On families of weakly cross-intersecting set-pairs

Zoltán Király[a,1], Zoltán L. Nagy[a], Dömötör Pálvölgyi[a], Mirkó Visontai[b,2]

[a]*Department of Computer Science, Eötvös University Budapest, Pázmány Péter sétány 1/C. H-1117 Hungary*

[b]*Department of Mathematics, University of Pennsylvania, 209 S. 33rd Street, Philadelphia PA, 19104 USA*

1. Introduction

Let $\mathcal{F}$ be a family of pairs of sets. We call it an (a, b)-set system if for every set-pair (A, B) in $\mathcal{F}$ we have that $|A| = a,\ |B| = b,\ A \cap B = \emptyset$. The following classical result on families of cross-intersecting set-pairs is due to Bollobás [6]. Let $\mathcal{F}$ be an (a, b)-set system with the cross-intersecting property, i.e., for $(A_i, B_i), (A_j, B_j) \in \mathcal{F}$ with $i \neq j$ we have that both $A_i \cap B_j$ and $A_j \cap B_i$ are non-empty. The maximum possible size of such a set system is $\binom{a+b}{a}$, independent of the size of the ground set, and this bound is sharp. Surprisingly, the same upper bound holds even if we relax the cross-intersecting property, namely if for $(A_i, B_i), (A_j, B_j) \in \mathcal{F}$ we only require that $A_i \cap B_j \neq \emptyset$ when $i < j$, as was shown in [7]. Several further generalizations were investigated in [2, 8, 10, 12]. For more details on the history and applications of this problem we refer to the surveys [14, 15] and Chapter 1 of [1].

In this paper we consider the following variant of the problem, introduced by Tuza [13].

Definition 1. *Let $\mathcal{F}$ be an (a, b)-set system. $\mathcal{F}$ is* weakly cross-intersecting *if for any $(A_i, B_i), (A_j, B_j) \in \mathcal{F}$ with $i \neq j$ we have that $A_i \cap B_j$ and $A_j \cap B_i$ are not both empty.*

We investigate the maximum possible size of such a system, which we denote by $g(a, b)$. In Section 2 we give an explicit construction based on lattice paths and prove our main result that $\liminf_{a+b\to\infty} g(a, b)/\binom{a+b}{a} \geq 2 - o(1)$. Here, and throughout the paper, we use $o(1)$ to denote that an expression tends to zero as $a+b$ tends to infinity. This is the first construction giving this lower bound, which is stronger than the previously best known $\left(2 - \frac{ab}{(a+b)(a+b-1)}\right)$. In Section 3 we recall the known upper bounds and introduce a fractional version of the problem. We prove an upper bound for the fractional version that matches the best know bound for the original problem. Finally, in Section 4 we present some computational results

Email addresses: `kiraly@cs.elte.hu` (Zoltán Király), `nagyzoltanlorant@gmail.com` (Zoltán L. Nagy), `dom@cs.elte.hu` (Dömötör Pálvölgyi), `mirko@math.upenn.edu` (Mirkó Visontai)

[1]Research of Z.K. was partially supported by OTKA grants CNK 77780 and CK 80124, and TÁMOP grant 4.2.1./B-09/1/KMR-2010-0003.

[2]Research of M.V. was supported by the Benjamin Franklin Fellowship of the University of Pennsylvania

for small values of a and b, in particular, we show that $g(2,2) = 10$. We conclude by posing several interesting open problems.

Before we start, we mention some related problems that arise in Graph Theory.

1.1. Economical bipartite edge-coverings

Let K_N denote the complete graph on N vertices. Suppose we are given K_N and a positive integer k. We would like to cover the graph's edges with a collection of complete bipartite graphs, $\mathcal{G}_j = \{(X_j \cup Y_j, E_j)\}_{j=1}^m$, such that every edge of K_N is covered by at most k bipartite graphs. This problem was studied by Alon [3] and he showed that the minimum possible number of bipartite graphs to cover all the edges of K_N is $m = \Theta(kN^{\frac{1}{k}})$.

