Two-Level Designs with Good Projection Properties

Oddgeir Samset and John Tyssedal
Department of Mathematical Sciences
The Norwegian University of Science and Technology
N-7491 Trondheim, Norway

Abstract

We consider projection properties of Plackett-Burman designs and designs constructed from these. The designs are examined by computer search, and projection properties are summarised by means of their repeat and mirror-image patterns. In particular we found that Plackett-Burman designs with run sizes equal to 68, 72, 80 and 84 are of projectivity $P = 4$. Several methods for constructing designs with good projections properties are discussed, and some orthogonal as well as super-saturated designs with such properties are given.

Key Words: Plackett-Burman designs; Projectivity; Repeat and mirror-image patterns.

1 Introduction

In factor screening often only a small number of factors out of potentially many really affect a particular response. When choosing an appropriate design, it is then of importance to consider projections of the design on to small subsets of factors. This rationale for studying projection properties of screening designs was first pointed out by Box and Hunter (1961). Later a definition of *projectivity* for two-level designs was given by Box and Tyssedal (1996). A $n \times k$ design with n runs and k factors each at 2 levels is said to be of projectivity P if every subset of P factors out of possible k contain a complete 2^P factorial design, possibly with some points replicated. We will describe such a design as a (n, k, P) screen.

The concept of projectivity is related to *strength* of orthogonal arrays (Rao, 1947), but differs in that complete copies of a 2^P factorial are not required to have projectivity P. Projectivity is also related to the *resolution* of a design (Box and Hunter, 1961). A design is of resolution R if no p-factor effect is confounded with any other effect containing less than $R - p$ factors. For factorial two-level orthogonal arrays, also called *geometric* designs, it is always the case that $P = R - 1$. While saturated orthogonal two-level designs always are of resolution $R = 3$, it has been verified by computer search (Box and Bisgaard, 1993; Lin and Draper, 1992) and later shown mathematically (Cheng, 1995; Box and Tyssedal, 1996) that *non-geometric* Plackett-Burman (PB) designs, i.e. those for which the run size is not a power of two, most often have projectivity higher than 2. Therefore other criteria than resolution may be more appropriate for describing properties of these designs.

Statistically, projectivity P implies that all main effects and all interactions of any p factors are estimable with no bias if the other factors are inert. In addition, replicated runs may provide information about the error variance of the response values. The estimation of potential active effects is, however, not necessarily restricted to p factors. Cheng (1995) showed that for orthogonal two-level arrays, whose run sizes are not a multiple of eight, all main effects and two-factor interactions are estimable for any four factors if higher order interactions are negligible. If the design is of strength three and n is not a multiple of 16, Cheng (1998) also showed that this holds for any set of five factors. In general no two-factor interaction column is fully confounded with a main effect column in a projectivity $P = 3$ design. If the design is of projectivity $P = 4$, no two-factor interaction columns will be fully confounded in addition.

The main purpose of this paper is to point to specific designs with good projection properties. A natural set of arrays to investigate is the designs derived by Plackett and Burman (1946). These are two-level orthogonal arrays which exist when n is a multiple of 4 for $n \leq 100$, except for $n = 92$, and can be used to screen up to $n-1$ factors in n runs. Box and Tyssedal (1996) classified all the PB designs for $n \leq 84$ with respect to being of projectivity $P = 2$ or $P = 3$. Whether any of them might be of higher projectivity has so far been an open question. Several results about projection properties for orthogonal arrays have been provided by Cheng (1995, 1998). These results are, however, very general, and when a smaller class of designs is investigated, other projection results may exist. Motivated by this a computer search was conducted for the PB designs, and projection properties were summarised by means of the repeat and mirror-image (RMI) pattern introduced by Draper and Lin (1995). In particular we have focused on what information can be extracted from the RMI pattern of a design and how to use this in further deduction of design properties. This research enabled us to discover that four PB designs are of projectivity $P = 4$ and also some super-saturated designs with good projection properties. These results are given in Section 2. In Section 3 we give some ways of constructing designs of high projectivity from PB design, mainly $P = 3$ and $P = 4$ designs. We also found that some of the PB designs are 'near' projectivity $P = 4$ designs. We discuss these designs further in Section 4, and give another way of characterising them by means of the absolute value of the averaged dot product between contrast columns.

