$. We obtain

$$\sum_{\substack{(i_1, \ldots, i_N) \text{ distinct,} \\ (i_{N+1}, \ldots, i_{2N-1}) \text{ distinct,} \\ i_{2N} : i_{2N} \neq i_p \ \forall p \in P}} \prod_{r=1}^{4} b_{\{(i_p, i_q)\}_{(p,q) \in A_r}}$$

$$- \sum_{q \in Q} \sum_{\substack{(i_1, \ldots, i_N) \text{ distinct,} \\ (i_{N+1}, \ldots, i_{2N-1}) \text{ distinct}}} b_{\{(i_p, i_q)\}_{(p,q) \in A_4(2N \to q)}} \prod_{r=1}^{3} b_{\{(i_p, i_q)\}_{(p,q) \in A_r}},$$

where $A_4(2N \to q)$ is produced from A_4 by replacing $2N$ with q. We have

$$\sum_{\substack{(i_1, \ldots, i_N) \text{ distinct,} \\ (i_{N+1}, \ldots, i_{2N-1}) \text{ distinct,} \\ i_{2N} : i_{2N} \neq i_p \ \forall p \in P}} \prod_{r=1}^{4} b_{\{(i_p, i_q)\}_{(p,q) \in A_r}} = 0,$$

by (23), so the first term above is zero; none of the other terms involve the 'bad' index $2N$. Hence, we have

$$\sum_{\substack{(i_1, \ldots, i_N) \text{ distinct,} \\ (i_{N+1}, \ldots, i_{2N}) \text{ distinct}}} \prod_{r=1}^{4} b_{\{(i_p, i_q)\}_{(p,q) \in A_r}}$$

$$= - \sum_{q \in Q} \sum_{\substack{(i_1, \ldots, i_N) \text{ distinct,} \\ (i_{N+1}, \ldots, i_{2N-1}) \text{ distinct}}} b_{\{(i_p, i_q)\}_{(p,q) \in A_4(2N \to q)}} \prod_{r=1}^{3} b_{\{(i_p, i_q)\}_{(p,q) \in A_r}}.$$

We now have $|Q|$ sums to deal with, but each has one less bad index than the original sum. By repeating this process until there are no more bad indices remaining in any sum, we can express the original sum (26) as a (± 1)-linear combination of at most $O_t(1)$ terms of the form:

$$\sum_{(i_1,\ldots,i_K)\in\mathcal{R}} \prod_{r=1}^{4} b_{\{(i_p,i_q)\}_{(p,q)\in\tilde{A}_r}}, \tag{27}$$

where

- $K \le 2N \le 8t$,

- $\mathcal{R} \subset [n]^K$ is a subset defined by constraints of the form $i_k \ne i_l$,

- For each $r \in [4]$, we have $\tilde{A}_r \subset [K]^2$, and each $k \in [K]$ occurs in at most one ordered pair in $\tilde{A}_r$,

- Each index $k \in [K]$ is 'good', meaning that it appears in at least two of the $\tilde{A}_r$'s.

Note that the $\tilde{A}_r$'s vary with the term we are looking at. Crudely, (27) is at most

$$\sum_{(i_1,\ldots,i_K)\in[n]^K} \prod_{r=1}^{4} |b_{\{(i_p,i_q)\}_{(p,q)\in\tilde{A}_r}}|, \tag{28}$$

where we define $b_{\{(i_k,j_k)\}_{k\in[t]}} = 0$ if $\{(i_k,j_k)\}_{k\in[t]} \notin \mathcal{D}(n,t)$, i.e. if $i_k = i_l$ or $j_k = j_l$ for some $k \ne l$. To bound (28), we need the following claim.

Claim 4. *Let X be a finite set, let $m \ge 2$, and let $a_1,\ldots,a_m, L \in \mathbb{N}$. Suppose $f_j\colon X^{a_j} \to \mathbb{R}$ for $j = 1, 2, \ldots, m$. Suppose $\sigma_j\colon [a_j] \to [L]$ are injections, such that each $\ell \in [L]$ lies in the image of at least two of these injections. Then*

$$\left(\sum_{i_1,\ldots,i_L\in X} \prod_{j=1}^{m} f_j(i_{\sigma_j(1)},\ldots,i_{\sigma_j(a_j)}) \right)^2 \le \prod_{j=1}^{m} \left(\sum_{i_1,\ldots,i_{a_j}\in X} f_j(i_1,\ldots,i_{a_j})^2 \right).$$

Proof. This follows immediately from [18, Lemma 3.2], which is a weighted version of Shearer's entropy lemma. (We apply the latter with $r = m$ and $t = 2$ to the hypergraph (V, E) where $V = [L] \times X$ and $E = \{\{(1,x_1),(2,x_2),\ldots,(L,x_L)\} \ : \ x_1,\ldots,x_L \in X\}$, with $F_j = \sigma_j([a_j]) \times X$ for each $j \in [m]$ and with

$$w_j(\{(l,x_l) : l \in \sigma_j([a_j])\}) = f_j(x_{\sigma_j(1)},\ldots,x_{\sigma_j(a_j)})$$

for each $j \in [m]$.) $\qquad\square$

We shall apply Claim 4 to the function $F\colon [n]^{2t} \to \mathbb{R}$ defined by

$$F(i_1,\ldots,i_t,j_1,\ldots,j_t) = \begin{cases} b_{\{(i_k,j_k)\}_{k\in[t]}} & \text{if } i_1,\ldots,i_t \text{ are all distinct} \\ & \text{and } j_1,\ldots,j_t \text{ are all distinct,} \\ 0 & \text{otherwise.} \end{cases}$$

For each $r \in [4]$, choose any ordering $((p_1, q_1), \ldots, (p_t, q_t))$ of the ordered pairs in $\tilde{A}_r$, and define the injection $\sigma_r \colon [2t] \to [K]$ by

$$\sigma_r(w) = \begin{cases} p_w & \text{if } 1 \le w \le t, \\ q_{w-t} & \text{if } t+1 \le w \le 2t. \end{cases}$$

Applying Claim 4 with $m = 4$, $L = K$, $X = [n]$, $a_j = 2t$ for each $j \in [4]$, $f_j = F$ for each $j \in [4]$, and $\sigma_1, \sigma_2, \sigma_3, \sigma_4$ as above, yields

$$\left| \sum_{(i_1, \ldots, i_K) \in [n]^K} \prod_{r=1}^{4} |b_{\{(i_p, i_q)\}_{(p,q) \in A_r}}| \right| \le \left(t! \sum_T b_T^2 \right)^2.$$

