Journal of Statistical Planning and Inference 66 (1998) 345–361

journal of statistical planning and inference

Aliasing in 2^{n-k} fractions in the absence of within group interactions

G. Constantine*[,1], L. Xue

Department of Mathematics and Statistics, University of Pittsburgh, Pittsburgh, PA 15260, USA

Abstract

It is often known in advance that certain subsets of factors act independently upon a response. Such information can be used to estimational advantage by aliasing low-order effects with such zero interactions. We find the best 2^{n-k} fractions for the case when the factors can be partitioned into two classes such that non-zero interactions may exist only between classes but not within a class.

AMS classification: 62K10; 05B05

1. Introduction

Fractional 2^{n-k} designs are frequently used in experimentation. There is rich literature on the selection and analysis of such fractions. We refer to Box and Hunter (1961a, b) for basic properties of these designs, including the design resolution. Recent contributions focus on the selection of designs in the presence of prior information regarding certain interactions. Deterministic knowledge often informs that certain factors, or subsets of factors, act independently (additively, to be exact) upon the response. Interactions between such factors are therefore non-existent. This prompts exploitation of the aliasing set, whereupon fractions would be selected in which important low-order effects are aliased with the zero interactions.

In this paper we provide a complete study of the case in which the set of factors can be partitioned into two classes, such that no interactions are present between factors in the same class, but they may be present between factors in different classes. A useful way to interpret such situations is by thinking that factors are partitioned according to "gender" with possible interactions due only to the "gender difference".

* Corresponding author.

[1] Supported by a Fulbright grant.

0378-3758/98/$19.00
PII S0378-3758(97)00084-0

If factors $F_1, F_2, \ldots, F_n$ are believed to influence the response y, and each factor F_i has two levels, high and low, we denote the two values that F_i can take by 1 and -1, the former when F_i is at high level and the latter when it is at low level. Let H be a subgroup of rank k of the group of interactions. The elements in H are said to be *aliased* with the identity I. The full interaction group is, therefore, partitioned into 2^{n-k} *aliasing sets* (or cosets of H). Two interactions are *aliased* (or in the same coset of H) if and only if their product is contained in H. A 2^{n-k} *fractional factorial design* (*ffd*) is thus produced. We call H the *subgroup of defining relations* (*sdr*). A set of generators of H is called its *defining relations*. For convenience, we write $A = B$ if the interactions A and B are aliased. The smallest order among the non-identity interactions in H is called the *resolution* of the design. We denote by $R(n, k)$ the maximum resolution possible to be achieved among all 2^{n-k} fractional factorial designs.

Denote by $l(h)$ the order of interaction h. For a subgroup of defining relations H let $n_i(H)$ be the number of interactions of order i in H and the vector

$$V(H) = (n_1(H), n_2(H), \ldots)$$

be its *order pattern*. We say that H_1 has lesser *aberration* than H_2 if the first unequal components, say the ith components, of $V(H_1)$ and $V(H_2)$ satisfy $n_i(H_1) < n_i(H_2)$. A sdr H is said to have minimum aberration if there is no sdr which has lesser aberration than H. Note that the *minimum aberration* design has maximum resolution. We refer to Fries and Hunter (1980) and Franklin (1984) for more information on design aberration.

The *projection* of an interaction h onto factors $\{i_1, i_2, \ldots, i_r\}$ is the interaction obtained from h by ignoring the factors not from the set $\{i_1, i_2, \ldots, i_r\}$. For example, the projection of the interaction 125679 onto factors $\{1, 2, 3, 4, 5\}$ is 125 and the projection of the interaction 125679 onto factors $\{6, 7, 8, 9\}$ is 679. The projection of a sdr H onto factors $\{i_1, i_2, \ldots, i_r\}$ is the subgroup obtained by projecting each interaction of H onto $\{i_1, i_2, \ldots, i_r\}$. For example, the projection of

