TECHNOMETRICS, NOVEMBER 1992, VOL. 34, NO. 4

Projection Properties of Plackett and Burman Designs

Dennis K. J. Lin
Statistics Department
University of Tennessee
Knoxville, TN 37996

Norman R. Draper
Statistics Department
University of Wisconsin
Madison, WI 53706

The projection properties of the 2_R^{q-p} fractional factorials are well known and have been used effectively in a number of published examples of experimental investigations. The Plackett and Burman designs also have interesting projective properties, knowledge of which allows the experimenter to follow up an initial Plackett and Burman design with runs that increase the initial resolution for the factors that appear to matter and thus permit efficient separation of effects of interest. Projections of designs into 2–5 dimensions are discussed, and the 12-run case is given in detail. A numerical example illustrates the practical uses of these projections.

KEY WORDS: Adding runs; Hadamard matrices; Response surfaces; Screening designs; Two-level designs.

Plackett and Burman (1946, pp. 323–324) provided a series of two-level fractional factorial designs for examining $(n - 1)$ factors in n runs, where n is a multiple of four and $n \leq 100$. They omitted the case $n = 92$, which was later provided by Baumert, Golomb, and Hall (1962). When only main effects exist, these "PB designs" estimate all of them. PB designs are thus extremely useful in screening situations in which we examine many factors (e.g., 11) but believe only a few of these (e.g., 3 or 4 or 5) are of any consequence. We hope that the PB analysis will reveal those "real" effects, which we associate with the larger estimates. The $(n - 1)$ estimates available are the contrasts obtained by taking each column of signs of the PB design, multiplying them by the corresponding response values, and dividing the resulting sum by the divisor $\frac{1}{2}n$.

If main effects and two-factor iterations (*2fi*'s) exist, the estimation is blurred because the alias structure connects every main effect with a long linear combination of *2fi*'s (see Lin and Draper 1991a). One possibility to eliminate such blurring is to *fold over* the PB design (Box, Hunter, and Hunter 1978, pp. 340, 399)—that is, to repeat the PB design with all signs reversed. (Sometimes an extra factor is inserted using the I column of the PB design.) Foldover always converts a resolution III design into a resolution IV design (see Box et al. 1978, p. 398), but it doubles the run size—a major disadvantage in some contexts. In folding over a PB design, we must reset all $(n - 1)$ factors, whether they are "real" or not, for the second folded half of the design. That might not be judged necessary. In many 2_R^{q-p} examples, the projection properties of the initial design are used. For example, a 2_V^{5-1} design or a 2_{IV}^{8-4} design can be collapsed to a replicated 2^3 design if only three (*any* three) of the factors seem to be real. There are many such examples, such as those of Box et al. (1978, pp. 416–417) or Hare (1988). Not much effort has gone into examining the projection properties of the PB designs, however, both because of their irregular structure and because the existing numbers and types of the different projection designs have not been known. We examine some of these properties in this article, building on our recent work (Draper 1985; Draper and Lin 1990).

1. PROJECTIONS OF THE 12-RUN PLACKETT AND BURMAN DESIGNS

To see what the projection of the 12-run design is in any k of the 11 factor dimensions, we select k columns and examine the design that results by ignoring the other $11 - k$ columns. For example, suppose that $k = 3$; the reduced 12-point design consists of a 2^3 design plus a 2^{3-1} design with $I = \pm ABC$, where A, B, and C represent the three selected columns *no matter which three factors are designated as the survivors*. This is a very desirable arrangement, because it provides complete coverage of all of the factorial effects plus additional pure-error information obtained at four different locations well spread out over the experimental region.

Table 1 summarizes the situation for $n = 12$ runs and $k \leq 5$. An important aspect of this table is that, for $k \leq 4$, only one projected design type is obtained apart from variations caused by changes of signs in

Table 1. Projection of a 12-Run Plackett and Burman Design Into k Dimensions

k	*Design number*	*Description*
2	2.1	$2^2 \times 3$ (2^2 design with 3 replicates)
3	3.1	$2^3 + \frac{1}{2} \times 2^3$ (2^3 design plus 2^{3-1} design)
4	4.1	Add one more run to obtain a 2^{4-1}_{IV} design
		Add five more runs to obtain a 2^4 design
5	5.1	Add two more runs to obtain a 2^{5-2}_{III} design
		Add six more runs to obtain a 2^{5-1}_{V} design
	5.2	Add two more runs to obtain a 2^{5-2}_{III} design
		Add eight more runs to obtain a 2^{5-1}_{IV} design
		Add ten more runs to obtain a 2^{5-1}_{V} design

Table 2. Regrouping of 12 Plackett and Burman Runs Based on I = ±12 = ±134

Group	*Run no.*	*1*	*2*	*3*	*4*
I = 12 = 134	5	−	−	+	−
(= 234)	6	−	−	−	+
	8	+	+	−	−
I = −12 = 134	2	−	+	+	−
(= −234)	11 = 3	+	−	+	+
	4	−	+	−	+
	7	+	−	−	−
I = 12 = −134	1	+	+	−	+
(= −234)	9	+	+	+	−
	12	−	−	−	−
I = −12 = −134	10	−	+	+	+
(= 234)					

the columns. See also Box and Bisgaard (in press).

