TECHNOMETRICS, AUGUST 1993, VOL. 35, NO. 3

Another Look at First-Order Saturated Designs: The *p*-efficient Designs

Dennis K. J. Lin

Department of Statistics
University of Tennessee
Knoxville, TN 37996

The problem of constructing first-order saturated designs that are optimal in some sense has received a great deal of attention in the literature. Since these saturated designs are frequently used in screening situations, the focus will be on the potential projective models rather than the full model. This article discusses some practical concerns in choosing a design and presents some first-order saturated designs having two desirable properties, (near-) equal occurrence and (near-) orthogonality. These saturated designs are shown to be reasonably efficient for estimating the parameters of projective submodels and thus are called *p*-efficient designs. Comparisons with the efficiency of *D*-optimal designs are given for designs for all *n* from 3 to 30.

KEY WORDS: *D*-optimal designs; Equal occurrence; Orthogonality; Plackett and Burman designs.

We are constantly faced with certain constraints when running experiments. For example, the experiment has to be completed within a certain time period, equipment (such as the number of machines available) to conduct the experiment is limited, and so on. In one study at Oak Ridge National Laboratory, for example, an experiment was conducted to test irradiation effects on heavy-section steels. The project was aimed at obtaining fracture toughness data based on two weldments with high-copper contents to determine the shift and shape of the actual fracture toughness (K_{IC}) curve as a sequence of irradiation. Radiation experimentation is extremely expensive so the experiment *must* be run with minimum size. This is a common problem in many industries—the number of runs in the experiment is fixed or has to be minimized. In this article, I consider first-order saturated designs for such problems.

Consider an experimental situation in which a response y depends on k factors $x_1, \ldots, x_k$ with the first-order relationship of the form $E(\mathbf{y}) = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k = \mathbf{X}\boldsymbol{\beta}$, where $\mathbf{y}$ is an $n \times 1$ vector of observations; the design matrix $\mathbf{X}$ is $n \times (k + 1)$ whose jth row is of the form $(1, x_{1j}, x_{2j}, \ldots, x_{kj})$, $j = 1, 2, \ldots, n$; and $\boldsymbol{\beta}$ is the $(k + 1) \times 1$ vector of coefficients to be estimated. In a two-level factorial design, each x_i can be coded as ± 1. The design is then determined by the $n \times k$ matrix of elements ± 1. The ith column gives the sequence of factor levels for factor x_i; each row constitutes a *run*. When $k = n - 1$, the design is called a *saturated* design and the design matrix $\mathbf{X}$ is an $n \times n$ square matrix. Note that $n = k + 1$ is the minimal number of points (rows) required to estimate all coefficients of interest (the β_i's).

Much theoretical work has been done in this area to select designs that meet certain optimization criteria. Note that a typical preliminary investigation contains many potentially relevant factors, but often only a few are believed to have *actual* effects. This is sometimes called *effect-sparsity*. Once these actual effects are identified, the initial design is then projected into a much smaller dimension. In such a screening situation, considering the optimality properties based on the full model is irrelevant. In this article, we focus on the potential projective models and construct a series of designs that are quite efficient in terms of the projective model.

In Section 1, *D*-optimal designs are briefly reviewed and discussed. Note that *D*-optimal designs are used only for comparison with designs given here. Other previously known criteria can be viewed in a similar manner. In Section 2, some practical considerations for choosing a design are discussed and the construction method is formulated. In Section 3, a computer algorithm to construct the proposed designs is described. In Sections 4, 5, 6, and 7, comparisons with *D*-optimal designs are given for $n \equiv 0 \pmod 4$, $n \equiv 1 \pmod 4$, $n \equiv 2 \pmod 4$, and $n \equiv 3 \pmod 4$. In Section 8, I further discuss some properties for the new designs when the full design is projected into $p = 2, 3, 4$, and 5 dimensions.

1. D-OPTIMAL DESIGNS

Several criteria have been advanced for the purpose of comparing designs and for constructing optimal designs. One of the most popular is the D-optimality criterion, which seeks to maximize $\|\mathbf{X}'\mathbf{X}\|$, the determinant of the $\mathbf{X}'\mathbf{X}$ matrix. Recall that for any $n \times n$ matrix, $\mathbf{H}$, consisting entirely of elements ± 1, the maximum determinant possible is $\|\mathbf{H}'\mathbf{H}\| = n^n$ (Hadamard 1893). Following Kiefer (1959), I define d efficiency as

$$d \text{ efficiency} = \left[\frac{\|\mathbf{X}'\mathbf{X}\|}{\|\mathbf{H}'\mathbf{H}\|}\right]^{1/n} = \frac{\|\mathbf{X}'\mathbf{X}\|^{1/n}}{n}.$$

Clearly, D-optimal designs will yield the highest d-efficiency values, this being the reason for their name. Under our setting, the d efficiency is equal to 1 only for Plackett and Burman designs; usually it is less than 1.

