PREPRINT PDMI — 13/2004

An Improved Lower Bound on the Size of k-rankwise Independent Families of Permutations

Victor Bargachev [1]

St.Petersburg State University,
Department of Mathematics and Mechanics,
Universitetski prospekt, 28, Peterhof,
St.Petersburg, 198504, Russia

September 7, 2004

Abstract

Informally, a family $\mathcal{F} \subseteq S_n$ of permutations is k-rankwise independent if, for any $X \subseteq [n]$ with $|X| \leq k$, all elements of X are mapped in any possible order equally likely. It has been shown that $|\mathcal{F}| \geq \frac{n^d}{d!}(1+o(1))$ when $k = 2d+1$ is odd, and $|\mathcal{F}| \geq \frac{n^d}{d!2^{d-1}}(1+o(1))$ when $k = 2d$ is even. In this paper we show that $|\mathcal{F}| \geq \frac{!d+!(d+1)}{d!}n^d(1+o(1))$ when $k = 2d+1$ is odd, and $|\mathcal{F}| \geq \frac{!d}{d!}n^d(1+o(1))$ when $k = 2d$ is even. Here $!k$ denotes the subfactorial of k given by $!k = k!\sum_{i=0}^{k}\frac{(-1)^i}{i!}$. In particular, we improve previously known bound by an extra multiplicative factor exponential in k and independent of n.

[1] Supported in part by INTAS (grant no. 2000-447)

1 Introduction and the main result

The notion of k-rankwise independence was introduced by Itoh, Takei, and Tarui [1] as a notion stonger than k-restricted min-wise independence.

Throughout this paper we use the following notation. Let $[n, m]$ denote the set $\{n, n+1, \ldots, m\}$ and let $[n]$ denote $[1, n]$. Let S_n denote the symmetric group on $[n]$. For any finite set Y, let $|Y|$ denote size of Y. For $X \subseteq [n]$ we define the rank of element $x \in X$ as follows: $\operatorname{rank}(x, X) := |\{x' \in X : x' \leq x\}|$. Let $!n$ denote a subfactorial of n given by $!n = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}$. For a vector space V and a set of vectors $W \subseteq V$, let $\operatorname{span}(W)$ denote the linear hull of W, i.e., the minimal linear subspace of V containing W. Let us recall the main definition from [1]:

Definition 1 ([1]). *A family $\mathcal{F} \subseteq S_n$ of permutations is called* k-rankwise independent *if, for any subset $X = \{x_1, x_2, \ldots, x_k\} \subseteq [n]$ and any k distinct values $r_1, r_2, \ldots, r_k \in [k]$, we have*

$$\Pr_{\mathcal{F}}\left(\bigwedge_{i=1}^{k} \operatorname{rank}\big(\pi(x_i), \pi(X)\big) = r_i\right) = \frac{1}{k!}$$

when a permutation π is chosen at random from $\mathcal{F}$ with some specified probability distribution.

Remark. The original definition from [1] was given only for uniform distributions.

The following theorem provides a lower bound on the size of k-rankwise independent family:

Theorem 1 ([1]). *For any pair of integers $n \geq k \geq 1$, let $s = \lfloor k/2 \rfloor$. If a family $\mathcal{F} \subseteq S_n$ of permutations is k-rankwise independent, then $|\mathcal{F}| \geq m(n-1, k-1)$ for $k \leq 3$; $|\mathcal{F}| \geq \max\{m(n-1, k-1), \lfloor n/s \rfloor \binom{n - \lfloor n/2 \rfloor}{s-1}\}$ for $k \geq 4$. Here $m(n, d)$ denotes $\sum_{j=0}^{d/2} \binom{n}{j}$ if d is even and $\sum_{j=0}^{(d-1)/2} \binom{n}{j} + \binom{n-1}{(d-1)/2}$ if d is odd.*

In this paper we obtain an improved lower bound given by the next theorem:

Theorem 2. *Let $\mathcal{F} \subseteq S_n$ be a k-rankwise independent family. Then*

$$|\mathcal{F}| \geq \sum_{i=0}^{d} !i \binom{n}{i}$$

when $k = 2d$ is even and

$$|\mathcal{F}| \geq \sum_{i=0}^{d} !i \binom{n}{i} + !(d+1) \binom{n-1}{d}$$

when $k = 2d + 1$ is odd.