A related question is when we require every vertex – instead of every edge – to be covered by only a limited number of bipartite graphs. More precisely, suppose we are given K_N and two positive integers a and b. We would like to cover K_N's edges with a collection of complete bipartite graphs, $\mathcal{G}_j = \{(X_j \cup Y_j, E_j)\}_{j=1}^m$, such that for every vertex $v \in K_N$ there are at most a bipartite graphs such that v belongs to X_j, and at most b bipartite graphs such that v belongs to Y_j. Note that such a covering might not exist, if N is large. A natural question is to determine the largest N for which we still have a covering. The following proposition exposes the connection to the maximal size cross-intersecting families.

Proposition 2. *If $\mathcal{G}_j = \{(X_j \cup Y_j, E_j)\}_{j=1}^m$ is a collection of complete bipartite graphs, that covers the edges of K_N in a way that every vertex is contained in at most a of the X_j's and in at most b of the Y_j's, then $N \leq g(a,b)$. Furthermore, this bound is sharp.*

Proof. We show how to construct a weakly cross-intersecting (a,b)-set system from such a covering. Let the elements of the ground set $\{p_j\}_{j=1}^m$ correspond to the bipartite graphs in the covering. And let the pairs of sets $\{A_i, B_i\}_{i=1}^N$ correspond to the vertices of the complete graph. Set $p_j \in A_i$ if $v_i \in X_j$ and set $p_j \in B_i$ if $v_i \in Y_j$. This way we ensured that all the sets A_i have cardinality at most a, and the sets B_i have cardinality at most b. Also, any two pairs of sets weakly cross-intersect at the element corresponding to the bipartite graph that covered the edge connecting their corresponding vertices. Finally, adding some additional arbitrary points to the sets containing less than a (respectively, b) elements, we get a weakly cross-intersecting (a,b)-set system.

Similarly, we can also construct a covering from a weakly cross-intersecting (a,b)-set system. This shows that the two problems are equivalent, so the upper bound is sharp. □

2. Lower bounds

We start with the following two claims made by Tuza [13].

Claim 3. $g(a,1) \geq 2a+1$.

Proof. Let $B_i = \{i\}$ and $A_i = \{i+j \mod (2a+1) \mid 1 \leq j \leq a\}$, for $i = 0,\ldots,2a$. □

Claim 4. $g(a,b) \geq g(a-1,b) + g(a,b-1)$.

Proof. Suppose we have a construction $\{(A'_i, B_i)\}$ of cardinality $g(a-1,b)$ and another construction $\{(A_j, B'_j)\}$ of cardinality $g(a,b-1)$. Let x be an element not contained in any of these sets. Then, $\{(A'_i \cup \{x\}, B_i)\} \cup \{(A_j, B'_j \cup \{x\})\}$ gives a construction for a (a,b)-set system of cardinality $g(a-1,b) + g(a,b-1)$. □

From these claims one can easily obtain the following corollary.

Corollary 5. $g(a,b) \geq 2\binom{a+b}{a} - \binom{a+b-2}{a-1} = \left(2 - \frac{ab}{(a+b)(a+b-1)}\right)\binom{a+b}{a}$.

Proof. We prove by induction. Let $\widetilde{g}(a,b) = 2\binom{a+b}{a} - \binom{a+b-2}{a-1}$. From Claim 3 we have that $g(a,1) \geq \widetilde{g}(a,1)$ for all a, and by symmetry that $g(1,b) \geq \widetilde{g}(1,b)$ for all b, which settles the base case. The inductive step follows from Claim 4, and the fact that $\widetilde{g}(a,b) = \widetilde{g}(a-1,b) + \widetilde{g}(a,b-1)$. □

From this, one immediately gets that $\lim_{a\to\infty} \widetilde{g}(a,a)/\binom{2a}{a} = \frac{7}{4}$, and hence an asymptotic lower bound on the ratio $g(a,a)/\binom{2a}{a}$. The following simple construction shows that the lower bound can be improved to $\liminf_{a\to\infty} g(a,a)/\binom{2a}{a} \geq 2 - o(1)$. Note that we use lim inf here as the limit might not exist.

Claim 6. $g(a,a) \geq (2 - \frac{1}{a+1})\binom{2a}{a}$.

