Report No. 116

Projective Properties of
Certain Orthogonal Arrays

George E.P. Box and John Tyssedal

May 1994

Projective Properties of
Certain Orthogonal Arrays

George Box

Center For Quality and
Productivity Improvement

University of Wisconsin
Madison, Wisconsin

John Tyssedal

Division of
Mathematical Sciences

University of Trondheim
Trondheim, Norway

ABSTRACT

The projective properties of two-level orthogonal array designs are important in factor screening. General results are given which, in particular, allow the designs derived by Plackett and Burman to be categorized in terms of these properties. The following results are given:

1. every saturated fractional factorial design is of projectivity $P=2$;

2. a design obtained by doubling is always of projectivity $P=2$;

3. any saturated two-level design obtained from a orthogonal array constructed by cyclic generation is either a factorial orthogonal array with $P=2$ or else has projectivity $P=3$;

4. any saturated two-level design obtained from an orthogonal array containing $n=4m$ runs, with m odd, is of projectivity $P=3$.

KEY WORDS: *Design projectivity; Design resolution; Factor screening; Fractional factorial design; Orthogonal array.*

Projective Properties
of Certain Orthogonal Arrays

George E.P. Box and John Tyssedal

The projective properties of two-level orthogonal array designs are important in factor screening. General results are given which, in particular, allow the designs derived by Plackett and Burman to be categorized in terms of these properties. The following results are given:

1. *every saturated fractional factorial design is of projectivity P=2;*

2. *a design obtained by doubling is always of projectivity P=2;*

3. *any saturated two-level design obtained from a orthogonal array constructed by cyclic generation is either a factorial orthogonal array with P=2 or else has projectivity P=3;*

4. *any saturated two-level design obtained from an orthogonal array containing n=4m runs, with m odd, is of projectivity P=3.*

1. INTRODUCTION

In what follows a two-level orthogonal design refers to an $n \times k$ matrix **D** in which the elements of each column consist of equal numbers -1's and $+1$'s and for which $\mathbf{D'D} = n\mathbf{I}_k$. If the ± 1's in each row are associated with the two levels of each of k factors, **D** can define n factor combinations to be run in an experiment to help determine how these factors affect some response y. Such designs include the two-level factorials (Fisher, 1935), the fractional factorials (Finney, 1945), and arrangements obtained from orthogonal arrays not necessarily of the factorial type (Plackett and Burman, 1946; Rao, 1947; see also for example Hall, 1961, 1965; Raghavarao, 1971; Hedayat and Wallis, 1978).

Experience shows that such designs can be of particular value for factor screening. A principle, called by Robert Grosseteste the law of parsimony (Crombie, 1953) and also known as Occam's razor and the Pareto principle, assumes that among a large number of possibly causal entities there will be a "vital few" and a "trivial many" (Juran, 1988). We approximate this concept (see for example Box and Meyer, 1986) by supposing that among the k factors introduced into the experiment, only a few will be *active* (in the same sense that they individually and/or interactively produce changes in the response) while the remainder are *inert*. A characteristic of importance (Box and Hunter, 1961) in choosing a design for factor screening will here be called the *projectivity* of the design. A $n \times k$ design **D** will be said to be of *projectivity P* if it is such that a complete factorial (possibly with some points replicated) is produced in *every* subset of P of the k factors. The resulting design will then be called an (n, k, P) screen.

An important source of such screening designs are the two-level orthogonal arrays. Such an array H_n is an $n \times n$ orthogonal matrix with $n=4m$ and m a positive integer with a first column of $+1$'s and the remaining $n-1$ *contrast* columns consisting of an equal number of $+1$'s and -1's. Arrays obtained by renumbering rows, renumbering contrast columns, or switching all the signs in a contrast column are regarded as equivalent. An orthogonal array (OA) can be used to generate a statistical design by associating k of its contrast columns with experimental factor levels. When $k=n-1$ the design is sometimes called a *saturated* design (Box and Wilson, 1951) or a *main effect* plan (Addelman and Kempthorne, 1961). In such a saturated design every pair of contrast columns must contain adjacent pairs of elements with the signs: $(--), (+-), (-+), (++)$ replicated m times. It thus produces m replicates of a 2^2 factorial design in any pair of factors we isolate. It is thus an $(n, n-1, P)$ screen with P at least 2.

