Canad. Math. Bull. Vol. 27 (3), 1984

BLOCK SIZES IN PAIRWISE BALANCED DESIGNS

BY
CHARLES J. COLBOURN,* KEVIN T. PHELPS
AND VOJTĚCH RÖDL†

ABSTRACT. The number of sets of integers which are realizable as block sizes of a pairwise balanced design of order n is between $\exp(c_1\sqrt{n})$ and $\exp(c_2\sqrt{n})$; in contrast, when the multiplicity of each block size is also specified, the number of multisets which can be realized is between $\exp(c_1\sqrt{n}\log n)$ and $\exp(c_2\sqrt{n}\log n)$. Although this gives a reasonable bound on the number of multisets which can be realized, a good characterization is not likely to exist; deciding whether a multiset can be so realized is NP-complete.

1. **Introduction.** A pairwise balanced design (PBD) of order n is an n-set V of elements together with a collection B of subsets of V called blocks. Each unordered pair of elements appears in precisely one block. The profile of a PBD is the multiset (i.e., set with multiplicities) of the sizes of its blocks; the profile set is simply the set of these sizes.

The problem of determining $f(n)$, the number of profiles of n-element PBD's, was first discussed by Erdös [1]. He further remarked that "it is perhaps not reasonable to expect to obtain a necessary and sufficient condition for a sequence $x_1, \ldots, x_n$ that there should be a block design (=PBD)" with these block sizes. We establish that $\exp(c_1\sqrt{n}\log n) \le f(n) \le \exp(c_2\sqrt{n}\log n)$. Furthermore, we confirm Erdös' remark on the difficulty of characterizing profiles by showing that the recognition of profiles is NP-complete. We also consider the related problem of determining the number $g(n)$ of profile sets of n-element PBD's; here we establish $\exp(c_1\sqrt{n}) \le g(n) \le \exp(c_2\sqrt{n})$.

2. **The number of profile sets.** In this section, we establish upper and lower bounds on the number $g(n)$ of profile sets of n-element PBD's. We first establish the lower bound:

LEMMA 2.1. *There exists a fixed constant c_1 for which* $\exp(c_1\sqrt{n}) \le g(n)$.

Proof. Partition the n points into $k = \sqrt{n}/2$ groups, each of size $2\sqrt{n}$. Select k distinct integers $i_1, \ldots, i_k$ between 1 and $\sqrt{n}$; on the jth group, place two

---

Received by the editors August 22, 1983.
* Research supported by NSERC Canada Grant A5047.
† On leave at Georgia Institute of Technology.
AMS Subject Classification (1980): 05B30, 05B05.

blocks, one of size $\sqrt{n}+i_j$, the other of size $\sqrt{n}-i_j$. All other pairs are covered by blocks of size 2. There are $\exp(c_1\sqrt{n})$ distinct ways to select k integers, and each leads to a different profile set. ■

The upper bound is more complicated, and requires a preliminary result:

LEMMA 2.2. *Let $A_1, \ldots, A_p$ be subsets of an n-set such that $|A_i \cap A_j| \le 1$ for all $1 \le i < j \le p$ (i.e., a partial* PBD). *Suppose that $n/2^{t+1} \le |A_i| \le n/2^t$ for some fixed integer t, $n > 2^{2t+3}$. Then $p \le 2^{t+2}$.*

Proof. Since blocks intersect in at most one element,

$$\sum_{i=1}^{p} |A_i| - \binom{p}{2} \le n.$$

Substituting the smallest possible value for $|A_i|$ and simplifying gives,

$$p^2 - \frac{pn}{2^t} + 2n \ge 0.$$

Since possible values of p form an interval, p must not exceed the smaller of the two roots of this inequality. (The existence of two real roots follows from our assumption, $n > 2^{2t+3}$.) Then

$$2p \le \frac{n}{2^t} - \sqrt{\left(\frac{n^2}{2^{2t}} - 8n\right)}$$

and hence

$$p \le \frac{4n}{n/2^t} = 2^{t+2}. \quad ■$$

This lemma provides the basis for establishing the upper bound:

LEMMA 2.3. *There is a fixed constant c_2 for which $g(n) \le \exp(c_2\sqrt{n})$.*

Proof. We partition $(1, n)$ into two subintervals; $(1, n/2^{t_0})$, where t_0 is the largest integer such that $n/2^{t_0} > \sqrt{8n}$, and $(n/2^{t_0}, n)$. In the interval $(1, n/2^{t_0})$ there are at most $2^{\sqrt{32n}}$ possible sets of block sizes. This follows as $n/2^{t_0} \le 2\sqrt{8n}$. We subdivide the interval $(n/2^{t_0}, n)$ into intervals

$$I_t = (n/2^{t+1}, n/2^t) \quad \text{for} \quad t = 0, 1, 2, \ldots, t_0 - 1.$$