The problem of finding a D-optimal design has been very thoroughly explored mathematically; two early articles were by Hotelling (1944) and Mood (1946). The practical value of saturated D-optimal designs, apart from Plackett and Burman (1946) designs, is worthy of further investigation. First, these designs do not contain an equal number of high-level and low-level values. This can be quite disturbing to experimenters. The nonequal-occurrence property implies that the factor is being partially confounded with the constant term (the column with all +'s), which is usually significantly different from 0. Second, unlike orthogonal designs, these designs lack similarity relationships among all the columns. For example, the correlations between every pair of columns are not necessarily the same. This raises questions about which factors to assign to which columns and whether it matters.

2. SOME PRACTICAL CONCERNS

In the screening situation, a D-optimal design for the full model is not necessarily D-optimal for the submodel that contains only the active factors. Since we do not know in advance which factors will be important, it is reasonable to have designs that are *balanced* in all factors. This naturally leads to the desirability of the *(near-) equal occurrence* and *(near-) orthogonality* properties as explained in the following:

1. *(Near-) equal occurrence.* For each factor, both high-and low-level values are usually of equal interest, and each experimental result, y_i, should have equal influence. This leads to the equal-occurrence property—an equal number of high-level and low-level points for each factor in a design. When n is odd, however, the equal-occurrence property is unattainable. We thus seek a design that will be as near to equal occurrence as possible by specifying that the numbers of +'s and −'s should differ by no more than 1. Without loss of generality, for odd n, we assume that there are $(n + 1)/2$ +'s and $(n - 1)/2$ −'s. Denote the largest absolute correlation with the constant term among all factors by c as a measure of equal occurrence. A large value for c is undesirable. The designs given here have $c = 0$ for even n and $c = 1/n$ for odd n.

2. *(Near-) orthogonality.* Orthogonality as an important design principle was pointed out by the work of R. A. Fisher and F. Yates back in the 1920s. The degree of nonorthogonality between factors x_i and x_j can be measured by $s_{ij} = \Sigma_{u=1}^{n} x_{iu}x_{ju}$ ($s_{ij} = 0$ implies orthogonality). Even if circumstances are such that exact orthogonality is unattainable, it is still preferable to make the design as *nearly orthogonal* as possible. Denote the largest $|s_{ij}|$ among all pairs of factors for a given design by s ($s \geq 0$). We thus desire a design to have a minimum value for s. Under the equal-occurrence assumption, it is shown in the appendix that $s_{ij} \equiv n \pmod 4$—namely, that the smallest $s = |s_{ij}|$ possible are 0, 1, 2, and 1 for $n \equiv 0$, 1, 2, and 3 (mod 4). If two designs have the same value of s, we prefer the one in which the frequency of such s is smaller. Thus we minimize the average of s^2, denoted by ave(s^2). For a specific design, ave(s^2) is computed by $\Sigma s_i^2 f_i / \binom{n-1}{2}$, where f_i is the frequency of s_i of all $\binom{n-1}{2}$ pairs of columns. This criterion was first proposed by Booth and Cox (1962) in the context of supersaturated designs. It is similar to the S criterion (Kiefer 1974), which minimizes tr$(\mathbf{X}'\mathbf{X})^2$. We see that when $n \equiv 0 \pmod 4$ the Plackett and Burman designs are optimal in the sense of meeting both of these requirements.

Suppose that such a design is used for screening purposes. Now, consider its projective property; that is, consider the submodel that contains only the p ($\leq k = n - 1$) active factors. The projective design in any p of the k factor dimensions will always preserve the original (near-) equal-occurrence property and (near-) orthogonality *no matter which p factors are designated as the survivor columns*. Moreover, as we shall see in Section 8, these designs have high d efficiency in terms of the reduced model when p is small—for example $p \leq 5$. Because of this property, we call them *p-efficient* designs.

Consider the case $n = 7$ to investigate six factors, as an example. Suppose that two factors are found to be important (the first two columns, say). The D-optimal design for the full model (see Williamson 1946) could then project into $(+ - + - + + +)'$ and

*Table 1. The Vectors **r** and **q**′ to Construct p-Efficient Designs*

n	*Vectors **r** and **q**′*
5	**r** = (+ + +)
	q′ = (+ + + −)
9	**r** = (+ − − − − + +)
	q′ = (− − − + + + + +)
13	See Table 2.
17	**r** = (− − − − + + + + + − + − − − −)
	q′ = (− − − − − − − + + + + + + + + +)
21	**r** = (− − − − + + + + + + + + + − − − − − −)
	q′ = (− − − − − − − − − + + + + + + + + + + +)
25	**r** = (− − + − + + − − + + + − + + − + − − − − − − −)
	q′ = (− − − − − − + − − + − + + + + + + + + + − − + +)
29	**r** = (− + − − + − + + − + + + + + + + + − − + − − − + + − +)
	q′ = (− − − − + + − − + − + + + + + + − + + + − − + − − + + −)

$(+++-----)'$ whose

$$\mathbf{X}'_{\text{sub}}\mathbf{X}_{\text{sub}} = \begin{bmatrix} 7 & 3 & 1 \\ 3 & 7 & 1 \\ 1 & 1 & 7 \end{bmatrix}$$

is no longer D-optimal for the submodel. On the other hand, we shall see that the p-efficient design results in $\mathbf{X}'_{\text{sub}}\mathbf{X}_{\text{sub}} = 8\mathbf{I}\text{-}\mathbf{J}$, which is optimal in many senses (including D-optimality), no matter which two columns are selected.

