行政院國家科學委員會專題研究計畫成果報告

部分重複的 2-變級部分複因子設計之研究

Partially replicated 2-level fractional factorial designs

計畫類別: ☑個別型計畫 □ 整合型計畫

計畫編號: **NSC-90-2118-M-002-007**

執行期間: **90** 年 **8** 月 **1** 日 至 **91** 年 **7** 月 **31** 日

計畫主持人: 廖 振 鐸

處理方式: ☑可立即對外提供參考
□ 一年後可對外提供參考
□ 二年後可對外提供參考

執行單位: 國立台灣大學農藝學研究所試驗統計組

中華民國 **91** 年 **10** 月 **20** 日

中 文 摘 要

本研究探討部分重複 2-變級部分複因子設計之建構。對於試驗者有興趣估計之平均(grand mean)，所有的主效應(main effects)，及某些任意給定的 2-因子效應(2-factor interactions)。我們建構一個試驗大小 $N = 4 \times 2^{n-p}$ 的高效能(high efficiency)的設計，其中 2^{n-p} 個處理組合(treatment combination)有兩重複。設計之建構是基於 4-PFDs (four parallel-flats designs)，我們提出滿足可估設計(nonsingular designs)及直交設計(orthogonal designs)的充分必要條件，並且建構一系列 $N = 16$ 的設計來說明定理的應用。詳細研究成果請參考英文報告。

Partially Replicated Two-level Fractional Factorial designs

C. T. Liao

Division of Biometry
Institute of Agronomy
National Taiwan University
Taipei, Taiwan

Abstract

In a two-level factorial experiment, we consider the designs with partial duplication which permits the estimation of the grand mean, all main effects and some specified nonnegligible two-factor interactions, assuming the other effects are negligible. The best advantage of this kind of experiments is that we are capable of conducting a lack-of-fit test based on pure error. In this study, we consider the construction of the desired designs from the class of four parallel-flats designs in which two parallel flats are identical. It is shown that the designs obtained can have a very simple covariance structure and high D-efficiency. A series of practical designs with run-size $N = 16$ is generated from an algorithm proposed.
Key words: Parallel-flats designs; Fractional factorial designs; Pure error; Lack-of-fit test.

1 Introduction

During the initial stages of experimentation, the unreplicated two-level fractional factorial designs are widely used to estimate a specified set of factorial effects. The effects of interest are often specified according to the knowledge of the investigator and some practical principles of experimentation. See Wu and Hamada (2000). Since each treatment combination is run once in this kind of experiments, it is customary to obtain an experimental error estimate from the remaining degrees of freedom after fitting the specified effects; or the small effects identified by the normal plotting. In many situations, the investigator may not be sure whether this biased error estimate is reliable used to test the lack-of-fit of the model. Therefore, it would be desirable to construct an economical two-level factorial design with some duplicated treatment combinations to meet the requirement that the pure error is estimable, and without scarifying much the efficiency of the specified effects.

There were some two-level factorial designs with partial duplication presented in Dykstra (1959). These designs were constructed from combining a full factorial or a fraction of high resolution, at least resolution V, and a fraction of resolution III. Most of these designs are not practical due to their large run size. An example was given in Pigeon and McAllistar (1989) to show that it is sometimes possible to include partial replication of the design, allowing for an estimate of pure error, without scarifying the orthogonality of the main effects. This special design is actually a design combined from three fractions belonging to a family of regular fractional factorial design with the same defining relations, and one duplicate of the three fractions, i.e. the design is just a particular case of the class of four parallel-flats designs (4-PFDs). This motivates us to develop a systematic approach to construction the designs including partial duplication, for estimating any specified set of effects from the 4-PFDs.

The construction of orthogonal designs for any arbitrary set of effects based on the class of parallel-flats designs can be seen in Srivastava and Li (1966); and Liao, Iyer and Vecchica (1996). However, Chai and Liao (2001) investigate the construction of nonorthogonal designs of user-specified resolution for the class of three parallel-flats designs (3-PFDs). In this study, we modify the results of the paper to tackle our problem of interest. In the next section, we introduce the notation used in the article and review the definition of PFDs. Section 3 discusses the properties of 4-PFDs with two identical parallel flats; and presents an algorithm based on the theorem given. Finally, we generate a series of practical designs with $N = 16$ and compare their efficiency with the computer-aided designs according to the DETMAX algorithm of Mitchell (1974) in section 4.

2 Preliminaries

Let $F_1, F_2, \ldots, F_n$ denote the n two-level factors. As is common practice, $F_1, F_2, \ldots, F_n$ will also represent the main effects. The expression $F_1^{e_1} F_2^{e_2} \ldots F_n^{e_n}$ will represent a general factorial effect with e_i being 0 or 1. If e_i is 1, then F_i appears in the factorial effect, otherwise it does not. The vector $\boldsymbol{e} = [e_1, \ldots, e_n]$ is called the *defining vector* for the general factorial effect $F_1^{e_1} F_2^{e_2} \ldots F_n^{e_n}$. The defining vector $\boldsymbol{e} = [0, 0, \cdots, 0]$ is for the grand mean μ. We will sometimes use the defining vectors to represent the factorial effects for convenience. A treatment combination or run will be represented by an n-tuple whose entries are 0 or 1 depending on whether the corresponding factor occurs at the *low level* or at the *high level*, respectively.

2.1 PFDs

Single flat two-level designs are traditional 2^{n-k} fractional factorial designs defined by appropriately chosen aliasing relations. These aliasing relations can be expressed as linear equations over *Galois Field* of order 2, GF[2]. A single flat 2^{n-k} fractional factorial design $\mathcal{D}_i$ defined by the *alias matrix* $\boldsymbol{A}$ and the *coset representative vector* $\boldsymbol{c}_i$ consists of all treatment combinations $\boldsymbol{t} = (t_1, t_2, \cdots, t_n)$ which satisfy the equation

$$\boldsymbol{A}\boldsymbol{t} = \boldsymbol{c}_i$$

over GF[2], where $\boldsymbol{A}$ is a $k \times n$ matrix of rank k and $\boldsymbol{c}_i$ is any $k \times 1$ vector over GF[2].