Here, the factor of $t!$ comes from the fact that we are summing over all possible orderings of the ordered pairs in each $Z \in \mathcal{D}(n, t)$. Since there are only $O_t(1)$ terms of the form (27), we have

$$\sum_{(T_1, T_2, T_3, T_4) \in \mathcal{S}} b_{T_1} b_{T_2} b_{T_3} b_{T_4} \le O_t(1) \left(\sum_T b_T^2 \right)^2.$$

Using (25), together with the fact that there are only $O_t(1)$ different equivalent classes, yields

$$\sum_{\substack{(T_1, T_2, T_3, T_4) \text{ not all equal:} \\ T_1 \cap T_2 \cap T_3 \cap T_4 \ne \emptyset}} b_{T_1} b_{T_2} b_{T_3} b_{T_4} \mathbb{E}[1_{T_1 \cap T_2 \cap T_3 \cap T_4}] \le O_t(1/(n)_{t+1}) \left(\sum_T b_T^2 \right)^2,$$

completing the proof of Lemma 16. $\qquad \square$

Combining Proposition 13 and Lemma 15 shows that $\sum_T b_T^2$ is close to c:

Proposition 17.

$$(1 - 2\psi - O_t(1/n))c \le \sum_T b_T^2 \le (1 + O_t(1/n))c,$$

where $\psi = \sqrt{\epsilon} + O_t(\sqrt{c/n})$.

$\qquad \square$

Similarly, combining Lemma 16 and (20) yields

$$\sum_T b_T^4 + O_t(1/n) \left(\sum_T b_T^2 \right)^2 \ge (1 - 4(1 + \psi^2)\psi)c.$$

Using the fact that $\sum_T b_T^2 = O_t(c)$ yields

$$\sum_T b_T^4 \ge (1 - 4(1 + \psi^2)\psi)c - O_t(c^2/n) = (1 - 4(1 + \psi^2)\psi - O_t(c/n))c = (1 - \psi')c, \quad (29)$$

28

where $\psi' := 4(1+\psi^2)\psi + O_t(c/n)$. Since $\epsilon \le 1/2$ and $c = O_t(\sqrt{n})$, we have $\psi = O_t(1)$, and so
$$\psi' = O_t(\psi + c/n) = O_t(\sqrt{\epsilon} + \sqrt{c/n}).$$

By (12) and Proposition 17, we have
$$\sum_T b_T^4 \le (1 + O_t(c/\sqrt{n})) \sum_T b_T^2$$
$$\le (1 + O_t(c/\sqrt{n}))(1 + O_t(1/n))c$$
$$= (1 + O_t(c/\sqrt{n}))c,$$

using the fact, from (19), that $c = \Omega_t(1)$. Combining this with (29) shows that $\sum_T b_T^4$ is close to c:

Proposition 18.
$$(1 - \psi')c \le \sum_T b_T^4 \le (1 + O_t(c/\sqrt{n}))c,$$

where $\psi' = O_t(\sqrt{\epsilon} + \sqrt{c/n})$.

$\square$

Let $x_1, \ldots, x_N$ denote the entries $(b_T^2)_{T \in \mathcal{C}(n,t)}$ in non-increasing order. By Proposition 17, we have
$$\sum_{k=1}^{N} x_k \le (1 + O_t(1/n))c, \tag{30}$$

and by Proposition 18, we have
$$(1 - \psi')c \le \sum_{k=1}^{N} x_k^2 \le (1 + O_t(c/\sqrt{n}))c. \tag{31}$$

Subtracting (31) from (30) yields:
$$\sum_{k=1}^{N} x_k(1 - x_k) \le (\psi' + O_t(1/n))c = \psi''c, \tag{32}$$

where $\psi'' := \psi' + O_t(1/n) = O_t(\sqrt{\epsilon} + \sqrt{c/n})$. By an appropriate choice of the implicit constant in the definition of ψ'', we may ensure that
$$\sum_{k=1}^{N} x_k \ge (1 - \psi'')c,$$

by Proposition 17.

Let m be the largest integer k such that $x_k \ge 1/2$ (recall that the x_k are arranged in non-increasing order). Then
$$\sum_{k=m+1}^{N} x_k \le 2 \sum_{k=m+1}^{N} x_k(1 - x_k) \le 2\psi''c,$$

29

by (32). Therefore,

$$(1 - 3\psi'')c \leq \sum_{k=1}^{m} x_k \leq (1 + O_t(c/\sqrt{n}))m. \tag{33}$$

On the other hand, we have

$$\sum_{k=1}^{m} (1 - x_k) \leq 2 \sum_{k=1}^{m} x_k(1 - x_k) \leq 2\psi''c.$$

Rearranging, we have

$$\sum_{k=1}^{m} x_k \geq m - 2\psi''c. \tag{34}$$

Since $2x_k - 1 \leq x_k^2$, we have

$$(1 + O_t(c/\sqrt{n}))c \geq \sum_{k=1}^{m} x_k^2 \geq 2 \sum_{k=1}^{m} x_k - m \geq m - 4\psi''c. \tag{35}$$

Equation (35) shows that $m \leq (1 + O_t(\sqrt{\epsilon} + c/\sqrt{n}))c$. Since $\epsilon \leq 1/2$ and $c = O_t(\sqrt{n})$, this implies that $m = O_t(c)$. Combining (33) and (35) now yields

$$|m - c| \leq O_t(\epsilon^{1/2} + c/\sqrt{n})c. \tag{36}$$

Let $T_1, \ldots, T_m$ be the t-cosets corresponding to $x_1, \ldots, x_m$, and let

$$\mathcal{C}' = \bigcup_{k=1}^{m} T_k$$

denote the corresponding union of m t-cosets of S_n. Provided we choose c_t to be sufficiently small depending on t, our assumption $c \leq c_t\sqrt{n}$ implies, via (12), that $b_T > -1/\sqrt{2}$ for all t-cosets T. Hence, $b_T \geq 0$ whenever $b_T^2 \geq 1/2$. We have

$$\sum_{k=1}^{m} |\mathcal{A} \cap T_k|/(n - t)! = \sum_{k=1}^{m} a_{T_k}$$

$$\geq \sum_{k=1}^{m} b_{T_k} - mO_t(c/\sqrt{n})$$

$$\geq \frac{1}{1 + O_t(c/\sqrt{n})} \sum_{k=1}^{m} b_{T_k}^2 - O_t(c^2/\sqrt{n})$$

$$\geq (1 - 3\psi'' - O_t(c/\sqrt{n}))c$$

$$\geq (1 - O_t(\epsilon^{1/2} + c/\sqrt{n}))c,$$

using (11), (12) and (33). Since $|T_i \cap T_j| \leq (n - t - 1)!$ for each $i \neq j$, we have

$$|\mathcal{A} \cap \mathcal{C}'| \geq \sum_{k=1}^{m} |\mathcal{A} \cap T_k| - \binom{m}{2}(n - t - 1)! \geq (1 - O_t(\epsilon^{1/2} + c/\sqrt{n}))c(n - t)!,$$