For $k = 4$, one run can be added in the four selected design factors only, ignoring the other seven for the moment, to complete a 2^{4-1}_{IV} design; there are five additional runs as well. The added point is always the foldover complement of the duplicate point that arises and is uniquely determined. The other six runs of the 2^{4-1} design consist of three mirror-image (foldover) pairs. If a full 2^4 design is wanted, only five runs need be added. A minor difficulty arises over the question of which of the original duplicate points should be used in the analysis of the 2^{4-1}_{IV} design or whether the average should be used. The latter has half the variance of the individual observations and should be weighted accordingly. If the assumptions that the ignored effects are negligible are correct, the duplicate observations should be almost the same anyway.

What about the levels of the other seven variables that (it would have been decided at this stage) are not active ones? If the assumption made about the identity of the inactive factors is true, the choice of their levels should not be crucial. One possibility is to set quantitative variables at their central value 0, halfway between the minus and plus levels; or they could be set to either of their plus or minus values, perhaps their levels in the original design. Choices vary with circumstances.

Another possibility is to link into the $\frac{3}{4}$ replicate designs of John (1961, 1962). Consider columns 1–4 of Table 2, obtained from a particular projection of a 12-run PB design. The 12 runs can be categorized into the four groups shown.

If we identify our 1, 2, 3, 4 with John's (1961, p. 319) x_2, x_4, x_3, x_1, we see that we are two runs short of a $\frac{3}{4}$ replicate of a 2^4 design using the first three segments of a 2^4. The additional runs needed are obviously (+ + + +) in the first segment, and (− − + +) in the third. This two-run addition provides a compromise between adding one run and five runs. Adding these runs permits estimation of all main effects and two-factor interactions, providing we can assume that three- and four-factor interactions can be ignored. To analyze $\frac{3}{4}$ replicates of a 2^4, we can follow John (1961), using only one of the duplicate runs in the analysis. The extra runs (e.g., runs 3 or 11 and 10 in Table 2), can be used to check predictions from the model implicitly fitted. It would also be possible to analyze all of the data available if this were felt to be preferable. Typically this would be done by the least squares fitting of a suitable linear model—for example, the one appropriate to the design completed by the added runs.

For $k = 5$, two types of designs are possible, one with a repeat-run pair and one with a mirror-image pair; see Draper (1985, table 2). Several possibilities exist for supplementing these designs. If we write the two distinct 12-run, $k = 5$, reduced designs in the standard forms shown in Table 3, the results specified in our Table 1 are obtained by adding the additional runs listed in Table 3. This principle always works, but the actual runs to be added depend on the specific 12-point projection one has, and that depends on the choice of columns. All 462 possibilities correspond exactly to one or the other of the designs in Table 3, but sign changes in columns will occur. To determine which of the two design types has been obtained via projection, one must see if the specific design has a repeat run pair, or a mirror-image run pair, an easy thing to do, and then mentally convert it, via column sign changes and perhaps by rearranging the rows and columns appropriately, to one of the design types of Table 3.

In all such cases, the model appropriate to the completed 2^{q-p}_R design could be fitted by least squares, and the runs already made in addition to the 2^{q-p}_R runs will provide some residual degrees of freedom in an analysis of variance table.

Table 3. Additional Runs Needed to Convert Five Plackett and Burman Columns to the Designs Indicated