A PFD is obtained by taking several, say f, single flat designs $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_f$ together, where $\mathcal{D}_i$ is determined by the alias matrix $\boldsymbol{A}$ and the coset representative vector $\boldsymbol{c}_i$. Such a design consists of $f \times 2^{n-k}$ treatment combinations. Let $\boldsymbol{C}$ be the $k \times f$ *coset indicator matrix* whose columns are $\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_f$. Thus a PFD is determined by the pair $(\boldsymbol{A}, \boldsymbol{C})$. Note that $\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_f$ are not necessarily distinct.

Suppose $\boldsymbol{B}$ is any $n \times (n-k)$ matrix of rank $(n-k)$ over GF[2] such that $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{0}$, where $\boldsymbol{0}$ is a matrix with all elements equal to 0 over GF[2]. Also let $\boldsymbol{z}_i = (z_{i1}, \cdots, z_{in})$ be a *particular* solution of $\boldsymbol{A}\boldsymbol{t} = \boldsymbol{c}_i$. Then $\mathcal{D}_i$ consists of all $\boldsymbol{t}$ satisfying

$$\boldsymbol{t} = \boldsymbol{z}_i + \boldsymbol{B}\boldsymbol{v}$$

where $\boldsymbol{v}$ ranges over all possible vectors of length $(n-k)$ over GF[2] (there are 2^{n-k} such vectors). Hence a PFD can also be determined by the pair $(\boldsymbol{B}, \boldsymbol{Z})$, where $\boldsymbol{Z}$ is the $n \times f$ matrix whose columns are $\boldsymbol{z}_1, \boldsymbol{z}_2, \cdots, \boldsymbol{z}_f$. Certainly, $\boldsymbol{z}_1, \boldsymbol{z}_2, \cdots, \boldsymbol{z}_f$ are not necessarily distinct. In this study, we only focus on the case that $f = 4$ and two of the four $\boldsymbol{z}_i$s are identical.

2.2 Linear model and nonsingular design

Let $\boldsymbol{\beta}$ denote the vector of factorial effects that are not assumed to be zero. The corresponding linear model for the observations from an experiment using a design $\mathcal{D}$ may be written in the form

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{2.1}$$

where the matrix $\boldsymbol{X}$ depends on the design $\mathcal{D}$ and is often called the *design matrix.* The vector $\boldsymbol{\epsilon}$ consists of noise random variables which are assumed to have zero means and are pairwise uncorrelated with a common variance σ^2. The design $\mathcal{D}$ is said to be *nonsingular* for $\boldsymbol{\beta}$ if the *information matrix* $\boldsymbol{X}'\boldsymbol{X}$ is a nonsingular matrix. The Best Linear Unbiased Estimates (BLUEs) of all factorial effects of $\boldsymbol{\beta}$ are available when the design $\mathcal{D}$ used is nonsingular. Particularly, the design is *orthogonal* for $\boldsymbol{\beta}$ if $\boldsymbol{X}'\boldsymbol{X}$ is a diagonal matrix. It is well known that orthogonal designs are optimal (A-optimal, D-optimal and E-optimal) for estimating $\boldsymbol{\beta}$, but orthogonal designs may not exist for every value of N.

3 Partially replicated 4-PFDs

The following proposition gives an expression for the general element of the information matrix $\boldsymbol{X}'\boldsymbol{X}$ of an arbitrary PFD.

Proposition 3.1. Let $\mathcal{D}$ be the PFD determined by the pair $(\boldsymbol{B}, \boldsymbol{Z})$, where $\boldsymbol{B}$ is a $n \times (n-k)$ matrix of rank $(n-k)$ and $\boldsymbol{Z}$ is a $n \times f$ matrix over GF[2]. Suppose the linear model in (2.1) holds.

The element of the information matrix $\boldsymbol{M} = \boldsymbol{X}'\boldsymbol{X}$ corresponding to the row indexed by the factorial effect whose defining vector is $\boldsymbol{e}_1 = [e_{11}, \ldots, e_{1n}]$ and the column indexed by the factorial effect whose defining vector is $\boldsymbol{e}_2 = [e_{21}, \ldots, e_{2n}]$ is $m(\boldsymbol{e}_1, \boldsymbol{e}_2)$ given by

$$m(\boldsymbol{e}_1, \boldsymbol{e}_2) = [\sum_{i=1}^{f}(-1)^{(\boldsymbol{e}_1+\boldsymbol{e}_2)\boldsymbol{z}_i}][\sum_{\boldsymbol{v}}(-1)^{(\boldsymbol{e}_1+\boldsymbol{e}_2)\boldsymbol{B}\boldsymbol{v}}] \tag{3.1}$$

where $\boldsymbol{v}$ ranges over all 2^{n-k} binary vectors of length $(n-k)$. The exponents of (-1) in both factors of (3.1) are computing using arithmetic modulo 2, but the sums of the powers of (-1) are not performed modulo 2.

