NORGES TEKNISK-NATURVITENSKAPELIGE UNIVERSITET

Graphical Aids for the Analysis of Two-level Non-regular Designs

by

John Tyssedal and Ranveig Niemi

PREPRINT
STATISTICS NO. 3/2011

NORWEGIAN UNIVERSITY OF SCIENCE AND TECHNOLOGY
TRONDHEIM, NORWAY

This preprint has URL
http://www.math.ntnu.no/preprint/statistics/2011/S3-2011.pdf
John Tyssedal has homepage: http://www.math.ntnu.no/~tyssedal/
E-mail: tyssedal@math.ntnu.no
Address: Department of Mathematical Sciences,
Norwegian University of Science and Technology,
NO-7491 Trondheim, Norway.

Graphical Aids for the Analysis of Two-level Non-regular Designs.

John Tyssedal

Department of Mathematical Sciences
The Norwegian University of Science
and Technology, 7491 Trondheim, Norway
(*tyssedal@stat.ntnu.no*)

Ranveig Niemi

Safetec Nordic AS
Sluppenveien 12
7037 Trondheim, Norway
(ranveig.niemi@safetec.no)

The complex alias pattern between main-effects and two-factor interactions for two-level non-regular designs has been considered a problem for analysing these designs. If only a few two-factor interactions are active, however, the pattern induced into contrasts from active interactions may be very structured. This is in particular the case for the 12 run and the 20 run PB designs, probably the two most important ones for physical experimentation. This paper presents a graphical method for the analysis of non regular two-level orthogonal arrays. The method consists of two steps. The first step is called contrast plot interpretation and is directed towards revealing the cause for the pattern observed in the contrast plots. The second step is called alias reduction and aims at simplifying the interpretation of the plots by reducing the aliasing caused by effects that with a high degree of certainty may be considered active. The method is tested out on the 12 run and the 20 run PB designs with good results even for cases were the heredity principle does not hold.

KEY-WORDS: Contrast plot, Non-regular designs, PB designs, Screening

1. INTRODUCTION

Two-level non-regular designs include all orthogonal two-level designs that don't belong to the 2^{k-p} family, i.e. regular $1/2^p$, $p = 0,1,\ldots k-1$, fractions of 2^k factorials also called regular designs or sometimes just two-level fractional factorial designs.

The importance of non-regular designs is mainly due to two reasons. One is their projective properties, i.e. their design properties when restricted to a subset of experimental factors, which clearly outperform the ones for the regular designs. (Lin and Draper 1993, Box and Tyssedal 1996, Cheng 1995). The other is that they exist for all n that fulfils $n \equiv 0 \mod 4$ and thereby provide us with a lot more design alternatives than can be obtained from the 2^{k-p} family.

Their drawback is that they have a more complicated alias relationship. For instance in the 12 run Plackett-Burman (PB) design every main effect may be partially aliased with 45 two-factor interactions and a single two-factor interaction will appear in the alias pattern of all main effects not involved with this two-factor interaction. As a result methods such as normal and half-normal plots or more quantitative methods such as Lenth's method often used for analysing non-replicated two-level designs may be of little value since these methods are based on being able to separate active contrasts from contrasts estimating only noise.

Still normal plots and Lenth's method may be the only ones available in computer packages for analysing unreplicated non-regular designs, though several methods have been proposed in the literature. These can mainly be classified as factor based or effect based search procedures. In a factor based search the goal is to identify a few active factors among many that are being considered and the performance of such a procedure depends heavily on the projective properties of the design used. Examples of such methods can be found in Box and Meyer (1993), Tyssedal and Samset (1997) and Box and Kulahci (2003). Examples of effect based procedures can be found in Hamada and Wu (1993) and Chipman et al (1996). These procedures assume effect sparsity and the heredity principle is a strong guidance in the search for active effects. The heredity principle requires excluding an interaction to be in a model unless at least one (weak heredity) or both (strong heredity) of the parent main effects also are included in the model. In addition in order for the Hamada and Wu (1993) procedure to work well there should be few two-factor interactions and their sizes should be small. It is the purpose of this paper to point out that graphical aids may be of great value also in analysing non-regular designs. The proposed graphical procedure will in many cases point to the most likely active factors or effects. In others it may point to the complexity of the problem and thereby guide the analyst how to proceed with more quantitative methods. For instance it should be of particular value for the regression based Hamada and Wu procedure (1993), since it will likely give a much better start to the identification of active effects than using normal plots. But it should also be beneficial for a factor based approach. These procedures typically measure the ability of all subsets of factors up to a given size, often determined by the projectivity of the design (see Section 2), to explain the variation in the data. The graphical procedure may point out how many main effects and two-factor interactions that likely are active or that certain effects should be included in every model under investigation. The search for the active factor space could then be changed to an investigation of all models of a given type, for instance all models with m_1 main effects and m_2 two-factor interactions and with some main effects or two-factor interactions included in every model. Thereby factor spaces of higher dimensions than the projectivity of the design could be identified. Probably it will always be wise to use more than one method for analysing non-regular designs.

We start this paper by giving a short review of how to obtain non-regular designs and also point out some properties. Section 3 is devoted to the effect of aliasing in two-level non-regular designs with particular emphasize on the 12 run PB design. The graphical procedure is explained in Section 4. In Section 5 we provide some examples and in Section 6 we give some orthogonal arrays for which the proposed procedure will be particular useful. Concluding remarks are given in Section 7.

2. TWO-LEVEL NON-REGULAR DESIGNS AND SOME POPERTIES

Two-level non-regular designs are special cases of orthogonal arrays. They can be constructed from Hadamard matrices. A Hadamard matrix is a $n \times n$ matrix with entries ± 1 where rows and columns are pairwise orthogonal. Given a Hadamard matrix, H, an equivalent matrix with all elements in the first column equal to +1 can be obtained by multiplying by -1 each element in every row of H whose first element is -1. The remaining $n-1$ columns must have half 1's and half -1's and constitute an orthogonal two-level design that can accommodate $n-1$ factors in n runs. Orthogonal designs with $k = n-1$ is called saturated. Except when $n \le 12$, there are in general several design alternatives for each n. A list of Hadamard matrices is available on Neil Sloane's web site: www.research.att.com/~njas/index.html#TABLES. This list includes both regular and non-regular designs.

Compared to regular designs two-level non-regular designs have very favourable properties when projected onto a subset of factor columns. Projectivity of two-level designs was introduced by Box and Tyssedal (1996) and defined as follows. A $n \times k$ design with n runs and k factors each at two levels is said to be of projectivity P if every subset of P factor columns out of possible k contains a complete 2^P factorial design, possibly with some points replicated. It is convenient to describe such designs as (n,k,P) screens. A thorough discussion of the importance of projective properties is beyond the scope of this paper. Here we just point out that using columns from regular designs can only provide us with $(n, n/2, 3)$ screens while most non-regular designs are $(n, n-1, 3)$ screens. This makes these designs attractive for screening purposes.

