Set Systems and Families of Permutations with Small Traces

Otfried Cheong, Xavier Goaoc, Cyril Nicaud

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Set Systems and Families of Permutations with Small Traces

Otfried Cheong — Xavier Goaoc — Cyril Nicaud

N° 7154

December 2009

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--7154--FR+ENG

INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE

INRIA

centre de recherche
NANCY - GRAND EST

Set Systems and Families of Permutations with Small Traces

Otfried Cheong* , Xavier Goaoc† , Cyril Nicaud‡

Thème : Algorithmique, calcul certifié et cryptographie
Équipe-Projet Végas

Rapport de recherche n° 7154 — December 2009 — 11 pages

Abstract: We study the maximum size of a set system on n elements whose trace on any b elements has size at most k. We show that if for some $b \geq i \geq 0$ the shatter function f_R of a set system $([n], R)$ satisfies $f_R(b) < 2^i(b-i+1)$ then $|R| = O(n^i)$; this generalizes Sauer's Lemma on the size of set systems with bounded VC-dimension. We use this bound to delineate the main growth rates for the same problem on families of permutations, where the trace corresponds to the inclusion for permutations. This is related to a question of Raz on families of permutations with bounded VC-dimension that generalizes the Stanley-Wilf conjecture on permutations with excluded patterns.

Key-words: Set systems, VC dimension, Sauer Lemma, Permutation pattern

Otfried Cheong was supported by the Korea Science and Engineering Foundation Grant R01-2008-000-11607-0 funded by the Korea government. The collaboration between Otfried Cheong and Xavier Goaoc was supported by the INRIA *Équipe associée* KI.

* Theory of Computation Lab, Dept. of Computer Science, KAIST, Korea. otfried@kaist.edu
† Loria, INRIA Nancy Grand Est, France. goaoc@loria.fr
‡ LIGM Université Paris Est, France. nicaud@univ-mlv.fr

Centre de recherche INRIA Nancy – Grand Est
LORIA, Technopôle de Nancy-Brabois, Campus scientifique,
615, rue du Jardin Botanique, BP 101, 54602 Villers-Lès-Nancy
Téléphone : +33 3 83 59 30 00 — Télécopie : +33 3 83 27 83 19

Set Systems and Families of Permutations with Small Traces

Résumé : Nous étudions la taille maximale d'un hypergraphe à n sommets dont la trace sur toute sous-famille de b sommets est de taille au plus k. Nous montrons que pour tous entiers $b \geq i \geq 0$, si la fonction de pulvérisation f_R d'un hypergraphe $([n], R)$ satisfait $f_R(b) < 2^i(b - i + 1)$ alors $|R| = O(n^i)$; cela généralise le Lemme de Sauer sur la taille des hypergraphes de dimension de Vapnik-Chervonenkis bornée. Nous utilisons ensuite cette borne pour séparer les principaux régimes de croissance pour une question analogue sur les familles de permutations, où l'opération de trace correspnd à l'inclusion de motifs. Cela est relié à une question de Raz sur les familles de permutations à dimension de Vapnik-Chervonenkis bornée qui généralise la conjecture de Stanley-Wilf sur les permutations à motifs exclus.

Mots-clés : Hypergraphes, Dimension de Vapnik-Chervonenkis, Lemme de Sauer, Motifs de permutation

1 Introduction

In this paper, we study two problems of the following flavor: how large can a family of combinatorial objects defined on $[n] = \{1, \ldots, n\}$ be if its number of distinct "projections" on any small subset is bounded? We consider set systems, where the "projection" is the standard notion of trace, and families of permutations, where the "projection" corresponds to the notion of inclusion used in the study of permutations with excluded patterns.

Set systems. A *set system*, also called a *range space* or a *hypergraph*, is a pair (G, R) where G is a set, the *ground set*, and R is a set of subsets of G, the *ranges*. Since we will only consider finite set systems, our ground set will always be $[n]$. Given $X \subset [n]$, the *trace* of R on X, denoted $R_{|X}$, is the set $\{A \cap X \mid A \in R\}$. Given an integer b, let $\binom{R}{b}$ denote the set of b-tuples of R, and define:

$$f_R(b) = \max_{X \in \binom{[n]}{b}} |R_{|X}|.$$

The function f_R is called the *shatter function* of $([n], R)$, and counts the size of the largest trace on a subset of $[n]$ of size b. The first problem we consider is the following:

> **Question 1.** Given b and k, how large can a set system $([n], R)$ be if $f_R(b) \leq k$?