In particular it is greater than the previously known bound by an extra multiplicative factor exponential in k and independent of n.

2 Partial permutations

We introduce and study the notion of partial permutations in this section. The main result of this section is Theorem 3. We deduce Theorem 2 from Theorem 3 in the next section.

Definition 2. *A partial permutation of degree n and width m ($0 \le m \le n$) is a pair $\sigma = (X, f)$, where $X \subseteq [n]$ with $|X| = m$ and f is a bijective function from $[m]$ to X. Function f induces another linear order on X given by $f(1) < f(2) < \cdots < f(m)$.*

Let P_n^k be the set of all partial permutations of degree n and width k. Let $P_n = \bigcup_{k=0}^{n} P_n^k$ be the set of all partial permutations of degree n. Note that $|P_n^k| = \binom{n}{k} k!$. In particular, we have only one "empty" partial permutation of width 0. On the other hand, there are $n!$ partial permutation of width and degree n and they are in a "natural" one-to-one correspondence to the elements of the symmetric group S_n.

For $\sigma = (X, f) \in P_n$ we define $\mathrm{set}(\sigma) = X$. For $X \subseteq [n]$, let

$$P_n^X = \{\sigma \in P_n : \ \mathrm{set}(\sigma) = X\},$$

i.e., P_n^X is a set of partial permutation which define linear order exactly on the set X.

We say that partial permutation $\pi = (Y, g) \in P_n$ extends $\sigma = (X, f) \in P_n$ if $X \subseteq Y$ and the order on Y induces the order on X, i.e.,

$$\forall u, v \in X \quad f^{-1}(u) < f^{-1}(v) \iff g^{-1}(u) < g^{-1}(v).$$

Let $E(\sigma, Y)$ denote the set of partial permutations in P_n^Y that extend σ, i.e.,

$$E(\sigma, Y) \ := \ \{\pi \in P_n^Y : \ \pi \text{ extends } \sigma\}.$$

The following two lemmas are obvious.

Lemma 1. *For any sets $X \subseteq Y \subseteq [n]$ and for any $\sigma \in P_n^X$, we have*

$$|E(\sigma, Y)| = \frac{|Y|!}{|X|!}.$$

Lemma 2. *For any sets $X \subseteq Y \subseteq [n]$ and for any $\sigma_1, \sigma_2 \in P_n^X$ with $\sigma_1 \neq \sigma_2$, we have*

$$E(\sigma_1, Y) \cap E(\sigma_2, Y) = \emptyset.$$

In other words, Lemma 1 counts the number of ways to extend the linear order on X to a larger set Y. Similarly, Lemma 2 states that two different orders on X cannot be extended to the same order on a larger set Y. The following property is a little bit more complicated.

Lemma 3. *For any sets $X \subseteq Y \subseteq Z \subseteq [n]$ and any $\sigma \in P_n^X$ we have*

$$E(\sigma, Z) = \bigcup_{\pi \in E(\sigma, Y)} E(\pi, Z)$$

and the union above is disjoint.

Informally, right-to-left inclusion comes from the fact that the extension relation is transitive, i.e., π extends σ and τ extends π together imply that τ extends σ. The disjointness of the union follows from Lemma 2. Finally, Lemma 1 implies that the sets in the left and right sides are of the same size, which provides the equality.

Let $V_n = \mathbb{C}^{P_n^n}$ be an $n!$-dimensional vector space with coordinates indexed by the elements of P_n^n. Define a mapping $\mathsf{H} : P_n \to V_n$ as follows:

$$\mathsf{H}(\sigma)_\pi = \begin{cases} 1, & \text{if } \pi \text{ extends } \sigma; \\ 0, & \text{otherwise.} \end{cases}$$

Lemma 3 with $Z = [n]$ immediately implies that

$$\mathsf{H}(\sigma) = \sum_{\pi \in E(\sigma, Y)} \mathsf{H}(\pi) \tag{1}$$

for every $X \subseteq Y \subseteq [n]$ and for every $\sigma \in P_n^X$.