The most well known two-level non-regular designs are the PB designs (Plackett and Burman 1946) with $n \ne 2^k$. PB designs where $n = 2^k$ coincide with the regular designs. One particular design is especially worth mentioning. The 12 run PB design is a $P = 3$ design in 11 factors. It has also the property that all the $\binom{11}{3}$ projections onto three factors are of just one type. The same applies to all the $\binom{11}{4}$ projections onto four factors. This very fair treatment of any set of k factors, $k \leq 4,$ is a property that no other PB design has. The projections onto any three or four factors are also very favourable. For any three factors we get a full 2^3 design plus the very best half fraction of a 2^3 design. The projections onto any four factors allow all main effects and two-factor interactions to be estimated. Hence the 12 run PB has very attractive screening properties and deserves to be frequently used for such purposes. A 12 run PB design is given in Figure 1.

Run	A	B	C	D	E	F	G	H	J	K	L
1	+	-	+	-	-	-	+	+	+	-	+
2	+	+	-	+	-	-	-	+	+	+	-
3	-	+	+	-	+	-	-	-	+	+	+
4	+	-	+	+	-	+	-	-	-	+	+
5	+	+	-	+	+	-	+	-	-	-	+
6	+	+	+	-	+	+	-	+	-	-	-
7	-	+	+	+	-	+	+	-	+	-	-
8	-	-	+	+	+	-	+	+	-	+	-
9	-	-	-	+	+	+	-	+	+	-	+
10	+	-	-	-	+	+	+	-	+	+	-
11	-	+	-	-	-	+	+	+	-	+	+
12	-	-	-	-	-	-	-	-	-	-	-

Figure 1. The 12 run PB design.

For more on projective properties of non-regular designs we refer to Box and Tyssedal (1996), Lin and Draper (1993), Cheng (1995, 1998) , Bulugotlu and Cheng (2003) and Tyssedal (2007).

3. THE ALIAS MATRIX of NON-REGULAR TWO-LEVEL DESIGN

For the use of non-regular designs it is important to be aware of their alias relationship. The alias matrix was introduced by Box and Wilson (1951). Suppose we fit the model

$$E(\boldsymbol{y}) = \boldsymbol{X\beta}$$

when the true expectation is given by $E(\boldsymbol{y}) = \boldsymbol{X\beta} + \boldsymbol{X}_1\boldsymbol{\beta}_1$. Then the expected value of the estimator for the regression coefficient vector, $\hat{\boldsymbol{\beta}} = \left(\boldsymbol{X'X}\right)^{-1}\boldsymbol{X'y}$, is given by

$$\boldsymbol{E}\left(\hat{\boldsymbol{\beta}}\right) = \boldsymbol{\beta} + \boldsymbol{A\beta}_1$$

where $\boldsymbol{A} = \left(\boldsymbol{X'X}\right)^{-1}\boldsymbol{X'X}_1$ is called the alias matrix. The alias relationship will vary from design to design. For screening purposes the aliasing between main effects and two-factor interactions has received the greatest attention. How to obtain this aliasing in an efficient way for a general two-level PB design is given in Lin and Draper (1993).

Every non-regular design for which the number of run $n = 4m$, m odd, shares the property that each main effect contrast may be partially aliased with all two-factor interactions for which it is not involved.

The 12 run PB design has a special position among the non-regular PB designs in the sense that any two-factor interaction is partially aliased with every other main effect for which it is not involved with the same amount in absolute value, namely $1/3$. Let the factors in a 12 run PB design be labelled $A,B,C,\ldots,K$ and suppose there is only one active two-factor interaction, say *AB*. Assume further there is no noise. Every two-factor interaction column is orthogonal to the main effects columns involved with the corresponding interaction. Thus, with the factor columns contained in the matrix $\boldsymbol{X}$, the 11×1 matrix, $\boldsymbol{A}$, has zeroes in the positions corresponding to factor *A* and factor *B*, and hence the estimated main effects of these factors are not biased. The rest of the entries in $\boldsymbol{A}$ are equal to $\pm1/3$ (Lin and Draper 1993). The main effect contrasts of $C,D,\ldots,K$ are then biased with the same amount in absolute value. Thereby the absolute values of the estimated contrasts of the inert factors are identical.

If two two-factor interactions are active, $\boldsymbol{A}$ is a 11×2 matrix with corresponding entries for the inert factors equal to $\pm 1/3$. Thereby the absolute values of the estimated contrasts for the inert factors will fall in two groups. More than two two-factor interactions will cause a more confuse pattern. For three two-factor interactions the absolute values of the bias for the inert factors will for the 12 run PB design fall in four groups.

Now let μ be the expected value of contrast with corresponding estimator $\hat{\mu}$. Whenever noise is present there will be a difference between $\left|E\left(\hat{\mu}\right)\right|$ and $E\left(\left|\hat{\mu}\right|\right)$. Assuming normally distributed data it can be shown that $E\left(\left|\hat{\mu}\right|\right)/\left|E\left(\hat{\mu}\right)\right| \approx 1.17$ when $\mu/\sigma_{\hat{\mu}} = 1$ and that $E\left(\left|\hat{\mu}\right|\right)$ decreases towards $\left|E\left(\hat{\mu}\right)\right|$ when $\mu/\sigma_{\hat{\mu}}$ increases. Hence provided the standard deviation of the contrasts is small compared to the size of the contrasts it is to be expected that the absolute values of the estimated contrasts of the inert factors behave quite similar to the non noise case.

We now have the basis to introduce our graphical procedure.

4. A GRAPHICAL PROCEDURE BASED on CONTRAST PLOTS and ALIAS REDUCTION

There are two basic elements in our graphical procedure. These are contrast plots interpretation and alias reduction. By a contrast plot we will mean a graphical representation of the absolute value of estimated contrasts. The contrasts for estimating the main effects are given by: $2\hat{\boldsymbol{\beta}} = 2\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{y} = \frac{1}{n/2}\boldsymbol{X}'\boldsymbol{y}$. A plot of the absolute values of these estimated contrasts will be called a *main effect contrast plot*. The bias in these contrasts are given by

$2A\boldsymbol{\beta}_1 = 2\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{X}_1\boldsymbol{\beta}_1 = \frac{1}{n/2}\boldsymbol{X}'\boldsymbol{X}_1\boldsymbol{\beta}_1$ where the columns for the active interactions are contained in $\boldsymbol{X}_1$. When only a few of the higher order effects are active this bias may create groups of contrasts that are approximately equal in absolute value as pointed out in Section 3. This can be used to anticipate the number of active two-factor interactions.