3. THE BASIC CONSTRUCTION PROCEDURE

In this section, I describe the computer search routine that I used to construct p-efficient designs. Once the number of runs n is specified, the routine generates all possible combinations of columns ($n/2$ of $+$'s and $-$'s, when n is even; $(n + 1)/2$ of $+$'s and $(n - 1)/2$ of $-$'s, when n is odd) and then randomizes their orders. At each stage, a potential column enters and the s value with all other columns retained is calculated to check whether it satisfies the requirement ($s = 0, 1, 2, 1$ for $n \equiv 0, 1, 2$, and $3 \pmod 4$)). If not, the potential column is dropped and the search continues.

One difficulty with this routine arises from the fact that the s-value property is not transitive; namely, when x_i, x_j produce s value $= s^*$ and x_j, x_k produce s value $= s^*$, this does *not* imply that x_i, x_k will have s value $= s^*$. To handle this, two modifications were made to the search procedure. These were as follows: (1) Those candidate columns that meet the requirement for all but one retained column were saved in a queue. Whenever two columns in the queue linked to the same retained columns, this retained column was removed and replaced by the two columns in the queue. (2) Allowance was made for another loop with different random order of entrance for each possible candidate if less than $n - 1$ columns were obtained (at most, three loops were tried). This search procedure worked for all cases I considered except

Table 2. Design for n = 13 ($\mathbf{X}'\mathbf{X} = 25 \times 12^{12}$)

I	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
+	−	−	−	−	−	−	−	−	−	+	+	+
+	−	−	−	−	−	−	+	+	+	−	−	−
+	−	−	−	+	+	+	−	−	−	−	−	−
+	−	+	+	−	+	+	−	+	+	−	+	+
+	−	+	+	+	−	+	+	−	+	+	−	+
+	−	+	+	+	+	−	+	+	−	+	+	−
+	+	−	+	−	+	+	+	+	−	+	−	+
+	+	−	+	+	−	+	−	+	+	+	+	−
+	+	−	+	+	+	−	+	−	+	−	+	+
+	+	+	−	−	+	+	+	−	+	+	+	−
+	+	+	−	+	−	+	+	+	−	−	+	+
+	+	+	−	+	+	−	−	+	+	+	−	+
+	+	+	+	−	−	−	−	−	−	−	−	−

Table 3. Comparisons With D-optimal Designs For n ≡ 1 (mod 4)

n	Design	c	ave(s^2)	d efficiency
5	D-optimal	1/5	1	.941
	p-efficient	1/5	1	.941
9	D-optimal	1/9	1.67	.932
	p-efficient	1/9	1.67	.932
13	D-optimal	1/13	1	.977
	p-efficient	1/13	1	.977
17	D-optimal	3/17	1.94	.966
	p-efficient	1/17	2.06	.954
21	D-optimal	5/21	1.63	.976
	p-efficient	1/21	2.26	.963
25	D-optimal	7/25	1.61	.974
	p-efficient	1/25	2.20	.969
29	p-efficient	1/29	2.27	.974

for $n = 9$, 17, 21, 22, and 25, for which it has been proved that designs with minimum s do not exist (for example, see Galil and Kiefer 1980). In these cases, only $n - 2$ columns were found. Instead of using a branch-and-bound feature, I used a brute-force search for the very last column based on the criterion ave(s^2) to ensure the optimal results. Details of the program are available from me.

Much work has been done in constructing D-optimal (+1, −1) designs. For example, see DETMAX (Mitchell 1974a), branch-and-bound (Welch 1982), and annealing algorithm (Haines 1987). Compared with these, the approach given here is much simpler and is computationally inexpensive, in part because of the special structure of our problem. Specifically, (a) due to the equal-occurrence concern, the candidate space is much smaller, which allows us to search for designs with large n (here we show the results for n up to 30); (b) unlike other algorithms that find a set of columns at each stage to evaluate $\|\mathbf{X}'\mathbf{X}\|$, here the design columns are added sequentially; and (c) rather than evaluating $\|\mathbf{X}'\mathbf{X}\|$, only $s_{ij} = \Sigma_{u=1}^{n} x_{iu}x_{ju}$ is computed here. This is a much easier computational task.

Table 4. Design for n = 6 ($\|\mathbf{X}'\mathbf{X}\| = 6^2 \times 4^4$)

I	1	2	3	4	5
+	−	+	−	−	−
+	−	−	−	+	+
+	+	+	−	+	+
+	+	−	+	+	−
+	−	+	+	−	+
+	+	−	+	−	−

Table 5. Design for n = 10 ($\|\mathbf{X}'\mathbf{X}\| = 70^2 \times 8^6$)

I	1	2	3	4	5	6	7	8	9
+	+	+	+	−	−	+	−	−	−
+	+	+	−	+	+	−	−	+	+
+	+	−	−	+	+	+	+	+	−
+	+	−	+	−	−	−	−	+	+
+	+	−	+	+	−	+	+	−	+
+	−	−	−	−	+	−	+	−	+
+	−	−	+	+	+	−	−	−	−
+	−	+	+	−	+	+	+	+	−
+	−	+	−	+	−	−	+	−	−
+	−	+	−	−	−	+	−	+	+

