Journal of Modern Applied Statistical Methods
May, 2008, Vol. 7, No. 1, 94-100

1538 – 9472/08/$95.00

Selection of Non-Regular Fractional Factorial Designs When Some Two-Factor Interactions are Important

Weiming Ke
South Dakota State University

Rui Yao
University of Missouri-Columbia

A new method is proposed for selecting the optimal non-regular fractional factorial designs in the situation when some two-factor interactions are potentially important. Searching for the best designs according to this method is discussed and some results for the Plackett-Burman design of 12 runs are presented.

Key words: Alias matrix, fractional factorial design, non-regular design, partial confounding, Plackett-Burman design.

Introduction

Non-regular two-level fractional factorial designs, such as Plackett–Burman designs, are becoming popular choices in many areas of scientific investigation due to their run size economy and flexibility. The run size of non-regular two-level factorial designs is a multiple of 4. They fill the gaps left by the regular two-level fractional factorial designs whose run size is always a power of 2 (4, 8, 16, 32, …). In non-regular factorial designs each main effect is partially confounded with all the two-factor interactions not involving itself. Because of this complex aliasing structure, non-regular factorial designs had not received sufficient attention until recently. A number of authors studied the projection properties of two-level non-regular factorial designs. This includes Box and Bisgaard (1993), Lin and Draper (1993, 1995), Cheng (1995, 1998), Box and Tyssedal (1996), and Dean and Draper (1999). More recently, Deng and Tang (1999) proposed generalized resolution and minimum aberration criteria for ranking non-regular two-level factorial designs in a systematic way. Their criteria were further studied by Tang and Deng (1999), Tang (2001), Xu and Wu (2001), Ma and Fang (2001), and Butler (2003). Based on the generalized minimum aberration criteria, Deng, Li, and Tang (2000) and Deng and Tang (2002) provided tables of non-regular designs with favorable aberration properties for n ≤ 24 runs, Cheng, Li and Ye (2004) studied optimal blocking schemes for non-regular designs. Despite the above important contributions, a basic problem in this area still remains unsolved. The problem is how to assess, compare, and rank non-regular factorial designs when some two-factor interactions are potentially important.

In practical applications of non-regular designs, it is often in the case that some of the two-factor interactions are important and need to be estimated in addition to the main effects. In this article, we consider how to select non-regular two-level fractional factorial designs when some of the two-factor interactions are presumably important. We propose and study a method to select the optimal non-regular two-level fractional factorial designs in the situation that some of the two-factor interactions are

Weiming Ke is an Assistant Professor of Statistics in the Department of Mathematics and Statistics, South Dakota State University. Research interests: design of experiments, survival analysis, stochastic models, and biostatistics. Email: weiming.ke@sdstate.edu. Rui Yao is an Assistant Professor in the Department of Personal Financial Planning, University of Missouri-Columbia, and was formerly at South Dakota State University Research interests: applied statistics, individual financial risk tolerance, financial ratios, portfolio allocation, debt management, retirement preparation, and household consumption patterns. Email: yaor@missouri.edu.

potentially important. We then discuss how to search for the best designs according to this method and present some results for the Plackett-Burman design of 12 runs.

Non-regular fractional factorial designs

Non-regular fractional factorial designs are commonly obtained from Plackett-Burman designs or Hadamard matrices in general by selecting a subset of the columns. Plackett and Burman (1946) provided a series of two-level fractional factorial designs for examining ($n - 1$) factors in n run, where n is a multiple of 4. For example, a 12-run Plackett-Burman design can be constructed by shifting the row (+ + − + + + − − − + −) one place to the right 10 times and then adding a vector of − as the last row. See Table 1 for the 12-run Plackett-Burman design.

