How Balanced Can Permutations Be?

Gal Beniamini Nir Lavee Nati Linial *

The Hebrew University of Jerusalem

Abstract

A permutation $\pi \in \mathbb{S}_n$ is *k-balanced* if every permutation of order k occurs in π equally often, through order-isomorphism. In this paper, we explicitly construct k-balanced permutations for $k \leq 3$, and every n that satisfies the necessary divisibility conditions. In contrast, we prove that for $k \geq 4$, no such permutations exist. In fact, we show that in the case $k \geq 4$, every n-element permutation is at least $\Omega_n(n^{k-1})$ far from being k-balanced. This lower bound is matched for $k = 4$, by a construction based on the Erdős-Szekeres permutation.

1 Introduction

A permutation $\tau \in \mathbb{S}_k$ occurs as a *pattern* in $\pi \in \mathbb{S}_n$, if there are indices $1 \leq s_1 < \cdots < s_k \leq n$ such that

$$\forall i, j \in \{1, \ldots, k\} : \ \pi(s_i) < \pi(s_j) \iff \tau(i) < \tau(j).$$

This simple concept of *order-isomorphism* gives rise to many intriguing problems.

The local structure of permutations. The k-profile of an n-element permutation π is the vector that counts the occurrences of every order-k pattern in π. There are still many things we do not know about the set of $k!$-dimensional vectors that arise in this way. The Erdős-Szekeres Theorem [ES35] states that every order-n permutation must contain a monotone pattern of order $\lceil \sqrt{n}\, \rceil$. Conversely, the *packing density* of a pattern $\tau \in \mathbb{S}_k$ is its maximal proportion within the k-profile of an n-element permutation, as $n \to \infty$. Packing densities have received considerable attention [AAH+02, SS18, Wil02], yet even for $\mathbb{S}_4$ some answers remain presently unknown. Patterns in random permutations have also received their share of attention [EZ20, JNZ13], as has the algorithmic problem of computing the k-profile [EZL21, DG20]. Both are relevant to certain basic questions in mathematical statistics. Every one of the above questions also has a graph-theoretic analogue. For instance, the counterpart to Erdős-Szekeres' Theorem is Ramsey's Theorem, and the graph-theoretic equivalent to packing density is *inducibility* [PG75], and so forth.

1.1 Our Contribution

In this paper we consider the permutation-theoretic analogue of *combinatorial designs*. For integers $n \geq k \geq 1$, we say that a permutation $\pi \in \mathbb{S}_n$ is *k-balanced* if every order-k pattern occurs in π equally often, i.e. exactly $\binom{n}{k}/k!$ times. By way of example, $\pi = 2413$ is 2-balanced. So, for which values of n and k does there exist a k-balanced permutation $\pi \in \mathbb{S}_n$? We answer this question fully.

Constructions for $k \leq 3$. It is not hard to see that any k-balanced permutation is $(k-1)$-balanced. Therefore, for an n-element permutation to be k-balanced, we must at least have $r! \mid \binom{n}{r}$, for every $r \leq k$. It is a straightforward exercise to see that these necessary divisibility conditions suffice in the case $k = 2$, that is, 2-balanced permutations exist for every *admissible* n. Is the same true for $k = 3$? Our first result answers this in the positive, resolving an open question of [CP08].

Theorem 1. *For $k \leq 3$ and every n, there exists a k-balanced permutation in $\mathbb{S}_n$ iff n is admissible.*

Our construction for $k = 3$ is explicit and relies heavily on *rotation-invariance* (see Section 3). The divisibility conditions for 3-balanced permutations permit six remainders in $\mathbb{Z}/36\mathbb{Z}$, and we provide infinite families corresponding to each. These families are all based on a single basic construction, and together cover all but a spurious collection of 19 admissible values of n, which we handle individually.

*Supported in part by a NSF-BSF research grant "Global Geometry of Graphs".

Theorem 1 proves the existence of 3-balanced "designs". This has many noteworthy counterparts. For example, a $(r, s; n)$-*Steiner system* is a collection of r-element subsets of $\{1, 2, \ldots, n\}$, such that every s-element subset is contained in the same number of members in the system. Such a system can exist only if n, r and s satisfy certain arithmetic conditions, and a major discovery of Keevash [Kee14] (see also [GKLO16]) says that for fixed $r > s > 1$ and for large enough n, if the above arithmetic conditions are satisfied, then a Steiner system exists. Similarly, in graph theory, Janson and Spencer [JS92] considered *proportional graphs*, in which every subgraph of a fixed size appears the exact number of times as expected. They showed that with respect to order-3 subgraphs, there exist infinitely many proportional graphs. Finally, with regards to permutations, Cooper and Petrarca [CP08] noted the existence of 3-balanced permutations for $n = 9$ (the least admissible n), and our Theorem 1 extends this to *every* admissible n.

Non-existence for $k \geq 4$. For our second result, we prove that there are no 4-balanced permutations. As the k-balanced condition is downwards closed in k, we obtain the following.

Theorem 2. *There are no k-balanced permutations for $k \geq 4$.*

The proof of Theorem 2 follows by establishing a simple polynomial identity relating entries of the r-profile of any permutation, for $r \leq 4$ (see Section 4). The 4-balanced profile violates this identity. This resolves another open question of [CP08], who carried out a large (non-exhaustive) computer search for $n = 64$, the smallest admissible 4-balanced cardinality. This explains why none were found.

Theorem 2 is closely related to a result of Naves, Pikhurko and Scott [NPS18], who proved the non-existence of proportional graphs, in which every order-4 subgraph appears exactly the expected number of times. Their proof similarly relies on a polynomial identity. Our result is also related to quasirandom permutations, i.e., infinite families in which the normalised k-profile converges *asymptotically* to uniform, as n tends to infinity. The theory of graph limits and graphons [Lov12] has been highly influential in graph theory in recent years, and an analogous theory concerning limits of permutations and the notion of permutons has been investigated as well, e.g., [Coo04, HKM+13, HKMS11]. Notions of pseudo-random graphs were introduced by Thomason [Tho87] and a remarkable result of Chung, Graham and Wilson [CGW89] shows that a graph is pseudo-random iff it has the "right" number of 4-cycles. Proving a conjecture of R. Graham (see [Coo04]), Kràl' and Pikhurko [KP13] proved that a permuton is quasirandom iff it is 4-symmetric. Our techniques in proving Theorem 2 differ from those of [KP13], and we point out the difficulty in applying the latter to the discrete setting in Section 4.3.

Minimum distance from k-balanced. If (as we show) k-balanced permutations do not exist for $k \geq 4$, how *close* to balanced can they be? Formally, we define the *distance* of $\pi \in \mathbb{S}_n$ from being k-balanced, to be the ℓ_∞-distance between π's k-profile and the uniform vector $(\binom{n}{k}/k!)\mathbb{1}$. We prove:

Theorem 3. *For $k \geq 4$, the distance of every n-element permutation from being k-balanced is $\Omega_n(n^{k-1})$.*

Our proof of Theorem 3 can be viewed as a robust version of our proof of Theorem 2, using the same polynomial identity (see Section 5). We prove that the bound in Theorem 3 is tight for $k = 4$. That is, we give a construction of permutations that attain this distance. This construction is based on a modification of the well-known Erdős-Szekeres permutation [ES35]. For larger k the tightness of our bound remains open. However, we note that *all* entries in the k-profile of a uniformly random permutation in $\mathbb{S}_n$ are, with probability $> 99\%$ (for large enough n), within distance $\Theta_n(n^{k-1/2})$ from $\binom{n}{k}/k!$ (see Section 5.2.2). So, in the remaining cases, our bound is at most $\mathcal{O}_n(\sqrt{n})$-far from tight.

Relation between profiles and permutations. Our last result is of a slightly different flavour, and is of interest only when $k = k(n)$ grows with n. Given a k-profile we seek properties which are common to all n-element permutations that have this profile.

Theorem 4. *There exists a set of $\widetilde{\Omega}(k^2/n)$ points in the $[n] \times [n]$ grid, such that any two n-element permutations with the same k-profile, coincide in their restriction to this set.*

Our proof of Theorem 4 is established by drawing a connection between polynomials and k-profiles (see Section 6). We introduce a notion of evaluating a bivariate polynomial on a permutation, and show that fixing the k-profile of a permutation, also fixes the evaluation of all bivariate polynomials of degree $< k$ on it. Using results from approximation theory, this allows us to construct a family of low-degree polynomials, whose evaluations express permutation points in terms of the k-profile alone.

Open questions There remain many interesting questions. For instance, is the distance lower bound tight? And for how many patterns simultaneously? We refer the reader to the discussion in Section 7.

2 Preliminaries

As usual, we denote the symmetric group of order n by $\mathbb{S}_n$. By default we write permutations in $\mathbb{S}_n$ in the *one-line notation*, and think of a permutation as a bijection from $[n]$ to itself, where $[n] := \{1, 2, \dots, n\}$. Any finite set of points in the plane, no two of which are axis-aligned, defines a permutation. Let $\mathcal{A} = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\} \subset \mathbb{R}^2$ be a set of points, where $x_1 < \dots < x_n$ and all y_i are distinct. Then, corresponding to $\mathcal{A}$ is the permutation $\sigma \in \mathbb{S}_n$ (denoted $\mathcal{A} \cong \sigma$), where $y_{\sigma^{-1}(1)} < y_{\sigma^{-1}(2)} < \dots < y_{\sigma^{-1}(n)}$.