4. DESIGNS FOR $n \equiv 0$ (MOD 4)

For $n \equiv 0$ (mod 4), the smallest s value possible is 0—that is, an orthogonal design. The designs given by Plackett and Burman (1946) are not only orthogonal but also are of equal occurrence and thus are highly recommended. For $n \leq 24$, all of their designs can be obtained by cyclic permutation row by row by using the following sequences: Initially, write down the first row, remove the rightmost sign, place it on the extreme left, and move all signs one place to the right; finally add a row with all −1s. For $n = 8$, 12, 16, 20, and 24, the signs for the first row are

$n = 8$: + + + − + − −

$n = 12$: + + − + + + − − − + −

$n = 16$: + + + + − + − + + − − + − − −

$n = 20$: + + − − + + + + − + − + − − − − + + −

$n = 24$: + + + + + − + − + + − − + + − − + − + − − − −.

For the 28-run case, the Plackett and Burman design is formulated by permuting three 9 × 9 blocks of signs and then adding a row of minus signs. These values of n are used for testing the algorithm given in Section 3. The program successfully produced these designs as required.

5. DESIGNS FOR $n \equiv 1$ (MOD 4)

The smallest s value possible is 1; such a design, if it exists, would be a D-optimal design because the determinant of $\mathbf{X}'\mathbf{X}$ reaches the upper bound. It is known, however, that such a design does not exist if $2n - 1$ is not the square of an integer. [In fact, even for $n = 25$, the D-optimal design given by-Raghavarao (1959, p. 302) does not have all $s = 1$—e.g., columns 10 and 11.] The search routine successfully produced the only possible designs for $n = 5$ and 13. For $n = 9$, 17, 21, and 25, as mentioned, a further search for the last column is made by letting

Table 6. Design for $n = 14$ ($\|X'X\| = 14^2 \times 12^{12}$)

I	1	2	3	4	5	6	7	8	9	10	11	12	13
+	−	+	+	+	−	−	−	−	+	+	−	+	−
+	+	−	−	−	−	+	−	+	−	+	+	+	−
+	+	−	−	−	−	−	+	+	+	+	−	−	+
+	+	−	−	−	+	−	−	−	+	−	+	+	+
+	−	+	−	+	−	+	−	−	+	−	+	−	+
+	−	−	+	+	−	−	+	+	−	−	+	+	−
+	−	+	+	−	+	−	−	+	−	+	−	−	+
+	+	+	+	+	+	+	+	+	+	+	+	+	+
+	−	−	+	−	−	+	+	−	+	+	+	−	−
+	−	−	−	+	+	+	−	−	−	+	−	+	+
+	−	+	−	−	+	−	+	+	+	−	−	+	−
+	+	−	+	+	+	+	−	+	−	−	−	−	−
+	+	+	+	−	−	+	+	−	−	−	−	−	+
+	+	+	−	+	+	−	+	−	−	−	+	−	−

s be the next smallest number (in this case 5), and choosing the column that produces the minimum number of $s = 5$ among all pairs associated with the last column. The frequency of $s = 5$ appears 1, 6, 10, 15, and 21 times for $n = 9, 17, 21, 25$, and 29, respectively. For these cases, the factor thought to be least likely to be retained should be assigned to the last column in practice.

As suggested by one referee, these designs can also be constructed by adding a row and a column to the *standard* Plackett and Burman designs of the form

$$\begin{bmatrix} PB_{n-1} & \mathbf{q} \\ \mathbf{r} & - \end{bmatrix},$$

where PB_{n-1} is an $(n - 1) \times (n - 2)$ Plackett and Burman orthogonal array, $\mathbf{r}$ is a $1 \times (n - 2)$ row vector of $\pm$'s, and $\mathbf{q}$ is an $(n - 1) \times 1$ column vector of $\pm$'s. The vectors $\mathbf{r}$ and $\mathbf{q}$ are given in Table 1 (p. 286). The case $n = 13$ does not follow this rule and is given separately in Table 2. Note that these vectors are obtained using the standard Plackett and Burman designs (see Sec. 4); if other equivalent Plackett and Burman type designs (permutation on rows or columns) are used, they need to be modified accordingly. The D-optimal designs for $n = 9, 17, 21$, and 25 were given by Mitchell (1974b), Moyssiadis and Kounias (1982), Chadjipantelis, Kounias, and Moyssiadis (1987), and Raghavaro (1959), respectively. The D-optimal design for $n = 29$ is not available.