The main result of this section is the following theorem.

Theorem 3. *For each n there exists a set $Q_n \subseteq P_n$ such that $\mathsf{H}(Q_n)$ is a basis of V_n and $|Q_n \cap P_n^X| = !|X|$ for every $X \subseteq [n]$.*

The remaining part of this section contains the proof of this theorem. First of all, we introduce the notion of waste partial permutations. We say that $\sigma = (X, f) \in P_n^k$ is waste if there exists such $i \in [k]$ that

$$\big(\forall j < i \ f(j) > f(i)\big) \text{ and } \big(i = k \text{ or } f(i) > f(i+1)\big). \tag{2}$$

We define a set of waste indexes of σ as follows:

$$\mathrm{wst}(\sigma) := \{i \in [k] : \quad (2) \text{ holds } \}$$

Hence, σ is waste if and only if $\mathrm{wst}(\sigma) \neq \emptyset$.

Lemma 4. *For every $X \subseteq [n]$ there are exactly $!|X|$ non-waste partial permutations in P_n^X.*

Proof. First of all, we show that the number of non-waste partial permuations in P_n^X depends only on the size of X. Let $X, X' \subseteq [n]$ with $|X| = |X'| = k$. Let $X = \{x_1, \ldots, x_k\}$ with $x_1 < \cdots < x_k$ and let $X' = \{x'_1, \ldots, x'_k\}$ with $x'_1 < \cdots < x'_k$. Define a mapping $h : X \to X'$ that maps x_i to x'_i. Now define a mapping $\mathsf{T} : P_n^X \to P_n^{X'}$ which maps $\sigma = (X, f)$ to $\sigma' = (X', f')$

with $f'(\cdot) = h(f'(\cdot))$. The monotonicity of h implies that $\mathrm{wst}(\sigma) = \mathrm{wst}(\sigma')$. In particular, σ is non-waste if and only if σ' is non-waste. It is also clear that T is a bijection. Hence, the number of non-waste permutations in P_n^X is the same as in $P_n^{X'}$.

So, it is enough to consider only sets $X = [k]$ for $k = 0, \dots, n$. Let s_k be a number of non-waste permutations in $P_n^{[k]}$. Clearly, $s_0 = 1$ since the "empty" partial permutation, the only element of P_n^0, is non-waste. However, $s_1 = 0$ since the only element of $P_n^{[1]}$ is waste.

Now fix some k and let $X = [k]$. For every $i, j \in X$, let q_{ij} denote the number of waste partial permutations $\sigma = (X, f) \in P_n^X$ with $\min(\mathrm{wst}(\sigma)) = i$ and $f(i) = j$. Let $q_i = \sum_{j=1}^k q_{ij}$. Fix some i and j and count q_{ij}. There are $\binom{k-j}{i-1}$ choices for the set $f([i-1])$. There are s_{i-1} ways to choose an order on this set since i should be minimal waste index. If $i < k$ then there are $(j-1)$ choices for $f(i+1)$ and $(k-i-1)!$ choices for order on set $f([i+2, k])$. Hence

$$q_{ij} = \begin{cases} \binom{k-j}{i-1} s_{i-1}(j-1)(k-i-1)!, & \text{when } i < k; \\ s_{k-1}, & \text{when } i = k \text{ and } j = 1; \\ 0, & \text{when } i = k \text{ and } j > 1. \end{cases}$$

Hence, for $i < k$ we have

$$\begin{aligned} q_i &= \sum_{j=1}^{k} \binom{k-j}{i-1} s_{i-1}(j-1)(k-i-1)! \\ &= s_{i-1}(k-i-1)! \sum_{j=1}^{k} \binom{k-j}{i-1}(j-1) \\ &= s_{i-1}(k-i-1)! \binom{k}{k-i-1} \\ &= s_{i-1} \frac{(k-i-1)!k!}{(k-i-1)!(i+1)!} \\ &= s_{i-1} \frac{k!}{(i+1)!}, \end{aligned}$$

while $q_k = s_{k-1}$. Hence, we get the following recurrent relation:

$$s_k = k! - \sum_{i=1}^{k} q_i = k! - k! \sum_{i=1}^{k-1} \frac{s_{i-1}}{(i+1)!} - s_{k-1}. \tag{3}$$