The *order-isomorphism* of permutations is in the focus of our work.

Definition 2.1 (order-isomorphism). *Let $\pi \in \mathbb{S}_n$ and $\tau \in \mathbb{S}_k$ be permutations, where $k \le n$. Let $S = \{s_1, \dots, s_k\} \subseteq [n]$, where $s_1 < \dots < s_k$. We say that π induced on S is order-isomorphic to τ if:*

$$\forall i, j \in [k] : \ \pi(s_i) < \pi(s_j) \iff \tau(i) < \tau(j)$$

and we denote this condition by $\pi(S) \cong \tau$. When this is the case, we say the pattern τ occurs in π. The number of occurrences of τ in π is denoted $\#\tau\,(\pi)$, where:

$$\#\tau\,(\pi) := \left|\left\{S \in \binom{[n]}{k} \ : \ \pi(S) \cong \tau\right\}\right|$$

Thus, e.g., $\#\mathtt{123}\,(\pi)$ indicates the number of ascending triples in π. We also define:

Definition 2.2 (k-profile of a permutation). *Let $\pi \in \mathbb{S}_n$ be a permutation and let $1 \le k \le n$ be an integer. The k-profile of π is defined as follows:*

$$\mathcal{P}_k(\pi) := \big(\#\tau\,(\pi)\big)_{\tau \in \mathbb{S}_k} \in \mathbb{R}_{\ge 0}^{\mathbb{S}_k}$$

This is a vector of $|\mathbb{S}_k| = k!$ non-negative integers that sum to $\binom{n}{k}$.

This brings us to our main object of study.

Definition 2.3 (k-balanced permutation). *We say that a permutation $\pi \in \mathbb{S}_n$ is k-balanced for some $1 \le k \le n$ if:*

$$\forall \tau \in \mathbb{S}_k : \ \#\tau\,(\pi) = \frac{\binom{n}{k}}{k!}$$

2.1 Basic Observations on Balanced Permutations

We first observe that the k-profile of a permutation uniquely determines its r-profile for every $r < k$, and in particular every k-balanced permutation is also r-balanced.

Proposition 2.4 (Downward induction of pattern distribution). *Let $n > k > r$ be positive integers. If $\pi \in \mathbb{S}_n$ and $\tau \in \mathbb{S}_r$, then*

$$\binom{n-r}{k-r} \cdot \#\tau\,(\pi) = \sum_{\sigma \in \mathbb{S}_k} \#\tau\,(\sigma) \cdot \#\sigma\,(\pi) \tag{1}$$

Proof. TOPROVE 0 □

Corollary 2.4.1 (k-balanced implies ($< k$)-balanced). *For $n > k > r$, every k-balanced permutation $\pi \in \mathbb{S}_n$ is also r-balanced.*

Proof. TOPROVE 1 □

Corollary 2.4.1, yields the following *divisibility conditions.*

Corollary 2.4.2 (divisibility conditions for k-balanced permutations). *If $\pi \in \mathbb{S}_n$ is k-balanced for some $1 \le k \le n$, then $r! \mid \binom{n}{r}$ for all $1 \le r \le k$.*

3 k-Balanced Permutations for $k \leq 3$

In this section we show that for $k = 2$ and $k = 3$, the divisibility conditions of Corollary 2.4.2 are not only necessary, but also *sufficient*. Namely, we show that a k-balanced permutation on n elements exists, whenever n satisfies that arithmetic condition. As a warmup, we first describe the case $k = 2$.

3.1 2-Balanced Family

The following is one way (of many) to construct such an infinite family.

Proposition 3.1 (2-balanced family)**.** *There exists a 2-balanced permutation in $\mathbb{S}_n$ if and only if $n \equiv 0 \pmod 4$ or $n \equiv 1 \pmod 4$.*

Proof. TOPROVE 2 □

The case $k = 2$ is deceptively simple: the plot thickens for larger k, and it is far from the truth that any collection of $k!$ non-negative integers summing up to $\binom{n}{k}$ is a k-profile of some permutation $\pi \in \mathbb{S}_n$. E.g, by the Erdős-Szekeres Theorem [ES35], we cannot have $\#\mathtt{123}\,(\pi) = \#\mathtt{321}\,(\pi) = 0$ whenever $n \geq 5$.

3.2 3-Balanced Family

To construct a 3-balanced family, we take a different approach. Consider the action $D_4 \curvearrowright \mathbb{S}_n$ of the dihedral group on $\mathbb{S}_n$, where we view any permutation $\pi \in \mathbb{S}_n$ as the set of points $\{(i, \pi(i))\}$ in $\mathbb{R}^2$, and act on the square $[1, n]^2$ in the standard way. This group action has the useful property that it respects pattern counts, in the following sense:

Lemma 3.2 (pattern counts under $D_4 \curvearrowright \mathbb{S}_n$)**.** *Let $\pi \in \mathbb{S}_n$ and $\tau \in \mathbb{S}_k$ be two permutations, and let $g \in D_4$. Then, $\#\tau\,(\pi) = \#g.\tau\,(g.\pi)$.*

Proof. TOPROVE 3 □

Consequently, if the permutation $\pi \in \mathbb{S}_n$ is 3-balanced, then so are all the permutations in π's orbit under the action of D_4. This orbit may include at most $|D_4| = 8$ permutations. Conversely, for $n > 1$, the orbit *must* include at least two permutations, since no permutation is identical to its reflections about the horizontal and vertical axes, respectively (they agree on no more than one point).

For any element $g \in D_4$ and permutation $\pi \in \mathbb{S}_n$, we say that π is *g-invariant* if $g.\pi = \pi$ (i.e., g is in the stabilizer of π). As noted, clearly no permutation is invariant to the involutions of the reflections about either axis. However, *rotation-invariant* permutations do exist. In other words, $r.\pi = \pi$ where $r \in D_4$ is the $90°$-rotation of the square (therefore, $g.\pi = \pi$ for all $g \in \langle r \rangle$). Rotation-invariant permutations are simply characterised, as follows.

Proposition 3.3 (characterisation of rotation-invariant permutations)**.** *Let $n > 1$ be even.*[1] *Then,*

1. *There exists a rotation-invariant $\pi \in \mathbb{S}_n$ if and only if $n = 4m$, for some natural m.*
2. *Let $A \sqcup B = [2m]$ with $|A| = |B| = m$, and let $\sigma : A \to B$ be a bijection between A and B. To every such A, B and σ there corresponds a rotation-invariant permutation in $\mathbb{S}_{4m}$. All rotation-invariant permutations in $\mathbb{S}_{4m}$ are generated in this way.*

Proof. TOPROVE 4 □

Consider the orbits in $\mathbb{S}_3$ under the action of the $90°$-rotation $r \in D_4$.

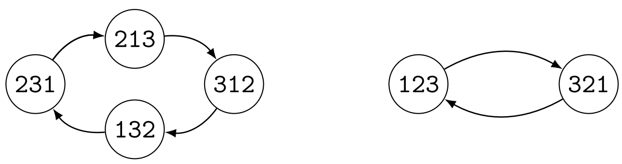

Figure 1: Orbits in $\mathbb{S}_3$ under the action $\langle r \rangle \curvearrowright \mathbb{S}_3$, where $r \in D_4$.

Referring to Figure 1 and Lemma 3.2, yields the following useful fact regarding rotation-invariant permutations.

[1]For odd n, any rotation-invariant permutation *must* include the point at the centre. The permutation induced on the remaining rows and columns is rotation-invariant, to which Proposition 3.3 now applies.

Lemma 3.4 (3-profile of rotation-invariant permutations)**.** *If $\pi \in \mathbb{S}_n$ is rotation-invariant, then:*

$$\#\mathtt{123}\,(\pi) = \#\mathtt{321}\,(\pi) \;\; \textit{and} \;\; \#\mathtt{132}\,(\pi) = \#\mathtt{231}\,(\pi) = \#\mathtt{213}\,(\pi) = \#\mathtt{312}\,(\pi)$$

In particular, π is 3-balanced if and only if $\#\mathtt{123}\,(\pi) = \#\mathtt{132}\,(\pi)$.

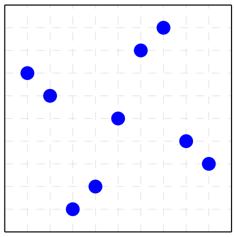

Figure 2: Plot of the 3-balanced permutation $\pi = \mathtt{761258943} \in \mathbb{S}_9$. By enumeration, the two shortest 3-balanced permutations are π and its inverse (see also [CP08]). Both are rotation-invariant.[3]

Before we proceed to describe our construction, we note the arithmetic implications of Corollary 2.4.2 on any 3-balanced permutation.

Lemma 3.5 (divisibility conditions for 3-balanced permutations)**.** *If $\pi \in \mathbb{S}_n$ is 3-balanced, then $n \equiv 0,\ 1,\ 9,\ 20,\ 28,$ or $29 \pmod{36}$.*

3.2.1 A Rotation-Invariant Construction

Lemma 3.4 suggests that we seek rotation-invariant permutations, while Proposition 3.3 provides a recipe for constructing such a permutation in terms of a bipartition and a bijection. By Lemma 3.5, there is a 3-balanced permutation in $\mathbb{S}_n$ only if $n \equiv 0$ or $1 \pmod 4$. For our construction, we fix the bipartition $A \sqcup B$ where $A = \{m+1, \ldots, 2m\}$ and $B = \{1, \ldots, m\}$. The planar diagram of π in $\mathbb{R}^2$ looks as follows.