$$H = \{I, 123479, 345689, 125678\}$$

onto factors $\{1, 2, 3, 4, 5\}$ is

$$H_1 = \{I, 1234, 345, 125\}.$$

As previously stated, from deterministic knowledge usually due to engineering reasons, the experimenter often knows that certain interactions between factors are zero. A graph G is therefore constructed as follows. Vertices are the factors. Join vertices i and j by an edge if *interactions which involve both factors F_i and F_j are zero.* Whereas we cannot assert that such interaction structure is always a reflection of what one finds in natural phenomena, it always is, on the other hand, less restrictive than working with main effects and second-order interactions only, by pretending that *all* the higher-order interactions are zero. The approach to the study of higher-order interactions is central to factorial experimentation. A detailed and clear discussion of the main thrusts of the problem can be found in Srivastava (1990, Section 8).

Given a 2^{n-k} ffd D and a graph G associated with D in the manner described above, we define the G-sets and G-estimable interactions under the pair (D, G). The *G-sets* under (D, G) are obtained by deleting the prescribed zero interactions from the aliasing sets of D. An interaction is *G-estimable* under (D, G) if it is either a prescribed zero interaction or the only interaction in a G-set.

Example. Let D be the 2^{4-2} ffd with defining relations $I = 123 = 34$, and let G be the graph of 4 vertices with edges 12, 23, 24, 34. Then the aliasing sets are

$$H = \{I, 123, 34, 124\},$$

$$1H = \{1, 23, 134, 24\},$$

$$2H = \{2, 13, 234, 14\},$$

$$3H = \{3, 12, 4, 1234\}.$$

The G-sets are

$$1H_G = \{1\},$$

$$2H_G = \{2, 13, 14\},$$

$$3H_G = \{3, 4\}.$$

The only G-estimable non-zero interaction under (D, G) is therefore the main effect 1.

We distinguish between the ffd (D, G) in which interactions containing edges of G are all zero, and the ffd D with no restrictions attached. Whenever necessary we emphasize the difference by calling D the *unconstrained* ffd.

For a ffd D and a graph G, *we define a* $(n + 1)$-dimensional vector,

$$m(D, G) = (m_1, m_2, \ldots, m_n; m_{n+1}),$$

where the ith coordinate m_i (written also as $m_i(D, G)$); $i = 1, 2, \ldots, n$, represents the number of G-estimable non-zero i-factor interactions, and m_{n+1} is the resolution of the design D.

Let D_1 and D_2 be 2^{n-k} fractional factorial designs and G a graph. We say that the design D_1 is *G-better* than D_2 if $m_1(D_1, G) > m_1(D_2, G)$ or there exists a positive integer t such that $m_i(D_1, G) = m_i(D_2, G)$ for $1 \leqslant i \leqslant t$ and $m_{t+1}(D_1, G) > m_{t+1}(D_2, G)$; $1 \leqslant t \leqslant n$. A 2^{n-k} fractional factorial design D is *G-best* for a given graph G if no other 2^{n-k} fractional factorial design is G-better than D.

This criterion leads most of the time to designs that are rather nice and whose resolution is either highest possible or close to it. Occasionally, however, the designs that are G-best turn out to have a resolution that is unacceptably low. We illustrate by an extreme example. Consider an experiment on four factors, each at two levels. We are interested in a half fraction, knowing that all interactions that contain either of 13, 14, 23, 24 are zero. In accordance to the above criterion the G-best half fraction is a design of resolution 1 with subgroup of relations $1 = I$. In this design all effects except 1 are G-estimable. Effect 1 is, of course, completely obliterated with no chance of being recovered even if some additional interactions are zero or negligible. In contrast, the fraction $1234 = I$, has both effect 1 and 234 not G-estimable, but its resolution is maximal. Should the 3-factor interaction 234 turn out to be zero, then all effects, including the main effect 1, are G-estimable. For all practical purposes, it is difficult to argue in favor of the first design eventhough the criterion prefers it; most practioners would opt for the latter, since it arguably offers at least a chance of rescuing effect 1.

