

---

---

## Block designs with near-minimal number of blocks

---

---

Oswaldo Marrero

Oswaldo Marrero studied mathematics at the University of Miami, and biometry and statistics at Yale University. After holding various positions in academia and in industry, he is now a professor at Villanova University. He has published papers in epidemiology, mathematics, medicine, and statistics.

### 1 Introduction

For our purposes, we can begin by saying that a *combinatorial design* is a binary matrix on which we have imposed some conditions; this is a simple and easy way to introduce this concept. The mathematical study of these designs was promoted by problems in the design and analysis of experiments in statistics, especially since the 1930s. For the mathematical aspects, two classic references are the books by Hall [2] and by Ryser [8]; as examples of more recent references, we have the books by Ionin and Shrikhande [3], and by Van Lint and Wilson [4]. For the statistical aspects, a classic reference is the book by Box, Hunter, and Hunter [1]. Binary matrices and vectors are of interest in combinatorics, in the theories of coding and of information, and in cryptology; some applications and several mathematical results are discussed in [7].

The central problems are the existence and the construction of such designs. In general, these problems remain unsolved, but there are partial results [2, 3, 4, 8].

Also important are problems concerned with minimal or maximal properties of combinatorial designs. A classic example is Fisher's inequality [2, p. 129]. More recently, Marrero [5] obtained a maximal property for binary nonsingular matrices in which each row has the same number of ones. Also, Marrero and Pasles [6] have found the best possible bounds

Block-Designs sind mathematisch äquivalent zu gewissen binären Matrizen. Die damit verbundenen Fragen können daher leicht verständlich gemacht werden. Die Hauptprobleme bestehen im Nachweis der Existenz und Eindeutigkeit solcher Designs; in voller Allgemeinheit sind beide Probleme allerdings nach wie vor ungelöst. In der vorliegenden Arbeit werden einige neue Ergebnisse zur Existenz einer Klasse von Block-Designs für den Fall bewiesen, dass die Block-Anzahl nahe dem minimal möglichen Wert ist. Die Beweise sind einfach, im Detail ausgeführt und leicht nachvollziehbar.

for the number of coincidences of ones between two distinct rows of a nonsingular binary matrix of constant row sum.

In this paper we consider two well-known combinatorial designs that are defined next.

Let  $S = \{x_1, \dots, x_v\}$  be a finite set. A collection  $X_1, \dots, X_b$  of subsets of  $S$  is called a  $(b, v, r, k, \lambda)$ -design if:

- $|X_i| = k$  for each  $i = 1, \dots, b$ ;
- for each  $j = 1, \dots, v$ , the element  $x_j$  is contained in exactly  $r$  of the subsets  $X_1, \dots, X_b$ ;
- for each pair of distinct  $i, j = 1, \dots, v$ , the elements  $x_i, x_j$  appear together in  $\lambda$  of the subsets  $X_1, \dots, X_b$ ; and
- $0 \leq \lambda, 0 < r$ , and  $0 < k < v$ .

The choice of  $b$  and  $v$  as symbols for two of the parameters is related to the use of these designs in agricultural experiments, where it is desired to compare  $v$  varieties of a crop using  $b$  blocks. In general, the varieties are compared with respect to several types of observations or responses. Statistically, the “blocks” correspond to a factor (explanatory or independent variable) which is not of primary interest, but which is known – or believed – to have an appreciable effect on the responses to be measured. The blocks factor is generally a natural one, such as: locations in an agricultural experimental field, hospitals where a clinical trial is going to be held, manufacturing plants that make the same kind of object, etc. In statistics, this design is known as a *balanced incomplete block design*. “Incomplete” refers to the fact that not every variety appears in each block; this is necessary sometimes, as would happen, for example, when the agricultural field is not large enough to plant all varieties in every block. “Balanced” refers to two facts:

- (1) each block contains the same number  $k$  of varieties, and
- (2) each pair of distinct varieties appear together in the same number  $\lambda$  of blocks.

Necessary conditions for the existence of a  $(b, v, r, k, \lambda)$ -design are that  $bk = vr$  and  $r(k - 1) = \lambda(v - 1)$  [8, p. 97].

The subsets  $X_1, \dots, X_v$  of  $\{x_1, \dots, x_v\}$  are called a  $(v, k, \lambda)$ -design if:

- $|X_i| = k$  for each  $i = 1, \dots, v$ ;
- $|X_i \cap X_j| = \lambda$  for each pair of distinct  $i, j = 1, \dots, v$ ; and
- $0 \leq \lambda < k < v$ .