As shown in Table 3, the p-efficient designs are identical to D-optimal designs for $n = 5, 9$, and 13. For $n = 17, 21$, and 25, the p-efficient designs have their d-efficiency values close to that of D-optimal

Table 7. Design for $n = 18$ (d efficiency $= .8913$)

I	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
+	+	−	+	+	−	+	+	+	+	+	+	−	+	+	−	−	−
+	+	−	−	+	+	−	+	+	+	+	−	−	−	−	+	+	−
+	+	+	+	+	+	−	−	−	−	−	+	−	+	−	+	−	+
+	+	−	+	−	−	+	+	−	−	−	+	+	−	−	+	+	−
+	+	−	−	+	+	+	−	−	−	−	−	+	+	+	−	+	−
+	+	+	−	−	−	+	−	+	+	−	−	−	+	−	−	+	+
+	+	+	−	−	+	−	+	+	−	+	+	+	−	+	−	−	+
+	+	+	+	−	−	−	−	−	+	+	−	+	−	+	+	−	−
+	−	−	+	−	−	−	−	+	−	+	−	+	+	−	−	+	+
+	−	−	+	−	+	+	−	+	−	−	−	−	−	+	+	−	+
+	−	−	−	−	+	−	+	−	+	−	−	+	−	−	−	−	+
+	−	−	−	+	−	−	−	−	+	+	+	−	−	+	+	+	+
+	−	+	+	+	−	+	+	−	−	+	−	−	−	−	−	+	+
+	−	+	−	+	−	−	+	+	−	−	−	+	+	+	+	−	−
+	−	+	+	−	+	−	+	−	+	−	+	−	+	+	−	+	−
+	−	−	−	−	+	+	−	−	−	+	+	−	+	−	+	−	−
+	−	+	−	+	−	+	−	+	+	−	+	+	−	−	−	−	−

Table 8. Design for $n = 22$ (*d* efficiency $= .8576$)

I	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
+	+	+	+	+	+	+	+	+	+	+	+	+	−	+	+	+	+	−	−	−	−
+	+	−	−	+	+	+	+	+	+	+	+	−	+	−	−	−	−	+	+	−	−
+	+	−	+	+	−	+	−	+	−	−	−	+	−	+	−	−	−	+	+	+	+
+	+	+	−	−	−	−	+	−	+	+	−	+	−	−	+	−	−	−	+	+	+
+	+	−	+	−	+	−	+	−	−	−	+	+	+	−	+	+	−	+	−	−	+
+	+	+	−	+	−	−	+	−	+	−	−	−	−	+	−	+	+	+	−	−	+
+	+	−	−	−	+	+	−	−	−	+	+	+	+	+	−	−	+	−	−	+	+
+	+	+	+	+	+	−	−	−	−	−	−	−	+	−	−	−	+	−	+	−	+
+	+	+	+	−	−	−	−	+	−	+	+	−	−	−	−	+	−	+	−	+	−
+	+	−	−	−	−	+	−	+	+	−	−	−	+	−	+	+	+	−	−	+	−
+	−	−	+	+	+	−	+	−	−	+	−	−	−	−	+	+	+	−	+	+	−
+	−	+	−	+	−	−	+	+	−	−	+	−	+	+	+	−	−	−	−	+	+
+	−	−	−	+	−	+	−	−	−	+	+	−	−	−	+	−	+	+	−	−	+
+	−	+	−	−	+	+	−	−	−	−	−	−	+	+	+	+	−	+	+	−	−
+	−	+	+	−	−	−	−	−	+	−	+	+	+	+	+	−	+	+	+	+	−
+	−	−	+	−	+	−	−	+	+	+	−	−	−	+	+	−	−	+	−	−	+
+	−	−	−	+	+	−	−	+	+	−	+	+	−	−	−	+	−	−	+	+	+
+	−	+	+	+	−	+	−	−	+	+	−	+	+	−	−	+	−	−	−	−	−
+	−	−	+	−	−	+	+	−	+	−	+	−	−	+	−	−	−	−	+	−	−
+	−	−	−	−	−	−	+	+	−	+	−	+	+	+	−	+	+	−	+	−	−
+	−	+	−	−	+	+	+	+	−	−	−	+	−	−	−	−	+	+	−	+	−

designs, and their c value is much smaller. Moreover, for $n = 21$ and 25, D-optimal designs have smaller $\text{ave}(s^2)$; this is partially due to their large c value (which is not desirable).

6. DESIGNS FOR $n \equiv 2$ (MOD 4)

The smallest s value possible is 2. These saturated designs are listed in Tables 4–9 (pp. 287–290) for $n =$ 6, 10, 14, 18, 22, and 26. Table 10 shows the comparison with the D-optimal designs given by Ehlich (1964). As one can see, the p-efficient designs have a slightly smaller d efficiency and a larger value for $\text{ave}(s^2)$.

For the D-optimal designs, all factors can be partitioned into two equal groups such that all factors have $s = 0$ between groups but $s = 2$ within each group. (This is certainly unattainable with the equal-occurrence property.) Thus their c, $\text{ave}(s^2)$, and $\|\mathbf{X}'\mathbf{X}\|$ are $2/n$, $2(n - 2)/(n - 1)$ and $4(n - 2)^{n-1}(n - 1)^2$, respectively. The case $n = 22$ is not available, however.

For the p-efficient designs, the $\text{ave}(s^2)$ is always $4(n - 2)/n$. The $\|\mathbf{X}'\mathbf{X}\|$ for these designs are $n(3n - 2)(n - 2)^{n-2}$, if $s_{ij} = 2$ for all pairs of columns, but often they have $s_{ij} = \pm 2$, which results in smaller d-efficiency values. The case $n = 22$ contains six elements $s = 6$ in its $\mathbf{X}'\mathbf{X}$ matrix because the program produced only 20 columns having all $s_{ij} = \pm 2$, as mentioned in Section 3.