30

i.e. $\mathcal{A}$ contains almost all of $\mathcal{C}'$. Equation (36) shows that

$$|\mathcal{C}'| \leq m(n-t)! \leq (1 + O_t(\epsilon^{1/2} + c/\sqrt{n}))c(n-t)!.$$

Since $|\mathcal{A}| = c(n-t)!$, we must have

$$|\mathcal{A}\triangle\mathcal{C}'| = |\mathcal{A}| + |\mathcal{C}'| - 2|\mathcal{A}\cap\mathcal{C}'| = O_t(\epsilon^{1/2} + c/\sqrt{n})c(n-t)!.$$

Crudely, we have

$$|m - \mathrm{round}(c)| \leq 2|m-c| = O_t(\epsilon^{1/2} + c/\sqrt{n})c.$$

By adding or deleting $|m - \mathrm{round}(c)|$ t-cosets to or from $\mathcal{C}'$, we may produce a family $\mathcal{C} \subset S_n$ which is a union of $\mathrm{round}(c)$ t-cosets, and satisfies

$$
\begin{aligned}
|\mathcal{A}\triangle\mathcal{C}| &\leq |\mathcal{A}\triangle\mathcal{C}'| + |\mathcal{C}'\triangle\mathcal{C}| \\
&= O_t(\epsilon^{1/2} + c/\sqrt{n})c(n-t)!.
\end{aligned}
$$

Since $|c - \mathrm{round}(c)| \leq |m-c| = O_t(\epsilon^{1/2} + c/\sqrt{n})c$, this completes the proof of Theorem 3.

5 An isoperimetric inequality for the transposition graph

In this section, we will apply Theorem 3 to obtain an isoperimetric inequality for S_n. We first give some background and notation on discrete isoperimetric inequalities.

Isoperimetric problems are of ancient interest in mathematics. In general, they ask for the smallest possible size of the 'boundary' of a set of a given 'size'. Discrete isoperimetric inequalities deal with discrete notions of boundary in graphs. There are two different notions of boundary in graphs, the vertex-boundary and the edge-boundary; here, we deal with the latter.

If $G = (V, E)$ is any graph, and $S, T \subset V$, we write $E_G(S, T)$ for the set of edges of G between S and T, and we write $e_G(S, T) = |E_G(S, T)|$. We write $\partial_G S = E_G(S, S^c)$ for the set of edges of G between S and $S^c := V \setminus S$; this is called the *edge-boundary of S in G*. An *edge-isoperimetric inequality for G* gives a lower bound on the minimum size of the edge-boundary of a set of size k, for each integer k.

The *transposition graph* T_n is the Cayley graph on S_n generated by the transpositions in S_n; equivalently, two permutations are joined if, as sequences, one can be obtained from the other by transposing two elements. In this section, we are concerned with the edge-isoperimetric problem for T_n.

It would be of great interest to prove an isoperimetric inequality for the transposition graph which is sharp for all set-sizes. Ben Efraim [3] conjectures that initial segments of the lexicographic order on S_n have the smallest edge-boundary of all sets of the same size. (The lexicographic order on S_n is defined as follows: if $\sigma, \pi \in S_n$, we say that $\sigma < \pi$ if $\sigma(j) < \pi(j)$, where $j = \min\{i \in [n] : \sigma(i) \neq \pi(i)\}$. The *initial segment of size k of the lexicographic order on S_n* simply means the smallest k elements of S_n in the lexicographic order.)

Conjecture 19 (Ben Efraim). *For any $\mathcal{A} \subset S_n$, $|\partial \mathcal{A}| \geq |\partial \mathcal{C}|$, where $\mathcal{C}$ denotes the initial segment of the lexicographic order on S_n of size $|\mathcal{A}|$.*

This is a beautiful conjecture; it may be compared to the edge-isoperimetric inequality in $\{0,1\}^n$, due to Harper [20], Lindsey [28], Bernstein [4] and Hart [21], stating that among all subsets of $\{0,1\}^n$ of size k, the first k elements of the binary ordering on $\{0,1\}^n$ has the smallest edge boundary. (Recall that if $x, y \in \{0,1\}^n$, we say that $x < y$ *in the binary ordering* if $x_j = 0$ and $y_j = 1$, where $j = \min\{i \in [n] \colon x_i \neq y_i\}$.)

In this section, we use eigenvalue techniques to prove an approximate version of Conjecture 19; this version is asymptotically sharp for sets of size $n!/\mathrm{poly}(n)$. We then combine eigenvalue techniques with Theorem 3 to prove the exact conjecture for sets of size $(n-t)!$, for n sufficiently large depending on t.

For sets of size $c(n-1)!$, where $c \in [n]$, Conjecture 19 follows from calculating the second eigenvalue of the Laplacian of T_n. We briefly outline the argument.

If $G = (V, E)$ is a finite graph and A is the adjacency matrix of G, the *Laplacian matrix* of G may be defined by

$$L = D - A, \tag{37}$$

where D is the diagonal $|V| \times |V|$ matrix with rows and columns indexed by V, and with

$$D_{u,v} = \begin{cases} \deg(v) & \text{if } u = v \\ 0 & \text{if } u \neq v. \end{cases}$$

The following well-known theorem, due independently to Dodziuk [10] and Alon and Milman [2], provides an edge-isoperimetric inequality for a graph G in terms of the smallest non-trivial eigenvalue of its Laplacian matrix (the 'spectral gap'):

Theorem 20 (Dodziuk / Alon–Milman). *Let $G = (V, E)$ be a finite graph, let L denote the Laplacian matrix of G, and let $0 = \mu_1 \leq \mu_2 \leq \ldots \leq \mu_{|V|}$ be the eigenvalues of L (repeated with their multiplicities). If $S \subset V(G)$, then*

$$e(S, S^c) \geq \mu_2 \frac{|S||S^c|}{|V|}.$$

If equality holds, then the characteristic vector 1_S of S satisfies

$$1_S - \frac{|S|}{|V|}\mathbf{f} \in \ker(L - \mu_2 I), \tag{38}$$

where $\mathbf{f}$ denotes the all-1's vector.