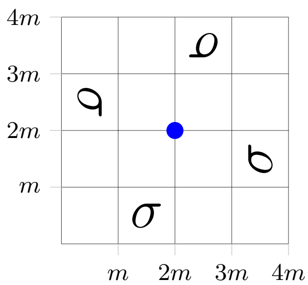

Figure 3: Schematic plot of a rotation-invariant permutation $\pi \in \mathbb{S}_n$, where we fix the bipartition $\{1, \ldots, m\} \sqcup \{m+1, \ldots, 2m\}$ (see Proposition 3.3). When n is odd, we add the blue point at the centre. Here π has the "external structure" of $\mathtt{3142} \in \mathbb{S}_4$.

We now express the pattern counts of π in terms of σ, as follows.

Lemma 3.6. *Let $\sigma \in \mathbb{S}_m$, and let $\pi \in \mathbb{S}_n$ be obtained by rotation as in Figure 3, where $n = 4m$. Then:*

$$\#\mathtt{123}\,(\pi) = 2 \cdot \#\mathtt{123}\,(\sigma) + 2 \cdot \#\mathtt{321}\,(\sigma) + 4m \cdot \#\mathtt{12}\,(\sigma) + 2m \cdot \#\mathtt{21}\,(\sigma)$$
$$\#\mathtt{132}\,(\pi) = \#\mathtt{132}\,(\sigma) + \#\mathtt{231}\,(\sigma) + \#\mathtt{213}\,(\sigma) + \#\mathtt{312}\,(\sigma) + m^3 + m \cdot \#\mathtt{12}\,(\sigma) + 2m \cdot \#\mathtt{21}\,(\sigma)$$

In particular, π is 3-balanced if and only if:

$$3 \cdot \#\mathtt{123}\,(\sigma) + 3 \cdot \#\mathtt{321}\,(\sigma) + 3m \cdot \#\mathtt{12}\,(\sigma) = \binom{m}{3} + m^3 \tag{2}$$

Proof. TOPROVE 5 □

[3] We note that for $n > 9$, there appear to be many 3-balanced permutations which are *not* rotation-invariant. These can be found, for instance, via a random-greedy search.

To construct an infinite 3-balanced family, it suffices to find permutations $\sigma \in \mathbb{S}_m$ that satisfy Equation (2). Initially, let us consider the following construction. Place three identical descending segments, each of length $\ell \geq 1$, in ascending order. As before, the patterns in σ can be counted through case analysis. For example, an ascending pair is formed by choosing two of the three segments, and then one element from each. We obtain:

$$\#12(\sigma) = 3\ell^2, \qquad \#123(\sigma) = \ell^3, \qquad \#321(\sigma) = 3\binom{\ell}{3}$$

These values *nearly* satisfy Equation (2). Indeed, both sides of the equation agree on the cubic and quadratic terms, and disagree *only* on the linear terms. To achieve equality, we amend the construction slightly, by inserting two additional points "in-between" the existing ones (i.e., placing them at non-integer coordinates). For a parameter $\ell < r < 3\ell/2$ to be chosen below, and a small constant $0 < \varepsilon < 1$, the new coordinates are the following (see Figure 4).

$$(x_1, y_1) := (r + 2 + \varepsilon, r + \ell + \varepsilon), \qquad (x_2, y_2) := (r + \ell + \varepsilon, r - \varepsilon)$$

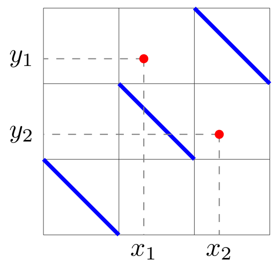

Figure 4: Amending the basic construction of $\sigma \in \mathbb{S}_m$ by inserting two points, and fixing $r = \lceil 4\ell/3 \rceil$.

Theorem 3.7 (3-balanced family)**.** *For every $n \geq 9$, there exists a 3-balanced permutation in $\mathbb{S}_n$ if and only if n satisfies the divisibility conditions. That is, $n \equiv 0,\ 1,\ 9,\ 20,\ 28,$ or $29 \pmod{36}$.*

Proof. TOPROVE 6 □

Remark 3.8. For n that fails the divisibility conditions, this construction still produces *nearly* balanced permutations. In particular, letting $\ell = 3t$ or $\ell = 3t + 2$, and taking $r = \lfloor 4\ell/3 \rfloor + 1$, the discrepancy in Equation (2) is at most ± 2.

4 Non-existence of k-Balanced Permutations for $k \geq 4$

In view of the results in Section 3, one may seek k-balanced permutations for $k > 3$. In this section we show that no such permutations exist. By the monotonicity proven in Corollary 2.4.1, it suffices to show that there exist no 4-balanced permutations.

4.1 Warmup: Ruling out $k(n) \geq \log n + (2 + \varepsilon) \log \log n$

For a permutation $\pi \in \mathbb{S}_n$ to be k-balanced, it clearly must have *at least* $|\mathbb{S}_k|$ k-tuples. By Stirling's approximation of the factorial this yields $k \lesssim e\sqrt{n}$. In fact, more is true: by Corollary 2.4.1 the number of r-tuples in π must be *divisible* by $r!$, for all $r \leq k$. This yields the following (see [CP08] for further discussions of these divisibility conditions).

Proposition 4.1 (ruling out $k \geq \log n + (2 + \varepsilon) \log \log n$)**.** *Let $k = k(n)$ be a function and let $\varepsilon > 0$ be a constant. If there exist $k(n)$-balanced permutations in $\mathbb{S}_n$, then for any sufficiently large n,*

$$k(n) < \log n + (2 + \varepsilon) \log \log n$$

Proof. TOPROVE 7 □

Remark 4.2. If we take $n = (k!)^2$, then $r! \mid \binom{n}{r}$ for all $r \in [k]$. It follows that divisibility alone does not imply $k(n) \leq o(\log n / \log \log n)$, so that Proposition 4.1 is tight up to $\mathcal{O}(\log \log n)$ factor.

4.2 Non-existence of 4-Balanced Permutations

The following simple lemma provides a polynomial identity relating the $\{2,3,4\}$-profiles of any permutation. It is a direct corollary of this lemma that there exist no 4-balanced permutations.

Lemma 4.3. *Every permutation $\pi \in \mathbb{S}_n$ satisfies the following identity:*

$$\begin{aligned}(\#12\,(\pi))^2 = {} & 6\cdot\#1234\,(\pi) + 4\cdot\#1243\,(\pi) + 4\cdot\#1324\,(\pi) + 2\cdot\#1342\,(\pi) \\ & + 2\cdot\#1423\,(\pi) + 4\cdot\#2134\,(\pi) + 4\cdot\#2143\,(\pi) + 2\cdot\#2314\,(\pi) \\ & + 2\cdot\#2413\,(\pi) + 2\cdot\#3124\,(\pi) + 2\cdot\#3142\,(\pi) + 2\cdot\#3412\,(\pi) \\ & + 6\cdot\#123\,(\pi) \quad + 2\cdot\#132\,(\pi) \quad + 2\cdot\#213\,(\pi) \quad + \#12\,(\pi)\end{aligned}$$

Proof. TOPROVE 8 □

The non-existence of 4-balanced permutations now follows directly.

Theorem 4.4. *There are no 4-balanced permutations.*

Proof. TOPROVE 9 □

4.3 Comparison to the Quasirandomness Proof of Kràl' and Pikhurko

In this paper we examine profiles of permutations and the conditions under which they may be balanced. This is somewhat related to *quasirandomness* of permutations, and the *asymptotic* convergence of the k-profile to uniform. Formally,

Definition 4.5 (quasirandom permutations)**.** *Let $\Pi = \{\pi_n\}$ be an infinite family of permutations of non-decreasing order. We say that Π is quasirandom if for every $k > 1$ and every $\tau \in \mathbb{S}_k$, we have:*

$$\frac{\#\tau\,(\pi_n)}{\binom{n}{k}} \to \frac{1}{k!}\ , \text{ as } n \to \infty$$

In an influential paper, [KP13] Kràl' and Pikhurko proved a conjecture of Graham (see [Coo04]), showing that every asymptotically 4-balanced infinite family of permutations is quasirandom. Namely, if Π has the property that $\#\tau\,(\pi_n)/\binom{n}{4} \to \frac{1}{4!}$ (as $n \to \infty$) for every $\tau \in \mathbb{S}_4$, then Π is quasirandom.

Permutons or permutation limits are central to the proof of [KP13] (see also [HKM+13, HKMS11]). A limit object in this framework is a *doubly stochastic* measure. I.e., a measure with *uniform marginals* on the unit square $[0,1]^2$. Any such measure μ gives rise to a sampling process that produces permutations: just pick k points uniformly at random from μ, and consider the corresponding planar pattern. With probability 1 these points define a permutation, since ties in any coordinate occur with zero probability. Kràl' and Pikhurko show that, up to sets of measure zero, the Lebesgue measure λ is the one and only 4-balanced, doubly stochastic measure on $[0,1]^2$.