It is well known that if a Hadamard matrix exists, then its order n has to be a multiple of 4. A Hadamard matrix H of order n is an $n \times n$ orthogonal matrix with the elements ±1 whose columns (and rows) are orthogonal to each other. That means $H^TH = nE$ where E is the identity matrix. One can always normalize a Hadamard matrix by sign changes within complete rows so that its first column consists of all 1's. Removing the first column, one obtains a saturated two-level deaign with n runs and (n − 1) columns, which is a non-regular design and called a Hadamard design. Plackett-Burman designs are special cases of Hadamard designs. Non-regular designs are useful for factor screening and they fill the gaps between regular designs in terms of various run sizes. Unlike regular two-level fractional factorial designs in which any two effects are either orthogonal or fully aliased, non-regular designs exhibit some complex aliasing structure. In a non-regular design, there exist two effects that are partially aliased, meaning that they are neither orthogonal nor fully aliased.

For example, in a non-regular two-level factorial design, a main effect is partially confounded with all the two-factor interactions not involving itself. Because of this complex aliasing structure, non-regular factorial designs were traditionally not advocated when some interactions are potentially important. However Hamada and Wu (1992) showed that some interactions could be detected using non-regular factorial designs. Hence the arising question is how to select non-regular fractional factorial designs when some interactions are potentially important and need to be estimated. In this article, a new method was proposed and studied to solve this problem.

Method for selecting optimal non-regular factorial designs

Suppose the interest is in estimating all the m main effects and some important two-factor interactions by using a non-regular two-level fractional factorial design. Then the fitted model should include all the m main effects and important two-factor interactions. The fitted model is given by

$$Y = \beta_0 I + X_1\beta_1 + \varepsilon \qquad (1)$$

where Y denotes the vector of n observations, β_0 is the grand mean and I the all +1 column, β_1 is the vector of parameters containing all the main effects and important two-factor interactions, X_1 is the corresponding design matrix, and ε is the vector of uncorrelated random errors, assumed to have mean 0 and a constant variance. Because other interactions may not be negligible, the true model can be written as

$$Y = \beta_0 I + X_1\beta_1 + X_2\beta_2 + X_3\beta_3 + \cdots + X_m\beta_m + \varepsilon \qquad (2)$$

where β_2 is the vector of parameters containing the remaining two-factor interactions and X_2 is the corresponding design matrix, β_k is the vector of parameters containing k-factors interactions and X_k is the corresponding design matrix. The least square estimator $\hat{\beta}_1 = (X_1^TX_1)^{-1}X_1^TY$ from the fitted model in (1) has expectation (under the true model in (2)),

$$E(\hat{\beta}_1) = \beta_1 + (X_1^TX_1)^{-1}X_1^TX_2\beta_2 + (X_1^TX_1)^{-1}X_1^TX_3\beta_3 + \ldots + (X_1^TX_1)^{-1}X_1^TX_m\beta_m \qquad (3)$$

So the bias of $\hat{\beta}_1$ for estimating β_1 is given by

Table 1. The 12-run Plackett-Burman design

Run	1	2	3	4	5	6	7	8	9	10	11	Response
1	+1	+1	−1	+1	+1	+1	−1	−1	−1	+1	−1	y_1
2	−1	+1	+1	−1	+1	+1	+1	−1	−1	−1	+1	y_2
3	+1	−1	+1	+1	−1	+1	+1	+1	−1	−1	−1	y_3
4	−1	+1	−1	+1	+1	−1	+1	+1	+1	−1	−1	y_4
5	−1	−1	+1	−1	+1	+1	−1	+1	+1	+1	−1	y_5
6	−1	−1	−1	+1	−1	+1	+1	−1	+1	+1	+1	y_6
7	+1	−1	−1	−1	+1	−1	+1	+1	−1	+1	+1	y_7
8	+1	+1	−1	−1	−1	+1	−1	+1	+1	−1	+1	y_8
9	+1	+1	+1	−1	−1	−1	+1	−1	+1	+1	−1	y_9
10	−1	+1	+1	+1	−1	−1	−1	+1	−1	+1	+1	y_{10}
11	+1	−1	+1	+1	+1	−1	−1	−1	+1	−1	+1	y_{11}
12	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	y_{12}

$$\text{Bias}(\hat{\beta}_1, \beta_1) = (X_1^T X_1)^{-1} X_1^T X_2 \beta_2 + (X_1^T X_1)^{-1} X_1^T X_3 \beta_3 + \ldots + (X_1^T X_1)^{-1} X_1^T X_m \beta_m \quad (4)$$