The proof follows straightforwardly from the formula

$$e(S, S^c) = |V|\langle 1_S, L 1_S \rangle,$$

where $\langle f, g \rangle := \frac{1}{|V|} \sum_{v \in V} f(v)\overline{g(v)}$.

If G is a finite d-regular graph (and note that T_n is $\binom{n}{2}$-regular), then (37) becomes $L = dI - A$. Therefore, if the eigenvalues of A are

$$d = \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{|V|},$$

32

then $\mu_i = d - \lambda_i$ for each i. In particular, $\mu_2 = d - \lambda_2$, so for any set $S \subset V(G)$,

$$e(S, S^c) \geq (d - \lambda_2)\frac{|S\|S^c|}{|V|}.$$

The transposition graph is a normal Cayley graph, and therefore the eigenvalues of its adjacency matrix are given by (9). Frobenius gave the following formula for the value of χ_α at a transposition, where $\alpha = (\alpha_1, \ldots, \alpha_m)$ is a partition of n.

$$\chi_\alpha((1\ 2)) = \frac{\dim(\rho_\alpha)}{\binom{n}{2}}\frac{1}{2}\sum_{j=1}^{m}((\alpha_j - j)(\alpha_j - j + 1) - j(j-1)) \quad (\alpha \vdash n).$$

Combining this with (9) yields the following formula for the eigenvalues of the adjacency matrix of the transposition graph.

$$\lambda_\alpha = \frac{1}{2}\sum_{j=1}^{m}((\alpha_j - j)(\alpha_j - j + 1) - j(j-1)) \quad (\alpha \vdash n). \tag{39}$$

Note that $\lambda_{(n)} = \binom{n}{2}$, $\lambda_{(n-1,1)} = \binom{n}{2} - n$ and $\lambda_{(n-2,2)} = \binom{n}{2} - 2n + 2$. Diaconis and Shashahani [8] verify that if α and α' are two partitions of n, then

$$\alpha \trianglerighteq \alpha' \quad \Rightarrow \quad \lambda_\alpha \geq \lambda_{\alpha'}. \tag{40}$$

Hence, if $n \geq 2$, then $\mu_2 = n$, and the μ_2-eigenspace of the Laplacian is precisely $U_{(n-1,1)}$. Theorem 20 therefore yields the following.

Theorem 21 (essentially due to Diaconis and Shahshahani). *If $\mathcal{A} \subset S_n$, then*

$$|\partial\mathcal{A}| \geq \frac{|\mathcal{A}|(n! - |\mathcal{A}|)}{(n-1)!}.$$

Equality holds only if

$$1_\mathcal{A} \in U_{(n)} \oplus U_{(n-1,1)}. \tag{41}$$

Observe that equality holds in Theorem 21 if and only if $\mathcal{A}$ is a disjoint union of 1-cosets of S_n. (The 'only if' part follows from (41) and Theorem 2.)

When $|\mathcal{A}| = o(n!)$, Theorem 21 merely implies $|\partial\mathcal{A}| \geq (1 - o(1))n|\mathcal{A}|$, whereas Conjecture 19 would imply that $|\partial\mathcal{A}| \geq (1-o(1))n(t+1)|\mathcal{A}|$ whenever $|\mathcal{A}| = o((n-t)!)$. Our first aim is to prove the latter when t is small; specifically, we prove the following.

Theorem 22. *Let $\mathcal{A} \subset S_n$ with $|\mathcal{A}| \leq (n-t+1)!$. Then*

$$|\partial\mathcal{A}| \geq (1 - \tilde{C}_t|\mathcal{A}|/(n-t+1)!)t(n-t+1)|\mathcal{A}|,$$

where $\tilde{C}_t > 0$ depends upon t alone.

Proof. Let $f = 1_\mathcal{A}$ and let L be the Laplacian of T_n. By Lemma 11, we have

$$\|f_{t-1}\|_2^2 \leq C'_t n^{t-1}(|\mathcal{A}|/n!)^2 = O_t(|\mathcal{A}|/(n-t+1)!)|\mathcal{A}|/n!.$$

Therefore, we have

$$\sum_{\substack{\alpha \vdash n: \\ \alpha_1 \leq n-t}} \|f_\alpha\|_2^2 = \|f\|_2^2 - \|f_{t-1}\|_2^2 \geq (1 - O_t(|\mathcal{A}|/(n-t+1)!))|\mathcal{A}|/n!.$$

Note that if $\alpha_1 \leq n-t$, then $\alpha \trianglelefteq (n-t,t)$, and therefore $\lambda_\alpha \leq \lambda_{(n-t,t)}$, by (40). By (39), we have

$$\lambda_{(n-t,t)} = \binom{n}{2} - t(n-t+1),$$

so $\lambda_\alpha \leq \binom{n}{2} - t(n-t+1)$ whenever $\alpha_1 \leq n-t$. In other words, $\mu_\alpha \geq t(n-t+1)$ whenever $\alpha_1 \leq n-t$. We have

$$\begin{aligned}
|\partial \mathcal{A}| &= n!\langle f, Lf \rangle \\
&= n! \sum_{\alpha \vdash n} \mu_\alpha \|f_\alpha\|_2^2 \\
&\geq n! \sum_{\alpha_1 \leq n-t} \mu_\alpha \|f_\alpha\|_2^2 \\
&\geq n! t(n-t+1)(1 - O_t(|\mathcal{A}|/(n-t+1)!))|\mathcal{A}|/n! \\
&\geq (1 - O_t(|\mathcal{A}|/(n-t+1)!))t(n-t+1)|\mathcal{A}|,
\end{aligned}$$

as required. $\qquad\square$

Note that when t is fixed and $\mathcal{A} \subset S_n$ with $|\mathcal{A}| = o((n-t+1)!)$, Theorem 22 yields

$$|\partial \mathcal{A}| \geq (1 - o(1))tn|\mathcal{A}|.$$

If

$$\mathcal{A} = \{\sigma \in S_n : \sigma(i) = i \ \forall i \in [t-1], \ \sigma(t) \in \{t, \ldots, t+m-1\}\},$$

then $|\mathcal{A}| = m(n-t)!$ and

$$\begin{aligned}
|\partial \mathcal{A}| &= (t(n-t) + t(t-1)/2 - m + 1)|\mathcal{A}| \\
&= (t(n - (t+1)/2) - m + 1)|\mathcal{A}|.
\end{aligned}$$

If $m = o(n)$ then for constant t we obtain

$$|\partial \mathcal{A}| = (1 - o(1))tn|\mathcal{A}|,$$

showing that Theorem 22 is asymptotically sharp in this case.