Reinterpreting the proof of [KP13]. Consider two experiments. In the first, sample a point uniformly from the unit square, then sample two points *independently* from the measure μ. In the second experiment, we first sample a point uniformly from $[0,1]^2$, then *one* point from μ, and another point sampled uniformly from the unit square. In both experiments we consider the event that the first sampled point lies to the top-right of both subsequent points. The success probabilities of these experiments can be expressed in terms of μ and λ's density functions $F, G : [0,1]^2 \to [0,1]$, respectively, on the bottom-left rectangles of the unit square. Concretely, for all $(a,b) \in [0,1]^2$,

$$F(a,b) := \mu([0,a]\times[0,b]), \text{ and } G(a,b) := \lambda([0,a]\times[0,b]) = ab$$

From here the proof of [KP13] proceeds by connecting between F, G and the probabilities of the aforementioned events, and then computing these probabilities using the fact that μ is 4-balanced. Rather than recount the proof, let us slightly delay the exposition and instead proceed directly to a discrete setting, where it is more convenient to provide full details.

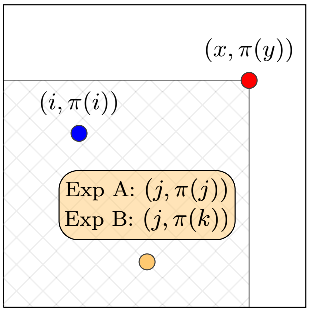

Figure 5: Two experiments. In the first, we sample a uniform grid point and two points from π, and consider the event that both π-points fall to the bottom-left of the initial point. The second experiment starts likewise, but the final point is sampled *uniformly*. In the continuous setting one cannot distinguish the hybrid distribution from the original one, whereas in the discrete setting this can be done.

The discrete setting. To recast [KP13] in a discrete setting, let us replace the measure μ with a permutation $\pi \in \mathbb{S}_n$. In other words, rather than sample from μ, we now sample points uniformly from the set $\{(i,\pi(i))\}_{i\in[n]}$. Instead of the two functions F, G on the unit square defined above, we have two functions on the grid, $u, v : [n]^2 \to [0,1]$, which are defined by:

$$\forall x, y \in [n]^2 : \; v(x,y) := \frac{xy}{n^2}, \text{ and } u(x,y) := \frac{|\{i \le x : \pi(i) \le y\}|}{n}$$

Here v is the bottom-left density function of the uniform doubly stochastic matrix $\frac{1}{n} \cdot \mathbb{1} \otimes \mathbb{1}$, and u is the normalised number of points in the bottom-left rectangles of the permutation matrix associated with π (i.e., the probability that a point chosen at random from π falls in any such rectangle). As before, we would like to compute the probabilities of the events corresponding to the two experiments outlined above. To do so, observe that sampling a uniform point on the grid can be "simulated" by sampling two points independently and uniformly at random from $\{(i,\pi(i))\}_i$, keeping only the x-coordinate from the first point, and the y-coordinate from the second (discarding the two remaining coordinates). Since any permutation has "uniform marginals", this process yields a uniformly random point in the grid $[n]^2$. Consequently, we can now conduct the two experiments, and present their relation to the functions u and v. By total probability, for the first experiment we have:

$$\begin{aligned}
\Pr_{i,j,x,y\sim[n]}[i,j \le x \wedge \pi(i),\pi(j) \le \pi(y)] &= \sum_{s,t\in[n]} \Pr_{x,y\sim[n]}[x = s, \pi(y) = t] \cdot \Pr_{i,j\sim[n]}[i,j \le s \wedge \pi(i),\pi(j) \le t] \\
&= \frac{1}{n^2} \sum_{s,t\in[n]} \Pr_{i\sim[n]}[i \le s \wedge \pi(i) \le t]^2 \\
&= \frac{1}{n^2} \sum_{s,t\in[n]} u(s,t)^2 = \frac{1}{n^2}\|u\|_2^2
\end{aligned}$$

and for the second experiment:

$$\begin{aligned}
\Pr_{i,j,k,x,y\sim[n]}[i,j \le x \wedge \pi(i),\pi(k) \le \pi(y)] &= \sum_{s,t\in[n]} \Pr_{x,y\sim[n]}[x = s \wedge \pi(y) = t] \cdot \Pr_{i,j,k\sim[n]}[i,j \le s \wedge \pi(i),\pi(k) \le t] \\
&= \frac{1}{n^2} \sum_{s,t\in[n]} \Pr_{i\sim[n]}[i \le s \wedge \pi(i) \le t] \cdot \Pr_{j,k\sim[n]}[j \le s \wedge \pi(k) \le t] \\
&= \frac{1}{n^2} \sum_{s,t\in[n]} u(s,t) \cdot v(s,t) = \frac{1}{n^2}\langle u, v\rangle
\end{aligned}$$

Similarly to the proof of Lemma 4.3, these two probabilities can *also* be directly expressed in terms of weighted sums of permutation patterns in π, over at most 5 points (in the continuous setting by applying Cauchy-Schwartz one can make do with only 4-point patterns). However, this is where the proof from the continuous setting no longer carries over to the discrete case.

In the continuous setting, both events can be shown to occur with the same probability, which is precisely $\frac{1}{9}$. Then, since $\|G\|_2^2 = \frac{1}{9}$, it follows that $\langle F, G\rangle^2 = \|F\|_2^2\|G\|_2^2$, and by Cauchy-Schwarz,

$F = G$ (up to a set of measure zero), and this implies $\mu = \lambda$. In the discrete setting this does not hold. The primary difference being that we must *also* consider the events in which ties occur, and unlike the continuous setting, these events have non-zero probabilities. Factoring in the possibility of ties and assuming that π is a 5-balanced permutation (and thus also 4-balanced, see Corollary 2.4.1), we obtain the following identities (once again, by a computation similar to the proof of Lemma 4.3):

$$\Pr_{i,j,x,y\sim[n]}[i,j \leq x \wedge \pi(i),\pi(j) \leq \pi(y)] = \frac{1}{9} + \frac{13}{36n} + \frac{7}{18n^2} + \frac{5}{36n^3}$$

$$\Pr_{i,j,k,x,y\sim[n]}[i,j \leq x \wedge \pi(i),\pi(k) \leq \pi(y)] = \frac{1}{9} + \frac{1}{3n} + \frac{13}{36n^2} + \frac{1}{6n^3} + \frac{1}{36n^4} = \frac{\|v\|_2^2}{n^2}$$

While the two probabilities agree on the leading term (which corresponds to the event of no ties), they *disagree* on the remaining terms. Therefore we cannot apply Cauchy-Schwarz to argue that $u = v$ (which would have indeed yielded a contradiction, since the function u associated with any permutation has precisely n different values, whereas v has $\Theta(n^2)$ different values and therefore does not correspond to any permutation), and must pursue a different proof.

5 The Minimal Distance from k-Balanced for $k \geq 4$

As we just saw, permutations cannot be k-balanced for any $k \geq 4$. But what is the *smallest possible distance* (in, say, ℓ_∞-norm) between an *attainable* profile and the uniform profile? Here is what we know:

Lower bound. For any $k \geq 4$, we show that the minimal distance from k-balanced is at least $\Omega\left(n^{k-1}\right)$. The proof follows from a robust version of Lemma 4.3, and is given in Section 5.1.

Upper bound. For $k = 4$, we provide an explicit construction of an infinite family of permutations whose members attain a distance $\mathcal{O}\left(n^3\right)$ from uniform. The construction is based on modification of the well-known Erdős-Szekeres permutation [ES35], and is given in Section 5.2. Consequently, our bounds on the 4-profile are asymptotically tight to within a constant factor. The remaining cases, where $k > 4$, are presently left open.

Concentration and anti-concentration The asymptotic distribution of the k-profile is a well-researched topic [EZ20, JNZ13, Hof18, Bón07]. We observe that these results imply that for any fixed $k \geq 2$, with probability $> 99\%$ the k-profile of a uniformly random $\pi \sim \mathbb{S}_n$ has distance $\Theta(n^{k-1/2})$ from balanced, as $n \to \infty$. Therefore, *if* our lower bound for $k \geq 4$ is not tight for any k, then it is off by a multiplicative factor of $\mathcal{O}(n^{1/2})$ (see Section 5.2.2 for a discussion).

5.1 A Lower Bound on the Distance

Notation 5.1 (distance from uniform k-profile). *Let $\pi \in \mathbb{S}_n$ be a permutation and let $1 \leq k \leq n$ be an integer. The distance of π from the uniform k-profile, in ℓ_∞-norm, is denoted as follows:*

$$\delta_{\pi,k} := \max_{\tau\in\mathbb{S}_k} \left| \#\tau\left(\pi\right) - \frac{\binom{n}{k}}{k!} \right|.$$

We also denote the smallest *distance over* all *n-element permutations by $\delta_k(n) := \min_{\pi\in\mathbb{S}_n} \delta_{\pi,k}$.*

Lemma 5.2 (low-distance k-profile implies low-distance $(k-1)$-profile). *For every $\pi \in \mathbb{S}_n$ and $1 < k \leq n$:*

$$\delta_{\pi,k-1} \leq \frac{k^2}{n-k+1}\delta_{\pi,k}$$

Proof. TOPROVE 10 □

Theorem 5.3 (lower bound on distance from k-balanced). *For every constant $k \geq 4$, there holds $\delta_k(n) = \Omega(n^{k-1})$.*

Proof. TOPROVE 11 □

5.2 Matching Upper Bound on the Distance for $k = 4$

In this subsection we show that Theorem 5.3 is asymptotically tight when $k = 4$. I.e., there exists an infinite family of permutations, the 4-profiles of which are only $\mathcal{O}(n^3)$ away from balanced. Our construction is a modification of the classical *Erdős-Szekeres permutation* [ES35].