In statistics, this design is known as a *symmetric balanced incomplete block design*. “Symmetric” refers to the fact that the number of varieties is the same as the number of blocks.

In these two definitions, the last statement concerning conditions on the parameters  $b, v, r, k$ , and  $\lambda$  is there to avoid cases that are generally considered uninteresting. In general, such designs are known as *block designs*, and the blocks correspond to the subsets.

The *incidence matrix* of a design is the binary matrix  $A := (a_{ij})$  defined by

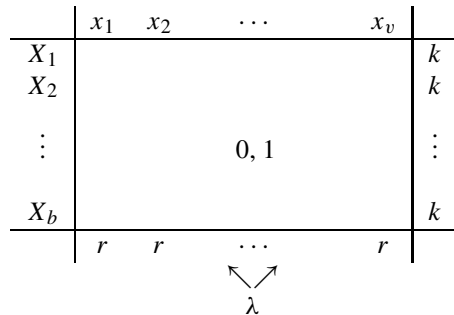
$$a_{ij} := \begin{cases} 1, & \text{if } x_j \in X_i, \\ 0, & \text{otherwise.} \end{cases}$$

A design is completely determined by its incidence matrix; that is why we began by saying that a combinatorial design is a certain binary matrix. In fact, if  $A^T$  is the transpose of the

matrix  $A$ ,  $I_m$  is the  $m \times m$  identity matrix, and  $J_{m,n}$  is the  $m \times n$  matrix having all its elements equal to 1, then

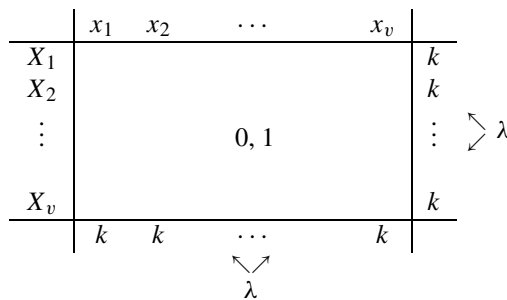
- a  $(b, v, r, k, \lambda)$ -design exists if and only if there exists a  $b \times v$  binary matrix  $A$  such that  $A^T A = (r - \lambda)I_v + \lambda J_{v,v}$  and  $A J_{v,1} = k J_{b,1}$ ; and
- a  $(v, k, \lambda)$ -design exists if and only if there exists a  $v \times v$  binary matrix  $A$  such that  $AA^T = (k - \lambda)I_v + \lambda J_{v,v}$ . It is not obvious, but it can be shown [8, Theorem 2.1, p. 103] that if  $A$  is the incidence matrix of a  $(v, k, \lambda)$ -design, then  $AA^T = A^T A$ . This means that, in  $A$ , the column sums and the row sums are all equal to  $k$ ; and the pairwise inner product of distinct column vectors and of distinct row vectors are all equal to  $\lambda$ . These properties are used in the proof of Theorem 5 below.

The diagram



helps to visualize the way in which the defining properties of a  $(b, v, r, k, \lambda)$ -design manifest themselves in the corresponding incidence matrix. The symbols  $x_1, x_2, \dots, x_v$  at the top indicate that the columns correspond to the elements of the set  $\{x_1, \dots, x_v\}$ ; the symbols  $X_1, X_2, \dots, X_b$  on the left show that the subsets  $X_1, \dots, X_b$  of  $\{x_1, \dots, x_v\}$  correspond to the rows; the 0, 1 inside indicate that the matrix is binary; the  $k, k, \dots, k$  on the right say that each row sum is equal to  $k$ , that is, each row has  $k$  ones in it; the  $r, r, \dots, r$  at the bottom say that each column sum is equal to  $r$ , that is, each column has  $r$  ones in it; finally, the symbols  $\begin{matrix} \swarrow \\ \lambda \\ \searrow \end{matrix}$  at the bottom are intended to mean that each pair of distinct columns have  $\lambda$  ones in common, that is, the inner product of each pair of distinct column vectors is equal to  $\lambda$ .

Likewise, the diagram



provides a way to visualize the principal properties of a  $(v, k, \lambda)$ -design through its incidence matrix.