Proof. Consider a ground set of $2a+1$ elements. We form a bipartite graph $(X \cup Y, E)$ the following way. Each vertex in X corresponds to an a-subset of the ground set, and Y is just a copy of X. We connect vertices $x \in X$ and $y \in Y$ if the corresponding sets are disjoint. The resulting graph is regular, so, by Kőnig's Theorem, it must have a perfect matching $\mathcal{M}$. For every edge uv of $\mathcal{M}$ we add a set-pair (A_i, B_i) to our system, where A_i corresponds to u in X and B_i corresponds to v in Y. It is easy to see that this family will be indeed weakly cross-intersecting, and has $\binom{2a+1}{a} = \frac{2a+1}{a+1}\binom{2a}{a} = (2 - \frac{1}{a+1})\binom{2a}{a}$ elements. □

In what follows, we give another explicit construction using lattice paths that improves on this bound and also works for general a and b. We obtain that $\liminf_{a+b\to\infty} g(a,b)/\binom{a+b}{b}$ is also at least $2 - o(1)$.

2.1. A construction using lattice paths

Now we describe the construction of an (a,b)-set system $\mathcal{F}$ for arbitrary a and b using lattice paths. Let $m = 2a + 2b - 1$ be the cardinality of the ground set. We will construct $\mathcal{F} = \{(A,B) : |A| = a,\ |B| = b,\ A, B \subseteq \{0,\ldots,m-1\}\}$.

Let $\mathcal{L}(a,b)$ denote a set of all lattice paths on an $a \times b$ grid from $(0,0)$ to (a,b), where each move is either to the right, or up; and the path is strictly below the diagonal except at the two endpoints. We will identify a path $\pi \in \mathcal{L}(a,b)$ with its binary representation, i.e., a permutation π of the multiset $\{0^a, 1^b\}$, using 0 and 1 for steps to the right and up, respectively. For each path $\pi \in \mathcal{L}(a,b)$ we create the following m set-pairs $\{(A_{\pi,i}, B_{\pi,i})\}_{i=0}^{m-1}$. Let $A_{\pi,i} = \{i + t \mod m \mid 1 \leq t \leq a+b,\ \pi(t) = 0\}$ and $B_{\pi,i} = \{i + t \mod m, \mid 1 \leq t \leq a+b,\ \pi(t) = 1\}$.

We have to show that the obtained $\mathcal{F}$ indeed satisfies Definition 1. Clearly, it is an (a,b)-set system by construction. To show the weakly cross-intersecting property, first observe that $A_{\pi,i} \cap B_{\sigma,i} \neq \emptyset$ for any $\pi \neq \sigma$ as the paths are different and hence there exists a t such that $\pi(t) = 0$ and $\sigma(t) = 1$. If $i \neq j$ then $A_{\pi,i} \cap B_{\sigma,j} \neq \emptyset$ or $A_{\sigma,j} \cap B_{\pi,i} \neq \emptyset$ follows from the fact that any proper suffix of a lattice path in $\mathcal{L}(a,b)$ cannot equal the prefix of another (or even the same) lattice path, and hence there is a t such that $\pi(t) \neq \sigma(t+i-j)$. One way to see this is to consider the ratio of the right and up steps: in the prefix this ratio has to be greater than a/b whereas in the suffix it has to be strictly less than that.

We note that for $\gcd(a,b) = 1$ this (a,b)-set system is *maximal*, in the sense that we cannot add another pair of sets to it without violating the weakly cross-intersecting property.

Unfortunately, there is no simple closed formula for $|\mathcal{L}(a,b)|$ for general (a,b). There is a generating function for the number of lattice paths that never cross the diagonal but are allowed to touch it [11]. We believe, however, that it is unlikely that one can obtain better asymptotics using that result. For our purposes, the following result for the special case of relative primes a and b will suffice.

Theorem 7 (Bizley [5])**.** *$|\mathcal{L}(a,b)| = \binom{a+b}{a}/(a+b)$ if $\gcd(a,b) = 1$.*

From Theorem 7 and the construction described above we obtain a better lower bound than in Corollary 5.

Theorem 8. *$g(a,b) \geq (2 - \frac{1}{a+b})\binom{a+b}{a}$, if $\gcd(a,b) = 1$.*

Proof. $g(a,b) \geq (2a+2b-1)|\mathcal{L}(a,b)| = (2a+2b-1)\binom{a+b}{a}/(a+b)$. □

Corollary 9. *$g(a,a-1) \geq (2 - \frac{1}{2a-1}) \cdot \binom{2a-1}{a}$.*

Using this we can improve Claim 6.

