國 立 高 雄 師 範 大 學
高雄師大學報 50, 2021. 6, 頁 1-10

Double Semi-Folding in the 2^{k-p} Designs

Pen-Hwang Liau*

Abstract

Two-level fractional factorial designs have been wildly used to investigate the effect of factors in many fields. For some designs, the main effects and the interactions are aliased in chain. The technique of foldover, in which requires the same size as an original experiment, is often used for conducting follow-up experiments to break the alias chain. In this article, for economic reasons, we give some examples to illustrate two folding techniques, semi-folding and double semi-folding, and compare the results based on the criterion of clear.

Keywords: Alias chain, Defining contrast subgroup, Minimum aberration, Semifolding

Submitted: 2020/11/11；Accepted: 2021/01/05

* Professor, Department of Mathematics, National Kaohsiung Normal University

No.50, June. 2021, pp. 1-10

2^{k-p}的雙半摺疊設計

廖本煌*

摘 要

2 層次的部分因子設計被廣泛應用於各個領域。對於一些設計其主效應與交互作用有別名鏈的關係。摺疊設計為用與原設計相同的實驗點，然後應用技巧來破壞設計中的別名鏈。此篇文章我們舉一些例子並用清晰的準則，說明並比較半摺疊設計與雙半摺疊設計。

關鍵詞：別名鏈、定義對比子群、最小偏離、半摺疊

投稿日期:2020/11/11;接受日期:2021/01/05

*國立高雄雄師範大學數學系教授

1. Introduction

Experiments play an important role for investigators in all fields studied either in discovering more about a particular process or in comparing the effects of several factors on some phenomena. Usually, experiments often involve several factors at the same time. A complete factorial experiment can solve problems that occur in the design of an experiment of which its goal is to determine suitable tolerances for the factors in a manufacturing process. As we know, the complete factorial experiments are quite resource consuming either in labor or in time if the number of runs required is large. In practice, what we need is methods that involved only the selection of a subset of the complete factorial design which the important effects are still estimated with the same accuracy. The orthogonal main effects plans permit the estimate of all main effects with minimum error in a factorial experiment without correlation given that, the interactions are negligible. These kinds of designs are considered by authors such as Plackett and Burman (1946), Bose and Bush (1952), Addelman (1961), Dey and Ramarkrishna (1977), Gupta, Nigam and Dey (1982), and Nigam and Gupta (1985).

Many experimenters believe it is better to introduce many factors, each with a high and low level, into the experiment rather than choosing arbitrarily a few factors and run many levels on each. In the 2^k factorial designs, as the number of factors increases, the number of runs required for a complete replicate of the design rapidly outgrows the resources of most experiments. For example, a complete 2^7design requires 128 runs. In this situation only 7 of the 127 degrees of freedom correspond to main effects, and only 21 degrees of freedom correspond to two-factor interactions. Finney (1945) introduced fractional factorial designs which call for few runs, that is, it is not necessary to run all possible combinations of factor levels. The fractional factorial designs have been widely used in industrial experiments in which several factors are involved, which in this case, the higher order factor interactions may be ignored. A regular 2^{k-p} fractional factorial design is determined by its defining contrast subgroup, which consists of 2^{p-1} defining words. These designs have p independent defining words, which a word consists of letters that are the names of the factors denoted by 1, 2, …, k. Let A_i denote the number of words of length i in its defining contrast subgroup. The vector $W = (A_1, \ldots, A_k)$ is called the worth-length pattern of the design. The resolution of a design is defined as the smallest r such that $A_r \geq 1$. In above example, if we concern the main effects only, then we may perform 2^{7-4} fractional factorial design with eight points. The generators for this design can be constructed by I = 124, I = 135, I = 236, and I = 1237, and thus the complete defining subgroup for this design is I = 124 = 135 = 2345 = 236 = 1346 = 1256 = 456 = 1237 = 347 = 257 = 1457 = 167 = 2467 = 3567 = 1234567 with the word-length pattern W = (0, 0, 7, 7, 0, 0, 1). This is a resolution *III* design because the smallest number of letters in any word of the defining contrast subgroup is three. Furthermore, if the experimenter cares about some two-factor interactions, then 2^{7-3} with resolution *IV* can be considered.