Definition 5.4 (Erdős-Szekeres permutations). *For integers $n, m \geq 1$ and $\theta > 0$, let $\mathcal{P}(\theta)$ be the set of points in the $[n] \times [m]$ grid, rotated by an angle of θ about the origin. Let $\delta > 0$ be the smallest angle such that some two points in $\mathcal{P}(\delta)$ reside on the same axis-parallel line. Pick some $0 < \varepsilon < \delta$. The positive (resp. negative) Erdős-Szekeres Permutation, denoted $\mathrm{ES}^+(n,m)$ (resp. $\mathrm{ES}^-(n,m)$) is the permutation associated with the point set $\mathcal{P}(\varepsilon)$ (resp. $\mathcal{P}(-\varepsilon)$). When $n = m$, we omit the second operand.*

Motivation. In [KP13] it was shown that for any $k \geq 4$, the *unique* measure corresponding to a limit permutation with balanced k-profiles is the Lebesgue measure on the unit square. This suggests that in search of permutations with nearly balanced k-profiles, one may consider "square-like" families whose members locally resemble this measure. While the Erdős-Szekeres permutations are natural candidates in this respect, their distance from uniform substantially exceeds the cubic bound.

Proposition 5.5 (Erdős-Szekeres is far from 4-balanced). *Let $n > 1$ be an integer and let $\pi = \mathrm{ES}^+(n)$ be the Erdős-Szekeres permutation over n^2 elements. Then $\#\mathtt{3142}\,(\pi) = \binom{n+2}{4}^2$, and in particular:*

$$\delta_{\pi,4} \geq \frac{1}{144}n^7 + \Omega(n^6)$$

Proof. TOPROVE 12 □

Remark 5.6. Apart from $\tau = \mathtt{3142}$, there is only one other pattern $\sigma = \mathtt{2413}$ in $\mathbb{S}_4$, for which the distance is $\Theta(n^7)$. The computation for this pattern proceeds identically to the proof of Proposition 5.5, the only difference being the set of strict inequalities imposed. All remaining entries of the 4-profile of $\mathrm{ES}^+(n)$ are indeed within $\mathcal{O}(n^6)$ of uniform.

5.2.1 A Modification of Erdős-Szekeres

We next define a modification of the Erdős-Szekeres permutation.

Definition 5.7 (two-sided Erdős-Szekeres). *With n, m and ε as in definition 5.4, the two-sided Erdős-Szekeres Permutation, denoted $\mathrm{ES}^\pm(n,m)$, is the permutation associated with the points $\mathcal{P}(\varepsilon) \sqcup \mathcal{P}(-\varepsilon)$.*

The 4-profile of this family of permutations is optimally balanced, up to a multiplicative constant.

Theorem 5.8 (4-profile of two-sided Erdős-Szekeres). *Let $n \geq 1$ and let $\pi = \mathrm{ES}^\pm(n) \in \mathbb{S}_{2n^2}$ be the two-sided Erdős-Szekeres permutation. Then,*

$$\delta_{\pi,4} = \frac{2n^6}{9} + \mathcal{O}(n^5)$$

Proof. TOPROVE 13 □

5.2.2 Profiles and Distance of Random Permutations

A simple probabilistic argument shows that for every *fixed* $k \geq 2$ and large n, the k-profile of almost every permutation in $\mathbb{S}_n$ is $\left(\binom{n}{k}/k! \pm o(n^k)\right)\mathbb{1}$. In this discussion we are interested in exactly how close to balanced the k-profile of a typical (random) permutation is, and in particular, whether this distance attains, or nearly attains, our lower bound of Theorem 5.3.

So fix some $k \geq 2$ and consider a pattern $\tau \in \mathbb{S}_k$. Associated with τ is the random variable $X_\tau := \#\tau\,(\pi)$ where π is uniformly sampled from $\mathbb{S}_n$. Clearly, $\mathbb{E}[X_\tau] = \binom{n}{k}/k!$. The distribution of X_τ, its moments, and even the pairwise joint distributions of patterns have received considerable attention (e.g., [EZ20, JNZ13, Hof18, Bón07]). It is known in particular that X_τ satisfies a central limit theorem. Concretely, there exists a constant $\sigma_\tau > 0$ such that as $n \to \infty$,

$$\sqrt{n}\left(\frac{X_\tau}{\binom{n}{k}} - \frac{1}{k!}\right) \xrightarrow{d} N(0, \sigma_\tau)$$

This CLT implies asymptotic concentration and anti-concentration of k-profiles.

Proposition 5.9 (concentration and anti-concentration of k-profile)**.** *Let $k \geq 2$ be a constant. Then, for any $\alpha > 0$ we have:*

$$\Pr_{\pi \sim \mathbb{S}_n}\left[\delta_{\pi,k} \geq \alpha \cdot n^{k-1/2}\right] = 2 \cdot \Phi\left(-\frac{\alpha}{k!}\right) \pm o(1)$$

and conversely (by union over $\mathbb{S}_k$),

$$\Pr_{\pi \sim \mathbb{S}_n}\left[\delta_{\pi,k} \leq \alpha \cdot n^{k-1/2}\right] \geq 1 - 2k! \cdot \Phi\left(-\frac{\alpha}{k!}\right) \pm o(1)$$

In particular, this implies that $\delta_{\pi,k} \in \left(\frac{1}{100k!},\ 2k!\right) n^{k-1/2}$ with probability $> 99\%$ as $n \to \infty$.

6 An Asymptotic Relation Between Profiles and Permutations

So far we have considered k-profiles of order-n permutations with $k > 1$ *fixed*, and $n \to \infty$. But such profiles are interesting also when $k = k(n)$ grows with n. For example, as observed at the start of Section 4, when $k \gtrsim e\sqrt{n}$, at least some order-k permutations must be missing, i.e., the support of the k-profile is necessarily incomplete. Also, in the extreme case where $k = n$, the k-profile is a singleton. Our main discovery in this section is the following:

Profiles determine points. In the range $n \geq k(n) \geq \Omega(\sqrt{n}\log n)$, the k-profile of $\pi \in \mathbb{S}_n$ reveals a lot about π. Explicitly, we prove that there exists a set $\mathcal{D} \subset [n]^2$ of $\widetilde{\Omega}(k^4/n^2)$ points (consisting of four symmetric regions, of widths roughly k^2/n), such any two permutations in $\mathbb{S}_n$ with the same k-profile, *must* agree on their restriction to $\mathcal{D}$ (see Figure 6). In the extreme case where $k = n$, our Theorem is close to tight, as the set $\mathcal{D}$ nearly covers the entire grid $[n] \times [n]$ (up to a logarithmic factor).

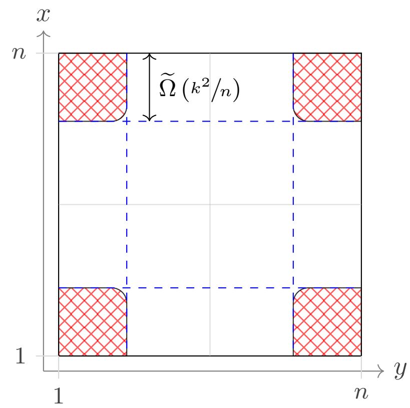

Figure 6: The k-profile fully determines the restriction of any $\pi \in \mathbb{S}_n$ with that profile, to the red region.

The main result of this section is Theorem 6.7. Its proof goes as follows: We define the *evaluation* $p(\pi)$ of a bivariate polynomial p over a permutation π. Then we show that if $\deg(p) < k$, then this real number $p(\pi)$ is uniquely defined by the k-profile of the permutation π. We subsequently use standard tools from approximation theory to construct a family of polynomials of degree $< k$, which allow us to uncover the points in $\mathcal{D}$.

6.1 k-Profiles Determine the Evaluation of Degree $< k$ Polynomials

Here is the main notion that we use in this section:

Notation 6.1 (evaluation of polynomial on permutation)**.** *Let $p \in \mathbb{R}[x, y]$ be a real bivariate polynomial and let $\pi \in \mathbb{S}_n$ be a permutation. The evaluation of p on π is denoted:*

$$p(\pi) := \sum_{i=1}^{n} p(i, \pi(i))$$

With this notation we show:

Proposition 6.2 (k-profile determines ($\deg < k$)-evaluations)**.** *Let $p \in \mathbb{R}[x, y]$ be a bivariate polynomial with $\deg(p) < k$ for some integer $k > 1$. Also, let $\pi \in \mathbb{S}_n$ be a permutation of order $n \geq k$. Then $p(\pi)$ can be efficiently computed, given π's k-profile.*

Proof. TOPROVE 14 $\square$

Remark 6.3. Proposition 6.2 can be extended by analysing a different sets of events. For example, for any $r < k$ and $\tau \in \mathbb{S}_r$, we could consider the event in which we sample r permutation points, and condition on their relative ordering so that they form an instance of τ in π. Then, using the remaining budget of at most $k - r$ points, we could sample from their marginals. Such events determine the evaluations of many more polynomials (albeit, on a modified and weighted pointset).

6.2 Determining Points using Approximate Indicators

Here is our method for "reading the bit" at position (x, y). We construct to this end a *low-degree* polynomial that is a good pointwise approximator of the indicator $\mathbb{1}_{(x,y)} : [n]^2 \to \{0, 1\}$. If the polynomial has degree $< k$ and the approximation error is small, then by evaluating it, we can determine the value of the corresponding bit. This means that either *every* permutation π with a given k-profile must contain this point, or *none* of them do, and the evaluation of the polynomial will reveal this.