7. DESIGNS FOR $n \equiv 3$ (MOD 4)

The smallest s value possible in absolute value is 1. These designs can be easily obtained by deleting one row and one column of the $(n + 1)$-run Plackett and Burman designs. Such p-efficient designs have $\|\mathbf{X}'\mathbf{X}\| = (n + 1)^{n-1}$ and are D-optimal only for the case $n = 3$. A comparison of p-efficient design with D-optimal designs is given in Table 11. The case $n =$ 7 is derived from Williamson (1946); $n = 11$ was given by Galil and Kiefer (1980); for $n \geq 15$, a D-optimal design is not available.

Table 11 also shows the d efficiency, $\text{ave}(s^2)$, and c of these designs, as well as the comparison with the D-optimal designs, if available. As one can see, the p-efficient designs have increasing d efficiencies as n increases and $\text{ave}(s^2)$ and c reach the minimum values. [In fact, the d efficiency here is equal to $(n + 1)^{(n-1)/n}/n$, which is a monotonic increasing function for n.] The relationships among all factors are identical, with the smallest correlation ($c = s/n = 1/n$).

8. PROJECTION INTO MORE THAN TWO DIMENSIONS

The designs just given seek to meet two criteria, (near-) equal occurrence and (near-) orthogonality. For near-orthogonality, we first control the s value at its minimum possible and then apply the criterion $\text{ave}(s^2)$ to designs that have the same value of s. Even though $\text{ave}(s^2)$ is computed by *all* possible pairs of factors, the value of s only measures the degree of nonorthogonality of *two* factors. Certainly, the true model may contain more than two factors. The performances of the p-efficient design when the true model contains more than two factors is thus an important issue to be addressed.

Table 9. Design for $n = 26$ ($\|X'X\| = 26^2 \times 24^{24}$)

I	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
+	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
+	+	−	−	−	−	+	−	−	+	+	+	+	−	+	−	+	−	+	−	+	+	+	+	−	−
+	−	+	−	−	−	+	+	−	−	+	−	+	+	−	+	+	+	−	+	−	−	+	+	+	−
+	−	−	+	−	−	+	+	+	−	−	+	−	+	+	−	−	+	+	−	+	−	−	+	+	+
+	−	−	−	+	−	−	+	+	+	−	−	+	−	+	+	+	−	+	+	−	+	−	−	+	+
+	−	−	−	−	+	−	−	+	+	+	+	−	+	−	+	−	+	−	+	+	+	+	−	−	+
+	+	−	−	+	+	+	+	−	+	−	+	−	+	−	+	+	+	+	−	−	+	−	−	−	−
+	+	+	−	−	+	−	+	+	−	+	+	+	−	+	−	−	+	+	+	−	−	+	−	−	−
+	+	+	+	−	−	+	−	+	+	−	−	+	+	−	+	−	−	+	+	+	−	−	+	−	−
+	−	+	+	+	−	−	+	−	+	+	+	−	+	+	−	+	−	−	+	+	−	−	−	+	−
+	−	−	+	+	+	+	−	+	−	+	−	+	−	+	+	+	+	−	−	+	−	−	−	−	+
+	+	+	−	+	−	+	−	+	−	+	+	+	+	−	−	+	−	−	−	−	+	−	−	+	+
+	−	+	+	−	+	+	+	−	+	−	−	+	+	+	−	−	+	−	−	−	+	+	−	−	+
+	+	−	+	+	−	−	+	+	−	+	−	−	+	+	+	−	−	+	−	−	+	+	+	−	−
+	−	+	−	+	+	+	−	+	+	−	+	−	−	+	+	−	−	−	+	−	−	+	+	+	−
+	+	−	+	−	+	−	+	−	+	+	+	+	−	−	+	−	−	−	−	+	−	−	+	+	+
+	+	−	+	−	+	+	+	+	−	−	+	−	−	−	−	+	−	−	+	+	+	+	−	+	−
+	+	+	−	+	−	−	+	+	+	−	−	+	−	−	−	+	+	−	−	+	−	+	+	−	+
+	−	+	+	−	+	−	−	+	+	+	−	−	+	−	−	+	+	+	−	−	+	−	+	+	−
+	+	−	+	+	−	+	−	−	+	+	−	−	−	+	−	−	+	+	+	−	−	+	−	+	+
+	−	+	−	+	+	+	+	−	−	+	−	−	−	−	+	−	−	+	+	+	+	−	+	−	+
+	+	+	+	−	−	+	−	−	−	−	+	−	−	+	+	+	+	−	+	−	+	−	+	−	+
+	−	+	+	+	−	−	+	−	−	−	+	+	−	−	+	−	+	+	−	+	+	+	−	+	−
+	−	−	+	+	+	−	−	+	−	−	+	+	+	−	−	+	−	+	+	−	−	+	+	−	+
+	+	−	−	+	+	−	−	−	+	−	−	+	+	+	−	−	+	−	+	+	+	−	+	+	−
+	+	+	−	−	+	−	−	−	−	+	−	−	+	+	+	+	−	+	−	+	−	+	−	+	+