For $k = 2^b - 1$, the answer is given by Sauer's Lemma [15] (also proven independently by Perles and Shelah [17] and Vapnik and Chervonenkis [18]), which states that:

$$|R| \leq \sum_{i=0}^{b-1} \binom{n}{i} = O(n^{b-1}). \tag{1}$$

The largest b such that $f_R(b) = 2^b$ is known as the *VC-dimension* of $([n], R)$. The theory of set systems of bounded VC-dimension, and in particular Sauer's Lemma, has many applications, in particular in geometry and approximation algorithms; classical examples include the epsilon-net Theorem [7] or improved approximation algorithms for geometric set cover [6].

For the case of graphs, that is, set systems where all ranges have size 2, Question 1 is a classical problem known as a *Dirac-type problem*: what is the maximum number $Ex(n, m, k)$ of edges in a graph on n vertices whose induced subgraph on any m vertices has at most k edges? These problems were extensively studied in extremal graph theory since the 1960's, and we refer to the survey of Griggs et al. [11] for an overview. In the case of general set systems, the only results we are aware of are due to Frankl [9] and Bollobás and Radcliffe [5]. Specifically, Frankl proved that

$$f_R(3) \leq 6 \Rightarrow |R| \leq t_2(n) + n + 1 \quad \text{and} \quad f_R(4) \leq 10 \Rightarrow |R| \leq t_3(n) + n + 1,$$

where $t_i(n)$ denotes the number of edges of the Turán graph $T_i(n)$. Bollobás and Radcliffe showed that:

$$f_R(4) \leq 11 \Rightarrow |R| \leq \binom{n}{2} + n + 1 \quad \text{except for } n = 6.$$

There has also been interest in the case where $b = \alpha n$ and $b = n - \Theta(1)$; we refer to the article of Bollobás and Radcliffe [5] for an overview of these results.

Permutations. The notion of VC-dimension was extended to sets of permutations by Raz [14] as follows. Let σ be a permutation on $[n]$ and X some subset of $[n]$. The *restriction* of σ to X is the permutation $\sigma_{|X}$ of X such that for any $u, v \in X$, $\sigma_{|X}^{-1}(u) < \sigma_{|X}^{-1}(v)$ whenever $\sigma^{-1}(u) < \sigma^{-1}(v)$; if we consider a permutation as an ordering, $\sigma_{|X}$ is simply the order induced on X by σ. This allows to define the *shatter function* of a set F of permutations similarly:

$$\phi_F(m) = \max_{X \in \binom{[n]}{m}} |F_{|X}|.$$

The VC-dimension of F is then the largest m such that $\phi_F(m) = m!$, and the analogue of Question 1 arises naturally for sets of permutations:

> **Question 2.** Given m and k, how large can a set F of permutations on $[n]$ be if $\phi_F(m) \leq k$?

Raz [14] showed that any family of permutations on $[n]$ such that $\phi_F(3) < 6$ has size at most exponential in n, and asked whether the same holds whenever $k < m!$.

This problem is related to classical questions on families of permutations with *excluded pattern*. A permutation σ on $[n]$ *contains* a permutation τ on $[m]$ if there exists $a_1 < a_2 < \ldots < a_m$ in $[n]$ such that $\sigma^{-1}(a_i) < \sigma^{-1}(a_j)$ whenever $\tau^{-1}(i) < \tau^{-1}(j)$. If no permutation in a family F contains τ then F *avoids* τ and τ is an *excluded pattern* for F. The study of families of permutations with excluded patterns goes back to a work of Knuth [12], motivated by sorting permutations using queues, and received considerable attention over the last decades. In particular, Stanley and Wilf asked whether for any fixed permutation τ the number of permutations on $[n]$ that avoid τ is at most exponential in n, a question answered in the positive by Marcus and Tardos [13]. If a family of permutations has VC-dimension at most $m-1$ then for any m-tuple $X \subset [n]$ there is a permutation $\sigma(X)$ on $[m]$ which is forbidden for restrictions *to* X. In that sense, Raz's question generalizes that of Stanley and Wilf.