We supplement the above criterion by constructing designs of either maximal resolution, or resolution at least five, in addition to the designs that the criterion selects as optimal. This is pointed out throughout the paper, whenever the situation requires such corrective measures.

A graph is *complete* if there is an edge between any pair of distinct vertices. A graph is a *direct sum of two complete graphs* if its vertices can be partitioned into two classes such that there are no edges between vertices from different classes but there is an edge between any pair of vertices in the same class. For example, the graph with edges 12, 13, 14, 23, 24, 34, and 56 is a direct sum of two complete graphs. For the direct sum of these two complete graphs the G-best 2^{6-2} ffd is the design with the sdr

$$H = \{I, 1236, 1245, 3456\},$$

in which the G-estimable non-zero interactions are all the main effects and the 2-factor interactions 15, 16, 25, and 26. We prove this in Section 3.2.2. If we know in advance, for example, that interactions 35 and 46 are important and that the zero interactions are given as above, then the G-best design to use has sdr

$$H^{(13)(24)} = \{I, 1346, 2345, 1256\}.$$

It is simply obtained by a suitable relabeling of factors. This is the G-best 2^{6-2} fraction in which, as before, all the main effects are G-estimable along with the 2-factor interactions 35, 36, 45, 46, the first and last of which are the two interactions we believe are particularly important.

2. An overview of cases

We assume that the factors $F_1, F_2, \ldots, F_n$ can be partitioned into *two* classes such that there is no interactions between the factors in the same class but there may be interactions between the factors from different classes. Without loss of generality, we can assume the partition of the factors to be $\{F_1, F_2, \ldots, F_m\}$ and $\{F_{m+1}, \ldots, F_n\}$ with $m \leqslant n - m$. Therefore, the graph G described in the introduction is the direct sum of the two complete graphs with vertices $\{1, 2, \ldots, m\}$ and $\{m+1, m+2, \ldots, n\}$, respectively. Since we assume that interactions which contain an edge of G are zero, it follows that *the only possibly non-zero interactions are the main effects* $1, 2, \ldots, n$ *and* the 2-factor interactions ij, $i = 1, 2, \ldots, m$, $j = m + 1, m + 2, \ldots, n$. We, therefore, need only investigate the G-estimability of these interactions, since zero interactions are G-estimable by the definition of G-estimability.

For different resolutions and different values of fractions of designs the methods of constructing the G-best design are different. We, for convenience, separate the results into three parts: designs of resolution at least 5, designs of resolution at most 4 with $k > 6$, and designs of resolution at most 4 with $k \leqslant 6$. These three parts will be discussed in Sections 3–5, respectively. A useful technical result is the following.

Lemma 2.1 (The 3-factor lemma – TFL). *Let D be a 2^{n-k} ffd with sdr H. If each non-identity interaction in H involves at least three factors from $\{1, 2, \ldots, m\}$ or at least three factors from $\{m + 1, m + 2, \ldots, n\}$, then every interaction is G-estimable under (D, G).*

Proof. Since every non-identity interaction in H involves at least three factors from $\{1, 2, \ldots, m\}$ or at least three factors from $\{m + 1, m + 2, \ldots, n\}$, every interaction in iH, except main effect i, involves either at least two factors from $\{1, 2, \ldots, m\}$ or at least two factors from $\{m + 1, m + 2, \ldots, n\}$; $i = 1, 2, \ldots, n$. Hence every interaction in iH is a zero interaction, except main effect i. Therefore, main effect i is the only interaction in the G-set iH_G. Hence, all the main effects are G-estimable. Similarly, every interaction in ijH, $i = 1, 2, \ldots, m$, $j = m + 1, m + 2, \ldots, n$, is a zero interaction, except the 2-factor interaction ij. Therefore, all the 2-factor interactions ij, $i = 1, 2, \ldots, m$; $j = m + 1, m + 2, \ldots, n$ are G-estimable. This completes the proof of the lemma. □