Table 10. Comparisons with D-Optimal Designs for $n \equiv 2 \pmod 4$

n	Design	c	ave(s^2)	d efficiency
6	D-optimal	2/6	1.60	.905
	p-efficient	0	2.67	.763
10	D-optimal	2/10	1.78	.941
	p-efficient	0	3.20	.815
14	D-optimal	2/14	1.85	.957
	p-efficient	0	3.43	.876
18	D-optimal	2/18	1.88	.967
	p-efficient	0	3.56	.891
22	p-efficient	0	4.33	.858
26	D-optimal	2/26	1.92	.977
	p-efficient	0	3.69	.929
30	D-optimal	2/30	1.93	.980
	p-efficient	0	3.73	.938

Table 11. Comparisons With D-Optimal Designs for $n \equiv 3 \pmod 4$

n	Design	c	ave(s^2)	d efficiency
3	D-optimal	1/3	1	.840
	p-efficient	1/3	1	.840
7	D-optimal	3/7	2.14	.878
	p-efficient	1/7	1	.849
11	D-optimal	3/11	2.02	.915
	p-efficient	1/11	1	.870
15	p-efficient	1/15	1	.887
19	p-efficient	1/19	1	.899
23	p-efficient	1/23	1	.909
27	p-efficient	1/27	1	.917

The d efficiencies for *all* possible projective designs when $p = 2, 3, 4$, and 5 dimensions were evaluated for both D-optimal and p-efficient designs for $n \leq 25$. Table 12 shows their minimum and maximum d efficiencies. If the minimum and maximum values were identical, only one value was reported. The p-efficient designs are seen to be superior to D-optimal designs, even judged by the d-efficiency criterion. The only exception occurs for the case $n = 6$ ($k = 4$ and 5). It should be emphasized, however, that the D-optimal designs are optimal for the full model.

Some values of n are absent from Table 12. These are the cases in which either the D-optimal designs are not available or they are identical to the p-efficient designs. The p-efficient design retains its basic properties of (near-) equal occurrence and (near-) orthogonality and thus preserves the structure of $\mathbf{X}'\mathbf{X}$ matrix no matter which columns were projected. Therefore, only one value is found in Table 6, except for some $n \equiv 2 \pmod 4$ cases, where the signs of s_{ij} vary, which results in different d efficiencies.

9. CONCLUDING REMARKS

The fact that saturated designs are often used in screening situations in which it is expected that there will only be a few important factors leads to the practical value of these designs. The p-efficient designs discussed here are attractive because their (near-) equal-occurrence property and (near-) orthogonality are preserved when projecting into $p(\leq k)$ dimensions. Furthermore, for estimating from a submodel, it is shown that these designs are more efficient than D-optimal designs. Even for the full model, the D-optimal designs are not substantially more efficient than the p-efficient designs as shown in Tables 3, 10, and 11.

Table 12. Comparisons on D Efficiencies When Projected Into p = 2, 3, 4, and 5 Dimensions

		p			
n	Design	2	3	4	5
6	D-optimal	(.905, .962)	(.928, .943)	.920	.905
	p-efficient	.962	.928	.901	.763
7	D-optimal	(.926, .977)	(.894, .961)	(.901, .923)	(.882, .897)
	p-efficient	.977	.961	.939	.907
10	D-optimal	(.964, .987)	(.951, .980)	(.941, .970)	(.951, .964)
	p-efficient	.987	(.964, .973)	(.929, .961)	(.906, .937)
11	D-optimal	(.970, .991)	(.950, .986)	(.937, .979)	(.928, .972)
	p-efficient	.991	.986	.979	.972
14	D-optimal	(.981, .993)	(.974, .990)	(.968, .985)	(.962, .981)
	p-efficient	.993	.986	.979	(.959, .968)
17	D-optimal	.979	.976	.974	.973
	p-efficient	.997	.995	.994	.993
18	D-optimal	(.988, .996)	(.984, .994)	(.980, .991)	(.976, .988)
	p-efficient	.996	.991	.987	(.973, .983)
21	D-optimal	(.962, .998)	(.957, .997)	(.954, .996)	(.961, .995)
	p-efficient	.998	.997	.996	.995
25	D-optimal	(.990, .998)	(.989, .998)	(.988, .997)	(.987, .996)
	p-efficient	.998	.998	.997	.996

D optimality has an appealing invariant property under a nonsingular linear transformation, a property that is clearly possessed by the p-efficient design. This is not true for most of the other optimalities (including A, E, G, L, and R optimality; see Kiefer 1959, p. 294). We note that blindly following a single optimality criterion is dangerous, although as pointed out by one referee, the p-efficient designs should perform well in general because of their better balance property.

The designs given here are *not* unique. This is also true for most of the D-optimal designs. The contribution of this article is not to find a class of designs that have higher d efficiency in terms of the projective model. Rather, it is to illustrate a simpler and more realistic approach to the class of screening designs when the number of runs is limited. D efficiency is used here to illustrate that these designs generally have high efficiency in terms of the projective models. On the other hand, if a slightly larger run is possible, the Plackett and Burman type design (where n must be a multiple of 4) has been proven to be optimal for many reasons and thus is strongly recommended.

