A Catalogue of Two-level and Three-level Fractional Factorial Designs with Small Runs

Jiahua Chen, **D.X. Sun** **and** **C.F.J. Wu**

Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

Summary

Fractional factorial (FF) designs with minimum aberration are often regarded as the best designs and are commonly used in practice. There are, however, situations in which other designs can meet practical needs better. A catalogue of designs would make it easy to search for "best" designs according to various criteria. By exploring the algebraic structure of the FF designs, we propose an algorithm for constructing complete sets of FF designs. A collection of FF designs with 16, 27, 32 and 64 runs is given.

Key Words: Defining contrast subgroup; Minimum aberration design; Resolution; Word length pattern; Letter pattern.

1 Introduction

An outstanding problem in experimental design theory is the choice of "good" two-level and three-level fractional factorial designs which are commonly used in practice. A key question is how to choose a fraction of the full factorial design for a given run size and number of factors. Box and Hunter(1961) first approach the problem by introducing the notion of resolution as a goodness criterion for designs. Since designs of the same resolution may not be equally good, Fries and Hunter(1980) suggest the minimum aberration criterion to further discriminate designs. The minimum aberration criterion was already used implicitly in the construction of designs in the classic work at the National Bureau of Standards(1957, 1959). As argued and demonstrated in Section 2, when there is no design with resolution V or higher, maximum resolution and minimum aberration do not always lead to best designs. Different situations call for use of different designs. Since we cannot anticipate all the goodness criteria for designs, it seems impractical to give optimal designs for each criterion. A more realistic approach, adopted in this paper, is to give a catalogue of designs which are judged to be good by the minimum aberration criterion. Our rationale is that useful designs are in most cases good according to the minimum aberration criterion. For designs with 16 and 27 runs, we give a complete catalogue. For 32 and 64 runs, the number of designs is too large to be all included. Only five to ten designs are given in most cases. An algorithm for enumerating designs is presented in Section 3. Some comments on the designs in the catalogue are given in Section 4.

2 Definitions and Motivations

A s^{n-k} fractional factorial design, which has n factors of s-levels and s^{n-k} runs, is uniquely determined by k independent **defining words**. A word consists of **letters** which are names of factors denoted by $1, 2, \ldots, n$ or $A, B, \ldots$. The number of letters in a word is called **word-length** and the group formed by the k defining words is the **defining contrast subgroup**. The vector

$$W = (A_1, \ldots, A_n) \tag{1}$$

is called the **word-length pattern**, where A_i denotes the number of words of length i in the defining contrast subgroup. The concept of **resolution**, proposed by Box and Hunter (1961), is defined as the smallest r such that $A_r \geq 1$. It is a useful and convenient criterion for selecting practical designs.

Goodness of a design, however, cannot be fully judged by its resolution. Consider, for example, the following two 2^{7-2} designs:

$$d_1 : I = 4567 = 12346 = 12357,$$

$$d_2 : I = 1236 = 1457 = 234567.$$

Both have resolution IV, but have different word-length patterns

$$W(d_1) = (0, 0, 0, 1, 2, 0, 0), \text{ and } W(d_2) = (0, 0, 0, 2, 0, 1, 0).$$

The design d_1 has three pairs of aliased two-factor interactions (2fi's), e.g., 45&67, 46&57, 47&56, while d_2 has six pairs. This is because d_1 has one 4-letter words while d_2 has two. To further characterize or discriminate fractional factorial designs, Fries and Hunter (1980) propose the following criterion. For two designs d_1 and d_2 with r being the smallest value such that $A_r(d_1) \neq A_r(d_2)$, we say that d_1 has less aberration than d_2 if $A_r(d_1) < A_r(d_2)$. If there is no design with less aberration than d_1, then d_1

has **minimum aberration**(MA). Obviously, for given n and k, a MA design always exists. However, we do not know whether it is unique in general. See Chen (1992).