2. FACTORIAL ORTHOGONAL ARRAYS

When $n = 2^r$ it is well known that an $n \times n$ orthogonal array can be generated from the corresponding 2^r factorial design as follows: first write down a column of $+1$'s, denoted by I, followed by the r columns c_1, c_2, . c_i, . . ., c_j, . . c_j, of ± 1's of the complete 2^r factorial design. Now represent by $c_h = c_i c_j$ an operation whereby c_h is an entry-wise product of columns $c_i c_j$. Then a further $n{-}r{-}1$ columns corresponding to the interaction columns of the original factorial can be obtained from all possible products of the individual columns, thus

$$c_{r+1} = c_1 c_2 \; c_{r+2} = c_1 c_3 \; . . \; c_{n-1} = c_1 c_2 \ldots c_r \quad (1)$$

The resulting $n \times n$ matrix will be called a *factorial* orthogonal array (FOA). For such an array $c_i c_l = I$, $l = 1, 2, . . , n{-}1$, and the n columns form a closed group with respect to the entry-wise operation defined above. Following Finney (1945), we note that a design produced by the $n{-}1$ contrast columns of a FOA is a $2^{-(n-r-1)}$ fraction of a 2^{n-1} factorial with the $n{-}r{-}1$ generating relations given by the identities (1). These may be conveniently written with I on the left and the generating "word" on the right as

$$I = c_1 c_2 c_{r+1} \; I = c_1 c_3 c_{r+2} \ldots I = c_1 c_2, \ldots, c_r c_{n-1}. \quad (2)$$

Multiplying these generators together in all possible ways produces the *defining relation* for the fractional factorial with I on the left and $2^{n-r-1} - 1$ words of the identity on the right from which the alias relationships between the effects can be constructed. For example, for $r=4$, $n = 2^r = 16$, there are 11 interaction columns (six two-factor interaction columns, four three-factor interaction columns, and one four-factor interaction column) which may be employed to accommodate 11 additional factors and to produce a 16×16 factorial orthogonal array. The $n{=}16$, $k{=}15$ design formed by the 15 contrast columns is a 2^{-11} fraction of a 2^{15} factorial with the 11 generators giving rise to defining relations containing $2^{11} - 1$ words. When the number of factors to be accommodated is such that $r < k < n-1$, then only $k{-}r$ of the generators are involved and the defining relation will contain only $2^{(k-r)} - 1$ words, producing less aliasing.

A useful criterion for the classification of such fractional factorials is their *resolution R* defined as the length of the shortest word in the defining relation. It was shown (Box and Hunter, 1961) that for any fractional factorial design, what is here called

the design projectivity is one less than its resolution. That is, for any fractional factorial $P = R-1$. In particular a saturated design obtained from an FOA is always of resolution $R{=}3$ and hence of projectivity 2. It thus provides an $(n, n{-}1, 2)$ screen. As a further example, when $n = 2^r$, a design of resolution 4 (and hence of projectivity $P{=}3$) for $1/2n$ variables can always be generated by dropping the $1/2(n-1)$ contrast columns containing an even number of letters from the generating factorial of the FOA. Designs of this kind can alternatively be produced by the "fold over" principle discussed by Box and Wilson (1951). This $(n, 1/2n, 3)$ screen is a $2^{-((1/2)n-r)}$ fraction of a $2^{(1/2)n}$ factorial. Thus for example, with $r = 4$ and $n = 16$, a $1/16$ fraction of a 2^8 design of resolution 4 that is a $(16, 8, 3)$ screen is produced by dropping from the generating factorial, all six two-factor interactions columns and the one four-factor interaction column. This valuable design screens $k{=}8$ factors at projectivity $P{=}3$ producing a duplicated 2^3 factorial in every one of the 56 choices of three factors out of eight. Notice, however, that this 16-run design provides the screen of greater projectivity $P{=}3$ at a cost of screening only $(1/2)n = 8$ factors instead of the $n{-}1{=}15$ factors possible with the saturated 16-run design of lesser projectivity $P{=}2$.