ACKNOWLEDGMENTS

I am most grateful to the past editor, Vijay Nair, and two anonymous referees for their extensive and constructive remarks, which have led to major improvements in the presentation. My thanks also go to Toby Mitchell (Oak Ridge National Laboratory), who gave many valuable comments on an earlier version of this manuscript. This research was partially supported by the National Science Foundation via Grant DMS-9204007 and by the Professional Development Award, the University of Tennessee.

APPENDIX: PROOF FOR $s_{ij} \equiv n$ (MOD 4)

For the cross-product of any two columns, x_i and x_j, assume that $(+,+)$, $(+,-)$, $(-,+)$, $(-,-)$ appears b_1, b_2, b_3, and b_4 times, respectively.

Case 1 (n is even). Because of the equal-occurrence property, we have $b_1 + b_2 = b_3 + b_4 = n/2$ (for column x_i) and $b_1 + b_3 = b_2 + b_4 = n/2$ (for column x_j). Therefore, $b_2 = b_3 = n/2 - b_1$ and $b_4 = b_1$. This implies that $s_{ij} = (b_1 + b_4) - (b_2 + b_3) = 4b_1 - n$. In other words, $s_{ij} + n \equiv 0 \pmod 4$ or $s_{ij} \equiv n \pmod 4$ because n is even.

Case 2 (n is odd). Without loss of generality, also assume that each column contains $(n + 1)/2$ of $+1$s and $(n - 1)/2$ of -1s. We thus have $b_1 + b_2 = b_3 + b_4 + 1 = (n + 1)/2$ (for column x_i) and $b_1 + b_3 = b_2 + b_4 + 1 = (n + 1)/2$ (for column x_j). Therefore, $b_2 = b_3 = (n + 1)/2 - b_1$ and $b_4 = b_1 - 1$. This implies that $s_{ij} = (b_1 + b_4) - (b_2 + b_3) = 4b_1 - n - 2$. In other words, $s_{ij} + n \equiv 2 \pmod 4$ or $s_{ij} \equiv n \pmod 4$ because n is odd.

Thus, the candidate sets for s_{ij} are $n \equiv 0 \pmod 4$: $s_{ij} \in \{\ldots -8, -4, 0, 4, 8, \ldots\}$, $n \equiv 1 \pmod 4$: $s_{ij} \in \{\ldots -7, -3, 1, 5, \ldots\}$, $n \equiv 2 \pmod 4$: $s_{ij} \in \{\ldots -6, -2, 2, 6, \ldots\}$, and $n \equiv 3 \pmod 4$: $s_{ij} \in \{\ldots -5, -1, 3, 7, \ldots\}$.

[*Received January 1991. Revised December 1992.*]

REFERENCES

Booth, K. H. V., and Cox, D. R. (1962), "Some Systematic Supersaturated Designs," *Technometrics*, 4, 489–495.

Chadjipantelis, T., Kounias, S., and Moyssiadis, C (1987), "The Maximum Determinant of 21 × 21, (±1, −1)-Matrices and D-Optimal Designs," *Journal of Statistical Planning and Inference*, 16, 167–178.

Ehlich, H. (1964), "Determinantenabschätzungen für binäre Matrizen," *Mathematische Zeitschrift*, 83, 123–132.

Galil, Z., and Kiefer, J. (1980), "*D*-Optimum Weighing Designs," *The Annals of Statistics*, 8, 1293–1306.

Hadamard, J. (1893), "Résolution d'une question relative aux déterminants," *Bulletin des Sciences Mathémathques*, 17, 240–246.

Haines, L. M. (1987), "The Application of the Annealing Algorithm to the Construction of Exact Optimal Designs for Linear-Regression Models," *Technometrics*, 29, 439–447.

Hotelling, H. (1944), "Some Improvements in Weighing and Other Experimental Techniques," *The Annals of Mathematical Statistics*, 15, 297–306.

Kiefer, J. (1959), "Optimal Experimental Designs" (with discussion), *Journal of the Royal Statistical Society*, Ser. B, 21, 272–319.

——— (1974), "General Equivalence Theory for Optimal Designs (Approximate Theory)," *The Annals of Statistics*, 2, 849–979.

Mitchell, T. J. (1974a), "An Algorithm for the Construction of '*D*-Optimal' Experimental Designs," *Technometrics*, 16, 203–210.

——— (1974b), "Computer Construction of '*D*-Optimal' First-Order Designs," *Technometrics*, 16, 211–220.

Mood, A. M. (1946), "On Hotelling's Weighing Problem," *The Annals of Mathematical Statistics*, 17, 432–446.

Moyssiadis, C., and Kounias, S. (1982), "The Exact *D*-Optimal First Order Saturated Design With 17 Observations," *Journal of Statistical Planning and Inference*, 7, 13–27.

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

Raghavarao, D. (1959), "Some Optimum Weighing Designs," *The Annals of Mathematical Statistics*, 30, 295–303.

Welch, W. J. (1982), "Branch-and-Bound Search for Experimental Designs Based on *D*-Optimality and Other Criteria," *Technometrics*, 24, 41–48.

Williamson, J. (1946), "Determinants Whose Elements Are 0 and 1," *American Mathematics Monthly*, 53, 427–434.