The interval I_t has length $l_t = n/2^{t+1}$ and Lemma 2.2 ensures that at most $r_t \le 2^{t+2}$ block sizes are chosen in this interval. Then N_t, the number of possible sets of block sizes in the interval I_t, satisfies

$$N_t \le \binom{l_t}{r_t} \le \left(\frac{l_t e}{r_t}\right)^{r_t} = \left(\frac{ne}{2^{2t+3}}\right)^{2^{t+2}}$$

https://doi.org/10.4153/CMB-1984-057-0 Published online by Cambridge University Press

Then the number, N, of possible sets of block sizes in the interval $(n/2^{t_0}, n)$ is,

$$N=\prod_{t=0}^{t_0-1} N_t$$

Using the inequality $\binom{m}{a}\le\left(\frac{me}{a}\right)^a$, we have $N_t\le\left(\frac{ne}{2^{2t+3}}\right)^{2^{t+2}}$. Hence

$$N=\exp\left(\sum_{t=0}^{t_0-1} 2^{t+2}\log\left(\frac{ne}{2^{2t+3}}\right)\right)$$
$$=\exp\left(\sum_{t=0}^{t_0-1} 2^{t+2}(\log n+\log e-(2t+3))\right)$$

Since $\log n<2t_0+5$,

$$N\le\exp\left(2^3(t_0+2)\sum_{t=0}^{t_0-1} 2^t-8\sum_{t=0}^{t_0-1} t2^t\right)$$

The first summation yields $8(t_0+2)(2^{t_0}-1)$. In order to determine the second summation, let $s_a=2^3\sum_{t=0}^{a} t2^t$. Using the recurrence relation $s_a-2s_{a-1}=16(2^a-12)$, we find $s_{a-1}=a2^a-2(2^a-1)$. Hence $N\le\exp(2^{t_0}(16)-8(t_0+2))\le\exp(8\sqrt{n}-8\log n)$.

Lemmas 2.1 and 2.3 together establish:

THEOREM 2.4. *Let $g(n)$ be the number of profile sets of n-element* PBD'*s then* $\exp(c_1\sqrt{n})\le g(n)\le\exp(c_2\sqrt{n})$.

3. **The number of profiles.** In this section, we consider the problem suggested by Erdös [1] of determining the number $f(n)$ of profiles of n-element PBD's. We first establish a lower bound. This can be done rather easily if we allow blocks of size 2. If we insist that all blocks have size strictly larger than 2, the same lower bound can be achieved but the argument is more involved.

LEMMA 3.1. *There is a fixed constant c_1 for which* $\exp(c_1\sqrt{n}\log n)\le f(n)$.

Proof. It follows from the Prime Number Theorem that for any $\delta>0$, and $n>n_0(\delta)$, there exists a projective plane with at most $n(1+\delta)$ points. It is a simple matter to delete δn points so that each remaining block has at least $c\sqrt{n}$ points ($c\sim1-\delta$) and there are at least n blocks remaining as well. (Consider the affine plane and choose $\delta\sqrt{n}$ blocks from one parallel class.) Let $b_1, b_2, \ldots, b_m$, $m\ge n$ be the blocks which remain. From each b_j we choose a subset U_j of u_j points $3\le u_j\le c\sqrt{n}/4$ so that $|b_j|-u_j\equiv3 \pmod 6$. Thus, there are approximately $c\sqrt{n}/24$ choices of u_j. On the remaining $6t+3=|b_j\backslash U_j|$ points we construct a Kirkman triple system of order $6t+3$ and, selecting u_j different parallel classes, assign point $i\in U_j$ as a fourth point to each triple in the ith parallel class selected. The resulting collection of 4-tuples along with

https://doi.org/10.4153/CMB-1984-057-0 Published online by Cambridge University Press

the remaining triples and the set U_j cover all of the original pairs in b_j. For $j = 1, 2, \ldots, n$ we can choose U_j independently which gives at least $\exp(c'\sqrt{n}\log n)$ different sequences of block sizes.

Note, we have a similar lower bound even if we insist that all blocks have size at least k, for some fixed k; the argument would be almost identical except that we would use resolvable Steiner systems $S(2, p, v)$ for prime powers $p \geq k$. Though the spectrum for these designs is not completely known, the existence of affine geometries over GF(p) along with the fact that resolvability is PBD-closed is enough to ensure that the spectrum has positive density.

The upper bound can be established, as noted in [1] as follows:

LEMMA 3.2. *There is a fixed constant* c_2 *for which* $g(n) \leq \exp(c_2\sqrt{n}\log n)$.