The matrix $A_2 = (X_1^T X_1)^{-1} X_1^T X_2$ is called the second alias matrix, $A_k = (X_1^T X_1)^{-1} X_1^T X_k$ the kth alias matrix. The idea of alias matrix was originally introduced by Box and Wilson (1951). In other words, A_2 shows to which extent the estimates of the main effects and the important two-factor interactions in the model will be biased by the remaining two-factor interactions, A_k shows to which extent the estimates of the effects in the model will be biased by the k-factor interactions.

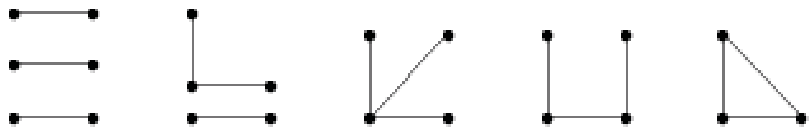
Note that $A_2\beta_2$ is the contribution of β_2 to the bias and $A_k\beta_k$ is the contribution of β_k to the bias. Because β_k is unknown, we have to work with A_k and minimize $A_k\beta_k$ through minimizing A_k. One size measure for a matrix $A = (a_{ij})$ is given by $\|A\|^2 \overset{def}{=} \text{trace}(A^T A) = \sum_{i,j} a_{ij}^2$. Under the hierarchical assumption that lower order effects are more important than higher order effects, to minimize the bias of $\hat{\beta}_1$ we should sequentially minimize $\|A_2\|^2, \ldots, \|A_m\|^2$. Here $\|A_k\|^2$ can be viewed as a confounding index which is a measure of the partial confounding between j-factor interactions not in the model and the effects in the model. For regular two-level fractional factorial designs, the entries of A_k are 0 or 1, and thus $\|A_k\|^2$ is simply the number of k-factor interactions not in the model confounded with the effects in the postulated model (Ke & Tang, 2003). For non-regular two-level fractional factorial designs, the entries of A_k are usually not integers because of the partial confounding structure. Now let $N_k = \|A_k\|^2$. Based on the above results, we can select optimal non-regular two-level fractional factorial designs by sequentially minimizing $N_2, \ldots, N_m$ where N_k is a measure of the bias contributed by the k-factor interactions. The design selection criterion is given below.

Optimal design selection criterion: Suppose the interest is in estimating all the m main effects and some important two-factor interactions by using a non-regular two-level fractional factorial design. Let A_k, $k = 2, 3, \ldots, m$ be the kth alias matrix of the model and let $N_k = \text{trace}(A_k^T A_k)$ which is a measure of A_k. The optimal design is selected by sequentially minimizing $N_2, \ldots, N_m$.

To gain further insight into the criterion, examine the criterion in detail. The postulated model consists of all the main effects and important two-factor interactions. If the effects not in the postulated model cannot be

Figure 1. Graph for model with one 2-factor interaction.

Figure 2. Graphs for models with two 2-factor interactions.



(a) (b) (c) (d) (e)

Figure 3. Graphs for models with three 2-factor interactions.

completely ignored, they will bias the estimates of the effects in the model. To solve this problem, the key issues are to permit estimation of the main effects and important two-factor interactions in the postulated model and to minimize the bias caused by the effects not in the model. Those two-factor interactions not in the model and higher-order interactions generally cause a bias on the estimation of the effects in the model. The measure of this bias, as given by N_k, is a measure of the bias caused by the k-factor interactions. Under the hierarchical principle that lower-order effects are more important than higher-order effects (Wu & Hamada, 2000), to minimize the bias, we should sequentially minimizing N_2, N_3, …, N_m. The vector (N_2, N_3, …, N_m) is called the confounding index pattern of a design. The optimal design should be selected such that to sequentially minimize the bias caused by those non-negligible interactions. Therefore this criterion selects the optimal non-regular design that has minimum N_2. If several designs have the same number of N_2, it selects optimal design that has minimum N_3 among the designs that have minimum N_2, and so on.