We now use a similar argument, together with Theorem 3, to prove Theorem 4, thus verifying Conjecture 19 for sets of size $(n-t)!$, when n is sufficiently large depending on t.

Proof of Theorem 4. Let $\mathcal{A} \subset S_n$ with $|\mathcal{A}| = (n-t)!$. Note that

$$|\partial T_{(1,2,\ldots,t)(1,2,\ldots,t)}| = t(n - (t+1)/2)(n-t)!.$$

34

Assume that $|\partial\mathcal{A}| \le |\partial T_{(1,2,\ldots,t)(1,2,\ldots,t)}|$; we will show that $\mathcal{A}$ must be a t-coset.

Let $f = 1_{\mathcal{A}}$; then $\|f\|_2^2 = 1/(n)_t$. Write

$$\|f - f_t\|_2^2 = \phi\|f\|_2^2 = \phi|\mathcal{A}|/n!,$$

where $\phi \in [0,1]$. Our first aim is to show that ϕ must be small.

By Lemma 11, we have

$$\|f_{t-1}\|_2^2 \le C_t' n^{t-1}(|\mathcal{A}|/n!)^2 = O_t(1/n)|\mathcal{A}|/n!.$$

Therefore, we have

$$\sum_{\substack{\alpha \vdash n:\\ \alpha_1 = n-t}} \|f_\alpha\|_2^2 = \|f_t\|_2^2 - \|f_{t-1}\|_2^2 = (1 - \phi - O_t(1/n))|\mathcal{A}|/n!.$$

Writing L for the Laplacian of T_n, we have

$$
\begin{aligned}
|\partial\mathcal{A}| &= n!\langle f, Lf\rangle\\
&= n!\sum_{\alpha \vdash n}\mu_\alpha\|f_\alpha\|_2^2\\
&\ge n!\sum_{\alpha_1 = n-t}\mu_\alpha\|f_\alpha\|_2^2 + n!\sum_{\alpha_1 \le n-t-1}\mu_\alpha\|f_\alpha\|_2^2\\
&\ge n!t(n-t+1)(1 - \phi - O_t(1/n))|\mathcal{A}|/n! + n!(t+1)(n-t)\phi|\mathcal{A}|/n!\\
&= [t(n-t+1)(1 - \phi - O_t(1/n)) + (t+1)(n-t)\phi](n-t)!\\
&= (tn + \phi n - O_t(1))(n-t)!.
\end{aligned}
$$

Since we are assuming that $|\partial\mathcal{A}| \le t((n-(t+1)/2)(n-t)!$, it follows that $\phi = O_t(1/n)$. Hence, we may apply Theorem 3 (provided n is sufficiently large depending on t) with $\epsilon = O_t(1/n)$. We see that there exists a t-coset $\mathcal{C} \subset S_n$ such that

$$|\mathcal{A}\triangle\mathcal{C}| = O_t(1/\sqrt{n})(n-t)!.$$

Our aim is to show that $\mathcal{A} = \mathcal{C}$. Let us write

$$|\mathcal{A}\setminus\mathcal{C}| = \psi(n-t)!, \quad \psi = O_t(1/\sqrt{n}).$$

Let $\mathcal{E} = \mathcal{A}\setminus\mathcal{C}$, and let $\mathcal{M} = \mathcal{C}\setminus\mathcal{A}$; then

$$|\mathcal{E}| = |\mathcal{M}| = \psi(n-t)!.$$

Let $\mathcal{X} = \mathcal{A}\cap\mathcal{C}$ and $\mathcal{Y} = S_n\setminus(\mathcal{A}\cup\mathcal{C})$, so that $\mathcal{A} = \mathcal{E}\cup\mathcal{X}$ and $\mathcal{C} = \mathcal{M}\cup\mathcal{X}$.

Observe that

$$
\begin{aligned}
|\partial\mathcal{A}| &= e(\mathcal{A},\mathcal{Y}) + e(\mathcal{A},\mathcal{M})\\
&= [e(\mathcal{C},\mathcal{Y}) + e(\mathcal{E},\mathcal{Y}) - e(\mathcal{M},\mathcal{Y})] \ + \ [e(\mathcal{E},\mathcal{M}) + e(\mathcal{X},\mathcal{M})]\\
&= [e(\mathcal{C},\mathcal{Y}) + e(\mathcal{C},\mathcal{E})] + [e(\mathcal{C},\mathcal{E}) + e(\mathcal{E},\mathcal{Y})] - 2e(\mathcal{C},\mathcal{E})\\
&\quad - [e(\mathcal{M},\mathcal{Y}) + e(\mathcal{M},\mathcal{E})] + 2e(\mathcal{M},\mathcal{E}) + e(\mathcal{X},\mathcal{M})\\
&= |\partial\mathcal{C}| + |\partial\mathcal{E}| - 2e(\mathcal{C},\mathcal{E}) - e(\mathcal{M}, S_n\setminus\mathcal{C}) + 2e(\mathcal{E},\mathcal{M}) + e(\mathcal{M},\mathcal{X}).
\end{aligned}
$$

In particular,

$$|\partial\mathcal{A}| \geq |\partial\mathcal{C}| + |\partial\mathcal{E}| - 2e(\mathcal{C},\mathcal{E}) - e(\mathcal{M}, S_n \setminus \mathcal{C}) + e(\mathcal{M}, \mathcal{X}). \qquad (42)$$