Similarly we can make contrast plots of absolute values of two-factor interactions involved with each factor, *two-factor interaction contrast plots*. The $n-2$ two-factor interaction columns associated with a factor, say *A*, are all mutually orthogonal and orthogonal to the main effect effect column for factor A, and will together with this column make a new essentially equivalent PB design. For an orthogonal saturated design it is possible to construct n-1 such plots.

Two-factor interaction contrast plots can be generated with very little additional effort. Let $\boldsymbol{aX}$ be the design matrix obtained by multiplying entrywise each column in $\boldsymbol{X}$ with the column for factor *A*. Then from $\frac{1}{n/2}\left(\boldsymbol{aX}\right)'\boldsymbol{y}$ we get an estimate of the $n-2$ two-factor interactions involved with factor *A* and $2\,\overline{y}$ since the first column in $\boldsymbol{aX}$ now will contain only pluses. Exactly the same result, however, may be obtained by multiplying $\boldsymbol{y}$ entrywise by $\boldsymbol{a}$ to obtain $\boldsymbol{ay}$ and then compute $\frac{1}{n/2}\boldsymbol{X}'\left(\boldsymbol{ay}\right)$.

For PB designs the aliasing between two-factor interactions with no common factors follows the same basic pattern as the aliasing between main effects and two-factor interactions not involved with these main effects and the possible values for the aliasing are identical in both cases (Samset and Tyssedal 1999). In particular for the 12 run and 20 run PB designs this means that the same pattern induced into inert main effects by active two-factor interactions will also be induced into inert two-factor interactions by active main effects and two-factor interactions. The simple way to interpret contrast plots for main effects and two-factor interactions is to consider contrasts that cluster in groups likely to be due to aliasing and those who don't cluster in groups to represent the active factors.

With many estimated contrasts, cluster analysis may be a valuable tool to identify the clusters. For two estimated pairs of contrasts $(x_i, y_j), i = 1,2,\ldots,k,\quad j = 1,2,\ldots,k$ their Euclidean distance is defined as

$$d_{ij} = \left| x_i - y_j \right|$$

Let the distance matrix be a $k \times k$ matrix with the (i, j) element equal to d_{ij}. Single linkage (Johnson and Wichern, 2007) is a hierarchical clustering method that starts with merging the two closest items. These form a cluster that is treated as one object. Thereby the number of objects is reduced by one. Distances between the other objects and this new object are updated as the smallest distance to an object in the cluster. The procedure continues until all objects are clustered and the obtained result is pictured in a dendrogram.

Contrast plots and two-factor interaction plots supported by dendrograms will in many cases provide significant aid in identifying active factors and effects. However, several situations may arise that causes ambiguity in the interpretation. Large noise may cause inert effects with the same bias to appear unequal. The amount of aliasing can be approximately equal to an active effect. More than two active interactions may cause many groups. This complicates the interpretation of these plots. The orthogonalization procedure given below is a way to reduce the aliasing and thereby simplify the interpretation. We shall call this procedure for *alias reduction.*

Consider again a model where $E(\boldsymbol{y}) = \boldsymbol{X\beta} + \boldsymbol{X}_1\boldsymbol{\beta}_1$. This model can be rewritten as follows:

$$\boldsymbol{E}(\boldsymbol{y}) = \boldsymbol{X}(\boldsymbol{\beta} + \boldsymbol{A\beta}_1) + (\boldsymbol{X}_1 - \boldsymbol{XA})\boldsymbol{\beta}_1$$

where $\boldsymbol{A}$ is the alias matrix. The columns in $\boldsymbol{X}_1 - \boldsymbol{XA}$ are orthogonal to the columns in $\boldsymbol{X}$. Thereby an unbiased estimate of $\boldsymbol{\beta}_1$ may be obtained by regression of y onto $\boldsymbol{X}_1 - \boldsymbol{XA}$. Denoting $\hat{\boldsymbol{\beta}}_1$ the corresponding estimator, an unbiased estimate of $\boldsymbol{\beta}$ may be obtained from $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\theta}} - \mathbf{A}\hat{\boldsymbol{\beta}}_1$ where $\hat{\boldsymbol{\theta}} = \left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{y}$, the least square estimator in the model $E\left(\boldsymbol{y}\right) = \boldsymbol{X\theta}$, or by interchanging the role of $\boldsymbol{X}$ and $\boldsymbol{X}_1$.

If the columns for the active main effects and interactions are contained in $\boldsymbol{X}$ and $\boldsymbol{X}_1$, the procedure above could be used to obtain unbiased estimate for these effects. There is a lot of freedom in choosing $\boldsymbol{X}$ and $\boldsymbol{X}_1$. $\boldsymbol{X}$ could contain main effects columns and $\boldsymbol{X}_1$ interaction columns or the other way around. A mixture of main effect and interaction columns could be used in both $\boldsymbol{X}$ and $\boldsymbol{X}_1$. The important thing is that there are no linear dependencies between the columns in $\boldsymbol{X}$ and $\boldsymbol{X}_1$. The total number of columns allowed can never exceed n-1, but will in general depend on the alias pattern for the design. The obvious choice of columns to put in $\boldsymbol{X}$ (or $\boldsymbol{X}_1$) would be the columns for the main effects and two-factor interactions that stand out as clear candidates for being active from the contrast-plots and then iteratively increase the number of columns in $\boldsymbol{X}$ (or $\boldsymbol{X}_1$), as the reduced aliasing enables us to find more potentially active candidates. Before we show some examples we summarize the basic steps in our procedure.

1. Perform a main effect contrast plot. Consider contrasts that cluster in groups likely to represent inert factors and those who don't likely to represent the active ones.
2. Perform two-factor interaction plots and interpret them the same way as for the main effect contrast plot. Normally these plots are made for interactions involved with main effects that potentially has been judged active from step 1. For step 1 and 2, be aware of that if a contrast who don't cluster in groups came out with a

value close to zero a possible explanation may be that the contrast is orthogonal to some active effect. Hence small values of contrasts may point to potentially active effects.

3. Whenever interpretation of main effect and two-factor interaction contrast plots becomes difficult due to masking of effects, large noise or many active two-factor interactions, perform alias reduction possibly in several steps. This is a very powerful procedure that in a few operation may reveal the complexity of the problem at hand.

5. EXAMPLES

For the 12 run PB design the pattern induced into contrasts from active two-factor interactions or main effects is especially simple if just one or two two-factor interactions are active. One active two-factor interaction causes expected absolute value of inert factor contrasts to be approximately equal. If there are two active two-factor interactions the expected absolute value of inert factor contrasts will approximately fall in two groups.