Let $\boldsymbol{e} = \boldsymbol{e}_1 + \boldsymbol{e}_2$, where $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ are the two defining vectors corresponding to two different elements of $\boldsymbol{\beta}$. If $\boldsymbol{e}\boldsymbol{B} = \boldsymbol{0}$, then $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ are said to be in the same *alias set*. Hence for the given $\boldsymbol{B}$ of rank $(n-k)$, 2^{n-k} distinct alias sets can be defined. If $\boldsymbol{e}\boldsymbol{B} \neq \boldsymbol{0}$, the term $\sum_{\boldsymbol{v}}(-1)^{(\boldsymbol{e}_1+\boldsymbol{e}_2)\boldsymbol{B}\boldsymbol{v}}$ is precisely 0 so that $m(\boldsymbol{e}_1, \boldsymbol{e}_2) = 0$. This implies that the information matrix can always be expressed as a block diagonal matrix given by

$$\boldsymbol{M} = \begin{bmatrix} \boldsymbol{M}_1 & & & \\ & \boldsymbol{M}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{M}_g \end{bmatrix},$$

where each submatrix $\boldsymbol{M}_j$ corresponds to an alias set, $j = 1, 2\cdots, g$, and $g = 2^{n-k}$ the number of distinct alias sets. Notice that a necessary condition ensures that a PFD is nonsingular for $\boldsymbol{\beta}$ is that the maximum dimension of $\boldsymbol{M}_j$ is the number of distinct parallel flats for all j. See Srivastava et al. (1984).

In this paper, we only consider the class of 4-PFDs with two identical parallel flats. We will call this class of 4-PFDs as 4-PFDRs hereinafter. The following theorem illustrates the structure of information matrix for this particular class of designs.

Theorem 3.1. Let the 4-PFDR $\mathcal{D}$ be determined by the pair $(\boldsymbol{B}, \boldsymbol{Z})$, where $\boldsymbol{Z} = [\boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3, \boldsymbol{z}_4]$. WLOG, also let $\boldsymbol{z}_3 = \boldsymbol{z}_4$. Then $\mathcal{D}$ is nonsingular for $\boldsymbol{\beta}$ if and only if each block matrix $\boldsymbol{M}_j$ of the information matrix must have one of the following forms.

(i) $\boldsymbol{M}_j$ is a scalar. Then $\boldsymbol{M}_j = N = 4 \times 2^{n-k}$.

(ii) $\boldsymbol{M}_j$ is a 2×2 matrix. Then $m(\boldsymbol{e}_1, \boldsymbol{e}_2) = 0$, $N/2$ or $-N/2$ if $\boldsymbol{e}_1 \neq \boldsymbol{e}_2$, otherwise it is N.

(iii) $\boldsymbol{M}_j$ is a 3×3 matrix. Then the elements of $m(\boldsymbol{e}_1, \boldsymbol{e}_2)$ have exactly one 0 and two $(N/2)$s; exactly one 0, one $N/2$ and one $-N/2$; or exactly one 0 and two $(-N/2)$s, for $\boldsymbol{e}_1 \neq \boldsymbol{e}_2$. The elements of $m(\boldsymbol{e}_1, \boldsymbol{e}_2) = N$, for $\boldsymbol{e}_1 = \boldsymbol{e}_2$.

Proof. There are exactly 3 distinct parallel flats included in a 4-PFDR. Hence, to ensure a 4-PFDR to be nonsingular, the dimension for each $\boldsymbol{M}_j$ must be less than or equal to 3. The results follow directly by proving the $\boldsymbol{M}_j s$ in (i), (ii) and (iii) are nonsingular.

(i) Trivial.

(ii) Suppose $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ are the two distinct factorial effects in the alias set corresponding to $\boldsymbol{M}_j$. If the $\boldsymbol{z}_i$ chosen such that $\boldsymbol{e}\boldsymbol{z}_i = (\boldsymbol{e}_1 + \boldsymbol{e}_2)\boldsymbol{z}_i = 0$, for all i, then $m(\boldsymbol{e}_1, \boldsymbol{e}_2) = N$. If $\boldsymbol{e}\boldsymbol{z}_i = 1$, for all i, then $m(\boldsymbol{e}_1, \boldsymbol{e}_2) = -N$. For these two cases, $\boldsymbol{M}_j$ is singular. For the remaining cases, $\boldsymbol{M}_j$ is nonsingular, i.e. the set $\{\boldsymbol{e}\boldsymbol{z}_1, \boldsymbol{e}\boldsymbol{z}_2, \boldsymbol{e}\boldsymbol{z}_3, \boldsymbol{e}\boldsymbol{z}_3\}$ has exactly two 0s and two 1s which makes $m(\boldsymbol{e}_1, \boldsymbol{e}_2) = 0$; exactly one 1 and three 0s which makes $m(\boldsymbol{e}_1, \boldsymbol{e}_2) = N/2$; or one 0 and three 1s which makes $m(\boldsymbol{e}_1, \boldsymbol{e}_2) = -N/2$.

(iii) Suppose $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$ are the three distinct factorial effects in the alias set corresponding to $\boldsymbol{M}_j$. Let $\boldsymbol{e}_1^* = \boldsymbol{e}_1 + \boldsymbol{e}_2$, $\boldsymbol{e}_2^* = \boldsymbol{e}_1 + \boldsymbol{e}_3$ and $\boldsymbol{e}_3^* = \boldsymbol{e}_2 + \boldsymbol{e}_3$. Also let the vectors $\boldsymbol{w}_l = (\boldsymbol{e}_l^*\boldsymbol{z}_1, \boldsymbol{e}_l^*\boldsymbol{z}_2, \boldsymbol{e}_l^*\boldsymbol{z}_3, \boldsymbol{e}_l^*\boldsymbol{z}_3)$, for $l = 1, 2, 3$. $\boldsymbol{M}_j$ is singular, if and only if $m(\boldsymbol{e}_l, \boldsymbol{e}_{l'}) = N$ or $-N$ for some $l \neq l'$. So none of $\boldsymbol{w}_l$ can be $(0, 0, 0, 0)$ or $(1, 1, 1, 1)$. Note that $\boldsymbol{e}_{12} + \boldsymbol{e}_{13} + \boldsymbol{e}_{23} = \boldsymbol{0}$. So the set $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3\}$ must be $\{(1, 0, 0, 0), (0, 1, 0, 0), (1, 1, 0, 0)\}$; $\{(1, 0, 0, 0), (1, 0, 1, 1), (0, 0, 1, 1)\}$; or $\{(1, 0, 1, 1), (0, 1, 1, 1), (1, 1, 0, 0)\}$. This completes the proof.