It is known that the regular 2^{k-p} fractional factorial design consists of $k-p$ basic factors and p independent generated factors. A reasonable criterion is to select the best generators such that the resulting design has the highest resolution. Although it is the criterion to choose a good design, sometime resolution alone is insufficient to distinguish among designs with the same resolution. Fries and Hunter (1980) proposed the criterion for the minimum aberration (MA) to discriminate designs with the same resolution. For any two designs D_1 and D_2 with r being the smallest value such that $A_r(D_1) \neq A_r(D_2)$, D_1 is said to have less aberration than D_2 if $A_r(D_1) < A_r(D_1)$, D_1 is referred ad the minimum aberration design when there is no design with less aberration than D_1 The minimum aberration criterion is commonly used for the selecting the optimal designs. However, other criterion, for example, *clear* can lead to better designs, especially in the irregular designs. We shall call a main effect or two-factor interaction "clear to estimate" if it can be estimated when other main effects or two-factor interactions appear in the same model. That is, a main effect or two-factor interaction clear if none of its aliases are main effects or two-factor interactions. The reader can refer to the paper, Wu and Chen (1992), to get more detail concept. Fold-over technique is the common application of the follow-up strategies for two-level fractional factorial experiments to run an additional experiment of the same size of the original experiment. Several authors are interested in this kind of research, such as Box and Wilson (1951), Box and Hunter (1961), Li and Mee (2002), Montgomery and Runger (1996), and Montgomery (2013). It poses an interesting question of this kind of design if the experimenter wants to run a second experiment when the analysis of the initial experiment did not reveal any specific type of significant effects. In this case, what action should she/he take? What kind of criterion should be applied to find the optimal follow-up design? We may refer this question to some papers that are described by Li and Lin (2003) and Chen and Cheng (2004). Using a half-run size of the original design as the follow-up plan was also considered by authors such as John (2000), Mee and Peralta (2000), and Liau and Huang (2007).

The organization of this article is as follows. In section 2, we present some construction methods of foldover design and semi-folding design to assist readers to understand the concept. Some examples that use the technique of double semi-folding will be mentioned in section 3, followed by our conclusions in section 4.

2. Foldover and semi-folding design

The regular 2^{k-p} fractional factorial is orthogonal, so the construction of the design and analysis of the data from the experiment are reasonably straightforward. The fraction is obtained from the original design by folding over, reversing the signs of factors – the original design – also known as the foldover design (Box and Hunter, 1961). If a foldover plan is only considered by folding the generated factors, then it is called a core foldover plan (Li and Lin, 2003). Base on the criterion of MA, they utilized computer for searching the optimal foldover plans for 16 and 32 runs and tabulated the results for practical use. As an example for interpreting the concept of core foldover plan, consider a MA2^{5-2} with resolution *III* design. We may select 124 and 135 as the independent defining words, hence the defining contrast subgroup of this plan is defined by I = 124 = 135 = 2345. We may choose the numbers 1, 2, and 3 as the basic factors and numbers 4 and 5 as the generated factors. That is, factor 4 can be constructed by 4 = 12 and factor 5 is constructed by 5 = 13. The four core foldover plans can be expressed by

(i) {0, 145, 24, 125, 35, 134, 2345, 123}, (ii) {4, 15, 2, 1245, 345, 13, 235, 1234},
(iii) {5, 14, 245, 12, 3, 1345, 234, 1235}, (iv) {45, 1, 25, 124, 34, 135, 23, 12345}.

In which the number 0 represents the identity element. Note that the first set (i) is a subgroup and the other three sets are the cosets of the first set. It is not hard to recognize that we should obtain the same experiment elements for folding factor 1 or factors 3 and 5. That is, if we perform the folding procedure on factor 1, we will obtain the eight experiment runs defined by I = -124 = -135 = 2345. Similarly, the same defining contrast subgroup will be obtained by folding factors 3 and 5.

In general, we describe a semi-folding plan the following way: foldover on ___; subset on ___ (Mee and Peralta, 2000). In fact, the combined design produced by the original 2^{k-p} design plus the 2^{k-p-1} design by semi-folding technique is an irregular design which three fractions are formed in this case. Hence the aberration criterion cannot be used for selecting the follow-up design in this situation. Alternatively, we may use other criterion, say clear, to pursue the better design. As a matter of convenience, we call the main effect or two-factor interaction "clear" if it can be estimated when other main effects or two-factor interactions appear in the same model. Suppose that two generators of a 2^{6-2} design are giving by 5 = 12 and 6 = 134, the complete defining contrast subgroup is I = 125 = 1346 = 23456. For example, suppose that "foldover on factors 5 and 6; subset on the high level of effect 13" is our consideration for the additional experiment. Abbreviating "foldover" with f_0 and "subset" with *ss*, we will denote this semi-folding design as "f_0 =56; ss=13^+". Consequently, we take the eight new points at the high level of the effect 13 defined by

$$I = 13 = -125 = -235 = -1346 = -46 = 23456 = 12456.$$

The original design can be partitioned into two parts based on the effect 13. The first fraction with eight runs at the high level of 13 is defined by

$$I = 13 = 125 = 235 = 1346 = 46 = 23456 = 12456.$$

The second fraction at the low level of the effect 13 is defined by

$$I = -13 = 125 = -235 = 1346 = -46 = 23456 = -12456.$$

The new experiment may be called the third fraction. The main effects 3, 4, 6 and two-factor interactions 23, 24, 26, 45, 35, 45, and 56 can be clear to estimate from the original plan. Combing the first and the third fractions, the fraction with sixteen runs is defined by