Type	*Design 5.1*						*Type*	*Design 5.2*					
	1	−	−	−	−	−		1	−	−	−	−	−
	2	−	−	−	−	−		2	+	+	+	+	+
	3	−	−	+	+	+		3	−	−	+	+	+
	4	−	+	−	+	+		4	−	+	+	−	+
	5	−	+	+	−	+		5	+	−	−	+	+
	6	−	+	+	+	−		6	+	+	−	+	−
	7	+	−	−	+	+		7	+	+	+	−	−
	8	+	−	+	−	+		8	+	−	+	−	−
	9	+	−	+	+	−		9	+	−	−	−	+
	10	+	+	−	−	+		10	−	+	−	+	−
	11	+	+	−	+	−		11	−	+	−	−	+
	12	+	+	+	−	−		12	−	−	+	+	−
2^{5-2}_{III}	13	−	+	−	−	+	2^{5-2}_{III}	13	−	−	−	+	−
(I = −134	14	+	−	−	+	−	(I = 124	14	−	+	+	−	−
= 1235)							= 1235)						
2^{5-1}_{V}	13	+	−	−	−	−	2^{5-1}_{V}	13	−	−	+	+	+
(I = 12345)	14	−	+	−	−	−	(I = 12345)	14	−	+	+	−	+
	15	−	−	+	−	−		15	+	−	−	+	+
	16	−	−	−	+	−		16	+	+	−	+	−
	17	−	−	−	−	+		17	+	+	+	−	−
	18	+	+	+	+	+		18	+	−	−	−	−
								19	−	+	−	−	−
								20	−	−	+	−	−
								21	−	−	−	+	−
								22	−	−	−	−	+
2^5	Add the 21 runs of a 2^5 that are other than rows 2–12 above.						2^5	Add the 20 runs of a 2^5 that are other than runs 1–12 above.					

Notes on Table 3.

1. A 2^{5-1}_{IV} with $I = 1235$ can be obtained by adding eight runs to design 5.2. Runs of design 5.2 such that $I = 1235$ already are numbers 1, 2, 3, 5, 6, 8, 9, and 11 in Table 3. The eight runs to be added will thus be (+ + − − −), (− − + − +), (− + + − −), (+ + + − +), (− − − + −), (− + − + +), (+ − + + −), and (− + + + −). As we have seen, however, a resolution V design can be produced by adding 10 runs, just two more.

2. A more general question, not investigated, is whether any five orthogonal, zero-sum columns with ±1 entries can be augmented by 14 more mutually orthogonal zero-sum columns with ±1 entries. Even if it is possible, the answer will not necessarily be unique because there are three different types of 20-by-20 Hadamard matrices (Hall 1965). Vijayan (1976, theorem 3) noted that, although recovery of one, two, or three missing columns of a Hadamard matrix can be uniquely made, recovery of four missing columns (and therefore of more than four) is not unique (subject, as usual, to sign changes in the columns and rearrangement of runs).

3. Point 2 raises other interesting questions. If we add more runs, we have, of course, to decide on the levels to be used for the variables (let us call them 6–11) judged not to be of interest after the first-stage analysis has indicated the five variables (1–5, say) requiring further experimentation. Considerations similar to those discussed for the $k = 4$ case arise once more. When we have design 5.2 as our chosen projected design, however, we can add 20 points forming five columns of a 20-run PB design to complete a 2^5 factorial in variables 1–5. This permits choice of six *additional* columns from the 20-run PB design to determine levels for factors 6–11 and to complete a 32-run, 11-variable, two-level design. Analysis of the 20-run portion alone allows the initial conjectures about variables 6–11 to be confirmed or denied separately, while permitting a full 2^5 analysis of factors 1–5 if the initial conjectures are confirmed. What is the best choice of 6 columns from the 14 available? What sorts of designs do the various choices give when considered as a 32-run design in 11 factors?

2. PROJECTIONS OF PLACKETT AND BURMAN DESIGNS WITH $n > 12$

A full investigation of all projections of the n = 20, 24, 28, 36, . . . run cases to lower dimensions is an enormous task and remains uncompleted. We content ourselves here principally with some remarks concerning $k = 2, 3, 4$, and 5, where k is the number of dimensions into which the n-run design is projected, for $n = 20$ and 24. For $k = 2$, the projection is always a 2^2 design, $n/4$ times over.

For $k = 3$ and $n = 20$, two types of projections can occur.

1. A 2:3 type. (This means two full 2^3 factorials and an additional 2^{3-1}. At the corners of the cube there are either two or three points.)
2. A 1:4 type. (This means a 2^3 factorial and three identical 2^{3-1} designs. At each corner of the cube there is either one point or there are four points.)

These and other possibilities for $k = 3$ appear in Table 4. In some cases, we can proceed from a three-column n-runs projection to a three-column $(n + 4)$-runs projection by simply adding a 2^{3-1} design. For $n = 20$, a (2:3) can be converted into either a (3:3) or a (2:4) with $n + 4 = 24$, depending on which 2^{3-1} is added. Similarly, a (1:4) can become a (2:4) or a (1:5); however, the latter is not a three-column projection of a 24-run PB design.