Our results. In this paper, we generalize Sauer's Lemma, and show that for any range space $([n], R)$, if $f_R(b) < 2^i(b - i + 1)$ for some $b > i \geq 0$ then $|R| = O(n^i)$ (Theorem 2). We then prove that the condition $f_R(b) = k$ is in fact *equivalent* to a Dirac-type problem on graphs for $k \leq 8 + 3\lfloor \frac{b-3}{2} \rfloor + s(b)$, where $s(b) = 1$ when b is even and 0 otherwise (Lemma 3). It follows that some conditions $f_R(b) = k$ lead to growth rates with fractional exponents (Corollary 4), a behavior not captured by Theorem 2. Finally, we give a reduction of the permutation problem to the set system problem (Lemma 5) from which we deduce the main transitions between the constant, polynomial and at least exponential behaviors for Question 2.

2 Set systems

In this section we give bounds on the size of a set R of ranges on $[n]$ with a given $f_R(b)$. Recall that a set system $([n], R)$ is *ideal*, also called *monotone decreasing*[1], if for any $B \subset A \in R$ we have $B \in R$. The next lemma was proven, independently, by Alon [1] and Frankl [9].

Lemma 1. *For any set system $([n], R)$ there exists an ideal set system $([n], \tilde{R})$ such that $|R| = |\tilde{R}|$ and for any integer b we have $f_{\tilde{R}}(b) \leq f_R(b)$.*

This can be shown by defining, for any $x \in [n]$, the operator (also called a *push-down* or a *compression*)

$$\tilde{T}_x(R) = \{A \setminus \{x\} \mid A \in R\} \cup \{A \mid A \in R \text{ such that } x \in A \text{ and } A \setminus \{x\} \in R\},$$

that removes x from any range in R where that does not decrease the total number of sets. Then,

$$\tilde{R} = \tilde{T}_1\left(\tilde{T}_2\left(\ldots(\tilde{T}_n\,(R)\right)\ldots\right)$$

is one such ideal set. We refer to Bollobás [4, Chapter 17] and the survey of Füredi and Pach [10] for more details. An immediate consequence of Lemma 1 is that we can work with ideal set systems when studying our first question.

2.1 Sauer's Lemma for small traces

Define $\binom{n}{-1} = 0$ and consider the sequence $\upsilon_i(b) = 2^i(b-i+1)$ that interpolates between $b+1 = \upsilon_0(b)$ and $2^b = \upsilon_{b-1}(b)$. Our first result is the following generalization of Sauer's Lemma.

Theorem 2. *Let $b > i \geq 0$ be two integers. Any range space $([n], R)$ with $f_R(b) < \upsilon_i(b)$ has size $|R| = f_R(n) < \sum_{j=0}^{i}(b-j+1)\binom{n}{j}$*

Proof. By Bondy's Theorem [4], for any $b+1$ distinct ranges there exist b elements on which they have distinct trace. It follows that if $f_R(b) < b+1$ we also have $f_R(n) < b+1$ for any n, and the statement holds for $i = 0$. Also, from

$$\sum_{j=0}^{i}(b-j+1)\binom{b}{j} \geq (b-i+1)\sum_{j=0}^{i}\binom{b}{j} \geq (b-i+1)2^i = \upsilon_i(b),$$

we have that the statement holds for $n = b$ and any i.

Now, we fix b and assume that we have

$$f_R(b) < \upsilon_k(b) \quad \Rightarrow \quad f_R(t) < \sum_{j=0}^{k}(b-j+1)\binom{t}{j}$$

whenever $k < i$ or $k = i$ and $t < n$. Let $R' = R_{|[n-1]}$ denote the trace of R on $[n-1]$ and let D denote the ranges in R' that are the trace of two distinct ranges from R. Notice that:

$$|R| = |R'| + |D| \quad \text{and} \quad f_{R'}(b) < \upsilon_i(b). \tag{2}$$

[1] An ideal set system is also an *abstract simplicial complex* to which the empty set was added.