Results

Searching method

Consider 12-run Plackett-Burman design as an example. Let k be the number of important two-factor interactions. For $k = 1$, there is only one non-isomorphic model, as represented by Figure 1. For $k = 2$ and 3, the number of non-isomorphic models is 2 and 5 and the graphs for these non-isomorphic models are given in Figure 2 and 3 respectively.

Because there are many choices for the assignment of the important two-factor interactions, the optimal Plackett-Burman design of 12 runs is not easy to select according to this criterion. A computer program is used to calculate the confounding index pattern for each choice of the designs for each model for the given number of main effects and important two-factor interactions. Then select the best one that has minimum N_2. If several designs have same N_2, we select the best one that has the minimum N_3, and so on.

For 12-run Plackett-Burman designs, Draper (1985) and Wang (1989) showed that except for $m = 5$ and 6, any $12 \times m$ designs are equivalent. Lin and Draper (1992) and Wang and Wu (1995) showed that the two non-

Table 2. Optimal 12-run Plackett-Burman designs for the model containing one two-factor interaction

m	design columns	2-f interaction	(N_2, N_3, N_4)
4	1 2 3 4	(1, 2)	(1.56, 1.01, 0.67)
5	1 2 3 4 5	(1, 2)	(4.48, 4.30, 3.26)
6	1 2 3 4 5 6	(1, 2)	(9.06, 13.16, 11.91)
7	1 2 3 4 5 6 7	(1, 6)	(18.22, 29.56, 30.44)
8	1 2 3 4 5 6 7 8	(1, 7)	(32.78, 62.67, 72)
9	1 2 3 4 5 6 7 8 9	(1, 2)	(59, 130.67, 160)
10	1 2 3 4 5 6 7 8 9 10	(1, 2)	(124, 320, 400)

Table 3. Optimal 12-run Plackett-Burman designs for the models containing two two-factor interactions

m	model	design columns	2-f interactions	(N_2, N_3, N_4)
4	2(a)	1 2 3 4	(1, 2)(3, 4)	(1.16, 3.34, 0.72)
	2(b)	1 2 3 4	(1, 2)(1, 3)	(1.65, 1.72, 1.04)
5	2(a)	1 2 3 4 5	(1, 3)(2, 4)	(5.46, 7.38, 4.54)
	2(b)	1 2 3 4 5	(1, 2)(2, 3)	(5.25, 8.08, 5.08)
6	2(a)	1 2 3 4 5 6	(1, 4)(2, 3)	(12.56, 18.78, 14.44)
	2(b)	1 2 3 4 5 6	(1, 2)(2, 4)	(12.22, 20.44, 15.33)

Table 4. Optimal 12-run Plackett-Burman designs for the models containing three two-factor interactions

m	model	design columns	2-f interactions	(N_2, N_3, N_4)
4	3(c)	1 2 3 4	(1, 2)(1, 3)(1, 4)	(1.41, 2.75, 1.67)
	3(d)	1 2 3 4	(1, 4)(4, 3)(3, 2)	(1.24, 4.05, 1.13)
	3(e)	1 2 3 4	(1, 2)(1, 3)(2, 3)	(1.41, 2.75, 1.67)
5	3(b)	1 2 3 4 5	(1, 2)(3, 5)(4, 5)	(6.33, 11.29, 6.33)
	3(c)	1 2 3 4 5	(2, 3)(2, 4)(2, 5)	(6.40, 10.16, 6.12)
	3(d)	1 2 3 4 5	(3, 2)(2, 1)(1, 4)	(6.20, 11.82, 6.76)
	3(e)	1 2 3 4 5	(1, 2)(1, 3)(2, 3)	(6.40, 10.16, 6.12)
6	3(a)	1 2 3 4 5 6	(1, 2)(3, 5)(4, 6)	(19.78, 33.33, 22.22)
	3(b)	1 2 3 4 5 6	(3, 6)(1, 4)(2, 4)	(16.39, 25.05, 17.17)
	3(c)	1 2 3 4 5 6	(1, 2)(2, 3)(2, 4)	(15.76, 28.19, 18.98)
	3(d)	1 2 3 4 5 6	(1, 4)(4, 2)(2, 3)	(15.82, 26.48, 18.13)
	3(e)	1 2 3 4 5 6	(1, 2)(1, 3)(2, 3)	(15.95, 27.05, 18.41)