3. PLACKETT BURMAN DESIGNS

Now two-level factorial orthogonal arrays exist only when n is a power of 2; however, other orthogonal arrays are available when $n{=}4m$ and m is an integer. In particular, Plackett and Burman (1946) derived, with one single exception, orthogonal arrays for all such cases when $n \leq 100$. For designs produced by orthogonal arrays that are not factorial arrays, the concept of design resolution is not helpful. However, by analogy with the fractional factorials it might have been expected that saturated designs obtained from such arrays would provide an $(n, n{-}1, P)$ screen with projectivity P only 2. It was therefore with some surprise that it was discovered (Bisgaard, 1987) that such designs could be of resolution higher than 2. In particular, a computer search showed that *every one of the* 165 three dimensional projections of the $k{=}11$, $n{=}12$ design is a $1\frac{1}{2}$ replicate factorial. Specifically, it is a full 2^3 factorial design with the additional four duplicated runs forming of themselves a half-replicate, main-effect plan. Further computer enumeration by Lin and Draper (1992) showed that similar results were possible for some of the saturated designs derived from larger OAs. They referred to a

projection in which half the vertices of the projected factorial cube were replicated t times and the remainder of the vertices $s=m-t$ times a (t,s) projection. Thus each of the 165-three dimensional projections of the 12-run OA design was a (1:2) projection.

4. SOME GENERAL RESULTS

This paper gives four general results which provide a theoretical basis for these empirical discoveries. Computer studies of other aspects of the projective rationale for orthogonal arrays have been made by Wang and Wu (1993) and by Lin (1993).

RESULT 4.1

Every saturated fractional factorial design is of projectivity $P=2$:

Proof

This follows from the closed-group properties of an FOA. Specifically, if for every i and j, $c_i c_j = c_l$, is a column in the design, then since $c_i c_i = I$, $c_i c_j c_j = I$ and the design is of Resolution $R = 3$ and Projectivity $P=2$.

RESULT 4.2

A design obtained by doubling is always of projectivity $P=2$.

Proof

As before let H_n be the $n \times n$ matrix of an orthogonal array, then a $2n \times 2n$ orthogonal array may be constructed by "doubling" as follows

$$H_{2n} = \begin{bmatrix} H_n & H_n \\ H_n & -H_n \end{bmatrix}$$

The $2n-1$ contrast columns of a saturated orthogonal two-level design are then given by the $2n-1$ last columns of H_{2n}. If these columns are denoted $c_1 \ldots ., c_{2n-1}$ then, for instance, $c_i c_n = c_{n+i}$ and the design is of projectivity $P = 2$.

The result in 4.1 may also be considered as a special case of that in 4.2 since the factorial orthogonal arrays may also be obtained by doubling as was mentioned by Plackett and Burman (1946).

RESULT 4.3

Any saturated two-level design obtained from an orthogonal array constructed by cyclic generation is either a factorial orthogonal array with $P=2$ or else

has projectivity $P=3$.

Proof

Many of the Plackett Burman OAs are of the cyclic type. For such a cyclic orthogonal array it is possible to write down all the $n-1$ contrast columns knowing only the sequence of signs to be applied in the first row. For example, the 12×12 OA may be written down knowing only the 11 signs in the row sequence $(+ + - + + + - - - + -)$. Shifting this row of signs cyclically one place 10 times and adding a final row of minus signs and an initial column of plus signs produces a 12×12 orthogonal array. By omitting the first column, a 12×11 saturated orthogonal design is obtained.

Recalling that any two contrast columns are regarded as identical if one can be obtained from the other by sign reversal, it follows from the cyclic generation property that if $c_i c_{i+j} = c_l$ is in the design, the entry-wise product of any two columns at distance j must be in the design. Hence $c_i c_{i+2j} = c_i c_{i+j} c_{i+j} c_{i+2j} = c_i c_{i+j}$ is in the design and more generally $c_i c_{i+hj}$ for $h = 1, 2, \ldots$ is in the design. In particular, if $c_1 c_2$ is in the design, the entry-wise product of *any* two columns is in the design.

Now suppose $c_1 c_2$ is not in the design. Then since c_1 will be the column that follows after c_{n-1} in the cyclic generation procedure it follows that $c_{n-1} c_1$ is not in the design either. But the distance between c_1 and c_{n-1} is an even number $n-2$ from which it follows that the entry-wise product of no column with distance 2 or in general with distance $2h$, $h = 1, 2, \ldots$ can be in the design. If $c_1 c_{1+j}$ is a column in the design for j odd then $c_{n-1} c_j$ must also be in the design. But since $n-(j+1)$ is an even integer, that is not possible. Hence, if $c_1 c_2$ is not in the design no entry-wise product of any two columns is in the design.