For details on all the projection possibilities into $k = 4$ dimensions for the 20- and 24-run cases, see Lin and Draper (1991b, table 7). For $n = 20$, only three types of projections exist (apart from sign changes in the columns, permutations of the columns, and rearrangements of the rows). These require the addition of 1, 4, and 4 more runs to provide a complete 2^4. Which projection is actually attained depends on the specific four factors retained after analysis. For $n = 24$, two types of projections provide a full 2^4 design plus a 2^{4-1}_R, where R = III or IV. Two other types of projection require two additional runs to complete a full 2^4 design.

For cases $k = 5$, $n = 20$, 24, see Lin and Draper (1991b, tables 8 and 9). There are nine possible projections for each n. An additional tenth projection for $n = 20$ arises from a Hadamard matrix not equivalent to a PB design. For a discussion of non-equivalent Hadamard matrices for $n = 16$ and 20, see Hall (1961, 1965) and Lin and Draper (1991b, sec. 6).

Table 4. Projections of n-Run Plackett and Burman designs into k = 3 Dimensions

n	*Types of projections**
12	1:2
16	2:2; 0:4
20	2:3; 1:4
24	3:3; 2:4
28	3:4; 2:5
32	4:4; 0:8
36	4:5; 3:6

*The notation "$r : s$" means r points lie at four of the "2^{3-1} locations" and s points lie at the other four. 2^{3-1} locations are always defined by $I = \pm$ the relevant three-factor interaction.

3. ILLUSTRATIVE EXAMPLE

PB-type projection analyses of the dimensions discussed in this article are new, and no published examples exist. Therefore, we provide a constructed example.

Suppose that an experiment is to be conducted on 10 factors, not more than half of which, it is anticipated, will have an effect on the response. A possible design that uses columns 1–10 of the (11-column) 12-run PB design is shown as runs 1–12 of Table 5, together with corresponding response values. (Ignore runs 13–18 for the moment.) Choice of *any* 10 of the original 11 columns would give a design that is essentially equivalent apart from sign changes in the columns and the reordering of column headings and run order (Draper and Lin 1990). In practice, runs would be performed in a randomized order.

From runs 1–12, estimates are obtained by taking the crossproduct of columns of signs times the y_i and dividing the results by 6. The corresponding 10 effects are, in order, 16.3, 1.1, 14.3, 2.6, 4.5, .7, 32.7, 23.0, .7, 42.7, and the unused eleventh column provides an estimate of 2.4, which can be tentatively called error. A first assessment indicates that columns 1, 3, 7, 8, and 10 may be connected with real effects. To explore this further, runs 13–18 are performed in factors 1, 3, 7, 8, and 10 only. (Although the levels of the other factors have been left blank in Table 4, some specific choice would have to be made for them, as previously explained.) Apart from runs 5 and 10, we now have a 2^{5-1}_{V} design defined by $I = -1378\overline{10}$. Runs 5 and 10 can now be considered as pseudorepeats. They are not repeats in all factors, merely in factors 1, 3, 7, 8, and 10. We have design 5.1 in fact.

A standard (e.g., Yates) analysis on all runs except 5 and 10 now gives the following estimates for main effects and two-factor interactions involving the five selected factors:

72.2 →	mean	22.6 →	8
−18.6 →	1	−2.6 →	18
−13.4 →	3	−1.7 →	38
−1.3 →	13	1.6 →	−7 10
33.4 →	7	5.6 →	78
1.6 →	17	.3 →	−3 10
1.1 →	37	.3 →	−1 10
.7 →	−8 10	−44.4 →	−10

Only the 78 interaction seems to be a candidate for significance (in addition to the five main effects). A 9 df estimate of the standard error of an effect can be made by taking the square root of the average

Table 5. A 10-Factor 12-Run Plackett and Burman Design With Six Added Runs

Run no.	1	2	3	4	5	6	7	8	9	10	y_i
1	+	+	−	+	+	+	−	−	−	+	70.19
2	−	+	+	−	+	+	+	−	−	−	57.12
3	+	−	+	+	−	+	+	+	−	−	63.17
4	−	+	−	+	+	−	+	+	+	−	99.43
5	−	−	+	−	+	+	−	+	+	+	90.72
6	−	−	−	+	−	+	+	−	+	+	110.37
7	+	−	−	−	+	−	+	+	−	+	120.36
8	+	+	−	−	−	+	−	+	+	−	40.15
9	+	+	+	−	−	−	+	−	+	+	81.38
10	−	+	+	+	−	−	−	+	−	+	88.89
11	+	−	+	+	+	−	−	−	+	−	9.63
12	−	−	−	−	−	−	−	−	−	−	36.25
13	−		−				−	+		+	104.18
14	+		−				+	−		−	50.28
15	−		+				−	−		+	71.74
16	−		+				−	+		−	43.66
17	+		+				−	+		+	67.76
18	−		+				+	+		+	129.10