Theorem 3.1. leads to the following results.

(1) From (ii), there are three choices of the nonsingular $\boldsymbol{M}_j s$. It is easy to see that the cases $m(\boldsymbol{e}_1, \boldsymbol{e}_2) = \pm N/2$ for $\boldsymbol{e}_1 \neq \boldsymbol{e}_2$ have "equivalent" information since their eigenvalues are identical and each eigenvalue also has the same multiplicity. The remaining case that $m(\boldsymbol{e}_1, \boldsymbol{e}_2) = 0$ for $\boldsymbol{e}_1 \neq \boldsymbol{e}_2$ is always desired to make $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ orthogonal. Similarly, the three choices of the nonsingular $\boldsymbol{M}_j s$ described in (iii) have equivalent information. Every $\boldsymbol{M}_j$ has the same three

eigenvalues N, $(1-\sqrt{2}/2)N$ and $(1+\sqrt{2}/2)N$. Therefore, there are at most two nonisomorphic nonsingular 4-PFDRs when each alias set consists of 2 or 3 effects.

(2) If a 4-PFDR is orthogonal for $\boldsymbol{\beta}$, then the maximum number of elements included in the $\boldsymbol{\beta}$ is $N/2$, i.e. none of its alias sets consists of three elements. Also note that if a 4-PFDR is nonsingular for $\boldsymbol{\beta}$, then the maximum number of elements included in the $\boldsymbol{\beta}$ is $3N/4$.

We now revisit the example given in Pigeon and McAllister (1989). They listed the following 16-run design; and claimed that the design is an orthogonal main-effect plan, with two duplicates of 4 treatment combinations, for the case that the number of factors $n = 7$.

	Factors						
Run	F_1	F_2	F_3	F_4	F_5	F_6	F_7
1	0	0	0	0	0	1	1
2	0	1	1	0	1	1	0
3	0	0	0	1	1	0	0
4	0	1	1	1	0	0	1
5	1	0	1	0	0	0	0
6	1	1	0	0	1	0	1
7	1	0	1	1	1	1	1
8	1	1	0	1	0	1	0
9	0	0	0	0	0	1	1
10	0	1	1	0	1	1	0
11	0	0	0	1	1	0	0
12	0	1	1	1	0	0	1
13	1	1	0	1	1	1	1
14	1	0	1	1	0	1	0
15	1	1	0	0	0	0	0
16	1	0	1	0	1	0	1

It can be verified that the above design is a 4-PFDR determined by the following pair $(\boldsymbol{A}, \boldsymbol{C})$.

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{C} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix}.$$

This 4-PFDR can also be determined by the following pair $(\boldsymbol{B}, \boldsymbol{Z})$.

$$\boldsymbol{B} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}, \quad \boldsymbol{Z} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

It can be easily checked, using Theorem 3.1, that the design is orthogonal for $\boldsymbol{\beta}$ given by

$$\boldsymbol{\beta} = \{\mu; F_1, F_2, F_3, F_4, F_5, F_6, F_7\}.$$

There are four alias sets given by $G_1 = \{\mu, F_1\}$, $G_2 = \{F_2, F_3\}$, $G_3 = \{F_4, F_6\}$ and $G_4 = \{F_5, F_7\}$. None of them consists of three nonzero effects. In particular, we can obtain another 4-PFDR, determined by the same matrix $\boldsymbol{B}$ but a different matrix $\boldsymbol{Z}$ given by

$$\boldsymbol{Z} = [\boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{0}, \boldsymbol{0}] = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

which is also orthogonal for this $\boldsymbol{\beta}$.

According to Theorem 3.1, WLOG, we can simply consider the 4-PFDR with $\boldsymbol{Z} = [\boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{0}, \boldsymbol{0}]$. The following procedure shows us how to choose the appropriate $\boldsymbol{B}$, $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ to obtain the highest efficiency 4-PFDRs for any given $\boldsymbol{\beta}$.

Step (1): Let m_i be the number of effects of alias set G_i, $i = 1, 2, \cdots, 2^{n-k}$. Choose the matrix $\boldsymbol{B}$ such that all m_i are less than or equal to 3 and as equal as possible, i.e. $m_i \leq 3$ for all i; and $0 \leq |m_i - m_j| \leq 1$ for $i \neq j$. A complete search algorithm for the eligible $\boldsymbol{B}$ can be obtained from directly modifying the algorithm presented in Franklin and Bailey (1977). The Franklin-Bailey algorithm is to search for all possible alias matrices $\boldsymbol{A}$ (equivalent to $\boldsymbol{B}$), such that the maximum number of m_i is equal to 1.

Step (2): Define a set of vectors $\boldsymbol{R}$ as

$$\boldsymbol{R} = \{\boldsymbol{e}_1 + \boldsymbol{e}_2 \mid \boldsymbol{e}_1, \boldsymbol{e}_2 \text{ are in an alias set consisting of exactly two effects.}\}.$$

Choose $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$ such that if $\boldsymbol{e}_1^*, \boldsymbol{e}_2^*, \boldsymbol{e}_3^* \in \boldsymbol{R}$ and $\boldsymbol{e}_1^* + \boldsymbol{e}_2^* + \boldsymbol{e}_3^* = \boldsymbol{0}$, then the set $\{\boldsymbol{e}_1^*\boldsymbol{z}_1, \boldsymbol{e}_1^*\boldsymbol{z}_2\}$ has exactly two 1s; and the set $\{\boldsymbol{e}_2^*\boldsymbol{z}_1, \boldsymbol{e}_2^*\boldsymbol{z}_2\}$ has exactly one 1 and one 0. For the remaining elements, say $\boldsymbol{e}^*$, the set $\{\boldsymbol{e}^*\boldsymbol{z}_1, \boldsymbol{e}^*\boldsymbol{z}_2\}$ has exactly two 1s.