2 Projection Properties from the Repeat and Mirror-Image Pattern

An useful way to describe the type of different projections is in terms of their RMI pattern (Draper and Lin, 1995). A mirror-image of a run is obtained if all signs in the run are switched. The RMI pattern describes how many single runs there are (i.e. neither repeats nor mirror-images involved), how many runs that occur in repeat pairs, how many runs occur as mirror-images to repeat pairs, and so on. We will use a slightly modified notation of the one introduced by Draper and Lin. Hence by a $s[r\ m]$ expression we shall mean that there are r replicates of s different runs and m replicates of the mirror-images of these s runs. A typical RMI pattern is then given as: $s_1[r_1\ m_1]\ s_2[r_2\ m_2] \ldots$. As an example, the most frequent projection of the 68 run PB design on to 4 factors, see Table 1, then has the following form: $1[6\,3]\ 3[5\,4]\ 3[5\,3]\ 1[4\,4]$. This is to be read as follows: There are six replicates of one run and three replicates of its mirror-image, five replicates of three runs and four replicates of each of their mirror-images, five replicates of three runs and three replicates of each of their mirror images and finally four replicates of one run and its mirror-image. When n is increasing, the list of all possible type of projections on to four and more factors, given by their RMI pattern, rapidly grows long. Then it may be useful to summarise the information in the RMI pattern further. Let the repeat and mirror-image pattern for a projection on to P factors be denoted

Table 1: Repeat and mirror-image pattern for the 68 run Plackett-Burman design.

Columns	Repeat and mirror-image pattern	Unique runs	No. of projections
1 2	2[17 17]	4	2211
1 2 3	4[9 8]	8	28743
1 2 6	4[10 7]	8	18425
1 2 39	4[11 6]	8	737
1 2 3 4	4[5 4] 1[5 3] 3[4 4]	16	55275
1 2 3 6	2[6 3] 1[6 2] 2[5 4] 1[5 3] 2[4 4]	16	132660
1 2 3 7	1[7 3] 1[6 4] 2[5 5] 2[5 2] 2[4 3]	16	112761
1 2 3 8	3[6 4] 1[5 5] 1[5 2] 3[4 3]	16	88440
1 2 3 9	1[6 4] 3[5 5] 4[4 3]	16	39798
1 2 3 10	1[6 3] 3[5 4] 3[5 3] 1[4 4]	16	159192
1 2 3 39	3[6 3] 1[6 2] 1[5 4] 3[5 3]	16	8844
1 2 3 42	1[7 2] 1[6 2] 3[5 4] 3[5 3]	16	75174
1 2 3 44	4[6 5] 1[4 2] 3[3 3]	16	2211
1 2 4 24	2[7 3] 1[6 4] 1[6 1] 1[5 5] 1[5 2] 2[4 3]	16	8844
1 2 4 34	2[7 4] 2[6 5] 1[5 1] 1[4 2] 2[3 3]	16	4422
1 2 5 25	1[7 2] 1[7 1] 2[6 3] 1[5 4] 2[5 3] 1[4 4]	16	8844
1 2 5 29	1[7 4] 3[6 5] 3[4 2] 1[3 3]	16	2948
1 2 5 31	1[7 3] 3[6 4] 1[6 1] 3[4 3]	16	35376
1 2 5 39	1[7 3] 3[6 4] 3[5 2] 1[4 3]	16	2948
1 2 6 17	1[8 1] 3[6 2] 3[5 4] 1[5 3]	16	2211
1 2 6 20	1[8 2] 2[6 4] 1[6 1] 1[5 5] 2[5 2] 1[4 3]	16	6633
1 2 6 31	1[7 1] 4[6 3] 3[4 4]	16	8844
1 2 6 33	1[8 3] 3[6 5] 1[5 1] 3[4 2]	16	2211
1 2 7 39	1[7 2] 1[6 3] 2[6 2] 2[5 4] 1[5 3] 1[4 4]	16	8844

as RMI_P, and let I_{m_i} be an indicator function equal to 1 if $m_i > 0$ and equal to 0 otherwise. Some information that can be extracted from the RMI pattern is given below.