isomorphic 12 × 5 designs are the sub-matrix of the saturated design in Table 1 consisting of columns 1, 2, 3, 4, and 5 and the one consisting of columns 1, 2, 3, 4, and 10, the two non-isomorphic 12 × 6 designs are the sub-matrix consisting of columns 1, 2, 3, 4, 5 and 6 and the one consisting of columns 1, 2, 3, 4, 5 and 7. The estimation capacity of the 12-run Plackett-Burman designs were studied by Li and Wang (2004). They proved that the design with $m = 4$ can estimate all two-factor interactions, the designs with $7 \leq m \leq 10$ can only estimate any one two-factor interactions, and the two designs of 12 × 5 and one design of 12 × 6 (with columns 1, 2, 3, 4, 5 and 6) can estimate any models with up to three two-factor interactions. The above information can help us to save time and effort for searching the optimal 12-run Plackett-Burman designs. We have found the optimal 12-run Plackett-Burman designs for the models containing up to three two-factor interactions using (N_2, N_3, N_4) instead of the entire vector (N_2, …, N_m) to further reduce the computing burden. Actually five-factor and higher order interactions are very small and usually negligible in practice.

Optimal 12-run Plackett-Burman designs

Tables 2, 3, and 4 present optimal 12-run Plackett-Burman designs for the models with one, two, and three important two-factor interactions respectively. In these tables, the entries under "*m*" give the number of factors, the entries under "model" indicate which model is under consideration, and for example, an entry of 2(a) denotes the model represented by Figure 2(a). The entries under "design columns" give the design columns of for the factors in the fitted model. Column *j* in these tables denotes the *j*-th column in the saturated 12-run Plackett-Burman design in Table 1. The entries under "2-f interaction" show how to assign the factors involved in the important two-factor interactions. The last column in these tables gives (N_2, N_3, N_4).

The optimal 12-run Plackett-Burman designs are listed in Tables 2−4. When planning to study several factors and some important two-factor interactions by using a 12-run Plackett-Burman design, choose an optimal design directly from these tables to satisfy the current needs. Now an example is employed to illustrate how to use these optimal design tables.

Suppose that in an experiment, the experimenter want to study six factors, temperature, moisture, light, nitrogen, phosphorus, and potassium. Supose a 12-run Plackett-Burman design is being considered. In addition to the main effects of these factors, suppose further there is the need to estimate the three two-factor interactions that are between temperature and nitrogen, between temperature and phosphorus, and between temperature and potassium. The graph for this model is 3(c) as in Figure 3. The optimal 12-run Plackett-Burman design for this model can be found at the row for $m = 6$ and model 3(c) in Table 4. From this row in Table 4, we see that the design columns are 1, 2, 3, 4, 5, and 6 in Table 1. To complete the specification of the optimal design, the six factors need to be appropriately assigned to the six columns. The 2-f interaction column in Table 4 says that we should assign temperature to column 2, and assign nitrogen, phosphorus, and potassium to column 1, 3, and 4 arbitrarily. Other two factors can be arbitrarily assigned to the remaining columns 5 and 6. This design has $N_2 = 15.76$ which is a measure of the bias caused by the two-factor interactions, meaning that this design is the best in the sense that no other designs have smaller N_2 than this one for the given model.