Step (3): For each alias set consisting of exactly 3 effects, say $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$, let $\boldsymbol{e}_1^* = \boldsymbol{e}_1 + \boldsymbol{e}_2$, $\boldsymbol{e}_2^* = \boldsymbol{e}_1 + \boldsymbol{e}_3$, $\boldsymbol{e}_3^* = \boldsymbol{e}_2 + \boldsymbol{e}_3$; and the vectors $\boldsymbol{w}_l = (\boldsymbol{e}_l^*\boldsymbol{z}_1, \boldsymbol{e}_l^*\boldsymbol{z}_2)$, for $l = 1, 2, 3$. Reset $\boldsymbol{z}_1$, $\boldsymbol{z}_2$ or both if necessary, such that the set $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3\}$ is $\{(1,1), (0,1), (1,0)\}$.

The following example is given to illustrate the algorithm.

Example 3.1. For a 2^4 experiment, suppose $\boldsymbol{\beta}$ consists of the following factorial effects:

$$\mu; F_1, F_2, F_3, F_4; F_1F_2, F_1F_3, F_1F_4, F_2F_3, F_2F_4, F_3F_4.$$

In this case, there are 11 effects to be estimated. Any factorial effect that is not in the span of the above effects is assumed to be zero. Note that the factorial effect F_1 has $(1,0,0,0)$ as its defining vector, F_1F_2 has the vector $(1,1,0,0)$ as its defining vector, etc. We consider a 4-PFDR with $N = 16$ for estimating the effects. Following the algorithm, we have

Step (1): Choose matrix $\boldsymbol{B}$ given by

$$\boldsymbol{B}(4 \times 2) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}.$$

There are four alias sets $G_1 = \{\mu, F_1\}$, $G_2 = \{F_2, F_1F_2, F_3F_4\}$, $G_3 = \{F_3, F_1F_3, F_2F_4\}$, and $G_4 = \{F_4, F_1F_4, F_2F_3\}$ determined by this $\boldsymbol{B}$.

Step (2): From the alias set G_1,

$$\boldsymbol{R} = \{F_1\}.$$

Then, first choose $\boldsymbol{z}_1 = (1,1,1,0)$ and $\boldsymbol{z}_2 = (1,0,0,0)$.

Step (3): For the remaining three alias sets G_2, G_3 and G_4, every alias set has the same vectors set $\{\boldsymbol{e}_1^*, \boldsymbol{e}_2^*, \boldsymbol{e}_3^*\}$ which is $\{F_1, F_2F_3F_4, F_1F_2F_3F_4\}$. Thus, it is needed to reset $\boldsymbol{z}_1$ to be $(1,1,1,1)$.

Consequently, the 4-PFDR determined by the matrix $\boldsymbol{B}$ of Step (1) and

$$\boldsymbol{Z}(4 \times 4) = \begin{bmatrix} \boldsymbol{z}_1 & \boldsymbol{z}_2 & \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

is a 16-run nonsingular design for the $\boldsymbol{\beta}$. It is interesting to point out that this 4-PFDR is also orthogonal for the $\boldsymbol{\beta}$ only consisting of effects μ, F_1, F_2, F_3, F_4, F_1F_2, F_1F_3, F_1F_4.

4 Designs of user-specified resolution with $N = 16$

In this section, we report a series of 4-PFDRs with $N = 16$ for the two-level factorial experiments. Suppose that n factors are divided into p subgroups, where the i^{th} subgroup has n_i factors in it. Let p' be the number of the subgroups of interest, where $p' = p$ or $p' = p - 1$. If $p' = p - 1$, the factors in the p^{th} subgroup are not involved in the interactions.

We consider the following two classes of designs:

(I) interactions within the subgroups of interest.
(II) interactions which involve factors from different subgroups.

For convenience, the parameters and symbols are listed below.

n =number of the factors under consideration,
p =number of subgroups,
p' = number of subgroups actually involved in the two-factor interactions,
n_i =number of factors in the i^{th} $(i = 1, 2, \cdots, p)$ subgroup,
v =number of the nonzero effects to be estimated.

Let $\tilde{\boldsymbol{B}} = \{\tilde{\boldsymbol{b}}_1, \tilde{\boldsymbol{b}}_2\}$, where $\tilde{\boldsymbol{b}}_i$ denotes the position number of nonzero entries of the i^{th} column of $\boldsymbol{B}$. Similarly, let $\tilde{\boldsymbol{z}}_1$ and $\tilde{\boldsymbol{z}}_2$ denote the position number of nonzero entries of the $\boldsymbol{z}_1$ and $\boldsymbol{z}_2$, respectively. Recall that we set $\boldsymbol{z}_3 = \boldsymbol{z}_4 = \boldsymbol{0}$. For simplicity, we also write $\boldsymbol{\beta} = \{\mu; F_1, \cdots, F_n; F_1F_2, \cdots, F_iF_j\}$ as $\boldsymbol{\beta} = \{0; 1, \cdots, n; 1 \cdot 2, \cdots, i \cdot j\}$. For example, $\boldsymbol{\beta} = \{\mu; F_1, F_2, F_3; F_1F_2, F_1F_3\} = \{0; 1, 2, 3; 1 \cdot 2, 1 \cdot 3\}$.