Theorem 2.1 *For the RMI pattern of any two-level design we have the following:*

i) *The number of unique runs for a projection on to P factors is given by*

$$\sum_{s_i \in RMI_P} s_i(1 + I_{m_i}).$$

ii) *A design is of projectivity P if and only if*

$$\sum_{s_i \in RMI_P} s_i = 2^{P-1}$$

and all $m_i > 0$ for all projections on to P factors.

iii) *Let P be an odd number. If the RMI_P contains terms of the form $s[r\ m]$ with $m > 0$, there is no defining relation between these P factors.*

iv) *Let P be an even number. If the RMI_P contains more than 2^{P-1} numbers of unique runs, there is no defining relation between these P factors.*

Proof.

i) It is a direct consequence of the definition of the RMI pattern.

ii) The result follows since each of the 2^P factorial designs has exactly 2^{P-1} repeat mirror-image pairs.

iii) A mirror-image run contained in a projection on to an odd number of factors prevents a defining relation to exist between these factors.

iv) If more than 2^{P-1} unique runs exist in the RMI_P, where P is an even number, there must be runs with different defining relation in the projection.

For orthogonal two-level designs it is possible to obtain an equivalence relation between the existence of a defining relation and the RMI_P when projections on to three or four factors are under consideration.

Theorem 2.2 *For the RMI pattern of an orthogonal two-level design we have:*

i) *A defining relation exists between three factors if and only if there is at least one projection with a RMI_3 of the form $4[r\ 0]$.*

ii) *A defining relation exists between four factors if and only if there is at least one projection with a RMI_4 solely of the form $4[r\ m]$ with $r = m$.*

Proof.

i) In order for a defining relation to exist between three factors in a two-level orthogonal array, the projection must be replicates of the four runs in a 2^{3-1}_{III} design, see Cheng (1995, Theorem 2.1). For three factors a 2^{3-1}_{III} design has a RMI_3 of the form $4[1\ 0]$.

ii) If the RMI_4 is solely of the form $4[r\ m]$ with $r = m$, the projection only contains repeat and mirror-image pairs and a defining relation exists. Suppose a defining relation exists. According to Cheng (1995, Theorem 2.1) the four factor columns must contain $2^{-3}n$ copies of a 2^3 design for any three factors. Otherwise the two-level design cannot be orthogonal. Once the level combinations for three factors are given, the signs for the fourth factor are also determined by the defining relation. This shows that the projection must be $2^{-3}n$ copies of a 2^{4-1}_{IV} design. Therefore $s = 4$ and the number of copies is equivalently given by r or m which then must be equal.

The number of unique runs for a projection on to P factors will tell us how many effects with these factors involved it may be possible to estimate simultaneously. It is also relevant information for the ability of the design to identify active factors. The more unique runs the projections on to P factors in general contain, the more equations have to be satisfied in order that different models with P factors shall have the same expected response values. Hence the easier the process of separating out one model from the others will normally be.

The RMI pattern for the 68 run PB design for projections on to four or less factors is given in Table 1. Here the design is constructed by taking the generator as the first column and then cycling this column downwards. A last row of -1's is finally added. We notice that the condition in Theorem 2.1 ii) is satisfied. Hence this design is of projectivity $P = 4$. The same is true for the cyclic generated PB designs with the number of runs equal to 72, 80 and

84. The RMI patterns for these designs on to four or less factors can be found in Samset and Tyssedal (1998).

Another result that can be found directly from the repeat and mirror-image pattern is the following. Cheng (1998, Corollary 2) showed that if no defining relation exists among any four as well as any three factors in two-level orthogonal arrays, this is equivalent to that all main effects and two-factor interactions for any four factors are estimable when higher order interactions are negligible. From Theorem 2.1 i) and ii) we get the following result.

Corollary 2.3 *A necessary and sufficient conditions that all main effects and two-factor interactions of any four factors are estimable in two-level orthogonal designs, assuming higher order interactions are negligible, is that no projection on to three factors has a* RMI_3 *equal to* $4[r\ 0]$ *and no projection on to four factors has a* RMI_4 *solely of the form* $4[r\ m]$ *with* $m = r$.

For PB designs we only found this to happen for the geometric designs and for the designs with run sizes equal to 40, 56, 88 and 96. These four PB designs are the only non-geometric PB designs for which $P = 2$ and only 2.

Lin (1993) introduced a new class of super-saturated designs constructed using half-fractions of Hadamard matrices. The next theorem is about the projectivity of designs constructed via half-fractions.

Theorem 2.4 *Any half-fraction constructed using one of the design columns as a branching column in a two-level projectivity* P *design is at least of projectivity* $P - 1$.