$$I = 13 = 23456 = 12456.$$

Consequently, the effects 2, 4, 5, 6, 24, 25, 26, 45, 46, and 56 can be clear to estimate from this fraction. Combing the second and the third fractions, we may obtain another fraction defined by

$$I = -235 = -46 = 23456.$$

Obviously, four effects, 1, 12, 13, and 15, are clear to estimate from this fraction. Hence, all six main effects plus eleven two-factor interactions are clear to estimate in the combing design with twenty-four runs. One thing worth mentioning is that we may consider using a MA 2^{6-2} defined I = 1235 = 1246 = 3456 as the initial design. The basic factors are 1, 2, 3, and 4 in this design. If the method of semi-folding, $f_0 = 5$; $ss = 1^+$, is applied, then all six main effects and nine two-factors, 13, 15, 23, 25, 34, 35, 36, 45, 56, are clear to estimate. In above design, if we perform folding on effects 5, 6, and 56 and the subset is taken from all the effects, then the optimal semi-folding plan is obtained with fifteen effects "clear" to estimate six main effects and nine two-factor interactions. In general, there are $(2^{p-1})X(2^{k-p}-1)$ possibilities in choosing a semi-folding design for a giving 2^{k-p} design. Here, we only consider taking the high level of the effect for the subset, because they have the same estimation properties for taking high or low level effect for the subset.

3. Double semi-folding design

Because of the run size economy, the 2^{k-p-2} replicate of the 2^{k-p} design is considered for the follow-up experiment. To perform the double semi-folding technique, we use the notation, $f_{0=X}{}^+$; ss=Y^+, to indicate the additional points. For illustration, we will give some examples to show the properties of double semi-folding.

Example 3.1. Suppose that the starting design is based on a 2^{7-2} design, two generators are giving by 6 =123 and 7 = 245, with the defining contrast subgroup I = 1236 = 2457 = 134567. The new fraction can be obtained by the way "f_0 =7^+ ss=1^+", it means that we take the additional eight points by the technique of double semi-folding. We may take the points at high level of the factors 1 and 7, then switch the sign of the high level of the factor 7. The original plan is divided by four sets of eight points each that have the same defining contrast subgroup, but with different signs based on the effects 1 and 7:

(i) I = 1 = 7 = 17 = 1236 = 236 = 12367 = 2367 = 2457 = 12457 = 245 = 1245 = 134567 = 34567 = 13456 = 3456;

(ii) I = -1 = -7 = 17 = 1236 = -236 = -12367 = 2367 = 2457 = -12457 = -245 = 1245 = 134567 = -34567 = -13456 = 3456;

(iii) I = -1 = 7 = -17 = 1236 = -236 = 12367 = -2367 = 2457 = -12457 = 245 = -1245 = 134567 = -34567 = 13456 = -3456;

(iv) I = 1 = -7 = -17 = 1236 = 236 = -12367 = -2367 = 2457 = 12457 = -245 = -1245 = 134567 = 34567 = -13456 = -3456.

The new eight points given by $f_0\ 7^+$; ss=1^+ is defined as

(v) I = 1 = -7 = -17 = 1236 = 236 = -12367 = 2367 = -2457 = -12457 = 245 = 1245 = -134567 = -34567 = 13456 = 3456.

Combing the first and the fifth designs, we obtain the new fraction defined by

$$I = 1 = 1236 = 236 = 245 = 1245 = 13456 = 3456.$$

Two-factor interactions 27, 37, 47, 57, and 67 can be clear to estimate from this combing plan. Adding set (iii) to the new points gives the fraction defined by

$$I = -17 = 1236 = -2367 = -12457 = 245 = -34567 = 13456$$

from which main effects 3, 6 and two-factor interactions 34, 35, and 56 are clear to estimate. Combing the set (iv) and (v), we obtain another new fraction defined by

$$I = 1 = -7 = -17 = 1236 = 236 = -12367 = -2367.$$

Obviously, two-factor interactions 24, 25, 34, 35, 45, 46, and 56 can be clear to estimate from this fraction. Consequently, all main effects and two-factor interactions 14, 15, 17, 24, 25, 27, 34, 35, 37, 45, 46, 47, 56, 57, and 67 are clear to estimate in this forty-run combined design. Note that the main effects, 1, 2, 3, 4, 5, 6, 7, and two-factor interactions, 14, 15, 17, 34, 35, 37, 46, 47, 67, are estimated from the original thirty-two runs by the method of the least squares; but the effects 24, 25, 27, 45, 56, 57 are not the least squares estimates. From another perspective, we add new eight points to break the alias chain and the combined design is clear to estimate additional six two-factor interactions.