For notational convenience, in what follows we consider (as in Section 3) the action $\langle r \rangle \curvearrowright [1, n]^2$ of the 90°-rotation. We denote by $O(a, b) = \{(a, b), (b, n+1-a), (n+1-a, n+1-b), (n+1-b, a)\}$ the r-orbit of $(a, b) \in [n]^2$. Similarly, for a set $\mathcal{D} \subset [n]^2$ we denote $O(\mathcal{D}) := \cup_{(a,b) \in \mathcal{D}} O(a, b)$ (i.e., the r-orbit of $\mathcal{D}$). The following fact is well-known and easy to verify (e.g., [MP69]).

Lemma 6.4 (symmetrisation)**.** *Let $p : \{0, 1\}^n \to \mathbb{R}$ be a real multilinear polynomial, and let $f : \{0, 1, \ldots, n\} \to \mathbb{R}$ be the function:*

$$\forall k \in \{0, 1, \ldots n\} : f(k) = \mathbb{E}_{|x| \sim k}[p(x)]$$

where the expectation is taken with respect to the uniform distribution over all $x \in \{0, 1\}^n$ of Hamming weight k. Then, f can be written as a real polynomial in k of degree at most $\deg(p)$.

Lemma 6.5 (approximate degree of symmetric boolean functions [Pat92])**.** *Let $f : \{0, 1\}^n \to \{0, 1\}$ be a symmetric Boolean function, and let:*

$$\Gamma(f) = \min_k \{|2k - n + 1| : f_k \neq f_{k+1}\}$$

where f_k is the value of f on inputs of Hamming weight k. Then, there exists a multilinear polynomial $g \in \mathbb{R}[x_1, \ldots, x_n]$ such that $\forall x \in \{0, 1\}^n : |g(x) - f(x)| \leq 1/3$, and furthermore:

$$\deg(g) \leq A \cdot \sqrt{n\,(n - \Gamma(f))}$$

where $A > 0$ is a universal constant.

Lemma 6.6 (one-sided approximation of $\mathbb{1}_{(a,b)}$)**.** *Let u, v and n be integers such that $1 \leq u, v < n/2$. Then, for any $(a, b) \in O(u, v)$, there exists a polynomial $\widetilde{\mathbb{1}}_{(a,b)} \in \mathbb{R}[x, y]$ such that:*

$$\deg\left(\widetilde{\mathbb{1}}_{(a,b)}\right) \leq C\left(\sqrt{n(2u+1)} + \sqrt{n(2v+1)}\right)\log n$$

where $C > 0$ is an absolute constant, and:

$$\forall (x, y) \in [n]^2 : \ \widetilde{\mathbb{1}}_{(a,b)}(x, y) \in \begin{cases} [1, \infty) & x = a \wedge y = b \\ [0, \frac{1}{2n}] & x \neq a \vee y \neq b \end{cases}$$

Proof. TOPROVE 15 $\square$

Theorem 6.7 (k-profiles determine points)**.** *Let $n \geq k > 1$ and let:*

$$\mathcal{D} = \left\{(a, b) \in [n/2]^2 \ : \ C\left(\sqrt{n(2a+1)} + \sqrt{n(2b+1)}\right)\log n < k\right\} \subset [n]^2$$

where C is the constant of Lemma 6.6. Then the k-profile of an order-n permutation $\pi \in \mathbb{S}_n$ uniquely determines the restriction of π to $O(\mathcal{D})$.

Proof. TOPROVE 16 $\square$

7 Discussion

In this paper we consider the existence of k-balanced permutations. For $k \le 3$ we show that such permutations exist whenever n satisfies the necessary divisibility conditions, and for $k \ge 4$, we show that no such permutations exist. Moreover, we prove that the k-profile of any n-element permutation must have an entry which is $\Omega_n(n^{k-1})$ away from uniform, whenever $k \ge 4$. This gives rise to several interesting open questions.

Is the lower bound tight? Recall that for $k = 4$ we provide an explicit construction of an infinite family (see Section 5.2.1) in which every pattern in $\mathbb{S}_4$ appears within additive distance of $\Theta(n^3)$ from uniform, i.e., matching the lower bound of Theorem 5.3. Conversely, we note (see Section 5.2.2) that *all* entries in the k-profile of a uniformly random permutation in $\mathbb{S}_n$ are, with probability $> 99\%$ (for large enough n), within distance $\Theta(n^{k-1/2})$ from uniform. In this view we ask what is the true behaviour for $k > 4$. Specifically, does our lower bound remain tight, or does the true bound change to $\Omega(n^{k-1/2})$, as attained by the majority of permutations?

How many k-patterns can appear the right number of times? We have ruled out the possibility that *every* entry in the 4-profile equals $\binom{n}{4}/4!$. However, for any *fixed* pattern $\tau \in \mathbb{S}_4$, we are able to construct a bespoke infinite family of permutations, in whose members τ appears exactly $\binom{n}{4}/4!$ times (these constructions are quite intricate, and are not included in this paper). So, we ask: how *many* entries in the k-profile of an n-element permutation may be precisely $\binom{n}{k}/k!$, simultaneously?

What is the maximal dimension of a k-balanced subspace? It makes sense to ask the same question with regards to linear subspaces. That is, what is the maximal dimension of a subspace $V_k \le \mathbb{R}^{\mathbb{S}_k}$ such that there exist infinitely many permutations $\pi \in \mathbb{S}_n$ for which:

$$\forall v = (\alpha_\tau)_{\tau \in \mathbb{S}_k} \in V_k : \ \frac{\binom{n}{k}}{k!} \langle v, \mathbb{1}_{\mathbb{S}_k} \rangle = \sum_{\tau \in \mathbb{S}_k} \alpha_\tau \# \tau\,(\pi)$$

In other words, unlike the previous question, here we allow any basis for V_k, not necessarily only the coordinate vectors. Clearly $\langle \mathbb{1}_{\mathbb{S}_k} \rangle \in V_k$, for any k. Also, since 3-balanced permutations exist, then by Proposition 2.4 there are $3! = 6$ linearly independent combinations in $\mathbb{S}_4$ that hold true, and $\dim(V_4) \ge 6$ ($\mathbb{1}_{\mathbb{S}_k}$ resides in their span). In general, we ask: what is the maximal dimension of V_k, for $k \ge 4$?

How many permutations are 3-balanced? 2- and 3-balanced permutations *exist* for every admissible value of n (see Section 3). In fact, they never appear "alone": as they are closed under the action of D_4, their entire orbit is also balanced and so there must at least be *two* balanced permutations, whenever one exists (no permutation is identical to its reflection about either axis). Therefore, we ask: what is the *exact count*, or even *asymptotic growth rate*, of 3-balanced permutations (restricted only to the admissible n)? We remark that for 2-balanced permutations these answers are already known (see [OEI18, A316775] and [OEI18, A000140]). However, interestingly, for $k = 3$ we presently only know that at $n = 9$ there are exactly two 3-balanced permutations (see Figure 2).

References

[AAH+02] Michael H Albert, Mike D Atkinson, Chris C Handley, Derek A Holton, and Walter Stromquist. On packing densities of permutations. *the electronic journal of combinatorics*, 9(1):R5, 2002.

[Bón07] Miklós Bóna. The copies of any permutation pattern are asymptotically normal. *arXiv preprint arXiv:0712.2792*, 2007.

[CGW89] Fan R. K. Chung, Ronald L. Graham, and Richard M. Wilson. Quasi-random graphs. *Combinatorica*, 9:345–362, 1989.

[Coo04] Joshua N Cooper. Quasirandom permutations. *Journal of Combinatorial Theory, Series A*, 106(1):123–143, 2004.

[CP08] Joshua Cooper and Andrew Petrarca. Symmetric and asymptotically symmetric permutations. *arXiv preprint arXiv:0801.4181*, 2008.

[DG20] Bartłomiej Dudek and Paweł Gawrychowski. Counting 4-patterns in permutations is equivalent to counting 4-cycles in graphs. *arXiv preprint arXiv:2010.00348*, 2020.

[ES35] Paul Erdős and George Szekeres. A combinatorial problem in geometry. *Compositio mathematica*, 2:463–470, 1935.

[EZ20] Chaim Even-Zohar. Patterns in random permutations. *Combinatorica*, 40(6):775–804, 2020.

[EZL21] Chaim Even-Zohar and Calvin Leng. Counting small permutation patterns. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2288–2302. SIAM, 2021.

[GKLO16] Stefan Glock, Daniela Kühn, Allan Lo, and Deryk Osthus. The existence of designs via iterative absorption: hypergraph F-designs for arbitrary F. *arXiv preprint arXiv:1611.06827*, 2016.

[HKM+13] Carlos Hoppen, Yoshiharu Kohayakawa, Carlos Gustavo Moreira, Balázs Ráth, and Rudini Menezes Sampaio. Limits of permutation sequences. *Journal of Combinatorial Theory, Series B*, 103(1):93–113, 2013.

[HKMS11] Carlos Hoppen, Yoshiharu Kohayakawa, Carlos Gustavo Tamm de Araújo Moreira, and Rudini Menezes Sampaio. Limits of permutation sequences through permutation regularity. *arXiv preprint arXiv:1106.1663*, 2011.