Proof. Let us take the $n/2$ rows that corresponds to either $+1$ or -1 in an arbitrary column in a projectivity P design. It is then obvious that for these rows any $P - 1$ different columns will contain a 2^{P-1} design. Hence these half-fractions are at least of projectivity $P - 1$.

The projectivity of a design is not always reduced with 1 when a half-fraction is constructed, but cyclic generated two-level orthogonal arrays represent an important class of designs for which this holds. For these both half-fractions are of the same projectivity independent of which column is used as a branching column. Assume that the projectivity of the half-fractions is P. It then follows that for any $P + 1$ columns in the original design where one column is the branching column, the P others will contain a 2^P design in their $n/2$ rows corresponding to $+$ and $-$ signs of the branching column. Hence the original orthogonal array is of projectivity $P + 1$. Therefore for cyclic generated designs we have a reduction in projectivity of 1 when a half-fraction is constructed using one of the design columns as a branching column. A further reduction in projectivity when additional fractioning is performed is, however, not true in general. It follows that all the super-saturated designs introduced by Lin (1993) are of projectivity $P = 2$, and the ones obtained as half-fractions of the four cyclic generated projectivity $P = 4$ designs mentioned above are also $(34, 66, 3)$, $(36, 70, 3)$, $(40, 78, 3)$ and $(42, 82, 3)$ screens respectively.

3 Constructing Designs with Good Projection Properties

In general we expect that it is possible to increase the projectivity of a design by removing columns. Removing columns from PB designs, however, is in general not an efficient way of obtaining designs with high projectivity. For instance a search for the maximum number of columns of projectivity $P = 4$ in PB designs for $n = 36$, 40, 44, and 48 ended up with 9, 20, 10 and 11 columns respectively. The promising result for the 40 run PB design indicated that folding over designs might be a more efficient way of obtaining $P = 4$ design. The fold-over of a $n \times k$ two-level design is the design obtained when all the mirror-images of the original design's runs are added in addition to a column where the first n entries are $+1$ and the last n entries are -1. Hence for a two-level design X its fold-over is given by

$$\begin{bmatrix} X & 1 \\ -X & -1 \end{bmatrix} \tag{1}$$

Seiden and Zemach (1966) showed that the fold-over of a $n \times k$ orthogonal two-level design with strength 2, is a $2n \times (k+1)$ design with strength 3, and therefore according to Cheng (1995, Corollary 2.4) also a projectivity $P = 4$ design in $k+1$ factors if n is not a multiple of 8. Let us start with a closer examination of the design obtained when just the mirror-images have been added, i.e. the design $[X' \quad -X']'$. As mentioned in Section 2 any four columns of projectivity $P = 4$ contain 8 repeat mirror-image pairs. Therefore, if a projection on to 4 factors contains one run from each of its mirror-image pairs, it will be of projectivity $P = 4$ when all its mirror-images have been added. This way of obtaining projectivity $P = 4$ designs can of course be generalised to arbitrary projectivity.

Theorem 3.1 *A necessary and sufficient condition that a $n \times k$ two-level design can be augmented to a $2n \times (k)$ design of projectivity P by adding all its mirror-images is that*

$$\sum_{s_i \in RMI_P} s_i = 2^{P-1}$$

for all projections on to P factors in the original $n \times k$ design.

Proof. If all the mirror-images are added to a two-level design containing one run from each of its 2^{P-1} repeat mirror-image pairs, the resulting design will have 2^P unique runs and hence be of projectivity P. If on the other side the $n \times k$ design does not contain runs from each of the possible 2^{P-1} mirror-image pairs, the augmented design cannot be of projectivity P.

As a consequence of Theorem 3.1 a projectivity P design is obtained if all mirror-images are added to a two-level array where all projections on to P factors contain a 2^{P-1}_{III} design. This follows from the fact that these respective runs contain no mirror-images and hence $\sum_{s_i \in RMI_P} s_i = 2^{P-1}$ for all these projections. In general the last statement is true as long as all projections on to P factors are in the general class of 2^{P-1}_R designs and R is an odd number.

If the original $n \times k$ two-level design is of projectivity $P-1$ we obtain the following result.

Theorem 3.2 *A necessary and sufficient condition that a $n \times k$ two-level design of projectivity $P-1$ can be folded over to a $2n \times (k+1)$ design of projectivity P is that*

$$\sum_{s_i \in RMI_P} s_i = 2^{P-1}$$

for any projection on to P factors in the original $n \times k$ design.