For small number of factors (up to 11) and run size (up to 128), Box, Hunter and Hunter (1978, p.410) provides a useful catalogue of 2-level fractional factorial designs with minimum aberration. Franklin (1984) constructs more minimum aberration designs. Chen and Wu (1991) and Chen (1992) investigate some theoretical properties of MA designs and construct MA 2^{n-k} designs for $k \leq 5$ and any n.

Both definitions of resolution and minimum aberration are based on the hierarchical assumption:

(i) lower order effects are more important than higher order effects,

(2)

(ii) effects of the same order are equally important.

The minimum aberration criterion can rank-order almost any two designs. In general it is a good design measure unless these two conditions are grossly violated. However, in some practical situations described later, the hierarchical assumption does not hold and better designs can be found. The second but more subtle point concerns its reliance on the word-lengths of the defining contrasts. Although minimizing the numbers of short-length words usually leads to the estimability of more lower order effects or under less stringent assumptions, combinatorial complexity of the defining contrasts makes the relation between lengths and estimability less certain. This point is best illustrated by the following example (due to C.F.J. Wu).

Consider the minimum aberration 2^{9-4} design, which has the word-length pattern $(0,0,0,6,8,0,0,1,0)$ and the defining contrast subgroup

$\mathrm{I} = 1236 = 1347 = 1389 = 2467 = 2689 = 4789$

$$= 12458 = 12579 = 14569 = 15678 = 23459 = 23578 = 34568 = 35679$$
$$= 12346789 \ .$$

Under the relative weak assumption of negligible 3-factor and higher order interactions, all the main effects and the eight 2fi's $(15, 25, 35, 45, 56, 57, 58, 59)$ are estimable. (Note that 5 does not appear in any of the words of length four.) In Wu and Chen (1992), any 2fi that is not aliased with any main effect or other 2fi's is called **clear**. So this design has eight clear 2fi's. Consider then the second best design in terms of the aberration criterion, which has the word-length pattern $(0, 0, 0, 7, 7, 0, 0, 0, 1)$ and the defining contrasts

$$\begin{aligned} \mathrm{I} &= 1236 = 1278 = 1347 = 1468 = 2348 = 2467 = 3678 \\ &= 12459 = 13589 = 15679 = 23579 = 25689 = 34569 \\ &= 123456789 \ . \end{aligned}$$

Although it has seven words of length four, one more than the MA design, both 5 and 9 are missing in these seven words. Therefore it has 15 clear 2fi's,

$$(15, 25, 35, 45, 56, 57, 58, 59, 19, 29, 39, 49, 69, 79, 89).$$

From the estimation point of view, it is far superior to the minimum aberration design. This illustrate the need of finding designs other than MA designs.

In some experimental situations the assumption 2(ii) does not hold. As argued in Wu and Chen(1992), there are practical situations in which certain interactions can be *a priori* identified as being potentially important and should be estimated clear of each other. In order to accommodate a set of specified interactions, one may have to choose a design with worse aberration. For example, consider the choice of a 2^{6-2} design, in which the following interactions $(13, 14, 16, 23, 34, 35, 36, 45, 56)$ can be estimated clear of each other and of the main effects (assuming the other 2fi's are negligible). By using a graph representation Wu and Chen (1992) show that the resolution III design with I=125=2346 meets the requirements while the MA design with I=1235=2346 does not. Broading the choice of designs will make it possible to find flexible graphs otherwise nonexistent.

There is indeed a whole class of problems that do not satisfy the assumption 2(i) and 2(ii). In parameter designs(Taguchi, 1987), the factors are divided into two types: control factors and noise factors. Since the noise factors are not controllable except when special efforts are made, estimability of the noise main effects is usually less important than that of the control-by-noise interactions. This violates 2(i). Similarly estimability of the noise-by-noise interactions is less important than that of the control-by-noise interactions, which violates 2(ii). As a result, neither the resolution nor the aberration criterion can guarantee a good statistical design for this type of experiments. A simple example is used to illustrate the point. Consider the resolution III design d_1 given by I=ABCr=rst=ABCst, and the resolution IV design d_2 given by I=ABCr=BCst=Arst, where A, B, C are three control factors and r, s, t are three noise factors. Under the assumption that 3-factor and higher order interactions are negligible, A, B, C, As, Bs, Cs, At, Bt, Ct are estimable in d_1, whereas only the main effects A, B, C, r, s, t are estimable in d_2. Since it is much less important to be able to estimate the three noise main effects r, s, t in d_2 than to estimate the six control-by-noise interactions in d_1, design d_1 is preferred in spite of its lower resolution. Further discussion on planning techniques for parameter designs can be found in Shoemaker, Tsui and Wu(1991).