References

Box, G. E. P. & Bisgaard, S. (1993). George's column. What can you find out from twelve experimental runs? *Quality Engineering*, *5*, 663-668.

Box, G. E. P. & Tyssedal, J. (1996). Projective properties of certain orthogonal arrays. *Biometrika*, *83*, 950–955.

Box, G. E .P. & Wilson, K. B. (1951). On the experimental attainment of optimum conditions. *Journal of the royal statistical society*, *Series B*, *13*, 1-45.

Butler, N. A. (2003). Minimum aberration construction results for non-regular two-level fractional factorial designs. *Biometrika*, *90*, 4, pp. 891–898.

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*, *85*, 491–495

Cheng, S. W., Li, W., & Ye, K. Q. (2004). Blocked non-regular two-level factorial designs, *Technometrics*, *46*, 269–279.

Dean, A. M. & Draper, N. R. (1999). Saturated main-effect designs for factorial experiments. *Statistics and Computing*, *9*, 179–185.

Deng, L.-Y., Li, Y., & Tang, B. (2000). Catalogue of small run non-regular designs from Hadamard matrices with generalized minimum aberration. *Communications and Statistics*, *A*, *29*, 1379–95.

Deng, L.-Y. & Tang, B. (1999). Generalized resolution and minimum aberration criteria for Plackett-Burman and other nonregular factorial designs. *Statistica Sinica*, *9*, 1071–1082.

Deng, L.-Y. & Tang, B. (2002). Design selection and classification for Hadamard matrices using generalized minimum aberration criteria. *Technometrics*, *44*, 173–184.

Draper, N. R. (1985). Small composite designs. *Technometrics*, *27*, 173-180.

Hamada,M. & Wu, C. F. J. (1992). Analysis of designed experiments with complex aliasing. *Journal of Quality Technology, 24*, 130-137.

Ke, W. & Tang, B. (2003). Selecting 2^{m-p} designs using a minimum aberration criterion when some two-factor interactions are important. *Technometrics*, *Volume 45, Number 4*, *November 2003, pp. 352-360.*

Li, Y. & Wang, J. (2004). Non-regular designs from Hadamard matrices and their estimation capacity. *Metrika*, *60*, 295-303.

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

Lin, D. K. J. & Draper N. R. (1993). Generating alias relationships for two-level Plackett and Burman designs. *Computational Statistics and data analysis*, *15*, 147-157.

Lin, D. K. J. & Draper N. R. (1995). Screening properties of certain two-level designs. *Metrika*, *42*, 99–118

Ma, C.-X. & Fang, K.-T. (2001). A note on generalized aberration in factorial designs. *Metrika*, *53*, 85-93.

Plackett, R. L. & Burman, J. P. (1946). The Design of optimum multi-factorial experiments. *Biometrika*, *33*, 305–325.

Tang, B. (2001). Theory of J-characteristics for fractional factorial designs and projection justification of minimum G_2-aberration. *Biometrika*, *88*, 401–7.

Tang, B. & Deng, L.-Y. (1999). Minimum *G*2-aberration for non-regular fractional factorial designs. *Annals of Statistics, 27*, 1914-1926.

Wang, J. C. (1989). Orthogonal arrays and nearly orthogonal arrays with mixed levels: Construction and applications. *Unpublished Ph: D: thesis. Department of Statistics, University of Wisconsin-Madison*.

Wang, J. C. & Wu, C. F. J. (1995) A hidden projection property of Plackett-Burman and related designs. *Statistica Sinica*, *5*, 235–250.

Wu, C. F. J. & Hamada, M. (2000). *Experiments: Planning, Analysis, Parameter Design Optimization*. New York: Wiley.

Xu, H. & Wu, C. F. J. (2001). Generalized minimum aberration for asymmetric fractional factorial designs. *Annals of Statistics*, *29*, 1066–1077.