According to Fisher's inequality [2, p. 129], the minimal value for the parameter  $b$  of a  $(b, v, r, k, \lambda)$ -design is  $b = v$ ; that is, a  $(b, v, r, k, \lambda)$ -design exists only if  $v \leq b$ . It has been known for a long time that a  $(b, v, r, k, \lambda)$ -design with  $b = v$  exists if and only if there exists a  $(v, k, \lambda)$ -design. In this paper we determine the existence of  $(b, v, r, k, \lambda)$ -designs when  $b = v + 1$  and when  $b = v + 2$ . For the case  $b = v + 3$ , we present new necessary conditions for the existence of a  $(b, v, r, k, \lambda)$ -design. Using a  $(v, k, \lambda)$ -design, and in a simple manner, we construct an infinite family of  $(b^*, v^*, r^*, k^*, \lambda^*)$ -designs. Moreover, we prove the nonexistence of an infinite family of  $(b, v, r, k, \lambda)$ -designs; the case  $b = v + 1$  belongs to this family.

## 2 $(b, v, r, k, \lambda)$ -designs when $b \in \{v, v + 1, v + 2, v + 3\}$

If a  $(b, v, r, k, \lambda)$ -design exists and the number  $b$  of blocks is equal to the smallest possible value  $v$ , then  $r = k$  and this design is in fact a  $(v, k, \lambda)$ -design. Thus, we can consider the  $(v, k, \lambda)$ -designs as particular cases of the  $(b, v, r, k, \lambda)$ -designs. To prove this, one can use a result [8, Theorem 2.1, p. 103] that says: if  $A$  is the incidence matrix of a  $(v, k, \lambda)$ -design, then  $AA^T = A^T A$ . Even though it has been known for a long time, this result is stated in the next theorem so as to make this paper more complete.

**Theorem 1.** *A  $(b, v, r, k, \lambda)$ -design with  $b = v$  exists if and only if there exists a  $(v, k, \lambda)$ -design.*

**Theorem 2.** *If  $b = v + 1$ , then a  $(b, v, r, k, \lambda)$ -design does not exist.*

*Proof.* Taking  $m = 1$  in Theorem 6 below, one obtains this result as a corollary.  $\square$

**Theorem 3.** *A  $(b, v, r, k, \lambda)$ -design with  $b = v + 2$  exists if and only if  $v = 2$  or  $v = 4$ . Moreover, given a  $(b, v, r, k, \lambda)$ -design with  $b = v + 2$ , then, in terms of the parameter  $v$ , the other parameters must be as follows:  $r = (v + 2)/2$ ,  $k = v/2$ , and  $\lambda = (v^2 - 4)/\{4(v - 1)\}$ .*

*Proof.* The proof is divided into two parts.

*Part I: To prove that if  $b = v + 2$  and a  $(b, v, r, k, \lambda)$ -design exists, then we must have  $v = 2$  or  $v = 4$ ; moreover, the values of the other parameters must be as in the statement of the theorem.*

Suppose there exists a  $(b, v, r, k, \lambda)$ -design with  $b = v + 2$ , and, to begin with, suppose also that  $v$  is odd; thus,  $v = 2n + 1$  for some positive integer  $n$ . Then the condition  $bk = vr$  implies  $r = (v + 2)k/v$ . Because  $r$  is a positive integer, it follows that  $(v + 2)k \equiv 0 \pmod{v}$ ; and because  $v = 2n + 1$ , we have  $2k \equiv 0 \pmod{2n + 1}$ , so that  $k \equiv 0 \pmod{v}$ , which contradicts  $0 < k < v$ . Therefore, a  $(b, v, r, k, \lambda)$ -design with  $b = v + 2$  and  $v$  odd cannot exist.

For the remainder of Part I of the proof, suppose that a  $(b, v, r, k, \lambda)$ -design with  $b = v + 2$  exists; then the integer  $v$  must be even.

In this case, as shown above when  $v$  is odd, we must have  $2k \equiv 0 \pmod{v}$ . Thus, it follows that  $2k \in \{0, \pm v, \pm 2v, \pm 3v, \dots\}$ . But, from the condition  $0 < k < v$ , there is just one possible value for  $2k$ , namely,  $2k = v$ , so that  $k = v/2$ . Moreover, the condition  $bk = vr$  gives  $(v+2)(v/2) = vr$ , from which we get

$$r = (v+2)/2.$$

With respect to the parameter  $\lambda$ , we use the condition  $r(k-1) = \lambda(v-1)$ , which becomes first  $\{(v+2)/2\}(v/2-1) = \lambda(v-1)$ , and then  $\lambda = (v-2)(v+2)/\{4(v-1)\}$ , or, in terms of the parameter  $k$ ,