For any given $\boldsymbol{\beta}$, there is no systematic construction method available in the literature for optimal partially replicated designs. However, a very natural method is to augment a nonsingular design for the $\boldsymbol{\beta}$ with some duplicates of its fraction. Therefore, to confirm how good the designs obtained from the class of 4-PFDRs are in estimating the user-specified effects, we compare D-efficiency of our designs with those generated by the following approach. For the given $\boldsymbol{\beta}$, we first generate a D-optimal design by the OPTEX procedure (GENERATE CRITERION=D METHOD=DETMAX ITER=100 N=12) of SAS/QC software. (SAS Institute, 2002). Here we choose the DETMAX algorithm of Mitchell (1974) for searching the D-optimal designs with $N = 12$ runs using 100 iterations. Then we augment those 12 runs generated with 4 of them. The final design we choose is the one having the highest D-efficiency among the $\binom{12}{4} = 495$ competing choices. D-efficiency, D_e, for each design is defined by

$$D_e = \frac{|\boldsymbol{X}'\boldsymbol{X}|^{\frac{1}{v}}}{N}.$$

Now we report the designs obtained and their D-efficiency for estimating $\boldsymbol{\beta}$, and the D-efficiency of the corresponding Mitchell designs.

Class I: Interactions within subgroups. The factorial effects for the interactions within subgroups are listed in Table 1. The 4-PFDRs obtained and their D_e are given in Table 2.

Table 1. Factorial effects for the interactions within subgroups.

Case	n	v	p	p'	$(n_1, \cdots, n_p)$	$\boldsymbol{\beta}$
1	4	8	2	1	(3, 1)	$\{0; 1, \cdots, 4; 1\cdot 2, 1\cdot 3, 2\cdot 3\}$
2	4	11	1	1	(4)	$\{0; 1, \cdots, 4; 1\cdot 2, 1\cdot 3, 1\cdot 4, 2\cdot 3, 2\cdot 4, 3\cdot 4\}$
3	5	10	2	2	(3, 2)	$\{0; 1, \cdots, 5; 1\cdot 2, 1\cdot 3, 2\cdot 3, 4\cdot 5\}$
4	5	12	2	1	(4, 1)	$\{0; 1, \cdots, 5; 1\cdot 2, 1\cdot 3, 1\cdot 4, 2\cdot 3, 2\cdot 4, 3\cdot 4\}$
5	6	8	2	1	(2, 4)	$\{0; 1, \cdots, 6; 1\cdot 2\}$
6	6	9	3	2	(2, 2, 2)	$\{0; 1, \cdots, 6; 1\cdot 2, 3\cdot 4\}$
7	6	10	2	1	(3, 3)	$\{0; 1, \cdots, 6; 1\cdot 2, 1\cdot 3, 2\cdot 3\}$
8	6	10	3	3	(2, 2, 2)	$\{0; 1, \cdots, 6; 1\cdot 2, 3\cdot 4, 5\cdot 6\}$
9	6	11	3	2	(3, 2, 1)	$\{0; 1, \cdots, 6; 1\cdot 2, 1\cdot 3, 2\cdot 3, 4\cdot 5\}$
10	7	8	1	0	(7)	$\{0; 1, \cdots, 7\}$
11	7	9	2	1	(2, 5)	$\{0; 1, \cdots, 7; 1\cdot 2\}$
12	7	10	3	2	(2, 2, 3)	$\{0; 1, \cdots, 7; 1\cdot 2, 3\cdot 4\}$
13	7	11	4	3	(2, 2, 2, 1)	$\{0; 1, \cdots, 7; 1\cdot 2, 3\cdot 4, 5\cdot 6\}$
14	7	11	2	1	(3, 4)	$\{0; 1, \cdots, 7; 1\cdot 2, 1\cdot 3, 2\cdot 3\}$
15	8	10	2	1	(2, 6)	$\{0; 1, \cdots, 8; 1\cdot 2\}$
16	8	11	3	2	(2, 2, 4)	$\{0; 1, \cdots, 8; 1\cdot 2, 3\cdot 4\}$
17	8	12	4	3	(2, 2, 2, 2)	$\{0; 1, \cdots, 8; 1\cdot 2, 3\cdot 4, 5\cdot 6\}$
18	8	12	2	1	(3, 5)	$\{0; 1, \cdots, 8; 1\cdot 2, 1\cdot 3, 2\cdot 3\}$
19	9	11	2	1	(2, 7)	$\{0; 1, \cdots, 9; 1\cdot 2\}$
20	9	12	3	2	(2, 2, 5)	$\{0; 1, \cdots, 9; 1\cdot 2, 3\cdot 4\}$

Table 2. D_e(4-PFDR) and D_e(MD).

Case	$\tilde{B}$	$\{\tilde{z}_1, \tilde{z}_2\}$	D_e(4-PFDR)	D_e(MD)
1	{24, 34}	{1234, 14}	1.000	1.000
2	{24, 34}	{123, 14}	0.828	0.823
3	{125, 345}	{135, 24}	0.871	0.890
4	{124, 345}	{134, 25}	0.794	0.809
5	{126, 3456}	{135, 246}	1.000	0.958
6	{126, 3456}	{135, 246}	0.926	0.914
7	{126, 3456}	{135, 246}	0.871	0.907
8	{246, 356}	{124, 13}	0.871	0.871
9	{256, 3456}	{135, 246}	0.828	0.837
10	{1346, 2356}	{1237, 4567}	1.000	0.968
11	{1346, 2356}	{1237, 4567}	0.926	0.957
12	{1346, 2356}	{1237, 4567}	0.871	0.901
13	{1346, 2356}	{1237, 4567}	0.828	0.862
14	{1346, 2356}	{1237, 4567}	0.828	0.882
15	{13468, 2356}	{12378, 45678}	0.871	0.928
16	{13468, 2356}	{12378, 45678}	0.828	0.882
17	{13468, 2356}	{12378, 45678}	0.794	0.827
18	{1346, 2356}	{145678, 17}	0.794	0.849
19	{13468, 23569}	{123789, 456789}	0.828	0.907
20	{35789, 25689}	{145679, 1478}	0.794	0.855