Proof. This follows from Theorem 3.1 and the fact that if $P-1$ columns are of projectivity $P-1$, they will together with the column of n $+1$'s and n -1's constitute a projection on to P factors with projectivity P.

Table 2: The $(16, 15, 2)$ screen which is also a $(16, 14, 3)$ screen if column 9 is removed. Add mirror-image runs to the first nine columns to obtain a $(32, 9, 4)$ screen.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
+	+	-	+	+	+	+	+	+	+	+	+	+	+	+
+	+	+	-	+	-	-	-	-	+	+	+	+	-	-
+	+	+	+	-	+	+	+	-	+	-	-	-	-	-
+	+	+	+	+	-	-	-	+	-	-	-	-	+	+
+	-	+	-	-	+	-	+	+	-	+	-	+	-	+
+	-	-	+	-	-	-	+	-	-	+	+	-	+	-
+	-	-	-	+	+	+	-	-	-	-	-	+	+	-
+	-	-	-	-	-	+	-	+	+	-	+	-	-	+
-	+	+	-	-	+	+	-	+	-	+	+	-	+	-
-	+	-	+	-	+	-	-	-	-	-	+	+	-	+
-	+	-	-	+	-	+	+	-	-	+	-	-	-	+
-	+	-	-	-	-	-	+	+	+	-	-	+	+	-
-	-	-	+	+	+	-	-	+	+	+	-	-	-	-
-	-	+	-	+	+	-	+	-	+	-	+	-	+	+
-	-	+	+	-	-	+	-	-	+	+	-	+	+	+
-	-	+	+	+	-	+	+	+	-	-	+	+	-	-

Inspection of the mirror-image patterns of the non-geometric PB designs shows that all of them, except for the ones with run sizes $n = 40$, 56, 88 and 96, satisfy $\sum_{s_i \in RMI_P} s_i = 8$ for all projections on to 4 factors. Hence the exception of being of projectivity $P = 4$ for the fold-over of non-geometric PB designs can be directly connected to these four designs. Especially we note that from the 12 and 20 run PB design, projectivity $P = 4$ designs may be constructed in 12 and 20 factors using 24 and 40 runs respectively.

It is well known that from the 32 run PB design it is possible to find 16 columns which constitute a $(32, 16, 3)$ screen, but only six columns will give a $(32, 6, 4)$ screen. It is shown in Box and Tyssedal (1995) that one of the 16 run orthogonal arrays discovered by Hall (1961) is a $(16, 14, 3)$ screen. From this design it is also possible to find nine factors which contain a 2^{4-1}_{III} design in every projection on to four factors. According to the comments after Theorem 3.1 a $(32, 9, 4)$ screen may be constructed by adding the mirror-images of its runs. The 16 run orthogonal array under consideration is given in Table 2. The nine first columns are the generators in the $(32, 9, 4)$ screen. The $(16, 14, 3)$ screen may be obtained by removing the ninth column from the design.

An effect of folding over a two-level design is to get rid of *odd* confounding, i.e. confounding between main effects and two-factor interactions, between two-factor and three-factor interactions and so on. Any *even* confounding, i.e. confounding between two-factor interaction columns, between main effects and three-factor interactions, etc. will not be changed. This leads to the following result.

Theorem 3.3 *For all the projections on to five factors in a two-level orthogonal projectivity $P = 4$ design constructed by folding over a design X, all the main effects and two-factor interactions are estimable under the assumption of negligible higher-order interactions.*

Proof. Since the fold-over is of projectivity $P = 4$, there can be no defining relation between any four factors as well as any three in the design X. According to Cheng (1998, Corollary 6) the result follows.

Doubling is another technique that focus on increasing the number of allowed columns with a certain projectivity P. From the original two-level design X the new design is then constructed as:

$$\begin{bmatrix} X & X & 1 \\ X & -X & -1 \end{bmatrix} \tag{2}$$

We have the following result.