Since $D \subset R'$, we have that $|D_{|X}| \leq |R'_{|X}|$ and thus $|D_{|X}| \leq \frac{1}{2}|R_{|(X \cup \{n\})}|$. It follows that $f_D(b-1) \leq \left\lfloor \frac{f_R(b)}{2} \right\rfloor$. Now, from $\upsilon_i(b) = 2\upsilon_{i-1}(b-1)$ we get that:

$$f_D(b-1) < \upsilon_{i-1}(b-1). \tag{3}$$

From Equations (2) and (3) and the induction hypothesis we obtain:

$$|R| < \sum_{j=0}^{i}(b-j+1)\binom{n-1}{j} + \sum_{j=0}^{i-1}(b-1-j+1)\binom{n-1}{j}.$$

This rewrites as

$$|R| < b+1+\sum_{j=1}^{i}(b-j+1)\binom{n-1}{j} + \sum_{j=1}^{i}(b-j+1)\binom{n-1}{j-1}$$

and with $\binom{n-1}{j} + \binom{n-1}{j-1} = \binom{n}{j}$ we get

$$|R| < b+1+\sum_{j=1}^{i}(b-j+1)\binom{n}{j} = \sum_{j=0}^{i}(b-j+1)\binom{n}{j},$$

and thus:

$$f_R(b) < \upsilon_i(b) \quad \Rightarrow \quad f_R(n) < \sum_{j=0}^{i}(b-j+1)\binom{n}{j}.$$

The statement follows by induction. □

Now, consider the following family of lower bounds. For $i = 1, \ldots, b$ let

$$\lambda_i(b) = \max_{b=b_1+\ldots b_i} \prod_{j=1}^{i}(b_j+1)$$

and consider the system $([n], R)$ where R is obtained by splitting $[n]$ into i roughly equal subsets and picking all i-tuples containing one element from each subset. Notice that $|R| = \Omega(n^i)$ and that $f_R(b) \leq \lambda_i(b)$. The same holds for $i = 0$ with $\lambda_0(b) = 1$. Thus, for any k such that $\lambda_i(b) \leq k < \upsilon_i(b)$, the maximum size of a set system $([n], R)$ with $f_R(b) = k$ is $\Theta(n^i)$.

b $\|R\|$	$\upsilon_0(b)$-1 $O(1)$	$\lambda_1(b)$ $\Omega(n)$	$\upsilon_1(b)$-1 $O(n)$	$\lambda_2(b)$ $\Omega(n^2)$	$\upsilon_2(b)-1$ $O(n^2)$	$\lambda_3(b)$ $\Omega(n^3)$	$\upsilon_3(b)-1$ $O(n^3)$	$\lambda_4(b)$ $\Omega(n^4)$	$\upsilon_4(b)-1$ $O(n^4)$	$\lambda_5(b)$ $\Omega(n^5)$
2	2	3								
3	3	4	5	6						
4	4	5	7	9	11	12				
5	5	6	9	12	15	18	23	24		
6	6	7	11	16	19	27	31	36	47	48

Table 1: The values $\upsilon_i(b)$ and $\lambda_i(b)$ for small b. Gaps appear in red.

In particular, the order of magnitude given by Theorem 2 is tight for all $b \leq 4$, with the exception of set systems with $f_R(4) = 8$.

Remark. Observe that the condition that $f_R(b) < \upsilon_i(b)$ does not imply that R has VC-dimension at most i. A simple example is given by

$$R = \{A \mid A \subset [i]\} \cup \{\{x\} \mid x \in [n]\},$$

which has VC-dimension i and for which $f_R(b) = 2^i + b - i - 1$ is smaller than $\upsilon_{i-1}(b) = 2^{i-1}(b - i)$ for b large enough.

2.2 Equivalence with Dirac-type problems

Recall that $Ex(n, m, k)$ denotes the maximum number of edges in a graph on n vertices whose induced subgraph on any m vertices has at most k edges. Let $\zeta(b) = 8 + 3\lfloor \frac{b-3}{2} \rfloor + s(b)$ where $s(b) = 1$ if b is even and 0 otherwise.

Lemma 3. *For any $b \geq 3$, the maximal size of a set system $([n], R)$ with $f_R(b) = \zeta(b) - 1$ is $Ex(n, b, \zeta(b) - b - 2) + n + 1$.*