For Case 1, we are interested in estimating the general mean μ, all main effects and all two-factor interactions within the subgroup $\{F_1, F_2, F_3\}$ for a 2^4 factorial experimental design. Thus we have $n = 4$, $v = 8$, $p = 2$, $p' = 1$, $(n_1, n_2) = (3, 1)$. The vector of the nonnegligible factorial effects $\boldsymbol{\beta}$ can be expressed as

$$\boldsymbol{\beta} = \{0; 1, 2, 3, 4; 1 \cdot 2, 1 \cdot 3, 2 \cdot 3\}.$$

One possible 4-PFDR for this $\boldsymbol{\beta}$ is the set of treatment combinations $\boldsymbol{t}$ satisfying $\boldsymbol{t} = \boldsymbol{z}_i + \boldsymbol{B}\boldsymbol{v}$, for $i = 1, 2, 3, 4$, where

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, \boldsymbol{Z} = \begin{bmatrix} \boldsymbol{z}_1 & \boldsymbol{z}_2 & \boldsymbol{z}_3 & \boldsymbol{z}_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix},$$

and $\boldsymbol{v}$ ranges over all possible vectors of length 2 over GF[2]. This 4-PFDR is a 16-run design and can be briefly described by $\tilde{\boldsymbol{B}} = \{24, 34\}$ and $\{\tilde{\boldsymbol{z}}_1, \tilde{\boldsymbol{z}}_2\} = \{1234, 14\}$ since $\boldsymbol{z}_3 = \boldsymbol{z}_4 = \boldsymbol{0}$. The rest of cases are similar to this case. The following is the class of designs for the cases of interactions between subgroups.

Class II: Interactions between subgroups. The factorial effects for the interactions between subgroup are given in Table 3. The 4-PFDRs obtained and their D-efficiency are given in Table 4.

Table 3. Factorial effects for the interactions between subgroups.

Case	n	v	p	p'	$(n_1, \cdots, n_p)$	$\boldsymbol{\beta}$
1	4	8	2	2	(1, 3)	$\{0; 1, \cdots, 4; 1\cdot 2, 1\cdot 3, 1\cdot 4\}$
2	4	9	2	2	(2, 2)	$\{0; 1, \cdots, 4; 1\cdot 3, 1\cdot 4, 2\cdot 3, 2\cdot 4\}$
3	5	8	3	2	(1, 2, 2)	$\{0; 1, \cdots, 5; 1\cdot 2, 1\cdot 3\}$
4	5	9	3	2	(1, 3, 1)	$\{0; 1, \cdots, 5; 1\cdot 2, 1\cdot 3, 1\cdot 4\}$
5	5	10	2	2	(1, 4)	$\{0; 1, \cdots, 5; 1\cdot 2, 1\cdot 3, 1\cdot 4, 1\cdot 5\}$
6	5	10	3	2	(2, 2, 1)	$\{0; 1, \cdots, 5; 1\cdot 3, 1\cdot 4, 2\cdot 3, 2\cdot 4\}$
7	5	12	2	2	(2, 3)	$\{0; 1, \cdots, 5; 1\cdot 3, 1\cdot 4, 1\cdot 5, 2\cdot 3, 2\cdot 4, 2\cdot 5\}$
8	6	9	3	2	(1, 2, 3)	$\{0; 1, \cdots, 6; 1\cdot 2, 1\cdot 3\}$
9	6	10	3	2	(1, 3, 2)	$\{0; 1, \cdots, 6; 1\cdot 2, 1\cdot 3, 1\cdot 4\}$
10	6	11	3	2	(1, 4, 1)	$\{0; 1, \cdots, 6; 1\cdot 2, 1\cdot 3, 1\cdot 4, 1\cdot 5\}$
11	6	12	2	2	(1, 5)	$\{0; 1, \cdots, 6; 1\cdot 2, 1\cdot 3, 1\cdot 4, 1\cdot 5, 1\cdot 6\}$
12	6	11	3	2	(2, 2, 2)	$\{0; 1, \cdots, 6; 1\cdot 3, 1\cdot 4, 2\cdot 3, 2\cdot 4\}$
13	7	10	3	2	(1, 2, 4)	$\{0; 1, \cdots, 7; 1\cdot 2, 1\cdot 3\}$
14	7	11	3	2	(1, 3, 3)	$\{0; 1, \cdots, 7; 1\cdot 2, 1\cdot 3, 1\cdot 4\}$
15	7	12	3	2	(1, 4, 2)	$\{0; 1, \cdots, 7; 1\cdot 2, 1\cdot 3, 1\cdot 4, 1\cdot 5\}$
16	7	12	3	2	(2, 2, 3)	$\{0; 1, \cdots, 7; 1\cdot 3, 1\cdot 4, 2\cdot 3, 2\cdot 4\}$
17	8	11	3	2	(1, 2, 5)	$\{0; 1, \cdots, 8; 1\cdot 2, 1\cdot 3\}$
18	8	12	3	2	(1, 3, 4)	$\{0; 1, \cdots, 8; 1\cdot 2, 1\cdot 3, 1\cdot 4\}$
19	9	12	3	2	(1, 2, 6)	$\{0; 1, \cdots, 9; 1\cdot 2, 1\cdot 3\}$
20	10	12	3	2	(1, 1, 8)	$\{0; 1, \cdots, T; 1\cdot 2\}$

Table 4. D_e(4-PFDR) and D_e(MD).