Example1. For the 12 run PB design given in Figure 1, let the response be given by $Y = 2A + 2C + AB + BC + \varepsilon,\ \varepsilon \sim N(0,0.25)$. Main effect contrast plot and dendrogram are shown in Figure 2. Both the main effect contrast plot and the dendrogram pick out two large main effects, *A* and *C*. The rest of the contrasts cluster in two groups. This indicates that in addition to the main effects of *A* and *C*, there are two active two-factor interactions.

Since *A* has the greatest estimated contrast in absolute value, one way to proceed is to make a two-factor interaction contrast plot of the two-factor interactions involved with factor *A*. This is shown in Figure 3 a). Now eight of the ten two-factor interactions seem to cluster in two groups each of size 4, while the two-factor interactions AB and AC separate out. The two clusters can be explained by the main effect of C and one more two-factor interaction. The

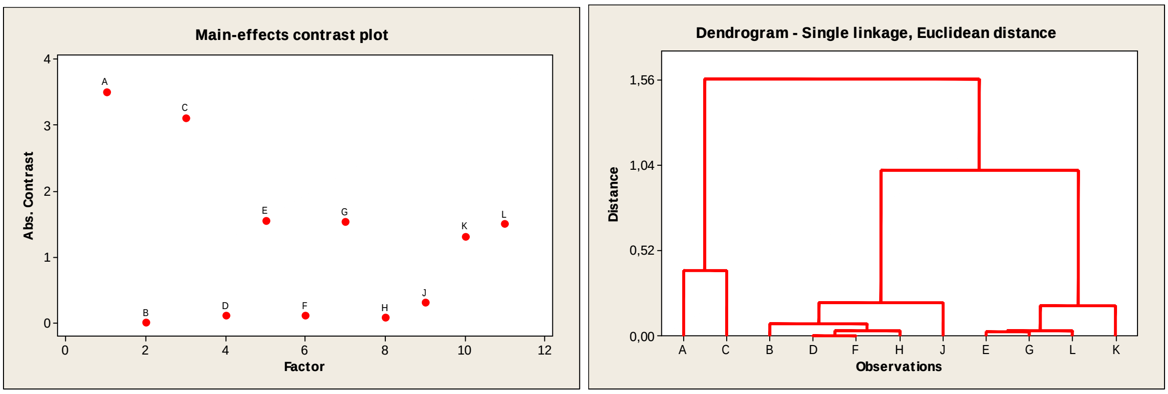


Figure 2. Main effect contrast plot and dendrogram for the model $Y = 2A + 2C + AB + BC + \varepsilon$, $\varepsilon \sim N(0, 0.25)$.

small value of AC indicates that this interaction is involved with factor C and that AC likely is inert. Performing alias reduction in two steps, first with $\boldsymbol{X} = [\mathbf{a,c}]$ and $\boldsymbol{X}_1 = [\mathbf{ab,ac,ad,ae,af}]$ and thereafter with $\boldsymbol{X} = [\mathbf{a,c,ab}]$ and $\boldsymbol{X}_1 = [\mathbf{ag,ah,aj,ak,al,}]$ gives us the results presented in Figure 3b). This plot supports that AB is active. We have then three likely active effects, A, C, and AB.

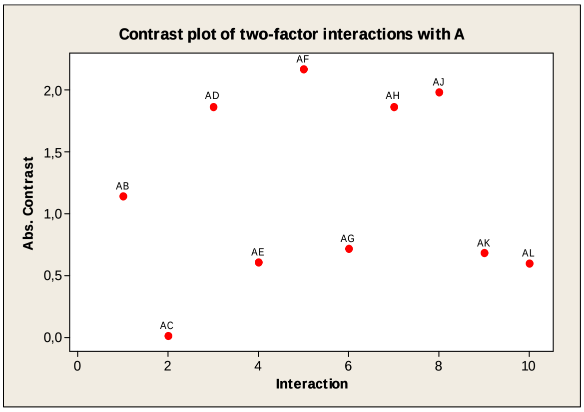


Figure 3a). Contrast plot of two-factor interactions involved with A.

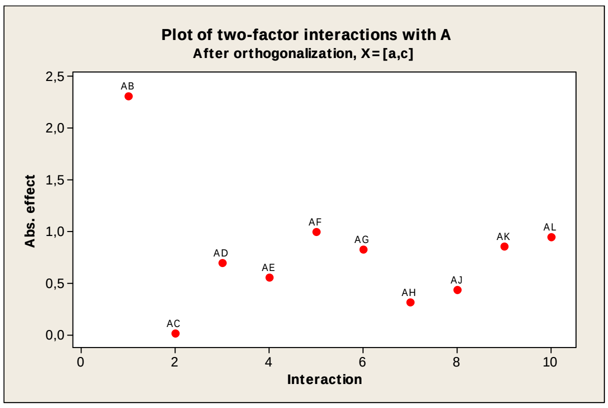


Figure 3b). Plot of two-factor interactions involved with factor A, after alias reduction starting with $\boldsymbol{X} = [\mathbf{a,c}]$.

When alias reduction is performed with $\boldsymbol{X} = [\mathbf{a,c,ab}]$ and $\boldsymbol{X}_1 = [\mathbf{ac,bc,cd,ce,cf,cg}]$ and thereafter with $\boldsymbol{X} = [\mathbf{a,c,ab,bc}]$ and $\boldsymbol{X}_1 = [\mathbf{ch,cj,ck,cl}]$ the results in Figure 4a is produced. Finally, with $\boldsymbol{X} = [\mathbf{ab,bc}]$, new estimates of main effects can be provided by performing alias reduction in two steps where **a** and **c** are included in $\boldsymbol{X}$ in step 2. These are shown in Figure 4b, and support that A, C, AB and AC are the active effects.

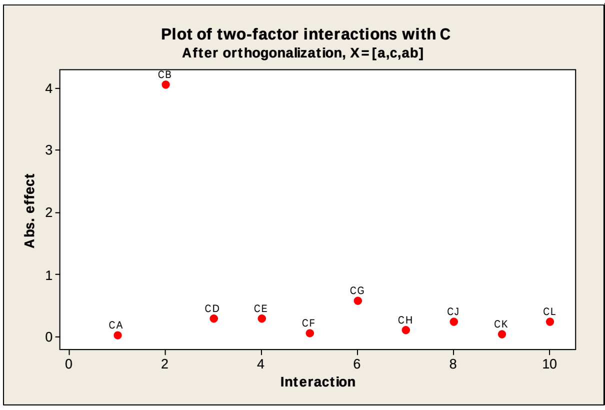


Figure 4a). Plot of two-factor interactions involved with factor C, after alias reduction starting with $\boldsymbol{X} = [\mathbf{a}, \mathbf{c}, \mathbf{ab}]$.