Proof. By Lemma 1 it suffices to prove the statement for ideal set systems. Let $([n], R)$ be an ideal set system with $f_R(k) < \zeta(b) - 1$ and maximal size. If R contains some range A of size 3, then $|R_{|A}| = 8$. Now, write $b = 3 + 2j + s$ with $s \in \{0, 1\}$. Let B denote the set A augmented by j pairs of elements that belong to R, and one single element of R if $s = 1$. The set B has size b and the trace of R on B has size at least $8 + 3j + s = \zeta(b)$. Thus, if $f_R(b) < \zeta(b)$ we get that R contains no triple, and can thus be decomposed into

$$R = \{\emptyset\} \cup V \cup E,$$

where V are the singletons and E the pairs in R; call the former the *vertices* of R and the latter its *edges*. If some element $x \in [n]$ is not a singleton of V then it is contained in no range of R, and we can delete it without changing the size of R; this contradicts the maximality of R. Now, notice that the trace of R on any b elements contains at most $f_R(b) - b - 1 = \zeta(b) - b - 2$ edges, since it contains the empty set and each of the b vertices. Conversely, let $G = ([n], E)$ be a graph whose induced graph on any b vertices has at most $\zeta(b) - b - 2$ edges. If $R = \{\emptyset\} \cup [n] \cup E$ then the set system $([n], R)$ satisfies $f_R(b) < \zeta(b)$ and the statement follows. $\square$

A graph whose induced subgraphs on any m vertices have at most $k < \lfloor \frac{m^2}{4} \rfloor$ edges cannot contain a $K_{\lfloor \frac{k}{2} \rfloor, \lfloor \frac{k}{2} \rfloor}$, and thus, by the Kővári-Sós-Turán Theorem, has at most $Ex(n, m, k) = O\left(n^{2 - \frac{1}{\lfloor \frac{k}{2} \rfloor}}\right)$ edges. It follows that:

$$Ex(n, 4, 3) = Ex(n, 5, 5) = O(n\sqrt{n}).$$

The classical constructions yielding bipartite graphs on n vertices with $\Theta(n\sqrt{n})$ edges and no $K_{2,2}$ show that this bound is best possible. Since $\zeta(4) = 9$, we get that the family of growth rates obtained by the conditions $f_R(b) = k$ does not only contain polynomial growth with integer exponents:

Corollary 4. *The largest set system $([n], R)$ with $f_R(4) = 8$ or $f_R(5) = 10$ has size $|R| = \Theta(n\sqrt{n})$.*

Note that Lemma 3 can be extended into an equivalence of Question 1 and Dirac's problem on r-regular hypergraphs for arbitrary large r.

3 Families of permutations

In this section we give bounds on the size of a family F of permutations on $[n]$ with a given $\phi_F(b)$.

Reduction to set systems. An *inversion* of a permutation σ on $[n]$ is a pair of elements $i < j$ such that $\sigma^{-1}(i) > \sigma^{-1}(j)$. The *distinguishing pair* of two permutations σ_1 and σ_2 is the lexicographically smallest pair $(i, j) \subset [n]$ that appears in different orders in σ_1 and σ_2, i.e. is an inversion for one but not for the other. If F is a family of permutations on $[n]$ we let I_F denote the set of distinguishing pairs of pairs of permutations from F. Given a permutation $\sigma \in F$, we let $R(\sigma)$ denote the set of elements of I_F that are inversions of σ, and let $R(F) = \{R(\sigma) \mid \sigma \in F\}$. Observe that $(I_F, R(F))$ is a range space and that R is a one-to-one map between F and $R(F)$. In particular $|F| = |R(F)|$.

Lemma 5. *$f_{R(F)}(\lfloor \frac{m}{2} \rfloor) \leq \phi_F(m)$ and $|I_F| \leq Ex(n, m, \phi_F(m) - 1)$.*

Proof. Consider $b = \lfloor \frac{m}{2} \rfloor$ elements $(p_1, \ldots, p_b)$ in I_F and assume there exists k ranges $R(\sigma_1), \ldots, R(\sigma_k)$ with distinct traces on $\{p_1, \ldots, p_b\}$. Then the restrictions of $\sigma_1, \ldots, \sigma_k$ on $X = \cup_{1 \leq i \leq b} p_i$ must also be pairwise distinct. Thus, $\phi_F(m) \geq k$ whenever $f_{R(F)}(\lfloor \frac{m}{2} \rfloor) \geq k$, and the statement follows.