Corollary 10. *$g(a,a) \geq (2 - \frac{1}{2a-1}) \cdot \binom{2a}{a}$.*

Proof. Claim 4 and Corollary 9 gives $g(a,a) \geq 2g(a,a-1) = (2 - \frac{1}{2a-1}) \cdot \binom{2a}{a}$. □

We can extend this idea to a general (a,b) by using Claim 4 repeatedly until we decrease $a+b$ to a prime p for which we automatically have $\gcd(p-q,q) = 1$ for $0 < q < p$ and then invoke Theorem 8.

Theorem 11. *$\liminf_{a+b\to\infty} g(a,b) \geq (2 - o(1))\binom{a+b}{a}$.*

Proof. We start with a well-known result in Number Theory. There exists $\delta > 0$ and x_0 such that for all $x > x_0$ the interval $[x - x^{1-\delta}, x]$ contains a prime [9]. We will use this result for $x = a+b$. If either a or b is smaller than $x^{1-\delta}$ then by Corollary 5 we have $g(a,b)/\binom{a+b}{a} \geq 2 - \frac{x \cdot x^{1-\delta}}{x(x-1)} = 2 - o(1)$. Otherwise, we will use the result that there is a prime p such that $0 \leq x - p \leq x^{1-\delta} \leq a, b$. By applying Claim 4 several times, and then Theorem 8, we obtain that

$$
\begin{aligned}
g(a,b) &\geq \sum_{i=0}^{x-p} \binom{x-p}{i} g(a-(x-p)+i,\ b-i) \\
&= \sum_{i=0}^{a+b-p} \binom{a+b-p}{i} g(p-(b-i),\ b-i) \\
&\geq \sum_{i=0}^{a+b-p} \binom{a+b-p}{i} (2p-1)\binom{p}{b-i}/p \\
&= \frac{2p-1}{p}\binom{a+b}{b}.
\end{aligned}
$$

And since $(2p-1)/p \geq 2 - 1/(x - x^{1-\delta}) = 2 - o(1)$, we have proven the theorem. $\square$

3. Upper bounds

For the sake of completeness we also mention the upper bounds which appeared in [13].

Claim 12. *$g(a, 1) = 2a + 1$.*

For general a and b we do not have matching upper bounds. In fact, the following theorem (also appeared as an exercise in [4, p. 12]) gives the best known upper bound.

Theorem 13 (Tuza, [13])**.** *$g(a,b) < \frac{(a+b)^{a+b}}{a^a b^b}$.*

Proof. The proof is a standard application of the probabilistic method. Let $\{(A_i, B_i)\}_{i=1}^N$ be a weakly cross-intersecting (a,b)-set system. Take a random partition of all the elements of the ground set. For every element v, with probability $\frac{a}{a+b}$ put v in X and with probability $\frac{b}{a+b}$ put it in Y. We say that the set-pair (A_i, B_i) is *contained* in (X,Y), if $A_i \subseteq X$ and $B_i \subseteq Y$. The probability that a set-pair is contained in (X,Y) is $\frac{a^a b^b}{(a+b)^{a+b}}$. On the other hand, at most one set-pair (A_i, B_i) can be contained in any partition (X,Y). Thus $N \leq \frac{(a+b)^{a+b}}{a^a b^b}$. Moreover, since not all random partitions can contain an (a,b) set-pair (e.g., when $X = \emptyset$), we have a strict inequality. $\square$

We can strengthen the above bound in some sense. Let $g(a,b;m)$ denote the maximum possible size of a weakly cross-intersecting (a,b)-set system on a ground set of size m, and let $k = k(a,b;m) := \lceil \frac{am-b}{a+b} \rceil$. Then concentrating the distribution to sets X of size k we obtain the following theorem.

Theorem 14. *$g(a,b;m) \leq \frac{\binom{m}{k}}{\binom{m-a-b}{k-a}}$.*

As $\frac{\binom{m}{k}}{\binom{m-a-b}{k-a}} < \frac{(a+b)^{a+b}}{a^a b^b}$ for $m \geq a+b$, we get back Theorem 13 as a corollary of this bound.

A reason why Theorem 13 is not that simple to improve on in general is, that this bound is optimal if we allow *fractional* weakly cross-intersecting families.