Case	$\boldsymbol{B}$	$\{\tilde{z}_1, \tilde{z}_2\}$	D_e(4-PFDR)	D_e(MD)
1	{24, 34}	{1234, 123}	1.000	1.000
2	{123, 24}	{23, 3}	0.869	0.871
3	{12, 13}	{14, 145}	1.000	0.965
4	{12, 13}	{14, 145}	0.926	0.938
5	{12, 13}	{14, 145}	0.871	0.917
6	{235, 245}	{123, 12}	0.871	0.874
7	{235, 245}	{12, 1}	0.794	0.794
8	{1235, 345}	{1346, 23}	0.926	0.949
9	{1235, 345}	{1346, 235}	0.871	0.917
10	{1235, 345}	{1346, 235}	0.828	0.876
11	{1235, 345}	{1346, 235}	0.794	0.855
12	{235, 245}	{123, 126}	0.828	0.857
13	{12357, 345}	{1346, 23}	0.871	0.916
14	{12357, 345}	{1346, 235}	0.828	0.879
15	{12357, 345}	{1346, 235}	0.794	0.855
16	{2357, 2457}	{1237, 126}	0.794	0.831
17	{12357, 3458}	{1346, 23}	0.828	0.883
18	{12357, 3458}	{1346, 23}	0.794	0.855
19	{123579, 34589}	{1346, 239}	0.794	0.861
20	{123579, 34589T}	{1346T, 239T}	0.794	0.891

†T denotes 10 .

Comments: (i) The 4-PFDRs obtained have high D-efficiency and their covariance structure is known and very simple. This appealing property ensures an easier statistical inference and interpretation. Note that the covariance structure of Mitchell's designs cannot be characterized. (ii) Mitchell's DETMAX algorithm is a powerful heuristic and very successful in finding D-optimal designs (see, Galil and Kiefer (1980) and Welch (1984)). It is shown that D-efficiency of the most 4-PFDRs obtained is close to that of the corresponding Mitchell's designs. It is interesting to notice that the 4-PFDRs obtained can be even better for some cases. This implies that the 4-PFRDs obtained are close to the best possible. In particular, the class of 4-PFDRs obviously is a good candidate for constructing partially replicated orthogonal designs of user-specified resolution. In practice, the desired 4-PFDRs can be easily obtained from the proposed algorithm in section 3.

5 Concluding remarks

There are several publications on identifying the location effects and dispersion effects by using the unreplicated two-level factorial designs. See Box and Meyer (1986), Bergman and Hynen (1997), Pan (1999), Liao and Iyer (2000) and McGrath and Lin (2001). As discussed in Pan (1999), the nature of the unreplicated designs inevitably leads to the problem that the dispersion effects are confounded with the location effects. Therefore, the dispersion effects cannot be identified efficiently. Pan (1999) strongly suggested that the dispersion effects should be estimated from the pure replicates; and illustrated that the dispersion effects can be identified using an economical twice-replicated factional experiment. So it is believed that the 4-PFDRs are practical and good choices for studying the location effects and the dispersion effects simultaneously. This will be our further study in the future.

References

[1] Bergman, B. and Hynên, A. (1997). Dispersion effects from fractional factorial designs in the 2^{k-p} series. *Technometrics*, **39**, 191-198.

[2] Box, G. E. P. and Meyer, R. D. (1986). Dispersion effects from fractional designs. *Technometrics*, **28**, 19-27.

[3] Chai, F. S. and Liao, C. T. (2001). Three parallel flats designs for two-level experiments. *Aust. N.Z. J. Stat.*, **43**, 113-119.

[4] Dykstra, O. Jr. (1959). Partial duplication of factorial experiments. *Technometrics*, **1**, 63-75.

[5] Franklin, M. F. and Bailey, R. A. (1977). Selecting defining contrasts and confounded effects in two-level experiments. *Applied Statistics*, **27**, 321-326.

[6] Galil, Z. and Kiefer, J. (1980). Time and space-saving computer methods, related to Mithchell's DETMAX, for finding D-optimal designs. *Technometrics*, **22**, 301-313.

[7] Liao, C. T. and Iyer, H. K. (2000). Optimal 2^{n-p} fractional factorial designs for dispersion effects under a location-dispersion model. *Commun. Statist.—Theory Meth.* **29**, 823-835.

[8] Liao, C. T., Iyer, H. K. and Vecchia, D. F. (1996). Orthogonal designs of two-level user-specified resolution where $N \neq 2^k$. *Technometrics*, **38**, 342-353.

[9] McGrath, R. N. and Lin, D. K. J. (2001). Testing multiple dispersion effects in unrepliacated fractional factorial designs. *Technometrics*, **43**, 406-414.

[10] Mitchell, T. J. (1974). An algorithm of D-optimal experimental designs. *Technometrics*, **16**, 203-210.

[11] Pan, G. (1999). The impact of unidentified location effects on dispersion effects identification from unreplicated factorial designs. *Technometrics*, **41**, 313-326.

[12] Pigeon, J. G. and McAllister, P. R. (1989). A note on partially replicated orthogonal main-effect plans. *Technometrics*, **31**, 249-251.

[13] SAS Institute. (2002). SAS/QC user's guide. Release 8.1 ed. SAS Inst., Cary, NC, USA.

[14] Srivastava, J. N., Anderson, D. A. and Mardekian, J. (1984). Theory of factorial designs of of parallel flats type. I. The coefficient matrix. *J. Statist. Plann. Inference*, **9**, 229-252.

[15] Srivastava, J. N. and Li, F. (1996). Orthogonal designs of parallel flats type. *J. Statist. Plann. Inference*, **53**, 261-283.

[16] Welch, W. J. (1984). Computer-aided design of experiments for response estimation. *Technometrics*, **26**, 217-224.

[17] Wu, C.F.J. and Hamada, M. (2000). Experiments: planning, analysis, and parameter design optimization. New York: John Wiley & Sons.