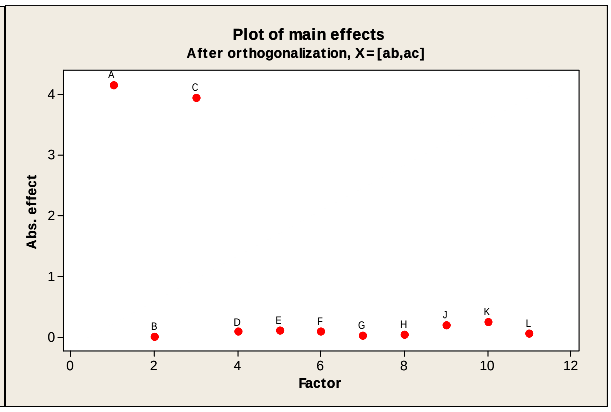


Figure 4b). Plot of main effects after alias reduction starting with $\boldsymbol{X} = [\mathbf{ab}, \mathbf{bc}]$.

Example 2: We will now try with a model that does not obey the heredity restrictions. Let $Y = 1.5A + AB + CD + \varepsilon$, $\varepsilon \sim N(0, 0.25)$. We notice that the model has four active factors. Main effect contrast plot and dendrogram are shown in Figure 5. The dendrogram roughly indicates three clusters. One with E, F, G, H and K, one with J and L and one with C and D. The contrast of A is larger than the others and also factor B seems to weakly separate out from the others. Several hypothesis about active factors are now possible.

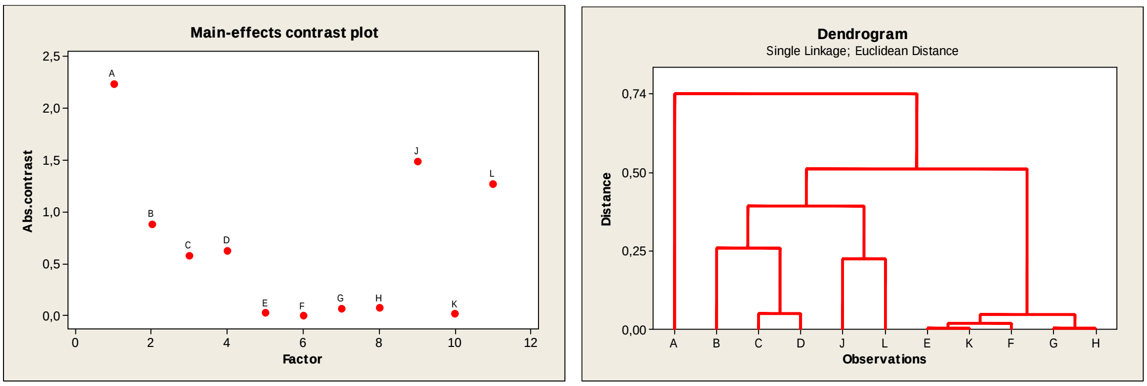


Figure 5. Main effects contrast plot and dendrogram for the model $Y = 1.5A + AB + CD + \varepsilon$, $\varepsilon \sim N(0, 0.25)$.

The safest way to proceed is to make a two-factor interaction contrast plot of the two-factor interactions involved with factor *A*. This is shown in Figure 6. Here AB, AC and AD separate out from the others which seemingly seem to be due to one other effect. Since AC and AD both are small, this effect is probably not aliased with these two-factor interactions.

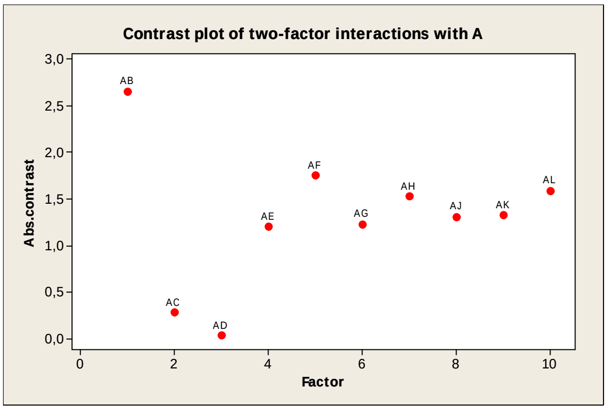
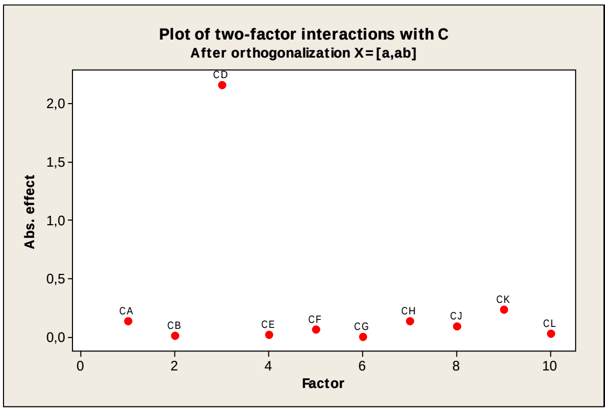


Figure 6. Contrast plot of two-factor interactions involved with factor A.

Now, assuming that A and AB are active, we may perform alias reduction with $\boldsymbol{X}=[\mathbf{a},\mathbf{ab}]$ and estimate two-factor interactions involved with factor C. As in the previous example the procedure had to done in two steps and **cd** was included in $\boldsymbol{X}$ after the first step. A plot of these is shown in Figure 7a). Then with $\boldsymbol{X}=[\mathbf{ab},\mathbf{cd}]$, new estimates of main effects can be provided, Figure 7b). These estimates were also calculated in two steps and **a** was



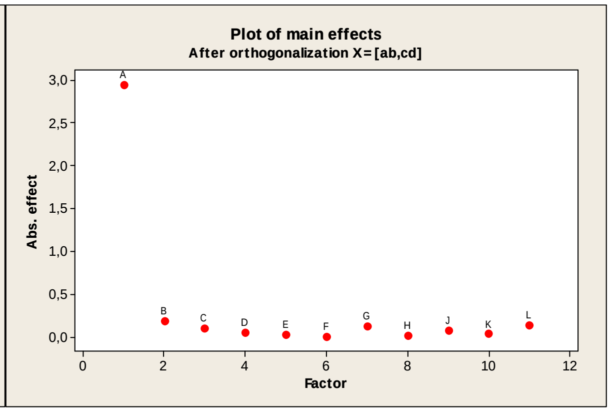


Figure 7a) Plot of estimated two-factor interactions involved with factor C after alias reduction starting with $\boldsymbol{X}=[\mathbf{a},\mathbf{ab}]$.