3.1. Fractional weakly cross-intersecting families

We now define the fractional relaxation of the weakly cross-intersecting property.

Definition 15. *A weighted (a,b)-set system is an (a,b)-set system $\{(A_i,B_i)\}_{i=1}^N$ with non-negative weights λ_i associated to each set-pair (A_i,B_i).*

Definition 16. *An (a,b)-set system $\{(A_i,B_i)\}_{i\in I}$ is* separable *if $\left(\bigcup_{i\in I} A_i\right) \cap \left(\bigcup_{j\in I} B_j\right) = \emptyset$.*

Definition 17. *A weighted (a,b)-set system $\{(A_i,B_i)\}_{i=1}^N$ with weights λ_i is* fractionally weakly cross-intersecting *if for every $I \subseteq \{1,\dots,N\}$ whenever $\{(A_i,B_i)\}_{i\in I}$ is separable, then $\sum_{i\in I}\lambda_i \le 1$.*

We denote the maximum total value $\sum_{i=1}^N \lambda_i$ of such a system on a ground set of size m by $g^*(a,b;m)$. Note that $\lambda_i \le 1$, otherwise $I=\{i\}$ would already violate the definition. If we require each λ_i to be an integer, then we get back Definition 1.

The following theorem shows that the bound of Theorem 14 is in fact optimal for the fractional case.

Theorem 18. *If $k = k(a,b;m)$ then $g^*(a,b;m) = \frac{\binom{m}{k}}{\binom{m-a-b}{k-a}}$.*

Proof. First we show that this is an upper bound on $g^*(a,b;m)$ as well. For a subset X with $|X| = \binom{m}{k}$, let $I_X = \{j \mid A_j \subseteq X,\ B_j \cap X = \emptyset\}$. By definition, $\sum_{i\in I_X}\lambda_i \le 1$ for all such X. A fixed j is in I_X for exactly $\binom{m-a-b}{k-a}$ different choices of X.

For a construction, let the weighted (a,b)-set system consist of all possible disjoint pairs $\{(A_i,B_i)\}_{i=1}^N$, where $|A_i| = a$, $|B_i| = b$, so we will have $N = \binom{m}{a}\binom{m-a}{b}$. For all possible values of i, let $\lambda_i = \lambda = \frac{\binom{m}{k}}{\binom{m-a-b}{k-a}\binom{m}{a}\binom{m-a}{b}}$. Let $I \subseteq \{1,\dots,N\}$ be such that the sub-system $\{(A_i,B_i)\}_{i\in I}$ is separable. We claim that $\sum_{i\in I}\lambda_i \le 1$, i.e., that $|I| \le \frac{\binom{m-a-b}{k-a}\binom{m}{a}\binom{m-a}{b}}{\binom{m}{k}}$. Let $X = \bigcup_{i\in I} A_i$ and $\ell = |X|$. By definition, $|I| \le \binom{\ell}{a}\binom{m-\ell}{b}$. An easy calculation shows that $\binom{\ell}{a}\binom{m-\ell}{b}$ takes its maximum at $\ell = k$, and clearly $\binom{m}{k}\binom{k}{a}\binom{m-k}{b} = \binom{m}{a}\binom{m-a}{b}\binom{m-a-b}{k-a}$. □

4. Proving bounds using computer programs

Next we present a straightforward method for computing $g(a,b)$ for fixed a and b with the help of a computer program. If we know m, the size of the ground set (or at least an upper bound on the size), then we can find a maximum size weakly cross-intersecting (a,b)-set system as follows. First, generate all possible (a,b) set-pairs. Next create a graph whose vertices represent set pairs and edges are drawn between vertices if the corresponding set-pairs weakly cross-intersect. Finally, find a maximum clique in this graph (since cliques in the graph correspond to set systems with weakly cross-intersecting property by construction).

Let us denote the size of the smallest ground set on which it is possible to realize a weakly cross-intersecting (a,b)-set system with N set-pairs by $h(a,b;N)$ (if it exists). We are only interested in the case where $N > \binom{a+b}{b}$, consequently, $h(a,b;N) > a+b$.

Claim 19. $\binom{h(a,b;N)}{2} \le ab \cdot N$.