We prove that $s_k = !k$ by induction on k. Note that $s_0 = !0$ and $s_1 = !1$. Suppose that for some $k_0 \geq 2$, equation $s_k = !k$ holds for every $k < k_0$. Let $k = k_0$. Then

$$\begin{aligned}
s_k &= k! - k! \sum_{i=1}^{k-1} \frac{s_{i-1}}{(i+1)!} - s_{k-1} \\
&= k! - k! \sum_{i=1}^{k-2} \frac{s_{i-1}}{(i+1)!} - s_{k-2} - s_{k-1} \\
&= k! - k! \sum_{i=1}^{k-2} \frac{s_{i-1}}{(i+1)!} - ks_{k-2} + (k-1)s_{k-2} - s_{k-1} \\
&= k\left((k-1)! - (k-1)! \sum_{i=1}^{k-2} \frac{s_{i-1}}{(i+1)!} - s_{k-2}\right) + (k-1)s_{k-2} - s_{k-1} \\
&= ks_{k-1} + (k-1)s_{k-2} - s_{k-1} \qquad \text{by (3)} \\
&= (k-1)\left(!(k-1) + !(k-2)\right) \\
&= !k,
\end{aligned}$$

where the last equality can be obtained from the definition of subfactorial. The proof is completed. $\square$

Remark. It is well-known that $!n$ is a number of derangements (permutations with no fixed points) in S_n. Moreover, for every $n \geq 1$ and every $k \in [n]$, the number of permutations from S_n that fix exactly k points is $!(n-k)\binom{n}{k}$. It can be shown that the number of partial permutation from P_n^n with exactly k waste indices is also $!(n-k)\binom{n}{k}$. Unfortunatelly, we were unable to find any natural bijection between S_n and P_n^n that maps derangements to non-waste permutations.

For every k define a (lexicographical) order on P_n^k as follows. Let $\sigma_1, \sigma_2 \in P_n^k$ with $\sigma_1 = (X_1, f_1)$ and $\sigma_2 = (X_2, f_2)$. Then

$$\sigma_1 <_{P_n^k} \sigma_2 \iff \exists i \in [k] : \; f_1(i) < f_2(i) \text{ and } \forall j < i \; f_1(j) = f_2(j). \tag{4}$$

Now extend these orders to a linear order on P_n by putting $P_n^0 < P_n^1 < \cdots < P_n^n$. For every $\sigma \in P_n$ let $P_{n,\sigma} = \{\sigma' \in P_n : \; \sigma' <_{P_n} \sigma\}$.

Lemma 5. *Let $\pi = (Y, g) \in P_n^k$ be a waste partial permutation and let $i \in \mathrm{wst}(\pi)$. Let $y = g(i)$, $X = Y \setminus \{y\}$, and let $\sigma \in P_n^X$ be a (unique) partial permutation such that π extends σ. Then π is the largest element in $E(\sigma, Y)$ with respect to the order on P_n^k defined above.*

Proof. First note that $E(\sigma, Y)$ contains exactly $k = |Y|$ partial permutations which differ only by the placement of y. Choose any $\tilde{\pi} = (Y, \tilde{g}) \in E(\sigma, Y)$ and let $\tilde{\pi} \neq \pi$. Let $j = \tilde{g}^{-1}(y)$ be a position of y in $\tilde{\pi}$. There are two cases. Suppose $j < i$. Then $\tilde{\pi} <_{P_n^k} \pi$ since $\tilde{g}(l) = g(l)$ when $l < j$ and $\tilde{g}(j) = y = g(i) < g(j)$ by (2). Suppose $j > i$. Again $\tilde{\pi} <_{P_n^k} \pi$ since $\tilde{g}(l) = g(l)$ when $l < i$ and $\tilde{g}(i) = g(i+1) < g(i)$ by (2). This implies that π is the largest element in $E(\sigma, Y)$ as required. $\square$