[Hof18] Lisa Hofer. A central limit theorem for vincular permutation patterns. *Discrete Mathematics & Theoretical Computer Science*, 19, 2018.

[JNZ13] Svante Janson, Brian Nakamura, and Doron Zeilberger. On the asymptotic statistics of the number of occurrences of multiple permutation patterns. *arXiv preprint arXiv:1312.3955*, 2013.

[JS92] Svante Janson and Joel Spencer. Probabilistic construction of proportional graphs. *Random Structures & Algorithms*, 3(2):127–137, 1992.

[Kee14] Peter Keevash. The existence of designs. *arXiv preprint arXiv:1401.3665*, 2014.

[KP13] Daniel Král' and Oleg Pikhurko. Quasirandom permutations are characterized by 4-point densities. *Geometric and Functional Analysis*, 23(2):570–579, 2013.

[Lov12] László Lovász. *Large networks and graph limits*, volume 60. American Mathematical Soc., 2012.

[MP69] Marvin Minsky and Seymour Papert. Perceptron: an introduction to computational geometry, 1969.

[NPS18] Humberto Naves, Oleg Pikhurko, and Alex Scott. How unproportional must a graph be? *European Journal of Combinatorics*, 73:138–152, 2018.

[OEI18] OEIS Foundation Inc. The On-Line Encyclopedia of Integer Sequences, 2018. Published electronically at http://oeis.org.

[Pat92] Ramamohan Paturi. On the degree of polynomials that approximate symmetric boolean functions (preliminary version). In *Proceedings of the twenty-fourth annual ACM symposium on Theory of computing*, pages 468–474, 1992.

[PG75] Nicholas Pippenger and Martin Charles Golumbic. The inducibility of graphs. *Journal of Combinatorial Theory, Series B*, 19(3):189–203, 1975.

[SS18] Jakub Sliacan and Walter Stromquist. Improving bounds on packing densities of 4-point permutations. *Discrete Mathematics & Theoretical Computer Science*, 19(Permutation Patterns), 2018.

[Tho87] Andrew Thomason. Pseudo-random graphs. In *North-Holland Mathematics Studies*, volume 144, pages 307–331. Elsevier, 1987.

[Wil02] Herbert S Wilf. The patterns of permutations. *Discrete Mathematics*, 257(2-3):575–583, 2002.

Appendix A 3-Balanced Constructions for all Remainders

In the proof of Theorem 3.7 we amended σ by inserting two points, yielding 3-balanced permutations for every $n > 56$ with $n \equiv 20 \pmod{36}$. The same strategy applies to the other residues as well, as we now describe. In the following discussion $\ell = 3t + 1$, and $0 < \varepsilon < 1$ is a small constant.

To prove the correctness of these constructions, we observe that all newly inserted coordinates are affine transformations of t. Consequently, each 2-pattern (resp. 3-pattern) count in σ is a quadratic (resp. cubic) polynomial in t. It follows that Equation (2) posits the vanishing of a cubic polynomial. This we can verify by checking only four distinct values of t. This is an alternative to the calculation presented in Theorem 3.7.

A.1 Constructions for even n

In Theorem 3.7, we amended $\sigma \in \mathbb{S}_m$ by inserting two additional points:

$$(x_1, y_1) := (r + 2 + \varepsilon,\ r + \ell + \varepsilon) \qquad (x_2, y_2) := (r + \ell + \varepsilon,\ r - \varepsilon)$$

and showed that for any $t \geq 2$, one can set $r = 4t + 2$ to satisfy Equation (2). The size of the resulting 3-balanced permutation $\pi \in \mathbb{S}_n$ obtained by rotation is $n = 4m = 36t + 20$, that is, $n \equiv 20 \pmod{36}$

The case $n \equiv 28 \pmod{36}$. Insert two more points to σ, so that $m = 3\ell + 4$:

$$(x_3, y_3) := (\varepsilon,\ \ell + \varepsilon) \qquad (x_4, y_4) := (1 + \varepsilon,\ \ell - \varepsilon)$$

The case $n \equiv 0 \pmod{36}$. Insert two more points, in addition to the above four:

$$(x_5, y_5) := (\ell + 2 + \varepsilon,\ \varepsilon) \qquad (x_6, y_6) := (\ell + 1 + \varepsilon,\ 5 + \varepsilon)$$

A.2 Constructions for odd n

For $t \geq 4$, inserting the following points (and a point at the centre of π) yields 3-balanced permutations.

The case $n \equiv 29 \pmod{36}$. Insert four points to σ, so that $m = 3\ell + 4$:

$$(x_1, y_1) := (-5 + \varepsilon,\ 1 + \varepsilon) \qquad (x_2, y_2) := (-3 + \varepsilon,\ t - 2 + \varepsilon)$$

$$(x_3, y_3) := (-2 + \varepsilon,\ 7t + 4 + \varepsilon) \qquad (x_4, y_4) := (\varepsilon,\ 7t + 3 + \varepsilon)$$

The case $n \equiv 1 \pmod{36}$. Insert two more points to σ, so that now $m = 3\ell + 6$:

$$(x_5, y_5) := (-4 + \varepsilon,\ 3t - 1 + \varepsilon) \qquad (x_6, y_6) := (2t - 2 + \varepsilon,\ 5t + 1 + \varepsilon)$$

The case $n \equiv 9 \pmod{36}$. Insert two last points to σ, in addition to the previous six. So, $m = 3\ell + 8$:

$$(x_7, y_7) := (-1 + \varepsilon,\ 3t - 2 + \varepsilon) \qquad (x_8, y_8) := (4t + 1 + \varepsilon,\ t - 1 + \varepsilon)$$

A.3 Small cases not covered by our construction

For completeness, we provide a list of 3-balanced permutations in Table 1, for these values of n that were not covered by the aforementioned constructions. This is because Theorem 3.7, Section A.1 and Section A.2 yield 3-balanced permutations for every residue modulo 36, starting only from some minimal value of t.