Proof. Suppose we have a construction of size N on a set of size m and $\binom{m}{2} > ab \cdot N$. Define a graph on the ground set in the following way. Connect x and y if there is an i such that A_i contains one of them and B_i the other one. By our assumption, there must be a pair of points that are not connected to each other. Contracting these points gives a smaller ground set, where still any two set-pairs weakly cross-intersect. It might happen that x and y are both in some A_j (or B_j), in this case just add to the shrunken set an arbitrary element that is not in $A_j \cup B_j$. Note that since $m > a + b$ we can always choose such an element. □

Theorem 20. $g(2, 2) = 10$.

Proof. From Claim 4 we have that $g(2, 2) \ge 10$. So, now we have left to prove that $g(2, 2) \le 10$. We prove by contradiction. Suppose, that there is a weakly cross-intersecting $(2, 2)$-set system with $N = 11$ set-pairs. Then Claim 19 implies that this $(2, 2)$-set system has a realization on a ground set with at most 9 elements. We performed an exhaustive search on the ground sets of sizes up to 9 with our computer program and found that no such $(2, 2)$-set system has cardinality 11 or more. Hence, we can conclude that the maximum size of a $(2, 2)$-set system is 10. □

Unfortunately, the running time of this brute force search grows superexponentially and we could not compute further values. We found a $(3, 2)$-set system of size 19, hence $g(3, 2) \ge 19$ (which also implies $g(3, 3) \ge 38$), thus our lower bound is not always optimal.

We would like to conclude with several interesting questions.

Problem 1. *Is* $g(a, a) = 2g(a - 1, a)$ *for all* $a \ge 2$*?*

Problem 2. *Is* $g(a, b) < 2\binom{a+b}{a}$*?*

Problem 3. *Is* $g(a, a) = o(2^{2a})$*?*

Acknowledgment. We thank Padmini Mukkamala for independently verifying the result that $g(2, 2) = 10$ and Christian Krattenthaler for bringing reference [11] to our attention.

References

[1] I. Anderson, Combinatorics of finite sets (2002), Dover.

[2] N. Alon, An extremal problem for sets with applications to graph theory, J. Combin. Theory Ser. A 40 (1985), pp. 82–89.

[3] N. Alon, Neighborly families of boxes and bipartite coverings, in: The Mathematics of Paul Erdős, R. L. Graham and J. Nesetril, eds., Springer Verlag, Vol II, Berlin (1997), 27–31.

[4] N. Alon and J. H. Spencer. The Probabilistic Method (2008), 3rd edition, Wiley.

[5] M. T. L. Bizley, Derivation of a new formula for the number of minimal lattice paths from $(0,0)$ to (km, kn) having just t contacts with the line; and a proof of Grossman's formula for the number of paths which may touch but do not rise above this line, J. Inst. Actuar. 80 (1954), pp. 55–62.

[6] B. Bollobás, On generalized graphs. Acta Math. Acad. Hungar. 16 (1965), pp. 447–452.

[7] P. Frankl, An extremal problem for two families of sets. European J. Combin. 3 (1982), pp. 125–127.

[8] Z. Füredi, Geometrical solution of a problem for two hypergraphs. European J. Combin. 5 (1984), pp. 133–136.

[9] G. Hoheisel, Primzahlenprobleme in der Analysis, Sitzungber. Berlin Math. Ges. (1930), pp. 550–558.

[10] L. Lovász, Combinatorial Problems and Exercises (1979), North-Holland.

[11] M. Sato, Generating functions for the number of lattice paths between two parallel lines with a fractional incline, Math. Japonica 24 (1989), pp. 123–137.

[12] J. Talbot, A new Bollobás-type inequality and applications to t-intersecting families of sets, Discrete Math. 285, 1-3, (2004), pp. 349–353.

[13] Zs. Tuza, Inequalities for two set systems with prescribed intersections, Graphs Combin. 3 (1987), pp. 75–80.

[14] Zs. Tuza, Applications of the set-pair method in extremal hypergraph theory. In: P. Frankl et al., Extremal Problems for Finite Sets, Bolyai Soc. Math. Stud. Vol. 3, (1994), pp. 479–514.

[15] Zs. Tuza, Applications of the set-pair method in extremal problems, II. In: D. Miklós et al., Combinatorics, Paul Erdős is Eighty, Bolyai Soc. Math. Stud. Vol. 2, (1996), pp. 459–490.