$$\lambda = (k-1)(k+1)/(2k-1).$$

We use this last equation to find the possible values for  $k$ , and thus for  $v = 2k$ . The parameter  $\lambda$  is a nonnegative integer; therefore, we must have  $(k-1)(k+1) \equiv 0 \pmod{2k-1}$ , or, equivalently,

$$k^2 \equiv 1 \pmod{2k-1}.$$

But, also, since  $4k^2 = (2k-1)(2k+1) + 1$ , we always have

$$k^2 \equiv 4^{-1} \pmod{2k-1},$$

since 4 and  $2k-1$  are relatively prime. Therefore,

$$1 \equiv 4 \pmod{2k-1},$$

which forces  $k = 1$  or  $k = 2$ . Thus, since we must have  $v = 2k$ , we see that if  $b = v+2$  and a  $(b, v, r, k, \lambda)$ -design exists, then we must have  $v = 2$  or  $v = 4$ . This concludes Part I of the proof.

*Part II: To prove that if  $v = 2$  or  $v = 4$ , then there exist  $(b, v, r, k, \lambda)$ -designs with  $b = v+2$ ,  $r = (v+2)/2$ ,  $k = v/2$ , and  $\lambda = (v^2-4)/\{4(v-1)\}$ .*

The proof is accomplished by exhibiting the pertinent incidence matrices. When  $v = 2$ , an example of the desired incidence matrix is given by

$$\begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix},$$

and when  $v = 4$ , an example is given by

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

This concludes Part II of the proof.

Therefore, a  $(b, v, r, k, \lambda)$ -design with  $b = v+2$  exists if and only if  $v = 2$  or  $v = 4$ . Moreover, the parameter values must be as in the statement of the theorem.  $\square$

**Theorem 4.** *If  $b = v + 3$  and a  $(b, v, r, k, \lambda)$ -design exists, then there must exist positive integers  $m$  and  $t$  such that*

$$v = 3t, \quad r = m(t + 1), \quad k = mt, \quad \text{and} \quad \lambda = m(t + 1)(mt - 1)/(3t - 1).$$

*Proof.* Let  $\mathbb{N} := \{0, 1, 2, 3, \dots\}$ , and suppose a  $(b, v, r, k, \lambda)$ -design exists. For the parameter  $v$ , there are three cases to consider:  $v = 3t + 1$ , or  $v = 3t + 2$ , or  $v = 3t$ , for some  $t \in \mathbb{N}$ .

If  $b = v + 3$ , then the condition  $bk = vr$  implies  $(v + 3)k = vr$ , so that

$$r = \frac{v + 3}{v}k,$$

which must be a positive integer.

*Case 1: Suppose  $v = 3t + 1$  for some  $t \in \mathbb{N}$ .*

In this case we must have

$$r = \frac{v + 3}{v}k = k + \frac{3}{3t + 1}k.$$

Thus,

$$\frac{3}{3t + 1}k = r - k$$

must be an integer. Therefore, since  $\gcd\{3, 3t + 1\} = 1$ , it follows that  $(3t + 1) \mid k$  and  $k \equiv 0 \pmod{3t + 1}$ . Thus, since  $k > 0$ , we must have  $k \geq 3t + 1 = v$ , a contradiction.

Therefore, it is not possible to have  $v = 3t + 1$  for some  $t \in \mathbb{N}$ .

*Case 2: Suppose  $v = 3t + 2$  for some  $t \in \mathbb{N}$ .*

This case is not possible either, by an argument similar to that in Case 1; we omit the details.

*Case 3: Suppose  $v = 3t$  for some  $t \in \mathbb{N}$ .*

Since  $v > 0$ , it follows that  $t > 0$ . In this case we must have

$$r = \frac{v + 3}{v}k = \frac{t + 1}{t}k,$$

a positive integer. Therefore, since  $\gcd\{t, t + 1\} = 1$ , it follows that  $t \mid k$  and  $k \equiv 0 \pmod{t}$ . Thus, since  $k > 0$ , there must exist a positive integer  $m$  such that  $k = mt$ .