Theorem 3.4 *From any $(n, k, 3)$ screen it is possible to construct a $(2n, 2k, 3)$ screen by doubling the design and removing the column with the n $+1$'s and the n -1's entries.*

Proof. Obviously the only two cases that need to be discussed are when two columns are picked from $[X' \;\; X']'$ and one from $[X' \;\; -X']'$ and vice versa. Let a and b be two columns from the $(n, k, 3)$ screen and let a typical such projection be written as:

$$\begin{bmatrix} a & b & x \\ a & b & -x \end{bmatrix} \tag{3}$$

If x is different from both a and b, a full 2^3 design can be extracted from these three columns since the original design is of projectivity $P = 3$. Let us assume $x = a$. Then the n first rows will contain the four combinations

$$\begin{bmatrix} -1 & -1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix} \tag{4}$$

replicated at least once. But the same must be true for the n last columns except that the signs in the last columns are reversed which ensures that the projection contains all the eight level combinations in a 2^3 design. If one column is picked from the k first columns and two from the k last the argumentation is similar.

In particular Theorem 3.4 shows that orthogonal $(40, 38, 3)$, $(56, 54, 3)$, $(88, 84, 3)$ and $(96, 94, 3)$ screens can be obtained just by removing one column from the original PB designs of projectivity $P = 2$ constructed by doubling. Similarly one may use the $(34, 66, 3)$, $(36, 70, 3)$, $(40, 78, 3)$ and $(42, 82, 3)$ non-orthogonal screens to construct the $(68, 132, 3)$, $(72, 140, 3)$, $(80, 156, 3)$ and $(84, 164, 3)$ screens respectively.

Designs obtained by doubling will have the specific property that a defining relation between four factors always exists. This follows since a possible projection on to four factors in the notation above is

$$\begin{bmatrix} a & b & a & b \\ a & b & -a & -b \end{bmatrix}$$

Therefore the fold-over constructed from these designs will be of projectivity $P = 3$ and not $P = 4$ as is usually obtained.

According to Theorem 3.4 an orthogonal $(32, 28, 3)$ screen can be constructed from the $(16, 14, 3)$ screen in Table 2. However, for $n = 32$ there exists another cyclic generated orthogonal two-level design that is different from the 32 run PB design. The generating row for this design is given by

$$++++-++-+++--+--+----+++-+-+---.$$

By Proposition 2 in Box and Tyssedal (1996) this design is of projectivity $P = 3$. A summary of orthogonal $P = 3$ and $P = 4$ designs and how to construct them is given in Table 3.

Table 3: List of orthogonal $P = 3$ and $P = 4$ screens.

Number of run $n =$	$P = 3$ screens	Way of construction	$P = 4$ screens	Way of construction
12	(12,11,3)	PB		
16	(16,14,3)	Hall	(16,5,4)	From PB 16
20	(20,19,3)	PB		
24	(24,23,3)	PB	(24,12,4)	F.o. PB 12
28	(28,27,3)	PB		
32	(32,31,3)	Not PB	(32,9,4)	From (16,14,3)
36	(36,35,3)	PB		
40	(40,38,3)	From PB 40	(40,20,4)	F.o. PB 20
44	(44,43,3)	PB		
48	(48,47,3)	PB	(48,24,4)	F.o. PB 24
52	(52,51,3)	PB		
56	(56,54,3)	From PB 56	(56,28,4)	F.o. PB 28
60	(60,59,3)	PB		
64	(64,62,3)	From (32,31,3)	(64,32,4)	F.o. (32,31,3)
68	(68,67,3)	PB	(68,67,4)	PB
72	(72,71,3)	PB	(72,71,4)	PB
76	(76,75,3)	PB		
80	(80,79,3)	PB	(80,79,4)	PB
84	(84,83,3)	PB	(84,83,4)	PB

4 Near Projectivity P Designs

Intuitively one will expect that designs for which the smallest number of unique runs in every projections on to P factors is close to 2^P, are not much inferior to a projectivity P design. Therefore a classification of designs according to their projectivity may be too crude. One way to get useful information of designs is by studying the alias matrix (Box and Wilson, 1951) which will inform about the estimation bias if certain effects can not be neglected. A similar, but less detailed information can be found from investigating the averaged dot products between main effect and interaction columns and between interaction columns. These dot products may be interpreted as a form for a 'correlation' between these columns and will hereafter be denoted as such. They are of importance when these designs are analysed. In particular they will affect a stepwise procedure even though a variable selection search carried out on subsets of variables as in Samset (1998) will reduce the problem. In Table 4 below we have given the two highest absolute values for the correlation between four categories of certain effect columns together with their percentage occurrences for all PB designs with $12 \leq n \leq 68$. We have also included the lowest number of unique runs that can be found in a projection on to four factors and the percentage occurrences of $P = 4$ projections. The correlation patterns for the projectivity $P = 4$ designs for $n = 72$, 80 and 84 are similar to the $n = 68$ case, see Samset and Tyssedal (1998).