Figure 7b) Plot of estimated main effects after alias reduction starting with $\boldsymbol{X}=[\mathbf{ab},\mathbf{cd}]$.

included in $\boldsymbol{X}$ after the first step. Clearly a very plausible explanation of the variation in the data is that, A, AB and CD are active.

Figure 8 and 9 show four plots for the same model with $\varepsilon \sim N(0,0.5)$. Except for factor A, very little can be said about the other factors from the main effect contrast plot. Still we easily end up with a correct suggestion for the active effects after a plot of two-factor interactions with A and alias reduction is performed.

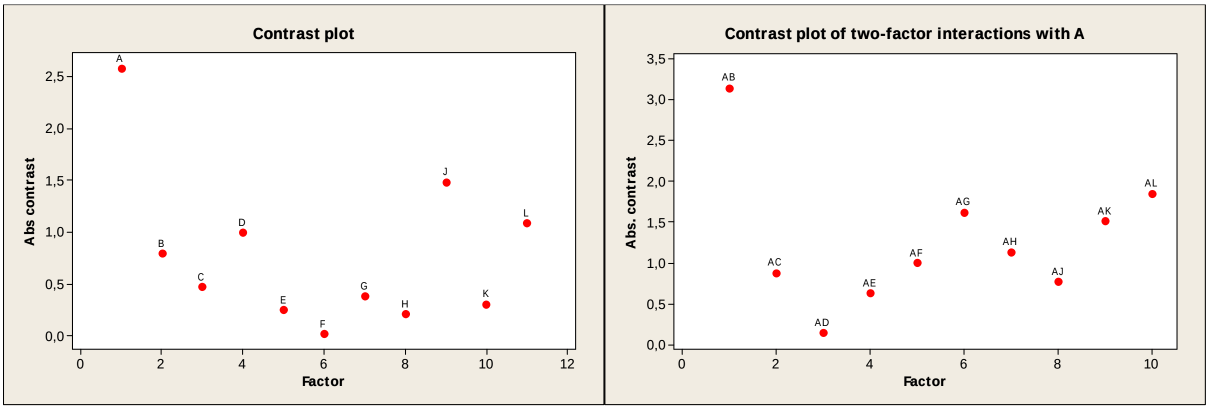


Figure 8a) Contrast plot of main effects and two-factor interactions involved with factor A for the model $Y = 1.5A + AB + CD + \varepsilon$, $\varepsilon \sim N(0,0.5)$.

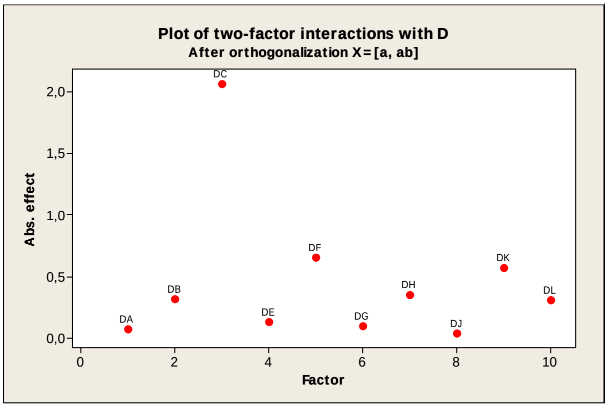


Figure 9a) Plot of estimated two-factor interactions involved with factor D after alias reduction starting with $\boldsymbol{X} = [\mathbf{a}, \mathbf{ab}]$.

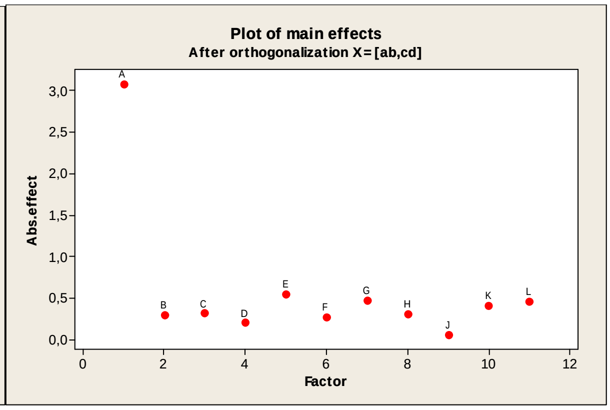


Figure 9b) Plot of estimated main effects after alias reduction starting with $\boldsymbol{X} = [\mathbf{ab}, \mathbf{cd}]$.

.

Example 3. A 20 run PB design is given in Figure 10. For this design any two-factor interaction is partially aliased with 17 main effects. For 16 of them it is aliased with the same amount, 0.2 in absolute value and for one the aliasing is 0.6. Hence the same arguing about effects clustering in groups can be applied to this design too. The only difference is that k two-factor interactions, $k = 1,2,\ldots$ creates k spurious main-effects. These may have to be considered as potentially active.

Run	A	B	C	D	E	F	G	H	J	K	L	M	N	O	P	Q	R	S	T
1	+	-	+	+	-	-	-	-	+	-	+	-	+	+	+	+	-	-	+
2	+	+	-	+	+	-	-	-	-	+	-	+	-	+	+	+	+	-	-
3	-	+	+	-	+	+	-	-	-	-	+	-	+	-	+	+	+	+	-
4	-	-	+	+	-	+	+	-	-	-	-	+	-	+	-	+	+	+	+
5	+	-	-	+	+	-	+	+	-	-	-	-	+	-	+	-	+	+	+
6	+	+	-	-	+	+	-	+	+	-	-	-	-	+	-	+	-	+	+
7	+	+	+	-	-	+	+	-	+	+	-	-	-	-	+	-	+	-	+
8	+	+	+	+	-	-	+	+	-	+	+	-	-	-	-	+	-	+	-
9	-	+	+	+	+	-	-	+	+	-	+	+	-	-	-	-	+	-	+
10	+	-	+	+	+	+	-	-	+	+	-	+	+	-	-	-	-	+	-
11	-	+	-	+	+	+	+	-	-	+	+	-	+	+	-	-	-	-	+
12	+	-	+	-	+	+	+	+	-	-	+	+	-	+	+	-	-	-	-
13	-	+	-	+	-	+	+	+	+	-	-	+	+	-	+	+	-	-	-
14	-	-	+	-	+	-	+	+	+	+	-	-	+	+	-	+	+	-	-
15	-	-	-	+	-	+	-	+	+	+	+	-	-	+	+	-	+	+	-
16	-	-	-	-	+	-	+	-	+	+	+	+	-	-	+	+	-	+	+
17	+	-	-	-	-	+	-	+	-	+	+	+	+	-	-	+	+	-	+
18	+	+	-	-	-	-	+	-	+	-	+	+	+	+	-	-	+	+	-
19	-	+	+	-	-	-	-	+	-	+	-	+	+	+	+	-	-	+	+
20	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Figure 10. The 20 run PB design.