Now assume a design constructed by cyclic generation contains $c_1 c_2$ and thereby all of their two factor interaction columns. If such a design contains more columns than c_1, c_2 and $c_1 c_2$, it must contain a column c_3 orthogonal to these, and since it also contains all the possible two factor interaction columns it must also contain $c_1 c_3$, $c_2 c_3$ and $c_1 c_2 c_3$ and none of these columns can be equal to any of the first three or to a column containing only $+1$'s. By induction the same argument now gives us that if a design generated by cyclic generation contains $c_1 c_2$ it must have 2^r-1 contrast columns for some r and must be a factorial orthogonal array.

RESULT 4.4

Any saturated two-level design obtained from an

orthogonal array containing $n=4m$ runs, with m odd is of projectivity $P=3$.

Proof

Consider a particular OA for which m is odd and let i be a column vector of m unities. Then $\mathbf{I}$ is a column of $n=4m$ unities which can be written in the partitioned form $(i, i, i, i)'$. Also, the rows of the orthogonal array can be re-ordered so that the array is partitioned into four $m \times n$ subarrays in which two *arbitrarily chosen* contrast columns $\mathbf{c}_i \, \mathbf{c}_j$ are partitioned in the form $(i, i, -i, -i)'$ and $(i, -i, i, -i)'$.

Now consider any other arbitrarily chosen column $\mathbf{c}_h = (\mathbf{c}_{h1}, \mathbf{c}_{h2}, \mathbf{c}_{h3}, \mathbf{c}_{h4})'$ after this re-ordering of rows. Orthogonality requires that $\mathbf{c}_h$ satisfies the linear equations $\mathbf{c}_h'\mathbf{I} = \mathbf{c}_h'\mathbf{c}_i = \mathbf{c}_h'\mathbf{c}_j = 0$, hence

$$\mathbf{i}'\mathbf{c}_{h1}=-a_h, \quad \mathbf{i}'\mathbf{c}_{h2}=+a_h, \quad \mathbf{i}'\mathbf{c}_{h3}=+a_h, \quad \mathbf{i}'\mathbf{c}_{h4}=-a_h .$$

Now let t_h represent the number of plus signs in $\mathbf{c}_{h1}$ and in $\mathbf{c}_{h4}$ or equivalently the number of minus signs in $\mathbf{c}_{h2}$ and in $\mathbf{c}_{h3}$. Then if $s_h=m-t_h$, $a_h=s_h-t_h$ and, if necessary, by switching signs in the whole column $\mathbf{c}_h$, a_h can be taken to be positive (so that $t_h < s_h$). Also since m is odd a_h cannot be zero and it cannot be equal to m for then $\mathbf{c}_h$ would correspond to the contrast $\mathbf{c}_i\mathbf{c}_j$ and no additional column could be simultaneously orthogonal to $\mathbf{c}_i$ $\mathbf{c}_j$ and $\mathbf{c}_h$. Thus, in the arbitrarily chosen space of $\mathbf{c}_j$, $\mathbf{c}_i$ and $\mathbf{c}_h$ the projected experimental points will lie on the vertices of a cube, and will be distributed as follows:

Replicated s_h times			Replicated t_h times		
−	−	+	−	−	−
+	−	−	+	−	+
−	+	−	−	+	+
+	+	+	+	+	−

The projected design will therefore consist of a full 2^3 factorial replicated t_h times with an additional half-replicate having defining relation $\mathbf{I} = \mathbf{c}_i\mathbf{c}_j\mathbf{c}_h$ replicated $a_h=s_h-t_h$, times. Note that the different choices of the columns $\mathbf{c}_j$, $\mathbf{c}_k$, $\mathbf{c}_h$ can result in different combinations of t and s and hence in different amounts of replication of the two design parts.

For illustration, Figure 1 shows a rearrangement of the 20×20 Plackett and Burman design with columns associated with factors $A, B, C, \ldots, T$ partitioned into four subarrays obtained by setting

$A = c_i$, $B = c_j$, respectively. Since the original design is obtainable by cyclic generation and also since $m = 5$ is odd, the design is a (20, 19, 3) screen with projections either of type (2, 3) or of type (1, 4). In all, there are 969 three-dimensional projections, 912 of which are of the (2,3) type and 57 of the (1, 4) type. The latter are generated by a single column with four like signs and one unlike sign in each quadrant (for the case illustrated this is column T). This ensures that 1/17 of the projections will be of the (1, 4) type and 16/17 will be of the (2,3) type. Notice, however, that the diagram does not imply that we can omit column T to produce a (20, 18, 3) design in which all the projections are of the (2, 3) type. This is because a different choice of two columns c_i, c_j to define the four subarrays would produce a different (1, 4) column.