We will show that $|\partial\mathcal{E}| - 2e(\mathcal{C},\mathcal{E}) - e(\mathcal{M}, S_n \setminus \mathcal{C}) + e(\mathcal{M}, \mathcal{X}) > 0$ unless $\psi = 0$. Let us bound each of these terms in turn. First, we bound $|\partial\mathcal{E}|$. Since $|\mathcal{E}| \leq (n-t)!$, Theorem 22 yields

$$|\partial\mathcal{E}| \geq (1 - O_t(1/n))t(n-t+1)|\mathcal{E}| = (1 - O_t(1/n))tn\psi(n-t)!.$$

Next, we bound $e(\mathcal{C},\mathcal{E})$. By definition, we have $\mathcal{E}\cap\mathcal{C} = \emptyset$. Without loss of generality, we may assume that $\mathcal{C} = \{\sigma \in S_n : \sigma(i) = i \ \forall i \in [t]\}$. For any $\sigma \in S_n \setminus \mathcal{C}$, choose $i \in [t]$ such that $\sigma(i) \neq i$. If $\sigma(i \ \sigma^{-1}(i)) \in \mathcal{C}$, then it is the unique neighbour of σ in $\mathcal{C}$; otherwise, σ has no neighbour in $\mathcal{C}$. It follows that

$$e(\mathcal{E},\mathcal{C}) \leq |\mathcal{E}| = \psi(n-t)!.$$

Next, we calculate $e(\mathcal{M}, S_n \setminus \mathcal{C})$. Observe that each $\sigma \in \mathcal{C}$ has exactly $t(n-t) + t(t-1)/2$ neighbours in $S_n \setminus \mathcal{C}$. Indeed, the neighbours of σ in $S_n \setminus \mathcal{C}$ are precisely $\{\sigma(i \ j) : i \in [t], j \notin [t]\} \cup \{\sigma(i \ j) : i, j \in [t], j \neq i\}$. It follows that

$$e(\mathcal{M}, S_n \setminus \mathcal{C}) = t(n - (t+1)/2)|\mathcal{M}| \leq tn|\mathcal{M}|.$$

Finally, we bound $e(\mathcal{M}, \mathcal{X})$. Since $T_n[\mathcal{C}]$ is isomorphic to T_{n-t}, applying Theorem 21 to T_{n-t} yields:

$$e(\mathcal{C} \setminus \mathcal{A}, \mathcal{A} \cap \mathcal{C}) \geq \psi(n-t)!(1-\psi)(n-t)!/(n-t-1)! = \psi(1-\psi)(n-t)(n-t)!.$$

Subsituting all of these bounds into (42), we obtain:

$$\begin{aligned}
|\partial\mathcal{A}| &\geq |\partial\mathcal{C}| + (1 - O_t(1/n))tn\psi(n-t)! \\
&\quad - 2\psi(n-t)! - tn\psi(n-t)! + \psi(1-\psi)(n-t)(n-t)! \\
&= |\partial\mathcal{C}| + \big[(1 - O_t(1/n))tn - 2 - tn + (1-\psi)(n-t)\big]\psi(n-t)! \\
&= |\partial\mathcal{C}| + [(1 - O_t(1/\sqrt{n}))n - O_t(1)]\psi(n-t)!.
\end{aligned}$$

If $\psi > 0$ and n is sufficiently large depending on t, then the right-hand side is greater than $|\partial\mathcal{C}|$, contradicting our assumption that $|\partial\mathcal{A}| \leq |\partial\mathcal{C}|$. It follows that $\psi = 0$ and therefore $\mathcal{A} = \mathcal{C}$, proving Theorem 4. $\qquad \square$

6   A note on t-intersecting families of permutations

In [15], we showed how a $t = 1$ version of Theorem 3 could be used to give a natural proof of the Cameron–Ku conjecture on large intersecting families of permutations. Let us briefly give the background to this result. Recall that a family of permutations $\mathcal{A} \subset S_n$ is said to be *intersecting* if for any two permutations $\sigma, \pi \in \mathcal{A}$, there exists $i \in [n]$ such that $\sigma(i) = \pi(i)$. Using a simple partioning argument, Deza and Frankl proved that if $\mathcal{A} \subset S_n$ is intersecting, then $|\mathcal{A}| \leq (n-1)!$, i.e. the 1-cosets are intersecting families of maximum size. Cameron and Ku [6] proved that equality holds only if $\mathcal{A}$ is a 1-coset of S_n. They made the following 'stability' conjecture.

Conjecture 23 (Cameron–Ku, 2003). *There exists $c > 0$ such that if $\mathcal{A} \subset S_n$ is an intersecting family with $|\mathcal{A}| \geq c(n-1)!$, then $\mathcal{A}$ is contained within a 1-coset.*

Conjecture 23 was first proved by the first author in [11], but the proof in [13] is, in a sense, more natural.

We say that a family of permutations $\mathcal{A} \subset S_n$ is *t-intersecting* if for any $\sigma, \pi \in \mathcal{A}$, there exist $i_1, \ldots, i_t \in [n]$ such that $\sigma(i_k) = \pi(i_k)$ for all $k \in [t]$. In 1977, Deza and Frankl [7] conjectured that if n is sufficiently large depending on t, then any t-intersecting family in S_n has size at most $(n-t)!$. This is proved in [15]. In [12], the first author proved the following analogue of Conjecture 23 for t-intersecting families of permutations.

Theorem 24. *For any $t \in \mathbb{N}$, there exists $c < 1$ and $n_0 \in \mathbb{N}$ such that if $n \geq n_0$, then any t-intersecting family $\mathcal{A} \subset S_n$ with $|\mathcal{A}| \geq c(n-t)!$ is contained within a t-coset.*

Almost exactly as in the $t = 1$ case, Theorem 3 can be used to give a more 'natural' proof of Theorem 24. We do not give the details here; it suffices to say that one simply replaces the adjacency matrix of the derangement graph (used in the $t = 1$ case) with the 'weighted' analogue constructed in [15].

We note that Theorem 24 implies that if n is sufficiently large depending on t, then any t-intersecting family $\mathcal{A} \subset S_n$ with $|\mathcal{A}| = (n-t)!$ must be a t-coset, i.e. the equality case of the Deza-Frankl conjecture. In [15], this is deduced from [15, Theorem 27], but as stated on page 4 (in the Introduction) of the current paper, that theorem is false. As mentioned in the Introduction, one may deduce the equality case of the Deza-Frankl conjecture from the Hilton-Milner type result in [12]. Alternatively, of course, one may also deduce it from Theorem 24, using Theorem 3 as a 'black box' to prove Theorem 24. While this alternative proof is arguably more natural than the proof in [12], it is not 'truly' shorter, as the proof of Theorem 3 in the current paper is longer than the proof of the Hilton-Milner type result in [12].