The regular fractional factorials for $n = 16$, 32 and 64 have a relatively high percentage of $P = 4$ projections, and their percentage of non-correlated contrast columns is also high. The problem with these designs is that they are resolution $R = 3$ designs and thereby there will always be a certain amount of two-factor interaction columns that are completely confounded with main effect columns. In addition there is full confounding between some of the contrast

Table 4: Maximum correlation and projectivity results for PB designs.

Design	m and 2fi		2fi		m and 3fi		2fi and 3fi		Min unique	% $P = 4$
PB 12	0.33	(81.8)	0.33	(66.7)	0.33	(72.7)	0.67	(7.3)	11	
	0	(18.2)	0	(33.3)	0	(27.3)	0.33	(43.6)		
PB 16	1.00	(6.7)	1.00	(5.8)	1.00	(6.2)	1.00	(6.2)	8	(61.5)
	0	(93.3)	0	(94.2)	0	(93.8)	0	(93.8)		
PB 20	0.60	(5.3)	0.60	(4.7)	0.60	(5.0)	0.60	(1.7)	12	0
	0.20	(84.2)	0.20	(75.3)	0.20	(79.3)	0.40	(20.6)		
PB 24	0.33	(39.1)	0.33	(35.7)	0.33	(37.3)	0.33	(37.3)	14	42.7
	0	(61.9)	0	(64.3)	0	(62.3)	0	(62.3)		
PB 28	0.43	(11.1)	0.43	(10.3)	0.43	(10.7)	0.57	(0.7)	13	50.0
	0.14	(81.5)	0.14	(75.4)	0.14	(78.2)	0.43	(2.5)		
PB 32	1.00	(3.2)	1.00	(3.0)	1.00	(3.1)	1.00	(3.1)	8	82.8
	0	(96.8)	0	(97.0)	0	(96.9)	0	(96.9)		
PB 36	0.33	(17.1)	0.33	(16.2)	0.33	(16.6)	0.67	(0.02)	11	84.6
	0.11	(77.1)	0.11	(72.7)	0.11	(74.8)	0.44	(2.36)		
PB 40	1.00	(0.2)	1.00	(0.19)	1.00	(0.19)	1.00	(0.03)	8	85.2
	0.60	(2.4)	0.60	(2.3)	0.60	(2.3)	0.60	(1.2)		
PB 44	0.27	(23.3)	0.27	(22.2)	0.27	(22.7)	0.36	(5.0)	15	92.7
	0.09	(72.1)	0.09	(68.7)	0.09	(70.3)	0.27	(3.24)		
PB 48	0.33	(6.4)	0.33	(6.1)	0.33	(6.2)	0.50	(0.1)	15	93.9
	0.17	(51.1)	0.17	(48.9)	0.17	(49.9)	0.17	(46.9)		
PB 52	0.85	(0.1)	0.85	(0.1)	0.85	(0.1)	0.85	(0.11)	12	95
	0.38	(3.7)	0.38	(3.6)	0.38	(3.6)	0.77	(0.02)		
PB 56	1.00	(0.1)	1.00	(0.1)	1.00	(0.1)	1.00	(0.01)	8	95.4
	0.43	(5.2)	0.43	(5.0)	0.43	(5.1)	0.57	(0.3)		
PB 60	0.20	(35.6)	0.20	(34.4)	0.20	(35.0)	0.40	(0.55)	15	99.6
	0.07	(61.0)	0.07	(59.0)	0.07	(60.0)	0.27	(9.6)		
PB 64	1.00	(1.6)	1.00	(1.5)	1.00	(1.6)	1.00	(1.56)	8	93.2
	0	(98.4)	0	(98.5)	0	(98.4)	0	98.4)		
PB 68	0.29	(1.5)	0.29	(1.5)	0.29	(1.5)	0.35	(1.1)	16	100
	0.18	(37.3)	0.18	(36.2)	0.18	(36.7)	0.29	(0.1)		