The proof of the next lemma is obvious:

Lemma 6. *Let V be a vector space. Let $S \subseteq V$ be some set of vectors with linear order $\prec$ on it. For every $s \in S$ let $S_s = \{s' \in S : s' \prec s\}$. Let $T = \{t \in S : t \in \operatorname{span}(S_t)\}$. Then $\operatorname{span}(S) = \operatorname{span}(S \setminus T)$.*

Now we are ready to prove Theorem 3. Let $W_n \subseteq P_n$ be a set of all waste partial permutations. Let $\pi = (Y, g) \in W_n$, $i \in \operatorname{wst}(\pi)$, $y = g(i)$, $X = Y \setminus \{y\}$, and let $\sigma \in P_n^X$ be such that π extends σ. Equation (1) can be rewritten as

$$\mathsf{H}(\pi) = \mathsf{H}(\sigma) - \sum_{\tilde{\pi} \in E(\sigma, Y),\ \tilde{\pi} \neq \pi} \mathsf{H}(\tilde{\pi}). \tag{5}$$

Note that Lemma 5 implies that π is a largest element in $E(\sigma, Y)$. Moreover, $\sigma < \pi$ since $|\operatorname{set}(\sigma)| < |\operatorname{set}(\pi)|$. Hence, equation (5) implies that $\mathsf{H}(\pi)$ belongs to $\operatorname{span}(\mathsf{H}(P_{n,\pi}))$. Applying Lemma 6 with $V = V_n$, $S = \mathsf{H}(P_n)$, $T = \mathsf{H}(W_n)$ we obtain

$$\operatorname{span}(\mathsf{H}(P_n)) = \operatorname{span}\left(\mathsf{H}(P_n) \setminus \mathsf{H}(W_n)\right). \tag{6}$$

It is evident that

$$\operatorname{span}\left(\mathsf{H}(P_n) \setminus \mathsf{H}(W_n)\right) \subseteq \operatorname{span}\left(\mathsf{H}(P_n \setminus W_n)\right) \subseteq V_n. \tag{7}$$

On the other hand,

$$V_n = \operatorname{span}(\mathsf{H}(P_n^n)) \subseteq \operatorname{span}(\mathsf{H}(P_n)). \tag{8}$$

Combining (6), (7) and (8) together obtain

$$V_n = \operatorname{span}(\mathsf{H}(P_n \setminus W_n)).$$

However, the number of non-waste partial permutations from P_n is $n!$ provided by Lemma 4 and

$$n! = \sum_{k=0}^{n} !k \binom{n}{k}.$$

This implies that $\mathsf{H}(P_n \setminus W_n)$ is a basis of V_n. Putting $Q_n = P_n \setminus W_n$ we obtain the first requirement of the theorem. The second one is given by Lemma 4. This completes the proof of Theorem 3.

3 Proof of Theorem 2

Now we show how the described properties of partial permutations can be used for proving Theorem 2.

Let $\mathcal{F} = \{\pi_1, \pi_2, \ldots, \pi_m\} \subseteq S_n$ be a k-rankwise independent family. Let Q_n be as in Theorem 2. Let

$$L = \bigcup_{i=0}^{\lfloor k/2 \rfloor} P_n^k.$$

If k is odd let

$$L' = \{\sigma \in P_n^{(k+1)/2} : 1 \in \mathrm{set}(\sigma)\}.$$

Put $Q = Q_n \cap L$ when k is even and $Q = Q_n \cap (L \cup L')$ when k is odd.