A recent discussion of an analysis which can be applied to such designs is given by Box and Meyer (1992). Also, if r_h and s_h are each at least equal to two, an additional analysis for dispersion effects (Box and Meyer, 1986) may be attempted.

The arrays tabled by Plackett and Burman for various values of n are not, of course, exhaustive. Other OAs exists and may be of interest to the practitioner. For example, five distinct OA's are given by Hall (1961) for the case $n = 16$. All of these are (16,15,2) screens. But we have found that if one of the columns is dropped, one of them is also a (16,14,3) screen. We plan to discuss the usefulness of these and other such designs in a separate paper.

5. CONCLUSION

To illustrate the usefulness of the present results we show in Table 1 how they characterize the projectivity of each of the Plackett and Burman designs for $n \leq 84$. Of the 20 saturated designs considered, fourteen are of projectivity 3, four are FOAs of projectivity $P=2$ and two obtained by doubling are also of projectivity $P=2$.

ACKNOWLEDGEMENTS

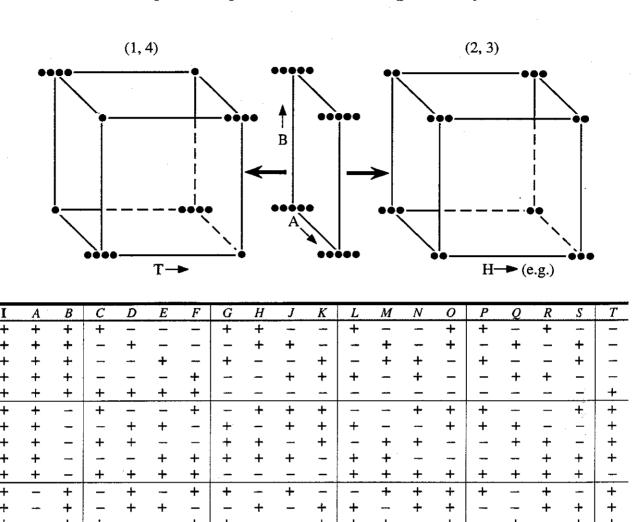| I	A	B	C	D	E	F	G	H	J	K	L	M	N	O	P	Q	R	S	T
+	+	+	+	−	−	−	+	+	−	−	+	−	−	+	+	−	+	−	−
+	+	+	−	+	−	−	−	+	+	−	−	+	−	+	−	+	−	+	−
+	+	+	−	−	+	−	+	−	−	+	−	+	+	−	+	−	−	+	−
+	+	+	−	−	−	+	−	−	+	+	+	−	+	−	−	+	+	−	−
+	+	+	+	+	+	+	−	−	−	−	−	−	−	−	−	−	−	−	+
+	+	−	+	−	−	+	−	+	+	+	−	−	+	+	+	−	−	+	+
+	+	−	−	+	+	−	+	−	+	+	+	−	−	+	+	+	−	−	+
+	+	−	+	+	−	−	+	+	−	+	−	+	+	−	−	+	+	−	+
+	+	−	−	−	+	+	+	+	+	−	+	+	−	−	−	−	+	+	+
+	+	−	+	+	+	+	−	−	−	−	+	+	+	+	+	+	+	+	−
+	−	+	−	+	−	+	+	−	+	−	−	+	+	+	+	−	+	−	+
+	−	+	−	+	+	−	−	+	−	+	+	−	+	+	−	−	+	+	+
+	−	+	+	−	−	+	+	−	−	+	+	+	−	+	−	+	−	+	+
+	−	+	+	−	+	−	−	+	+	−	+	+	+	−	+	+	−	+	+
+	−	+	+	+	+	+	+	+	+	+	−	−	−	−	+	+	+	+	−
+	−	−	−	+	−	+	−	+	−	+	+	+	−	−	+	−	−	−	−
+	−	−	−	−	+	+	+	+	−	−	−	−	+	+	−	+	−	−	−
+	−	−	+	−	+	−	−	−	+	+	−	+	−	+	−	−	+	−	−
+	−	−	+	+	−	−	+	−	+	−	+	−	+	−	−	−	−	+	−
+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	+	+	+	+