columns in the three other examined categories. Also for the four non-geometric PB designs that are of projectivity $P = 2$ only, there exists a defining relation between certain three and four factors, and there are contrast columns in all four categories for which the correlations between them are one. Their one column less $(40, 38, 3)$, $(56, 54, 3)$, $(88, 86, 3)$ and $(96, 94, 3)$ screens are then natural alternatives, but also for these designs a defining relation exists between certain four factors. For the other PB designs there are no completely correlated contrast columns in the four categories. As a result there can be no defining relations between five and less than five factors in these designs. If X is one of these designs, it follows that no defining relation will exist between any five factors in a design constructed by doubling X. Hence for none of the non-geometric PB designs there will be a defining relation between any five factors. Details can be found in Samset and Tyssedal (1998). In order to simplify identification of active factors, it is desirable with a low value for the maximum correlation between contrast columns. Among $P = 3$ designs, in addition to the low run size PB 12 design, we notice that also the PB designs for $n = 24$, 44, 48 and 60 have a very favourable correlation structure. These designs are also the ones having the highest minimum number of unique runs in projections on to four factors. However, there is probably no severe reason to avoid the PB 20 and PB 28 designs as screening designs. Also the $(32, 31, 3)$ screen introduced in Section 3 has excellent properties. For this design all the projections on to four factors contain at least 15 of the 16 possible unique runs in a 2^4 design, and the maximum correlation between main effects and two-factor interaction columns and between two-factor interaction columns is ≤ 0.25.

5 Summary

We have used the repeat and mirror-image pattern to investigate projection properties of PB designs. In particular we want to emphasise the following results:

- There are four PB designs of projectivity $P = 4$. Their run sizes are 68, 72, 80 and 84.
- The half-fractions of these four projectivity $P = 4$ PB designs are super-saturated projectivity $P = 3$ designs.
- Any non-geometric PB design can be folded over to become a projectivity $P = 4$ design except for $n = 40$, 56, 88 and 96.
- The PB designs for $n = 24$, 44, 48 and 60 are close to being of projectivity $P = 4$. Together with the 12 run PB designs these designs have a very favourable correlation structure.
- There exists a cyclic generated orthogonal array for $n = 32$ of projectivity $P = 3$ which is close to being of projectivity $P = 4$.

References

Box, G. E. P. and Bisgaard, S. (1993). What Can You Find Out from 12 Experimental Runs, *Quality Engineering* **5**, 663–668.

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

Box, G. E. P. and Tyssedal, J. (1995). Projective Properties of The Sixteen Run Two-Level Orthogonal Arrays, *Technical Report 135*, Center for Quality and Productivity Improvement. University of Wisconsin-Madison.

Box, G. E. P. and Tyssedal, J. (1996). Projective Properties of Certain Orthogonal Arrays, *Biometrica* **83**, 950–955.

Box, G. E. P. and Wilson, K. B. (1951). On the Experimental Attainment of Optimum Conditions, *Journal of the Royal Statistical Society, Series B* **13**, 1–38.

Cheng, C. S. (1995). Some Projection Properties of Orthogonal Arrays, *The Annals of Statistics* **23**, 1223–1233.

Cheng, C. S. (1998). Some Hidden Projection Properties of Orthogonal Arrays with Strength Three, *Biometrika* **2**, 491–495.

Draper, N. R. and Lin, D. K. J. (1995). Characterizing Projected Designs: Repeat and Mirror-Image Runs, *Communication in Statistics - Theory and Methods* **24**, 775–795.

Hall, M. J. (1961). Hadamard Matrix of Order 16, *Jet Propulsion Laboratory, Research Summary 1* pp. 21–36.

Lin, D. K. J. (1993). A New Class of Supersaturated Design, *Technometrics* **35**, 28–31.

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

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

Rao, C. R. (1947). Factorial Experiments Derivable from Combinatorial Arragements of Arrays, *Journal of the Royal Statistical Society Supplement* **9**, 128–139.

Samset, O. (1998). Stepwise Regression under Factor Sparsity using a new Stopping Criterion, *Technical Report 12*, Department of Mathematical Sciences. The Norwegian University of Science and Technology, Norway.

Samset, O. and Tyssedal, J. (1998). Repeat and Mirror-Image Patterns and Correlation Structures of Plackett-Burman Designs, their Foldovers and Half-Fractions, *Technical Report 13*, Department of Mathematical Sciences. The Norwegian University of Science and Technology, Norway.

Seiden, E. and Zemach, R. (1966). On Orthogonal Arrays, *Annals of Mathematical Statistics* **37**, 1355–70.