Now suppose that in a 20 run PB design the model is given by

$$Y = 1.5A + 1.5B + 2C + 1.5AB + 2AC + \varepsilon,\ \varepsilon \sim N(0,0.5)$$

In the main-effect contrast plot in Figure 11 the factors *A*, *C* and *N* have the largest estimated contrasts in absolute value. The other contrasts approximately cluster in two groups. *N* is a spurious effect created by the two two-factor interactions. The other is *R*. Both *B* and *R* is difficult to distinguish from the rest of the factors. The contrast plot of two-factor interactions involved with factor *C* indicates that both *AC* and *CO* may be active. Figure 12a) shows estimated two-factor interactions with factor C obtained by alias reduction with $\boldsymbol{X} = [\mathbf{a,c,n}]$ and $\boldsymbol{X}_1 = [\mathbf{ac,bc,cd,ce,cf,cg,ch,cj,ck}]$ and thereafter with $\boldsymbol{X} = [\mathbf{a,c,n,ac}]$ and

$X_1 = [\mathbf{cl,cm,cn,co,cp,cq,cr,cs,ct}]$. The effect of *CO* is now considerably smaller and not clearly distinguishable from *CK* and *CM*.

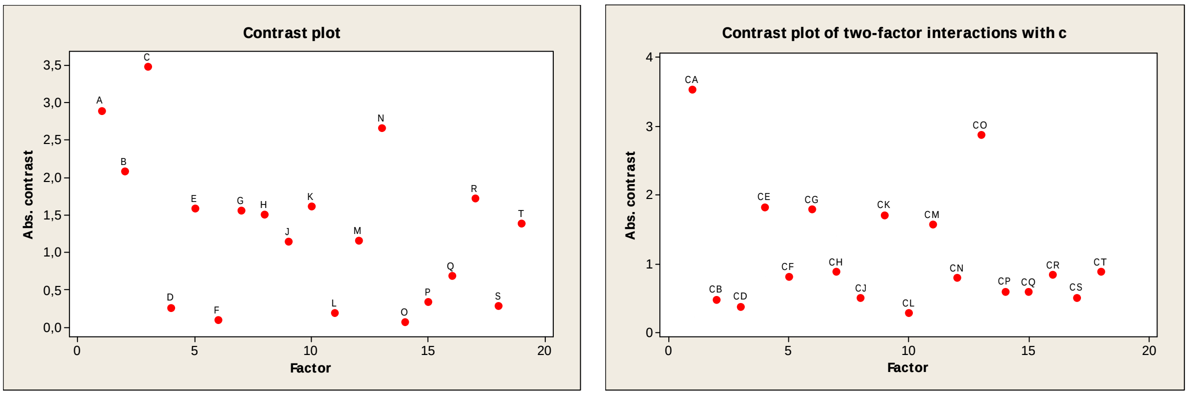


Figure 11. Contrast plot of main effects and two-factor interactions involved with factor C for the model $Y = 1.5A + 1.5B + 2C + 1.5AB + 2AC + \varepsilon,\ \varepsilon \sim N(0, 0.5)$.

A similar two-step procedure was performed for the two-factor interactions involved with factor A, Figure 12b). Again AC stands out as clearly active, but also AB seems to separate from the rest. Finally with $X = [\boldsymbol{ab}, \boldsymbol{ac}]$, new estimates for the main effects and

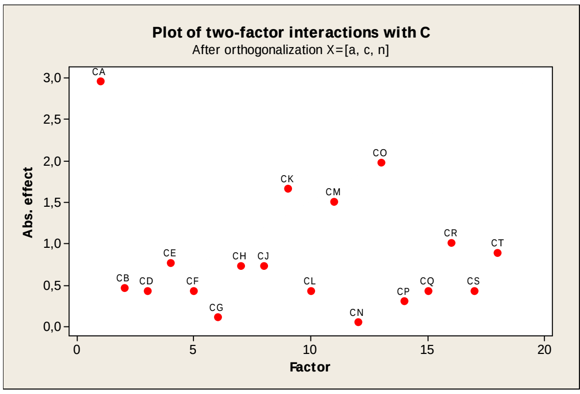


Figure 12a). Plot of estimated two-factor interactions involved with factor C after alias reduction starting with $X = [\mathbf{a,c,n}]$.

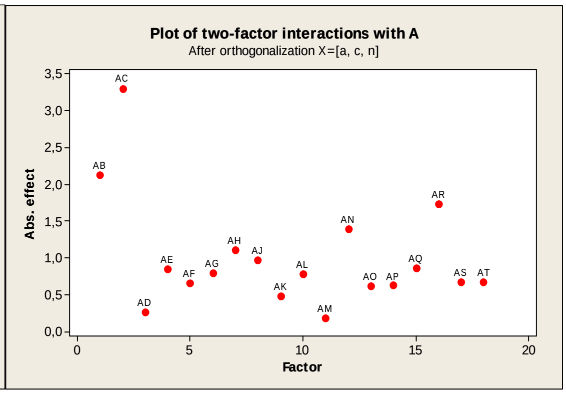


Figure 12b). Plot of estimated two-factor interactions involved with factor A after alias reduction starting with $X = [\mathbf{a,c,n}]$.

two-factor interactions with A can be computed in two steps. These are shown in Figure 13 a) and 13 b). The spurious effects are gone and the plots strongly support that the main effects of *A*, *B* and *C* and the two two-factor interactions *AB* and *AC* are active.

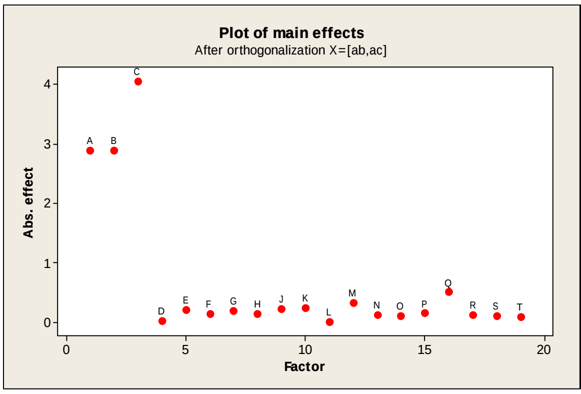


Figure13 a). Plot of estimated main effects after alias reduction starting with $\boldsymbol{X} = [\boldsymbol{ab}, \boldsymbol{ac}]$.