Let $s(t)$ denote the maximum number of distinguishing pairs in a family of t permutations (on $[n]$). From

$$s(2) = 1 \quad \text{and} \quad s(t) \leq 1 + \max_{1 \leq i \leq t-1} \{s(i) + s(t-i)\},$$

we get that $s(t) \leq t - 1$ by a simple induction. This implies that in the graph $G = ([n], I_F)$, any m vertices span at most $\phi_F(m) - 1$ edges, and it follows that

$$|I_F| \leq Ex(n, m, \phi_F(m) - 1),$$

which concludes the proof. $\square$

A subquadratic I_F is not always possible: every pair is a distinguishing pair of the family of all permutations on $[n]$ that restrict to the identity on some $(n-1)$-tuple. For that family, $\phi_F(m) = (m-1)^2 + 1$.

Main transitions. We can now outline the main transitions in the growth rate of families of permutations according to the value of $\phi_F(m)$. Let $b = \lfloor \frac{m}{2} \rfloor$.

- If $\phi_F(m) \leq \lfloor \frac{m}{2} \rfloor$ then, by Lemma 5, $f_{R(F)}(b) \leq b$ and Theorem 2 with $i = 0$ yields that $|F| = |R(F)| = O(1)$.
- Assume that $\lfloor \frac{m}{2} \rfloor < \phi_F(m) < 2\lfloor \frac{m}{2} \rfloor$. Then, by Lemma 5, $f_{R(F)}(b) < 2b$ and Theorem 2 with $i = 1$ yields that $|F| = |R(F)| = O(|I_F|) = O(Ex(n, m, m-2)) = O(n)$. A matching lower bound is given by the family

  F_1 : all permutations on $[n]$ that differ from the identity by the transposition of a single pair of the form $(2i, 2i + 1)$,

  of size $1 + \lfloor \frac{n}{2} \rfloor$ and with $\phi_{F_1}(m) = \lfloor \frac{m}{2} \rfloor + 1$.

- If $\phi_F(m) < 2^{\lfloor \frac{m}{2} \rfloor}$ then, by Lemma 5, $f_{R(F)}(b) < 2^b$ and $(I_F, R(F))$ has VC-dimension at most $b-1$. It follows, from Sauer's Lemma, that $|F| = |R(F)| = O(|I_F|^{b-1})$, and since $|I_F| = O(n^2)$, we get that $|F|$ is $O\left(n^{2\lfloor \frac{m}{2} \rfloor - 2}\right)$.

- If $\phi_F(m) \geq 2^{\lfloor \frac{m}{2} \rfloor}$ then the family

  F_2 : all permutations on $[n]$ that differ from the identity by the transposition of any number of pairs of the form $\{2i, 2i+1\}$,

  of size $2^{\lfloor \frac{n}{2} \rfloor}$ and with $\phi_{F_2}(m) = 2^{\lfloor \frac{m}{2} \rfloor}$ shows that $|F|$ can be exponential in n.

If $\phi_F(m) = m$ then $|I_F| = O(Ex(n, m, m-1))$, which is superlinear and $O(n^{1+\frac{1}{\lfloor \frac{m}{2} \rfloor}})$ [11]. We have not found any example showing that F could have superlinear size. The main transitions are summarized in Table 2.

	$\phi_F(m) \leq \lfloor \frac{m}{2} \rfloor$	$\lfloor \frac{m}{2} \rfloor < \phi_F(m) < 2\lfloor \frac{m}{2} \rfloor$	$2\lfloor \frac{m}{2} \rfloor \leq \phi_F(m) < 2^{\lfloor \frac{m}{2} \rfloor}$	$2^{\lfloor \frac{m}{2} \rfloor} \leq \phi_F(m)$
$\|F\|$	$\Theta(1)$	$\Theta(n)$	$\Omega(n)$ and $O\left(n^{2\lfloor \frac{m}{2} - 2 \rfloor}\right)$	$\Omega(2^{\lfloor \frac{n}{2} \rfloor})$

Table 2: Maximum size of a family F of permutations as a function of $\phi_F(m)$.

Exponential upper bounds. Raz [14] proved that if $\phi_F(3) \leq 5$ then $|F|$ has size at most exponential in n. The following simple observation derives a similar bounds for a few other values of $\phi_F(m)$.