The overall conclusion is that, practical situations can be different from one to the other and they may sometimes contradictory. Using a single criteria such that the minimum aberration criterion for selecting designs exclusively cannot meet practical needs. It is hence desirable to collect good designs in a catalogue.

3 Construction Method

3.1 Basic Idea

If a design d_1 can be obtained from d_2 by relabelling the factor numbers in the defining contrast subgroup or by change of signs, we say d_1 is isomorphic to d_2. Since isomorphic designs are essentially the same, it is sufficient to include only one of them in any catalogue of designs. To catalogue all possible designs, a straightforward approach does not work. For example, in a $32(=2^5)$ run design with 15 two-level factors, there are 5 independent factors, and 10 additional factors can be defined in $\binom{31-5}{15-5} = 5,311,735$ ways. It is impractical to identify isomorphic designs among all $5,311,735$ designs because of the difficulties in discriminating between non-isomorphic designs. This number becomes much larger as the run size and number of factors increase. By applying some algebraic and combinatorial methods, we are able to reduce the computations significantly. The basic idea of the proposed sequential construction method is to break the huge amount of combinatorial computations into a sequence of much smaller computations. At each step, the total number of designs are significantly reduced by keeping only non-isomorphic designs.

The 2^{n-k} designs given in Section 2 can be viewed as submatrices of regular Hadamard matrices. A *regular* Hadamard matrix of order 2^q is a $2^q \times 2^q$ orthogonal matrix of ± 1 with the additional property that the entrywise product of any two columns is among the 2^q columns. By replacing -1 by 1 and 1 by 0 and using addition over GF(2), these 2^q columns form an elementary Abelian group over GF(2), where

GF(2) is the Galois field with two elements. Except for the column corresponding to the identity element in the group, we may write the remaining columns as

$$\mathbf{C} = \{C_1, \cdots, C_{2^q-1}\}. \tag{3}$$

Within $\mathbf{C}$, we can find q independent columns that generate all the columns in $\mathbf{C}$. A 2^{n-k} design can now be viewed as a subset of $\mathbf{C}$ with n columns. Out of the n columns, $n-k(=q)$ are independent columns and the remaining k columns can be generated from the $n-k$ columns through the defining relations in its defining contrast subgroup. A similar matrix representation for three-level designs can be defined. The only difference is that its columns are grouped into pairs. For each pair of columns, one is a multiple of the other modulus three. This simple representation for 2^{n-k} and 3^{n-k} designs will be employed in the tabulation of designs.

Let $D^R_{n,k}$ be the set of non-isomorphic s^{n-k} designs with resolution $\geq R$, and $D_{n,k} = D^{III}_{n,k}$ for convenience. For given R, k, and $D^R_{n,k}$ we construct $D^R_{n+1,k+1}$ by assigning the additional factor to one of the unused columns of each design in $D^R_{n,k}$. There are at most $(s^{n-k}-1)/(s-1)-n$ ways to assign this factor. Therefore, we obtain a class of designs, denoted by $\tilde{D}^R_{n+1,k+1}$ with cardinality