Example 3.2. To illustrate the use of the technique of double semi-folding again, suppose we have nine factors and we are interested in estimating main effects and getting some information of two-factor interactions. The defining contrast subgroup of the original 2^{9-3}design is defined by I = 13457 = 2368 = 1245679 = 4569 = 13679 = 234589 = 12789. This is a resolution IV design with sixty-four points. Adding new sixteen points that is obtained by the way "f_0=79^+; ss=1^+". Finally, the combined design with eighty points provides all main effects and some two-factor interactions, 12, 13, 14, 15, 16, 17, 18, 19, 24, 25, 27, 29, 34, 35, 37, 39, 45, 46, 47, 48, 49, 56, 57, 58, 59, 69, 78, 79, 89, being clear to estimate.

Reconsidering Example 3.1 again, the experimenter may use the method of semi-folding "f_0=7 ; ss=1^+ " to get the new points. Then the forty-eight points are constructed from three fractions: I =

1236 = 2457 = 134567; I = 1 = 236 = 1236; I = 1236 = -12457 = -34567. After some algebraic operations, we can obtain that all main effects and some two-factor interactions 14, 15, 17, 24, 25, 27, 34, 35, 37, 45, 46, 47, 56, 57, and 67 are clear to estimate in this combined design. Similarly, if we use the method of semi-folding "f_0=79 ; ss=1" in Example 3.2 to get new points, then same clear effects will be obtained as the double semi-folding. The results in these examples reveal intrinsic structures of the semi-folding and double semi-folding that can have the same clear effects.

4. Conclusion

For the regular 2^{k-p} fractional factorial designs, the MA criterion is commonly used for choosing optimal plans. However, the criterion of MA does not seem to work in all situations, especially in irregular designs. In this paper, we apply the criterion of "clear" to judge the competitive plans. In the previous section, we give examples to illustrate the properties of the estimation of the double semi-folding, and discuss two constructive methods, semi-folding and double semi-folding. Two provided examples show that two methods have the same clear effects; that is, we can use the method of double semi-folding to obtain the appropriate additional points to replace the method of semi-folding which saves half of the cost. For large k, double semi-folding is significantly much better than the semi-folding because it requires only half the points of a semi-folding fraction to have the same clear effects. In general, the proper method to develop systematic techniques to obtain same information with the use of two approaches is worth further exploration.

Reference

Addelman, S. (1961). Irregular fractions of the 2^n factorial experiments. *Technometrics*, 3, 479-496.

Bose, R. C., Bush, K. A. (1952). Orthogonal array of strength two and three. *Annals of Mathematical Statistics*, 23, 508-524.

Box, G. E. P., Hunter, J. S. (1961). The 2^{k-p} fractional factorial designs. *Technometrics*, 3, Part I, 311-352; Part II, 449-458.

Box, G. E. P., Wilson, K. B. (1951). On the experimental attainment of optimum conditions. *J. Roy. Statist. Soc. Ser. B*, 13, 1-45..

Chen, H. H., Cheng, C. S. (2004). Aberration, estimation capacity and estimation index. *Statist. Sinica*, 14, 203-215.

Dey, A., Ramakrishna, G. V. S. (1977). A note on orthogonal main-effect plans. *Technometrics*, 19, 511-512.

Finney, D. J. (1945). The fractional replication of factorial arrangements. *Annals of Eugenics*, 12, 291-301.

Gupta, V. K., Nigam, A. K., Dey, A. (1982). Orthogonal main-effect plans for asymmetrical factorials.

Technometrics, 24, 135-137.

John, P. W. M. (2000). Breaking alias chains in fractional factorials. *Communications in Statistics Theory and Method*, 29, 2143-2155.

Li, H., Mee, R. W. (2002). Better foldover fractions for resolution III designs. *Technometrics*, 44, 278-283.

Li, W., Lin, D. K. J. (2003). Optimal foldover plans for two-level fractional factorial designs. *Technometrics*, 45, 142-149.

Liau, P. H., Huang, P.H. (2007). Clear effects in the 2^{n-p} combined design. *Communications in Statistics Theory and Method*, 36, 793-799.

Mee,, R. W., Peralta, M. (2000). Semifolding 2^{n-p} designs. *Technometrics*, 30, 122-134.

Montgomery, D. C., Runger, G. C. (1996). Foldovers of resolution IV experimental designs. *Journal of Quality Technology*, 4, 446-450.

Montgomery, D. C. (2013). *Design and Analysis of Experiments*. 8th ed. New York: John Wiley & Sons.

Nigam, A. K., Gupta, V. K. (1985). Construction of orthogonal main-effect plans using Hadamard matrices. *Technometrics*, 27, 37-40.

Plackett, R. L., Burman, J. P. (1946). The design of optimal multifactorial experiments. *Biometrika*, 33, 328-332.

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