For the parameter  $r$ , we have

$$r = \frac{t + 1}{t}k = m(t + 1),$$

and for the parameter  $\lambda$  we have

$$\lambda = \frac{r(k - 1)}{v - 1} = \frac{m(t + 1)(mt - 1)}{3t - 1}. \quad \square$$

The preceding result suggested a simple construction for a family of  $(b, v, r, k, \lambda)$ -designs. This construction is given in the next theorem, which shows one way to construct  $(b, v, r, k, \lambda)$ -designs in terms of  $(v, k, \lambda)$ -designs. We considered  $b^* = nv$ , instead of  $v^* = nv$ , for some integer  $n \geq 1$  because, by Fisher's inequality [2, p. 129],  $b^* \geq v^*$  is a necessary condition for the existence of a  $(b^*, v^*, r^*, k^*, \lambda^*)$ -design. Although it is simple, this construction appears to be new.

**Theorem 5.** *If a  $(v, k, \lambda)$ -design exists, then, for each integer  $n \geq 1$ , there exists a  $(b^*, v^*, r^*, k^*, \lambda^*)$ -design with parameters  $b^* = nv$ ,  $v^* = v$ ,  $r^* = nk$ ,  $k^* = k$ , and  $\lambda^* = n\lambda$ .*

*Proof.* Suppose a  $(v, k, \lambda)$ -design exists, and let  $A$  be its incidence matrix. Then [8, Theorem 2.1, p. 103]

$$A^T A = (k - \lambda)I_v + \lambda J_{v,v}, \quad A J_{v,1} = k J_{v,1}, \quad \text{and} \quad J_{1,v} A = k J_{1,v}.$$

For each integer  $n \geq 1$ , we use  $n$  copies of  $A$  to define the  $nv \times v$  binary matrix  $B$  by

$$B := \begin{pmatrix} A \\ A \\ \vdots \\ A \end{pmatrix}.$$

Then  $B$  is the incidence matrix of a  $(b^*, v^*, r^*, k^*, \lambda^*)$ -design with the desired parameters.  $\square$

### 3 A general nonexistence result

**Theorem 6.** *Let  $m \geq 1$  be an integer such that  $\gcd\{m, v\} = 1$ . If  $b = v + m$ , then a  $(b, v, r, k, \lambda)$ -design does not exist.*

*Proof.* Suppose  $m \geq 1$  is an integer relatively prime to  $v$ . If  $b = v + m$  and a  $(b, v, r, k, \lambda)$ -design exists, then the parameters must satisfy the condition  $(v + m)k = vr$ . Thus, because  $r$  is a positive integer, it follows that

$$(v + m)k \equiv 0 \pmod{v}.$$

Moreover, since  $\gcd\{m, v\} = 1$ , we now have

$$k \equiv 0 \pmod{v},$$

which contradicts the requirement that  $0 < k < v$ . Therefore, a  $(b, v, r, k, \lambda)$ -design with  $b = v + m$  and  $\gcd\{m, v\} = 1$  cannot exist.  $\square$

**Acknowledgement.** The author thanks the referee for helpful, constructive comments.

---

## References

- [1] Box, G.E.P.; Hunter, W.G.; Hunter, J.S.: *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building*. Wiley, New York 1978.
- [2] Hall, M., Jr.: *Combinatorial Theory*. 2nd ed., Wiley, New York 1986.
- [3] Ionin, Y.J.; Shrikhande, M.S.: *Combinatorics of Symmetric Designs*. Cambridge Univ. Press, Cambridge, UK, 2006.
- [4] Lint, J.H. van; Wilson, R.M.: *A Course in Combinatorics*. 2nd ed., Cambridge Univ. Press, Cambridge, UK, 2001.
- [5] Marrero, O: Une propriété maximale de matrices binaires et inversibles dont toutes les lignes contiennent le même nombre de répétitions du chiffre un. *European J. Combin.* 19 (1998), 503–505.
- [6] Marrero, O.; Pasles, P.C.: Row coincidences in nonsingular binary matrices. *European J. Combin.* 22 (2001), 217–224.
- [7] Marrero, O; Pasles, P.C.: Binary matrices. *Elem. Math.* 57 (2002), 158–167.
- [8] Ryser, H.J.: *Combinatorial Mathematics*. Mathematical Association of America, Buffalo, NY 1963.

Oswaldo Marrero  
Department of Mathematical Sciences  
Villanova University  
800 Lancaster Avenue  
Villanova, Pennsylvania 19085–1699, USA  
e-mail: [Oswaldo.Marrero@villanova.edu](mailto:Oswaldo.Marrero@villanova.edu)