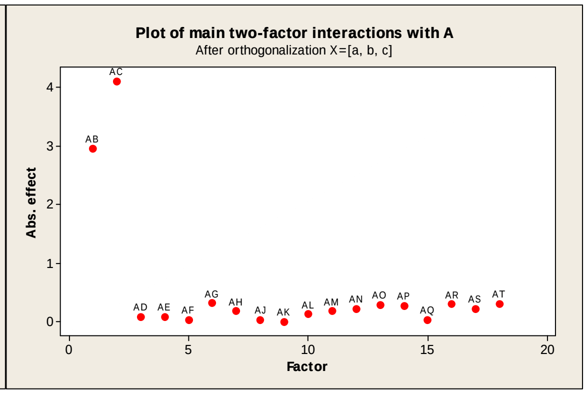


Figure 13b). Plot of estimated two-factor interactions involved with A after alias reduction starting with $\boldsymbol{X} = [\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}]$.

6. OTHER SUITABLE DESIGNS

From any $(n, n\text{-}1, 3)$ screen, D_n, a $(2n, 2(n-1), 3)$ screen can be constructed the following way $D_{2n} = \begin{bmatrix} D_n & D_n \\ D_n & -D_n \end{bmatrix}$. Such a design will have the same type of alias structure as the design D_n, part from some important differences, (Tyssedal and Kulahci 2005). Let $\boldsymbol{D}_1 = \begin{bmatrix} D_n \\ D_n \end{bmatrix}$ and $\boldsymbol{D}_2 = \begin{bmatrix} D_n \\ -D_n \end{bmatrix}$. Two-factor interactions between two factors in $\boldsymbol{D}_1$ or between two factors in $\boldsymbol{D}_2$ are only aliased with main effects in $\boldsymbol{D}_1$, and are orthogonal to interactions between one factor in $\boldsymbol{D}_1$ and one factor in $\boldsymbol{D}_2$. These two-factor interactions are only aliased with main effects in $\boldsymbol{D}_2$. Thereby the same rules for detecting potentially active effects that applies to D_n, can also be used for D_{2n}, treating $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ separately. For these designs, however, one has to be aware of that an interaction between two-factors in $\boldsymbol{D}_1$ is always fully confounded with a two-factor interaction in $\boldsymbol{D}_2$.

Non-regular PB designs with the number of runs equal to 24, 28, 36, 44, 60, all have two-factor interactions whose aliasing with main effects in absolute value can take two values. The 24 run PB design and the 32 run Paley design (1933) both obeys that one of these values are zero. If one two-factor interaction is present this means that the inert factors are divided into two groups. If more than one two-factor interaction is present, the pattern induced into inert contrasts may be difficult to interpret. However like in our Example 2 with $\sigma = 0.5$, experience has shown that the interpretation of this pattern, though useful, is seldom crucial for identifying the active factors.

7. CONCLUDING REMARKS

Non-regular two-level designs offer a lot more design alternatives than the regular two-level designs and have far better projective properties than these. Therefore they are very important screening designs that deserve to be used much more often than they are. An obstacle for the use of these designs has been the lack of methods offered by computer packages for their analysis. In this paper we have presented a graphical method based on contrast-plot interpretation and alias reduction for identifying the active effects in cases where it can be assumed that few two-factor interactions are active. The method is simple to implement and will in many cases be possible to use as a stand alone method. For more complex models it will provide valuable support for other more quantitative methods. The method has been tested out on the 12 run and the 20 run PB design with good results even for cases were the heredity principle does not hold. For these designs the pattern induced into contrasts from active interactions is very structured. But also for other types of two-level designs the suggested procedure should be beneficial since once an effect is found to likely be active its aliasing into other effects may be reduced.

REFERENCES

BOX, G. E. P., HUNTER, W. G. and HUNTER, J. S. (1978), *Statistics for Experimenters,* Wiley & Sons, New,York.

Box, G. E. P. and Meyer, R. D. (1993). Finding the Active Factors in Fractionated Screening Experiments, *Journal of Quality Technology* 25, 94-105.

Box, G. E. P. and Tyssedal, J. (1996). Projective properties of certain orthogonal arrays. *Biometrika, Vol* 83, No 4, 950-955.

Box, G. E. P. and Tyssedal, J. (2001). Sixteen run designs of high projectivity for factor screening. *Communication of Statistics - Simulations and computation,* 30(2): 217-228.

Bulutoglu, D. A. and Cheng, C. S. (2003). Hidden projection properties of some nonregular factorial designs and their applications. *The Annals of Statistics*, *Vol.* 31, *No*. 3, 1012-1026.

Cheng, C. S. (1995). Some projection properties of orthogonal arrays. *Ann. Statist.* 23. 1223-1233.

Cheng, C. S. Some hidden projection properties of orthogonal arrays with strength three *Biometrika* 1998; 85(2) 491-495.

Chipman, H., Hamada, H. and Wu, C. F. J. A Bayesian variable-selection approach for analyzing designed experiments. Technometrics 1997; 39: 372-381.

Daniels, C. Use of half-normal plots in interpreting factorial two-level experiments. *Technometrics* 1959; 1, 311-340.

Hamada, H. and Wu C. F. J. Analysis of designed experiments with complex aliasing. Journal of Quality Technology 1992; 24(3):130-137.

Kulahci, M. and Box, G. E. P. Catalysis of discovery and development in engineering and industry. Quality Engineering 2003; 15(3): 509-513.

Lin, D. K. J. and Draper, N. R. Projection properties of Plackett and Burman designs, Technometrics 1992; 34: 423-428.

Lin, D. K. J., Draper, N. R. Generating alias relationships for two-level Plackett and Burman designs. Computational Statistics & Data Analysis 1993; 15: 147-157.

Lenth, R.V. Quick and easy analysis of unreplicated factorials. Technometrics 1989; 31(4): 469-473.

Montgomery, D. C., Myers, R. H., Carter, W. H. and Vining, G. G. (2005): The Hierarchy Principle in Designed Industrial Experiments. *Quality and Reliability Engineering International*, Vol 21, 197-2001.

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

Paley, R. E. A. C. (1933). On orthogonal matrices. *Journal of mathematics and Physics,* 12, 311-320.

Rao, C. R. Factorial experiments derivable from combinatorial arrangements of arrays.

Journal of the Royal Statistical Society Supplement 9. 1947; 128-139.

Samset, O. and Tyssedal, J. (1999), ”Two-level designs with good projection properties,” *Technical Report 12,* Department of Mathematical Sciences, The Norwegian University of Science and Technology, Norway.

Tyssedal, J. (2008). Projectivity in Experimental Designs. *Encyclopedia of Statistics in Quality and Reliability.* John Wiley & Sons, 1516-1520.

Tyssedal, J. and Samset, O. (1997). Analysis of the 12 run Plackett and Burman design. *Technical Report 8*, Department of Mathematical Sciences. The Norwegian University of Science and Technology, Norway.