$$\{\ \#\text{ of designs in } D^R_{n,k}\} \times [(s^{n-k}-1)/(s-1)-n].$$

Clearly, $\tilde{D}^R_{n+1,k+1} \supset D^R_{n+1,k+1}$. However, some designs in $\tilde{D}^R_{n+1,k+1}$ are isomorphic and some may have resolutions less than R. To construct $D^R_{n+1,k+1}$, we need to eliminate these redundant designs. It is easy to eliminate designs with resolution smaller than R. To identify isomorphic designs, we divide all designs into different categories according to their word length patterns and letter patterns. The letter pattern counts the frequency of letters contained in the words of different lengths (Draper and Mitchell, 1970). Note that same letter pattern implies same wordlength pattern. Designs with different letter patterns are obviously non-isomorphic. Therefore we only need to examine the isomorphism of designs with the same letter pattern. This is done by using the following result in Chen (1992). Two designs of d_1 and of d_2 are isomorphic if there exists a one to one map M from the columns d_1 to the columns d_2 such that

$$M(C_{i1} + C_{i2} + \cdots + C_{il} \pmod 2) = M(C_{i1}) + M(C_{i2}) + \cdots + M(C_{il}) \pmod 2$$

for any l and $C_{i1}, C_{i2}, \cdots, C_{il} \in d_1$. After the elimination of isomorphic designs, we reduce $\tilde{D}^R_{n+1,k+1}$ to $D^R_{n+1,k+1}$.

Note that designs with the same letter pattern are not necessary isomorphic. See Chen and Lin (1991), which disproves a conjecture of Draper and Mitchell (1970).

This procedure will not only give us the complete set of $s^{(n+1)-(k+1)}$ designs, but also reduce the amount of computations for the subsequent step of constructing $s^{(n+2)-(k+2)}$ designs.

The rationale of this method is supported by the following facts.

FACT 1. *(Completeness)* $\tilde{D}_{n+1,k+1} \supset D_{n+1,k+1}$.

FACT 2. *(Monotonicity of resolution)* $\tilde{D}^R_{n+1,k+1} \supset D^R_{n+1,k+1}$.

The proofs are straightforward and omitted.

3.2 Implementation

Isomorphism Check:

Our approach to isomorphism check uses an idea which is illustrated by a simple example. To save space, the technical details are not given here.

Let us consider the 2^{7-3} designs, in which a, b, c, d denote four independent columns of the regular $2^4 \times 2^4$ Hadamard matrix. The set of columns C is then

$$\{a, b, ab, c, ac, bc, abc, d, ad, bd, abd, cd, acd, bcd, abcd\}.$$

To check isomorphism between the two 2^{7-3} designs:

$$\begin{aligned} d_1 &= \{a, b, c, d, ab, abd, bcd\}, \\ d_2 &= \{a, b, c, d, ac, acd, abcd\}, \end{aligned}$$

which have the same word-length pattern and letter pattern, we apply the following scheme:

1. Select four independent columns from d_2, say, $\{a, b, ac, acd\}$. There are $\binom{7}{4}$ choices.

2. Select a relabeling map from $\{a, b, ac, acd\}$ to $\{A, B, C, D\}$, i.e., $A = a, B = b, C = ac, D = acd$. There are $4!(= 24)$ choices.

3. Write the remaining columns $\{c, d, abcd\}$ in d_2 as interactions of $\{A, B, C, D\}$, i.e., $c = AC, d = CD, abcd = BD$. Then d_2 can be written as $\{A, B, C, D, AC, CD, BD\}$.

4. Compare the new representation of d_2 with that of d_1. If they match, d_1 and d_2 are isomorphic, and the process stops. Otherwise, return to step 2 and try another map of $\{A, B, C, D\}$. When all the relabeling maps are exhausted, return to step 1 and find another four columns.

If two designs are isomorphic, an isomorphic map will be found eventually. If two designs with the same letter pattern are nonisomorphic, it requires a complete search of relabeling maps. Fortunately, this happens rarely in our experience.

The isomorphism check for 3-level designs is similar but slightly more complicated. The details are omitted.

4 Tables of Designs

Using the method described in the last section, we obtain complete collections of designs with 16, 27, and 32 runs. We do not include 8- and 9-run designs because their number is small and can be found in standard texts. Since the total number of 64-run designs is too large, we only keep those with resolution IV or higher in the computer search. To save space, for 32 and 64 runs, we present only five to ten designs in most cases. The complete catalogue is available upon request. These designs are not chosen exclusively according to the minimum aberration criterion. Designs with worse aberration may be judged to be better by other properties, e.g. the number of clear 2fi's.