Proof. Consider the number of blocks p in the interval $(10\sqrt{n}, n)$. We know that $10p\sqrt{n} - \binom{p}{2} \leq n$, and hence $p \leq \sqrt{n}/5$. Thus, the number of possible selections of blocks in this interval is less than $n^{c\sqrt{n}} = \exp(c'\sqrt{n}\log n)$. In the interval $(1, 10\sqrt{n})$, the number of selections of block sizes is bounded by the number of multisets of at most $\binom{n}{2}$ elements chosen from the interval $(1, 10\sqrt{n})$; this is $\exp(c\sqrt{n}\log n)$. ∎

Lemmas 3.1 and 3.2 together establish the main result of this section:

THEOREM 3.3. *The number* $f(n)$ *of profiles of* n*-element* PBD's *is asymptotically* $\exp(c\sqrt{n}\log n)$. *Moreover the number* $f_k(n)$ *of profiles on* n*-element* PBD's *having all blocks of cardinality at least* k *is bounded from below by* $\exp(c_k\sqrt{n}\log n)$.

REMARK. We could not decide whether c_k has to tend to zero as $k \to \infty$.

4. **Characterizing profiles.** Theorem 3.3 demonstrates that the number of multisets of integers which satisfy the basic necessary conditions is on the same order as the number of profiles. Thus, one might hope for a good characterization of which multisets are profiles (as one has in the case of degree sequences of graphs, for example [2]). We show that a good characterization is unlikely, since the recognition of profiles is almost certainly a difficult computational problem:

THEOREM 4.1. *Deciding whether a multiset is a profile is NP-complete.*

Proof. Membership in NP is immediate. To show completeness, we reduce the problem of 3-PARTITION [3] to profile recognition. An instance of the 3-PARTITION problem is a set of $3m$ integers $a_1, \ldots, a_{3m}$ whose sum is mB.

The problem is to determine whether the numbers can be partitioned into m groups, each of whose sum is B. Moreover, each number lies in the range $B/4 < a_i < B/2$, and hence each group must contain exactly three of the numbers.

We transform the 3-PARTITION problem into a clique packing problem, as follows. Let B' be the smallest number of the form $2^t + 1$ satisfying $B' \geq 2B$. Modify the set of $3m$ integers from the 3-partition problem by adding m numbers of the form $B' - B + 6$, and change the bound to B'. Note that each group must contain exactly one of these large numbers, and hence each group will consist of four integers. In this revised problem, we ask: can one pack cliques of the sizes given by the $4m$ integers specified into m disjoint complete graphs, each with B' vertices. In this packing, cliques can overlap in a single vertex, but not in a pair; hence the bound differs by 6 from the corresponding "sum of integers" problem. It is straightforward to verify that this clique packing problem is equivalent to the 3-PARTITION problem from which it is produced.

To complete the transformation to recognition of profiles, we note that, given a clique packing problem constructed in this way, there is a projective plane of order n with block size $n+1 = B'$. We construct a multiset by taking the $4m$ integers from the clique packing problem. To this, we add $n^2 + n + 1 - m$ integers each equal to $n+1$, the plane's block size. Finally, we add the integer 2 to the multiset sufficiently many times to cover all remaining pairs. When the clique packing problem has a solution, the multiset constructed in this way is indeed a profile: one takes $n^2 + n + 1 - m$ blocks from the plane, and in the remaining m cliques of size B, the other blocks are packed.

It is necessary to show when a PBD exists, there is a solution to the clique packing problem. This is ensured here by a fact about projective planes, that a partial projective plane of order n with n^2 blocks specified has a unique completion [4], [5]. Hence any realization of the multiset as a PBD induces a clique packing as long as $m > B'$. This reduces clique packing to the recognition of profiles, and the reduction can easily be done in polynomial time. ∎

References

1. P. Erdös, *Problems and Results on Block Designs and Set Systems*, Proceedings of the Thirteenth Southeastern Conference on Combinatorics, Graph Theory, and Computing, 1982, pp. 3–18.

2. P. Erdös and T. Gallai, *Graphs with prescribed degrees of vertices* (Hungarian), Mat. Lapok **11** (1960), 264–274.

3. M. R. Garey and D. S. Johnson, *Computers and Intractability*, Freeman and Sons, San Francisco, CA, 1979.

4. J. I. Hall, *Bounds for equidistant codes and partial projective planes*, Discrete Mathematics **17** (1977), 85–94.

5. S. A. Vanstone, *The extendibility of* $(r, 1)$ *designs*, Proceedings of the Third Manitoba Conference on Numerical Mathematics and Computing, 1973.

Department of Computational Science
University of Saskatchewan
Saskatoon, Saskatchewan S7N 0W0
Canada

School of Mathematics
Georgia Institute of Technology
Atlanta, Georgia 30332
U.S.A.

FJFI, CVUT
Husova 5
Praha 1
Czechoslovakia

https://doi.org/10.4153/CMB-1984-057-0 Published online by Cambridge University Press