Assuming that S_n is identified with P_n^n in a natural way, we define a mapping $\mathsf{G} : Q \to \mathbb{C}^{|\mathcal{F}|}$ as follows:

$$\mathsf{G}(\sigma)_i = \begin{cases} \sqrt{\mathrm{Pr}_{\mathcal{F}}(\pi_i)}, & \text{if } \pi_i \text{ extends } \sigma; \\ 0, & \text{otherwise.} \end{cases}$$

We claim that

$$\forall \tau_1, \tau_2 \in Q \qquad \mathsf{G}(\tau_1) \cdot \mathsf{G}(\tau_2) = \frac{1}{n!} \mathsf{H}(\tau_1) \cdot \mathsf{H}(\tau_2), \tag{9}$$

where $\cdot$ denotes the dot product in both vector spaces. Fix $\tau_1, \tau_2 \in Q$. Let $X = \mathrm{set}(\tau_1) \cup \mathrm{set}(\tau_2)$. Note that $|X| \le k$ by the definition of Q. Rewriting (1) we obtain

$$\mathsf{H}(\tau_i) = \sum_{\sigma \in E(\tau_i, X)} \mathsf{H}(\sigma) \tag{10}$$

for $i = 1, 2$. Moreover, for any $\sigma_1, \sigma_2 \in P_n$ with $\mathrm{set}(\sigma_1) = \mathrm{set}(\sigma_2)$ and $\sigma_1 \ne \sigma_2$, we have $\mathsf{H}(\sigma_1) \cdot \mathsf{H}(\sigma_2) = 0$ by Lemma 2. Hence

$$\mathsf{H}(\tau_1) \cdot \mathsf{H}(\tau_2) = \sum_{\sigma \in E(\tau_1, X) \cap E(\tau_2, X)} \mathsf{H}^2(\sigma), \tag{11}$$

where $\mathsf{H}^2(\sigma)$ stands for scalar product $\mathsf{H}(\sigma) \cdot \mathsf{H}(\sigma)$. Note that $\mathsf{H}^2(\sigma)$ is the number of partial permutations from P_n^n that extend σ. By Lemma 1, $\mathsf{H}^2(\sigma) = \frac{n!}{|X|!}$. Hence

$$\frac{1}{n!} \mathsf{H}(\tau_1) \cdot \mathsf{H}(\tau_2) = \frac{1}{|X|!} \big| E(\tau_1, X) \cap E(\tau_2, X) \big|. \tag{12}$$

Similarly to (10) and (11) we have

$$\mathsf{G}(\tau_i) = \sum_{\sigma \in E(\tau_i, X)} \mathsf{G}(\sigma) \tag{13}$$

for $i = 1, 2$ and

$$\mathsf{G}(\tau_1) \cdot \mathsf{G}(\tau_2) = \sum_{\sigma \in E(\tau_1, X) \cap E(\tau_2, X)} \mathsf{G}^2(\sigma). \tag{14}$$

By the definition of G, for every $\sigma \in P_n^X$ we have

$$\mathsf{G}^2(\sigma) = \sum_{\substack{\pi \in \mathcal{F}, \\ \pi \text{ extends } \sigma}} \mathrm{Pr}_{\mathcal{F}}(\pi) = \frac{1}{|X|!}. \tag{15}$$

The second equality follows from k-rankwise independence of $\mathcal{F}$ since

$$|\,\mathrm{set}(\sigma)| = |X| \le k.$$

Combining (14) and (15) we obtain

$$\mathsf{G}(\tau_1) \cdot \mathsf{G}(\tau_2) = \frac{1}{|X|!}\big|E(\tau_1, X) \cap E(\tau_2, X)\big|. \tag{16}$$

Now (12) and (16) imply (9). From Theorem 3 it follows that $\mathsf{H}(Q)$ is a set of linearly independent vectors in V_n. By (9), the set of vectors $\mathsf{G}(Q)$ has the same Gram matrix up to a non-zero multiplicative constant. Hence, vectors of $\mathsf{G}(Q)$ are also linearly independent. The number of such vectors is not greater than the dimension of the corresponding vector space, i.e.,

$$|Q| = |\mathsf{G}(Q)| \leq |\mathcal{F}|.$$

Applying Lemma 4 for evaluating $|Q|$, we obtain the required lower bound for $|\mathcal{F}|$. The proof of Theorem 2 is completed.

References

[1] T. Itoh, Y. Takei, J. Tarui, *On Permutations with Limited Independence.* Proc. of the 11th Annual ACM-SIAM Symposium on Discrete Algorithms (2000), 137–146.