For each run size, we put the column set **C** (see (3)) in Yates order. The column numbers of the independent columns are indicated by bold face. A 2^{n-k} design is given by a subset of n columns of C, consisting of $n-k$ independent columns and k additional columns. Only the latter are specified in the tables. For clarity, we call it design n-k.i in the tables, where i denotes the i-th 2^{n-k} design in the catalogue. The word length pattern and the number of clear 2fi's are also provided. To save space, at most five non-zero components of the word length patterns are given. Also, we use the notation $19-22$ for columns 19 to 22. The three-level 27-run designs are given in the same vein. Note that in the corresponding design matrix, the three levels are denoted by 0, 1 and 2.

Usage of the tables is illustrated by the following example.

Example. 2^{6-2} *fractional factorial design*

The columns set C is presented in Table 1 with independent columns {1,2,4,8}. The first 2^{6-2} design in the table is {7, 11}, i.e., the design consists of columns {1,2,4,8,7,11}. To find the defining words, we name the corresponding factors A, B, C, D, E, F. Column 7 is the sum of columns 1, 2, and 4 (mod 2), i.e. the generator for factor E is $E = ABC$. Similarly, the generator for factor F is $F = ABD$.

Some comments on the tables:

1. If a design with resolution V or higher exists, we do not list any designs of resolution III or IV.

2. Among resolution IV designs, those with large numbers of clear 2fi's are not necessarily good according to the minimum aberration criterion. This phenomenon is especially pronounced for 64-run designs with n=14 to 17.

3. For the 32-run designs with n=10 to 16, none of the resolution IV designs has any clear 2fi's.

4. The numbering of designs is not strictly according to the minimum aberration criterion. Designs with worse aberration but with a much larger number of clear 2fi's may be placed ahead of others with less aberration. For example, designs 14-8.4 and 14-8.5 have worse aberration than designs 14-8.6 to 14-8.10.

Acknowledgements. The research was supported by the Natural Sciences and Engineering Research Council of Canada, General Motors of Canada Limited, and the Manufacturing Research Corporation of Ontario.

References

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

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

Chen, J. (1992). On minimum aberration 2^{n-k} fractional factorial designs. *Ann. Statist.* To appear.

Chen, J. and Wu, C.F.J. (1991). Some results on s^{n-k} fractional factorial designs with minimum aberration or optimal moments. *Ann. Statist.* **19**, 1028-1041.

Chen, J. and Lin, D.K.J. (1990). On the identity relationships of 2^{k-p} designs. *J. Statist. Plan. Infer.* **28**, 95-98.

Draper, N.R. and Mitchell, T.J. (1970). Construction of a set of 512-run designs of resolution ≥ 5 and the set of even 1024-run designs of resolution ≥ 6. *Ann. Math. Statist.* **41**, 876-887.

Franklin, M.F. (1984). Constructing tables of minimum aberration p^{n-m} designs. *Technometrics* **26**, 225-232.

Fries, A. and Hunter, W.G. (1980). Minimum aberration 2^{k-p} designs. *Technometrics* **22**, 601-608.

National Bureau of Standards (1957). *Fractional Factorial Experiment Designs for Factors at Two Levels, Applied Mathematics Series 48.* Washington, D.C.: U.S. Government Printing Office.

National Bureau of Standards (1959). *Fractional Factorial Experiment Designs for Factors at Three Levels, Applied Mathematics Series 54.* Washington, D.C.: U.S. Government Printing Office.

Shoemaker, A., Tsui, K.L. and Wu, C.F.J.(1992). Economical experimentation methods for robust design. *Technometrics* **33**, 415-427.

Taguchi, G.(1987). *Systems of Experimental Design.* New York: UNIPUB.

Wu, C.F.J. and Chen, Y.Y.(1992). A graph-aided method for planning two-level experiments when certain interactions are important. *Technometrics* **34**, 162-175.