7 Conclusion and open problems

We conjecture the following strengthening of Theorem 3.

Conjecture 25. *Let $\mathcal{A} \subset S_n$, and let $t \in \mathbb{N}$. Let f denote the characteristic function of $\mathcal{A}$, and let f_t denote the orthogonal projection of f onto U_t. If*

$$\mathbb{E}[(f - f_t)^2] \leq \epsilon \mathbb{E}[f],$$

then there exists $\mathcal{C} \subset S_n$ such that $\mathcal{C}$ is a union of t-cosets, and

$$|\mathcal{A} \triangle \mathcal{C}| \leq C_0(\epsilon + 1/n)|\mathcal{A}|,$$

where C_0 is an absolute constant.

We remark that the $1/n$ term would be best possible up to an absolute constant factor. For $t = 2$, this can be seen by considering the family

$$\mathcal{A} = \{\sigma \in S_n : \ |\sigma([3]) \cap [3]| = 0 \text{ or } 3\}.$$

It is easy to check that $f := 1_{\mathcal{A}} \in U_2$, that

$$\mathbb{E}[f] = 1 - \frac{9(n-3)}{n(n-1)} = 1 - O(1/n),$$

and that $|\mathcal{A} \Delta \mathcal{C}| = \Omega(1/n)|\mathcal{A}|$ for any family $\mathcal{C} \subset S_n$ such that $\mathcal{C}$ is a union of 2-cosets of S_n.

The ϵ-term would also be best possible up to an absolute constant factor, as can be seen by considering a family of the form

$$\mathcal{A} = (T_{(1,2,\ldots,t)(1,2,\ldots,t)} \cup \mathcal{E}) \setminus \mathcal{F}$$

where $\epsilon \leq 1/2$, $\epsilon(n-t)!/2 \in \mathbb{N}$, $\mathcal{F} \subset T_{(1,2,\ldots,t)(1,2,\ldots,t)}$ and $\mathcal{E} \subset S_n \setminus T_{(1,2,\ldots,t)(1,2,\ldots,t)}$ with $|\mathcal{E}| = |\mathcal{F}| = \epsilon(n-t)!/2$. (It is easy to check that $f = 1_{\mathcal{A}}$ satisfies $\mathbb{E}[(f - f_t)^2] \leq \epsilon\mathbb{E}[f]$, and that $|\mathcal{A} \triangle \mathcal{C}| \geq \epsilon|\mathcal{A}|$ for any $\mathcal{C} \subset S_n$ such that $\mathcal{C}$ is a union of t-cosets.)

In section 5, we gave an example of how Fourier-analytic arguments can yield sharp isoperimetric inequalities for relatively small sets, even when the classical 'eigenvalue gap' inequality in Theorem 20 is far from sharp. We speculate that this technique of combining eigenvalue arguments, Fourier-analytic arguments and stability arguments may be useful in obtaining other isoperimetric inequalities. Conjecture 19 remains very much open, however; we suspect that more combinatorial techniques are required to prove it in full generality.

It is instructive to contrast the situation for Boolean functions on $\{0,1\}^n$, with the situation for the symmetric group. For Boolean functions on $\{0,1\}^n$, the analogues of the t-cosets are the subcubes of codimension t, i.e. sets of the form $\{x \in \{0,1\}^n : x_i = c_i \ \forall i \in S\}$, where $S \in [n]^{(t)}$ and $c_i \in \{0,1\}$ for each $i \in S$. If $f : \{0,1\}^n \to \{0,1\}$, we write $\hat{f}$ for the Fourier transform of f, i.e.

$$\hat{f} : \mathcal{P}([n]) \to \mathbb{R}; \quad \hat{f}(S) = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} (-1)^{\sum_{i \in S} x_i} \quad \forall S \subset [n].$$

We equip $\mathbb{R}[\{0,1\}^n]$ with the inner product $\langle f, g \rangle = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x)g(x)$, and we let $\| \cdot \|_2$ denote the corresponding Euclidean norm. It is easy to see that if $f : \{0,1\}^n \to \{0,1\}$, then the square of the Euclidean distance from f to the subspace of $\mathbb{R}[\{0,1\}^n]$ spanned by the subcubes of codimension t is precisely $\sum_{S \subset [n]: \ |S| > t} \hat{f}(S)^2$, i.e. the ℓ^2-mass of the Fourier transform of f on sets of size greater than t.

It was proved in [17] that if $f : \{0,1\}^n \to \{0,1\}$ has Fourier transform with at most ϵ of its ℓ^2-mass on sets of size greater than t (i.e., f has Euclidean distance at most $\sqrt{\epsilon}$ from the subspace of $\mathbb{R}[\{0,1\}^n]$ spanned by the subcubes of codimension 1), then f is $O(\epsilon)$-close to the characteristic function of a subcube of codimension 1, or to a constant function. By contrast, for each even $t \geq 2$ and each $n \geq t+1$, the Boolean function

$$f : \{0,1\}^n \to \{0,1\}; \quad x \mapsto 1\{|x \cap [t+1] = 0 \text{ or } t+1\} \tag{43}$$

has Fourier transform supported on sets of size at most t, but has $\|f - g\|_2^2 \geq 2^{-t}$ for any $g : \{0,1\}^n \to \{0,1\}$ such that g is the characteristic function of a union of subcubes of codimension t. In addition, the function f in (43) has the same expectation

as the characteristic function of a subcube of codimension t. Theorem 3 rules out the existence of an analogous function on the symmetric group, for large n.

We can, however, say the following about Boolean functions on $\{0,1\}^n$. Kindler and Safra proved in [25] that if $f : \{0,1\}^n \to \{0,1\}$ has Fourier transform with at most ϵ of its ℓ^2-mass on sets of size greater than t, then f must be $O_t(\epsilon)$-close to a Boolean function $g : \{0,1\}^n \to \{0,1\}$ depending upon at most $j(t)$ coordinates, for some function $j : \mathbb{N} \to \mathbb{N}$. Combined with [31, Corollary 6.2], and an easy compactness argument, this yields the following.