n	**3-Balanced Permutation in $\mathbb{S}_n$**
9	(3, 4, 9, 8, 5, 2, 1, 6, 7)
20	(8, 3, 19, 16, 4, 11, 9, 20, 14, 6, 15, 7, 1, 12, 10, 17, 5, 2, 18, 13)
28	(1, 7, 18, 22, 19, 26, 9, 23, 28, 14, 6, 10, 24, 12, 13, 5, 17, 3, 20, 16, 8, 2, 4, 25, 15, 27, 21, 11)
29	(9, 3, 6, 17, 26, 14, 29, 13, 11, 22, 19, 23, 28, 25, 20, 18, 8, 12, 4, 2, 5, 1, 7, 10, 15, 16, 27, 24, 21)
36	(30, 12, 27, 4, 25, 15, 3, 10, 32, 5, 35, 33, 19, 17, 13, 24, 16, 1, 20, 9, 36, 21, 28, 8, 23, 14, 34, 18, 7, 26, 6, 22, 2, 29, 31, 11)
37	(16, 15, 1, 35, 9, 13, 29, 30, 33, 32, 12, 3, 34, 21, 14, 18, 8, 19, 10, 20, 36, 22, 31, 17, 24, 28, 11, 5, 2, 6, 4, 7, 37, 26, 25, 27, 23)
45	(42, 14, 15, 9, 36, 5, 44, 8, 26, 33, 17, 39, 21, 29, 11, 1, 25, 32, 20, 45, 35, 12, 27, 7, 3, 23, 22, 38, 24, 16, 31, 40, 19, 4, 10, 37, 41, 34, 2, 28, 6, 43, 30, 13, 18)
56	(7, 8, 38, 50, 42, 25, 27, 45, 10, 22, 15, 46, 17, 56, 21, 36, 12, 24, 54, 2, 48, 39, 14, 16, 30, 51, 55, 41, 5, 26, 31, 35, 49, 4, 43, 34, 13, 3, 1, 37, 29, 28, 40, 32, 11, 23, 18, 19, 20, 44, 6, 52, 33, 53, 9, 47)
64	(14, 8, 23, 48, 60, 38, 55, 50, 44, 26, 4, 62, 24, 53, 31, 35, 1, 47, 9, 28, 16, 33, 20, 40, 7, 63, 6, 43, 19, 29, 11, 13, 52, 54, 36, 46, 22, 59, 2, 58, 25, 45, 32, 49, 37, 56, 18, 64, 30, 34, 12, 41, 3, 61, 39, 21, 15, 10, 27, 5, 17, 42, 57, 51)
65	(59, 33, 5, 4, 55, 26, 44, 57, 37, 19, 7, 47, 20, 34, 9, 51, 27, 22, 50, 16, 12, 64, 36, 10, 56, 45, 40, 65, 39, 43, 54, 14, 28, 13, 15, 24, 30, 62, 32, 46, 63, 35, 41, 6, 53, 3, 49, 38, 11, 29, 1, 60, 2, 23, 18, 17, 8, 31, 52, 61, 48, 21, 42, 25, 58)
72	(11, 21, 16, 58, 45, 55, 3, 49, 9, 65, 31, 44, 69, 7, 60, 66, 2, 71, 42, 37, 35, 34, 24, 40, 29, 39, 54, 51, 8, 47, 25, 36, 56, 72, 38, 6, 13, 41, 48, 30, 4, 12, 46, 28, 62, 14, 63, 67, 10, 53, 32, 19, 43, 52, 22, 50, 1, 18, 5, 70, 64, 17, 57, 61, 23, 26, 15, 27, 68, 33, 59, 20)
73	(42, 68, 6, 13, 4, 70, 36, 38, 2, 29, 47, 24, 64, 51, 32, 48, 41, 54, 19, 20, 71, 60, 7, 53, 26, 67, 69, 37, 33, 18, 27, 44, 3, 30, 55, 72, 52, 49, 45, 16, 25, 14, 50, 12, 31, 5, 35, 61, 65, 46, 21, 9, 56, 10, 66, 8, 59, 43, 39, 1, 15, 22, 34, 23, 62, 73, 57, 11, 17, 63, 58, 28, 40)
81	(44, 31, 58, 10, 77, 49, 80, 79, 1, 52, 34, 15, 5, 25, 63, 42, 9, 47, 23, 38, 32, 73, 35, 4, 27, 48, 46, 69, 22, 68, 41, 2, 30, 59, 81, 66, 65, 40, 26, 53, 74, 29, 21, 28, 11, 54, 6, 19, 18, 55, 70, 75, 56, 43, 60, 13, 57, 14, 61, 37, 76, 64, 36, 3, 17, 8, 51, 67, 78, 72, 7, 33, 20, 39, 62, 12, 24, 16, 50, 71, 45)
101	(84, 58, 37, 76, 54, 38, 51, 20, 85, 83, 36, 22, 97, 7, 23, 39, 90, 6, 67, 88, 66, 11, 70, 93, 96, 4, 16, 2, 53, 34, 24, 3, 31, 48, 91, 13, 94, 74, 64, 14, 44, 21, 49, 46, 79, 18, 61, 43, 95, 17, 78, 56, 52, 69, 75, 57, 72, 41, 81, 32, 101, 73, 5, 98, 15, 42, 62, 89, 86, 87, 25, 68, 30, 50, 82, 33, 8, 60, 71, 45, 55, 27, 47, 63, 29, 1, 35, 9, 10, 65, 80, 19, 59, 92, 40, 12, 77, 26, 99, 28, 100)
109	(64, 39, 93, 24, 97, 65, 7, 54, 101, 52, 49, 62, 88, 35, 9, 81, 4, 8, 89, 13, 75, 87, 36, 41, 80, 71, 28, 3, 55, 73, 1, 78, 99, 104, 82, 14, 22, 57, 18, 63, 67, 29, 20, 109, 91, 31, 43, 100, 32, 106, 15, 95, 69, 16, 76, 79, 45, 53, 85, 103, 90, 92, 61, 27, 50, 23, 86, 66, 38, 26, 60, 34, 70, 25, 5, 105, 37, 46, 84, 83, 44, 6, 72, 10, 12, 51, 94, 68, 96, 74, 17, 59, 56, 47, 2, 30, 102, 11, 48, 40, 19, 21, 42, 98, 33, 77, 58, 108, 107)
117	(100, 99, 90, 107, 48, 67, 45, 88, 61, 50, 4, 21, 79, 89, 75, 23, 53, 1, 2, 56, 106, 26, 102, 87, 36, 96, 72, 3, 14, 8, 24, 47, 42, 60, 44, 93, 38, 81, 13, 69, 52, 85, 15, 83, 111, 27, 86, 113, 40, 108, 6, 77, 101, 55, 64, 98, 9, 34, 59, 84, 109, 20, 54, 63, 17, 41, 112, 10, 78, 5, 32, 91, 7, 35, 103, 33, 66, 49, 105, 37, 80, 25, 74, 58, 76, 71, 94, 110, 104, 115, 46, 22, 82, 31, 16, 92, 12, 62, 116, 117, 65, 95, 43, 29, 39, 97, 114, 68, 57, 30, 73, 51, 70, 11, 28, 19, 18)
137	(29, 66, 131, 65, 27, 44, 3, 119, 117, 61, 101, 47, 28, 79, 92, 18, 49, 122, 8, 31, 9, 48, 81, 103, 76, 104, 133, 125, 137, 40, 118, 42, 38, 26, 24, 87, 11, 105, 68, 108, 95, 106, 41, 132, 75, 15, 126, 116, 121, 74, 36, 83, 80, 60, 52, 71, 23, 53, 14, 84, 128, 25, 45, 50, 134, 136, 56, 99, 69, 39, 82, 2, 4, 88, 93, 113, 10, 54, 124, 85, 115, 67, 86, 78, 58, 55, 102, 64, 17, 22, 12, 123, 63, 6, 97, 32, 43, 30, 70, 33, 127, 51, 114, 112, 100, 96, 20, 98, 1, 13, 5, 34, 62, 35, 57, 90, 129, 107, 130, 16, 89, 120, 46, 59, 110, 91, 37, 77, 21, 19, 135, 94, 111, 73, 7, 72, 109)
145	(97, 26, 116, 136, 81, 11, 55, 32, 118, 4, 140, 50, 92, 103, 124, 69, 62, 23, 45, 44, 83, 15, 128, 67, 109, 144, 99, 9, 48, 3, 105, 138, 66, 75, 40, 94, 25, 59, 46, 111, 31, 95, 14, 126, 127, 107, 27, 117, 1, 134, 42, 36, 72, 13, 139, 57, 90, 60, 108, 88, 76, 129, 21, 78, 5, 113, 122, 64, 130, 61, 34, 93, 73, 53, 112, 85, 16, 82, 24, 33, 141, 68, 125, 17, 70, 58, 38, 86, 56, 89, 7, 133, 74, 110, 104, 12, 145, 29, 119, 39, 19, 20, 132, 51, 115, 35, 100, 87, 121, 52, 106, 71, 80, 8, 41, 143, 98, 137, 47, 2, 37, 79, 18, 131, 63, 102, 101, 123, 84, 77, 22, 43, 54, 96, 6, 142, 28, 114, 91, 135, 65, 10, 30, 120, 49)
153	(142, 69, 119, 121, 148, 5, 118, 57, 105, 80, 130, 1, 84, 83, 28, 47, 38, 50, 76, 44, 109, 61, 25, 11, 131, 115, 95, 139, 31, 55, 125, 89, 4, 73, 3, 7, 66, 137, 26, 96, 48, 102, 56, 134, 21, 87, 138, 113, 9, 136, 101, 42, 51, 75, 124, 111, 146, 40, 27, 86, 132, 90, 82, 62, 32, 117, 46, 60, 152, 13, 14, 63, 120, 10, 100, 135, 77, 19, 54, 144, 34, 91, 140, 141, 2, 94, 108, 37, 122, 92, 72, 64, 22, 68, 127, 114, 8, 43, 30, 79, 103, 112, 53, 18, 145, 41, 16, 67, 133, 20, 98, 52, 106, 58, 128, 17, 88, 147, 151, 81, 150, 65, 29, 99, 123, 15, 59, 39, 23, 143, 129, 93, 45, 110, 78, 104, 116, 107, 126, 71, 70, 153, 24, 74, 49, 97, 36, 149, 6, 33, 35, 85, 12)

Table 1: 3-balanced permutations for all values of n not covered by our constructions.

Appendix B Computation of Lemma 4.3

In the proof of Lemma 4.3 we fix $\pi \in \mathbb{S}_n$ and consider the event:

$$\Pr_{i,j,k,l\sim[n]}\left[i<j,\ \pi(i)<\pi(j),\ k<l,\ \pi(k)<\pi(l)\right]$$

The proof of Lemma 4.3 proceeds by showing that, by conditioning on the possible equalities between the sampled indices and on their total order, the aforementioned event can be expressed as a polynomial involving pattern-counts in π, of lengths at most 4. The complete details of this computation are presented in Table 2. The first column of the table corresponds to the partition over the indices, where two indices are identical if and only if they reside in the same set. The second column corresponds to a fixed linear order over the indices at play, and the third column lists the contributing patterns, conditioned over the first two events. The total probability is then computed by summing over the probabilities of each row. Every row containing a partition of cardinality k, a linear order, and a set of patterns $S \subset \mathbb{S}_k$, indicates an event occurring with probability:

$$\frac{\prod_{i=1}^{k}(n-i+1)}{n^4}\cdot\frac{1}{k!}\cdot\sum_{\tau\in S}\frac{\#\tau\left(\pi\right)}{\binom{n}{k}}$$

Index Partition	Linear Order	Contributing Patterns
$\{i\},\{j\},\{k\},\{l\}$	$i<j<k<l$	1234, 1324, 3412, 2413, 1423, 2314
	$i<k<j<l$	1324, 1234, 3142, 2143, 1243, 2134
	$i<k<l<j$	1342, 1243, 3124, 2134, 1234, 2143
	$k<i<j<l$	3124, 2134, 1342, 1243, 2143, 1234
	$k<i<l<j$	3142, 2143, 1324, 1234, 2134, 1243
	$k<l<i<j$	3412, 2413, 1234, 1324, 2314, 1423
$\{i,k\},\{j\},\{l\}$	$i<j<l$	123, 132
	$i<l<j$	123, 132
$\{j,l\},\{i\},\{k\}$	$i<k<j$	123, 213
	$k<i<j$	123, 213
$\{i,l\},\{j\},\{k\}$	$k<i<j$	123
$\{j,k\},\{i\},\{l\}$	$i<j<l$	123
$\{i,k\},\{j,l\}$	$i<j$	12

Table 2: The terms corresponding to the computation of the total probability of the event analysed in the proof of Lemma 4.3. The partitions and orderings which do not contribute are omitted.