Lemma 6. *If $|F|$ is at most exponential whenever $\phi_F(m-1) \leq k-1$ then $|F|$ is at most exponential whenever $\phi_F(m) \leq k$.*

Proof. Let $T(n, m, k)$ denote the maximum size of a family F such that $\phi_F(m) \leq k$. Assume that $\phi_F(m) = k = \phi_F(m-1)$ as otherwise the statement trivially holds. Let $X \in \binom{[n]}{m-1}$ such that $F_{|X} = \{\sigma_1, \ldots, \sigma_k\}$ has size k, and let:

$$F_i = \{\sigma \in F \mid \sigma_{|X} = \sigma_i\}.$$

Observe that F is the disjoint union of the F_i. Since $\phi_F(m) = k$, for any $e \in [n] \setminus X$ and for any $i = 1, \ldots, k$, there exists a unique permutation in $(F_i)_{|X \cup \{e\}}$ that restricts to σ_i on X. In other words, for every element in $[n] \setminus X$, the set $X \cup \{e\}$ appears in the same order in all permutations of F_i. It follows that

$$|F_i| = |(F_i)_{|[n] \setminus X}|,$$

that is, deleting X does not decrease the size of each F_i considered individually – although it may decrease the size of F. Now, let $G_i = (F_i)_{|[n] \setminus X}$ and consider the set system $([n] \setminus X, G_i)$. If $\phi_{G_i}(m-1) \leq k-1$ then $|G_i| \leq T(n-m+1, m-1, k-1)$, and otherwise $\phi_{G_i}(m-1) = k$ and we recurse. Altogether, we have the recursion

$$T(n, m, k) \leq k \max\left(T(n-m+1, m-1, k-1), T(n-m+1, m, k)\right),$$

and it follows that if $T(n, m-1, k-1)$ is at most exponential, so is $T(n, m, k)$. $\square$

It then follows, with Raz's result, that $|F|$ is at most exponential whenever $\phi_F(m) \leq m+2$. Table 3 tabulates our results for small values of m and $\phi_F(m)$, using the currently best known bounds on $Ex(n, k, \mu)$ we are aware of [11].

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$	$k=9$	$k=10$
$m=2$	$n!$	-	-	-	-	-	-	-	-
$m=3$	$2^{\Theta(n)}$	$2^{\Theta(n)}$	$2^{\Theta(n)}$	$2^{\Theta(n)}$	$n!$	-	-	-	-
$m=4$	2	$\lfloor\frac{2n}{3}\rfloor+1$	$2^{\Theta(n)}$	$2^{\Theta(n)}$	$2^{\Theta(n)}$	$2^{\Omega(n)}$	$2^{\Omega(n)}$	$2^{\Omega(n)}$	$2^{\Omega(n)}$
$m=5$	2	$\lfloor\frac{n}{2}\rfloor+1$	$2^{\Theta(n)}$	$2^{\Theta(n)}$	$2^{\Theta(n)}$	$2^{\Theta(n)}$	$2^{\Omega(n)}$	$2^{\Omega(n)}$	$2^{\Omega(n)}$
$m=6$	2	3	$\Theta(n)$	$\Theta(n)$	$O(n^3)$	$O(n^3)$	$2^{\Theta(n)}$	$2^{\Omega(n)}$	$2^{\Omega(n)}$
$m=7$	2	3	$\Theta(n)$	$\Theta(n)$	$O(n^2)$	$O(n^3)$	$2^{\Theta(n)}$	$2^{\Theta(n)}$	$2^{\Omega(n)}$
$m=8$	2	3	4	$\Theta(n)$	$\Theta(n)$	$\Theta(n)$	$O(n^{21/8})$	$O(n^{7/2})$	$O(n^{7/2})$
$m=9$	2	3	4	$\Theta(n)$	$\Theta(n)$	$\Theta(n)$	$O(n^2)$	$O(n^{7/2})$	$O(n^{7/2})$
$m=10$	2	3	4	5	$\Theta(n)$	$\Theta(n)$	$\Theta(n)$	$\Theta(n)$	$O(n^{27/10})$

Table 3: Maximum size of a family F of permutations on $[n]$ with $\phi_F(m) = k$.

4 Conclusion

A natural open question is the tightening of the bounds for both Questions 1 and 2. In particular, the first case where Lemma 5 no longer guarantees that the reduction from permutations to set systems leads to a ground set with linear size is $\phi_F(m) = m$; does that condition still imply that $|F|$ is $O(n)$ when m is large enough?

Raz's generalization of the Stanley-Wilf conjecture, that is, whether $\phi_F(m) < m!$ implies that $|F|$ is exponential in n, also appears to be a challenging question. Can it be tackled by a "normalization" technique similar to Lemma 1?