Corollary 26. *There exists $s : \mathbb{N} \to \mathbb{N}$ such that the following holds. If $f : \{0,1\}^n \to \{0,1\}$ is such that $\sum_{S \subset [n]:|S|>t} \hat{f}(S)^2 \leq \epsilon$, then there exists a function $h : \{0,1\}^n \to \{0,1\}$ such that $\|f - h\|_2^2 = O_t(\epsilon)$ and h is the characteristic function of a union of subcubes of codimension $s(t)$. Moreover, we may take $s(t) \leq t^3$.*

The function f in (43) shows that one cannot take $s(t) = t$ in Corollary 26, even if f has the same expectation as a subcube of codimension t. Theorem 3 provides an S_n-analogue of Corollary 26 with $s(t) = t$, albeit with $O_t(\sqrt{\epsilon} + c/\sqrt{n})c/(n)_t$ in place of $O_t(\epsilon)$. Conjecture 25 would provide such an analogue with $s(t) = t$, albeit with $O(\epsilon + 1/n)c/(n)_t$ in place of $O_t(\epsilon)$.

Acknowledgements

References

[1] N. Alon, I. Dinur, E. Friedgut and B. Sudakov, Graph products, Fourier analysis and spectral techniques, *Geometric and Functional Analysis* 14 (2004), 913–940.

[2] N. Alon and V. D. Milman, λ_1, isoperimetric inequalities for graphs, and super-concentrators, *Journal of Combinatorial Theory, Series B* 38 (1985), 73–88.

[3] L. Ben Efraim, personal communication.

[4] A. J. Bernstein, Maximally connected arrays on the n-cube, *SIAM Journal on Applied Mathematics* 15 (1967), 1485–1489.

[5] J. Bourgain, On the distribution of the Fourier spectrum of Boolean functions, *Israel Journal of Mathematics* 131 (2002), 269–276.

[6] P.J. Cameron and C.Y. Ku, Intersecting Families of Permutations, *European Journal of Combinatorics* 24 (2003) pp. 881–890.

[7] M. Deza and P. Frankl, On the maximum number of permutations with given maximal or minimal distance, *Journal of Combinatorial Theory, Series A* 22 (1977), 352–360.

[8] P. Diaconis and M. Shahshahani, Generating a random permutation with random transpositions, *Z. Wahrsch. Verw. Gebeite* 57 (1981), 159–179.

[9] I. Dinur, S. Knot, G. Kindler, D. Minzer and M. Safra, On non-optimally expanding sets in Grassmann graphs, *Electronic Colloquium on Computational Complexity* (2017), Report No. 94.
Preprint available at `https://eccc.weizmann.ac.il/report/2017/094/`.

[10] J. Dodziuk, Difference equations, isoperimetric inequality and transience of certain random walks, *Transactions of the American Mathematical Society* 284 (1984), 787–794.

[11] D. Ellis, A Proof of the Cameron–Ku Conjecture, *Journal of the London Mathematical Society* 85 (2012), 165–190.

[12] D. Ellis, Stability for t-intersecting families of permutations, *Journal of Combinatorial Theory, Series A* 118 (2011), 208–227.

[13] D. Ellis, E. Friedgut and Y. Filmus, A quasi-stability result for dictatorships in S_n, *Combinatorica* 35 (2015), 573–618.

[14] D. Ellis, E. Friedgut and Y. Filmus, A stability result for balanced dictatorships in S_n, *Random Structures and Algorithms* 46 (2015), 494–530.

[15] D. Ellis, E. Friedgut and H. Pilpel, Intersecting Families of Permutations, *Journal of the American Mathematical Society* 24 (2011), 649–682.

[16] P. Erdős, C. Ko and R. Rado, An Intersection Theorem for Systems of Finite Sets, *Quart. J. Math. Oxford, Ser. 2* 12 (1961), 313–320.

[17] E. Friedgut, G. Kalai and A. Naor, Boolean functions whose Fourier transform is concentrated on the first two levels, *Advances in Applied Mathematics* 29 (2002), 427–437.

[18] E. Friedgut, Hypergraphs, Entropy and Inequalities, *American Mathematical Monthly* 111 (2004), 749–760.

[19] Y. Filmus, A comment on 'Intersecting Families of Permutations', manuscript, available at `http://www.cs.toronto.edu/~yuvalf/EFP-comment.pdf`.

[20] L. H. Harper, Optimal assignments of numbers to vertices, *SIAM Journal on Applied Mathematics* 12 (1964), 131–135.

[21] S. Hart, A note on the edges of the n-cube, *Discrete Mathematics* 14 (1976), 157–163.

[22] H. Hatami and M. Ghandehari, Fourier analysis and large independent sets in powers of complete graphs, *Journal of Combinatorial Theory, Series B*, 98 (2008), 164–172.

[23] A.J.W. Hilton and E.C. Milner, Some intersection theorems for systems of finite sets, *Quart. J. Math. Oxford Series 2* 18 (1967), 369–384.

[24] A. J. Hoffman, On eigenvalues and colourings of graphs, *Graph Theory and its Applications*, Academic Press, New York (1970), 79–91.

[25] G. Kindler and S. Safra, Noise resistant Boolean functions are juntas, preprint. `http://www.cs.huji.ac.il/~gkindler/papers/noise-stable-r-juntas.ps`.

[26] G. James and A. Kerber, *The Representation Theory of the Symmetric Group*, Encyclopedia of Mathematics and its Applications 16, Cambridge University Press, Cambridge, 1985.

[27] B. Larose and C. Malvenuto, Stable sets of maximal size in Kneser-type graphs, *European Journal of Combinatorics* 25 (2004), 657–673.

[28] J. H. Lindsey, II, Assignment of numbers to vertices, *American Mathematical Monthly* 71 (1964), 508–516.

[29] J.-P. Serre, *Linear Representations of Finite Groups*, Graduate Texts in Mathematics, Volume 42, Springer-Verlag, New York, 1977.

[30] B. E. Sagan, *The Symmetric Group: Representations, Combinatorial Algorithms and Symmetric Functions*, Springer-Verlag, New York, 1991. [2nd revised printing, 2001.]

[31] A. Tal, Properties and applications of Boolean function composition. In *Proceedings of the 4th Conference on Innovations in Theoretical Computer Science*, ITCS 2013, pp. 441–454, 2013.