Figure 1. *Rearrangement of an L_{20} Plackett-Burman design partitioned into four subarrays defined by the signs of arbitrarily chosen columns labeled A and B. Addition to columns A and B of a further column such as H chosen from columns C through S results in a (2, 3) projection, but addition of column T to A and B produces a (1, 4) projection.*

Table 1.
Projectivity of Plackett and Burman designs for $n \leq 84$.

		$n =$	8	12	16	20	24	28	32	36	40	44	48	52	56	60	64	68	72	78	80	84
		$m =$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
Projectivity $P = 3$	m is odd			*		*		*		*		*		*		*		*		*		*
	cyclic and not FOA			*		*	*			*		*	*			*		*	*			*	*
$P = 2$ (factorial, doubling)			*		*			*		*			*		*							

REFERENCES

Addelman, S. and O. Kempthorne (1961), "Orthogonal Main-effect Plans," *ASTIA*; Arlington Hall Station, Arlington, VA. pp. 220, 224, 226.

Bisgaard, S. (1987), "Fractional Factorials and Other OAs," *An Explanation of Taguchi's Contributions to Quality Improvement*. Industrial Course Notes. Department of Engineering Professional Development; University of Wisconsin.

Box, G.E.P. and S. Bisgaard (1993), "What Can You Find Out from 12 Experimental Runs," *Quality Engineering*, Vol. 5, No. 4. pp. 663-668.

Box, G.E.P. and J.S. Hunter (1961), "The 2^{k-p} Fractional Factorial Designs," *Technometrics*, Vol. 3, 311-351. pp. 449-458.

Box, G.E.P. and R.D. Meyer (1986), "An Analysis for Unreplicated Fractional Factorials," *Technometrics*, Vol. 28. pp. 11-18.

Box, G.E.P. and R.D. Meyer (1992), "Finding the Active Factors in Fractional screening Experiments," *J. Qual. Technol.*, Vol. 25, No. 2. pp. 94-105.

Box, G.E.P. and K.B. Wilson (1951), "On the Experimental Attainment of Optimum Conditions," *J. Roy. Statist. Soc.* B13. pp. 1-38.

Crombie, A.C. (1953), *Robert Grosseteste and the Origins of Experimental Science*. Oxford; London.

Finney, D.J. (1945), "The Fractional Replication of Factorial Arrangements,"*Ann. Eugen.*, **12**. pp. 291-301.

Fisher, R.A. (1935), *The Design of Experiments*. Oliver and Boyd; Eninburgh.

Hall, M.J. (1961), "Hadamard Matrices of Order 16," *Jet Propulsion Laboratory*, Summary **1**. pp. 21-26.

Hall, M.J. (1965), "Hadamard Matrices of Order 20," *Jet Propulsion Laboratory*, Technical Report **1**. pp. 32-76.

Hedayat, A. and W.D. Wallis (1978), "Hadamard Matrices and Their Applications," *The Annals of Statistics*, Vol. 6, No. 6. pp. 1184-1238.

Juran, J.M. (1988), *Juran's Quality Control Handbook*. McGraw Hill; New York.

Lin, D.K. (1993), "Another Look at First-Order Saturated Designs: The *p*-efficient Designs," *Technometrics*, Vol. 35. pp. 284-294.

Lin, D.K.J. and N.R. Draper (1992), "Projection Properties of Plackett and Burman Designs," *Technometrics*, Vol. 34. pp. 423-428.

Plackett, R.L. and J.P. Burman (1946), "The Design of Optimum Multifactorial Experiments," *Biometrika* **33**. pp. 305-325.

Raghavarao, D. (1971), *Construction and Combinatorial Problems in Design of Experiments*. John Wiley; New York.

Rao, C.R. (1947), "Factorial Experiments Derivable from Combinational Arrangements of Arrays," *J. R.. Statist. Soc.*, Series B **9**. pp. 128-139.

Wang, J.C. and C.F. J. Wu (1993), "A Hidden Projection Property of Plackett-Burman and Related Designs," *IIQP Research Report RR93.08*. University of Waterloo; Ontario, Canada.