A line intersecting a collection C of pairwise disjoint convex sets in $\mathbb{R}^d$ induces two permutations, one reverse of the other, corresponding to the order in which each orientation of the line meets the set. The pair of these permutations is called a *geometric permutation* of C. One of the main open questions in geometric transversal theory [19] is to bound the maximum number of geometric permutations of a collection of n pairwise disjoint sets in $\mathbb{R}^d$ (see for instance [2, 3, 8, 16]). We can pick from each geometric permutation one of its elements so that the resulting family F has the following property: if any m members of C have at most k distinct geometric permutations then $\phi_F(m-2) \leq k$. One interesting question is whether bounds such as the one we obtained could lead to new results on the geometric permutation problem.

Acknowledgments

The authors thank Olivier Devillers and Csaba Tóth for their helpful comments.

References

[1] N. Alon. Density of sets of vectors. *Discrete Mathematics*, 46:199–202, 1983.

[2] B. Aronov and S. Smorodinsky. On geometric permutations induced by lines transversal through a fixed point. *Discrete & Computational Geometry*, 34:285–294, 2005.

[3] A. Asinowski and M. Katchalski. The maximal number of geometric permutations for n disjoint translates of a convex set in $\mathbb{R}^3$ is $\omega(n)$. *Discrete & Computational Geometry*, 35:473–480, 2006.

[4] B. Bollobás. *Combinatorics*. Cambridge University Press, 1986.

[5] B. Bollobás and A. J. Radcliffe. Defect Sauer results. *Journal of Combinatorial Theory. Series A*, 72(2):189–208, 1995.

[6] H. Bronnimann and M.T. Goodrich. Almost optimal set covers in finite VC-dimension. *Discrete & Computational Geometry*, 14:463–479, 1995.

[7] B. Chazelle. *The discrepancy method: randomness and complexity*. Cambridge University Press, 1986.

[8] O. Cheong, X. Goaoc, and H.-S. Na. Geometric permutations of disjoint unit spheres. *Computational Geometry: Theory & Applications*, 30:253–270, 2005.

[9] P. Frankl. On the trace of finite sets. *Journal of Combinatorial Theory. Series A*, 34:41–45, 1983.

[10] Z. Füredi and J. Pach. Traces of finite sets: Extremal problems and geometric applications. In *Extremal Problems for Finite Sets*, volume 3, pages 255–282. Bolyai Society, 1994.

[11] J. R. Griggs and M. Simonovits , G. Rubin Thomas. Extremal graphs with bounded densities of small subgraphs. *Journal of Graph Theory*, 3:185–207, 1998.

[12] Donald E. Knuth. *Art of Computer Programming, Volume 1: Fundamental Algorithms (3rd Edition)*. Addison-Wesley Professional, 1997.

[13] A. Marcus and G. Tardos. Excluded permutation matrices and the Stanley–Wilf conjecture. *Journal of Combinatorial Theory. Series A*, 107(1):153–160, 2004.

[14] R. Raz. VC-dimension of sets of permutations. *Combinatorica*, 20:255, 2000.

[15] N. Sauer. On the density of families of sets. *Journal of Combinatorial Theory. Series A*, 13:145–147, 1972.

[16] M. Sharir and S. Smorodinsky. On neighbors in geometric permutations. *Discrete Mathematics*, 268:327–335, 2003.

[17] S. Shelah. A combinatorial problem; stability and order for models and theories in infinitary languages. *Pacific J. Math*, 41(1):247–261, 1972.

[18] V. N. Vapnik and A. Ya. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory Probab. Appl.*, 16(2):264–280, 1971.

[19] R. Wenger. Helly-type theorems and geometric transversals. In Jacob E. Goodman and Joseph O'Rourke, editors, *Handbook of Discrete & Computational Geometry*, chapter 4, pages 73–96. CRC Press LLC, Boca Raton, FL, 2nd edition, 2004.

Centre de recherche INRIA Nancy – Grand Est
LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex
Centre de recherche INRIA Grenoble – Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq
Centre de recherche INRIA Paris – Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex
Centre de recherche INRIA Rennes – Bretagne Atlantique : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex
Centre de recherche INRIA Saclay – Île-de-France : Parc Orsay Université - ZAC des Vignes : 4, rue Jacques Monod - 91893 Orsay Cedex
Centre de recherche INRIA Sophia Antipolis – Méditerranée : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399