sum of squares of the small effects (Box et al. 1978, p. 327) to give $\{(1.6^2 + (-1.6)^2 + \ldots + .3^2)/9\}^{1/2} = 1.43$, by which standard we would judge the 78 interaction (value 5.6) to be significant. Now $1.43 = (4\hat{\sigma}^2/16)^{1/2}$, so $\hat{\sigma}^2 = 8$, approximately. From runs 5 and 10, we can also obtain a 1-df pseudo pure-error estimate of $s_e^2 = (90.72 - 88.89)^2/2 = 1.7$, lower but not significantly so. Thus the analysis seems to be satisfactory.

In examples like this, specific details are, of course, open to question; for example, why were runs 5 and 10 not used in the analysis? If they were included, would not biases be introduced? In general, in experimental work done with limited resources, no "perfect" analysis is possible. All analyses and conclusions are compromises based on one's combination of prior knowledge of the experimental situation and current examination of the data. Future experimental work planned to investigate that plausibility will provide information to confirm or deny the current opinion. The work we have given in this article exhibits the relative simplicity of the projection properties of PB designs and provides the experimenter with new ways to obtain and add information in a relatively few runs in somewhat complex experimental screening situations.

4. ADDITIONAL NOTES

In this article we focus on augmenting projected PB designs by adding a few more factorial runs. It is also often possible to study a few interactions in these types of designs *without* adding any runs; see Hamada and Wu (1991). If additional runs were needed, they could be added based on the optimization of some chosen variance criterion. Possible criteria include a "determinant of $\mathbf{X}'\mathbf{X}$ criterion" or a "rotatability criterion." For the former, see Hamada and Wu (1991); for the latter, see Draper and Pukelsheim (1990).

ACKNOWLEDGMENTS

N. R. Draper gratefully acknowledges partial support from the National Science Foundation under Grant DMS-8900426 and from the Wisconsin Alumni Research Foundation through the Graduate School, University of Wisconsin. D. K. J. Lin was partially supported by a Faculty Research Fellowship through the College of Business Administration, University of Tennessee. We are very grateful to the editor, the associate editor, and the referees for a number of pertinent and thoughtful comments.

[Received April 1990. Revised December 1991.]

REFERENCES

Baumert, L. D., Golomb, S. W., and Hall, M. J. (1962), "Discovery of an Hadamard Matrix of Order 92," *Bulletin of the American Mathematical Society*, 68, 237–238.

Box, G. E. P., and Bisgaard, S. (in press), "What Can You Find Out From 12 Experimental Runs," *Quality Engineering*, 4.

Box, G. E. P., Hunter, W. G., and Hunter, J. S. (1978), *Statistics for Experimenters*, New York: John Wiley.

Draper, N. R. (1985), "Small Composite Designs," *Technometrics*, 27, 173–180.

Draper, N. R., and Lin, D. K. J. (1990), "Using Plackett and Burman Designs With Fewer Than $N - 1$ Factors," Working Paper 253, University of Tennessee, College of Business Administration.

Draper, N. R., and Pukelsheim, F. (1990), "Another Look at Rotatability," *Technometrics*, 32, 195–202.

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

——— (1965), "Hadamard Matrix of Order 20," *Technical Report 1*, Jet Propulsion Laboratory, pp. 32–76.

Hamada, M., and Wu, C. F. J. (1991), "Analysis of Designed Experiments With Complex Aliasing," Research Report RR-91-01, University of Waterloo, Canada, Institute of Improvement in Quality and Productivity.

Hare, L. B. (1988), "In the Soup: A Case Study to Identify Contributors to Filling Variability," *Journal of Quality Technology*, 20, 36–43.

John, P. W. M. (1961), "Three-Quarter Replicates of 2^4 and 2^5 Designs," *Biometrics*, 17, 319–321.

——— (1962), "Three-Quarter Replicates of 2^n Designs," *Biometrics*, 18, 172–184.

Lin, D. K. J., and Draper, N. R. (1991a), "Generating Alias Relationships for the Plackett and Burman Designs," Technical Report 876, University of Wisconsin, Dept. of Statistics. (A shortened version will appear in 1993 in *Computational Statistics and Data Analysis*, 15.)

——— (1991b), "Projection Properties of Plackett and Burman Designs," Technical Report 885, University of Wisconsin, Dept. of Statistics.

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

Vijayan, K. (1976), "Hadamard Matrices and Submatrices," *Journal of the Australian Mathematical Society*, Ser. A, 22, 469–475.