3. Designs of resolution at least 5

Lemma 3.1. *If a 2^{n-k} design of resolution 5 exists, then the G-best 2^{n-k} ffd is the unconstrained maximum resolution 2^{n-k} design. In this design every interaction is G-estimable.*

Proof. Let D be a maximum resolution 2^{n-k} ffd with sdr H. Then each non-identity interaction in H has order at least 5 since $R(n, k) \geqslant 5$. Hence, each non-identity interaction in H involves either at least three factors from $\{1, 2, \ldots, m\}$ or at least three factors from $\{m + 1, m + 2, \ldots, n\}$. By TFL every interaction is G-estimable under (D, G). Therefore, the design (D, G) is the G-best design since it has maximum resolution. This completes the proof. □

4. Designs of resolution 4

In many practical situations, due to restrictions on cost or other considerations, a design of resolution at most 4 is either what is available or preferred. The rest of the paper explores and completely classifies the G-best designs whose resolution is at most four in the presence of the graph G described above. Construction of designs of high resolution appears in Addelman (1963) and Daniel (1962). Recent studies on aberration are found in Chen and Wu (1991). In this section we give the G-best 2^{n-k} $(k > 6)$ designs of resolution at most 4. We first give a preliminary result:

Lemma 4.1 (Resolution-four lemma – RFL). *If* $(n-k-1)(n-k-2)/2 \geqslant k$, *then* $R(n, k) \geqslant 4$.

Proof. Since $(n-k-1)(n-k-2)/2 \geqslant k$, we can choose k distinct interactions $\{h_1, h_2, \dots, h_k\}$ of order 3 from the set of the 3-factors interactions

$$\{123, 124, \dots, 12(n-k), 134, 135, \dots, 13(n-k), \dots, (n-k-2)(n-k-1)(n-k)\}.$$

Let H be the sdr generated by

$$\{1h_1, 2h_2, \dots, kh_k\}.$$

Then H has resolution 4. Hence $R(n, k) \geqslant 4$. □

Corollary 4.1. *If* $n \geqslant 2k$ *and* $k > 3$, *then* $R(n, k) \geqslant 4$.

Proof. It is easy to see that $n \geqslant 2k$ and $k > 4$ implies that $(n-k-1)(n-k-2)/2 \geqslant k$. Therefore, $R(n, k) \geqslant 4$ for $k > 4$ by RFL. Since $R(8, 4) = 4$, $R(n, k) \geqslant 4$ for $k > 3$. □

4.1. *The case:* $m \geqslant k > 6$

Below we describe the G-best designs and explain why they have this property.

Assume that $k > 6$ and $m \geqslant k$. Let $i + j = k$ such that $i = j$ or $j = i + 1$.

If $n - m > m$, or $m > k$, or k is an even number, then $R(n-m, j) \geqslant 4$ and $R(m, i) \geqslant 4$ for $i > 3$ by RFL. But for $i = 3$, $j = 4$ we have $R(m, i) = R(m, 3) \geqslant R(7, 3) = 4$ and $R(n-m, j) = R(n-m, 4) \geqslant R(8, 4) = 4$. Let H_m be the sdr of a maximum resolution 2^{m-i} unconstrained ffd on $\{1, 2, \dots, m\}$ and H_{n-m} be the sdr of a maximum resolution $2^{(n-m)-j}$ unconstrained ffd on $\{m+1, m+2, \dots, n\}$. Let D be the 2^{n-k} ffd with sdr the direct product of H_m and H_{n-m}. Then the design D has resolution 4 and every interaction in D is G-estimable by TFL, and therefore, the design D is the G-best design for $R(n, k) \leqslant 4$.

If $k = m = n - m = 2l + 1 > 6$, the $l \geqslant 3$. By RFL $R(2l + 2, l + 1) \geqslant 4$ for $l \geqslant 3$. Therefore, $R(n - m, l + 1) = R(2l + 1, l + 1) \geqslant 3$ for $l \geqslant 3$. Let K be the sdr of a maximum resolution $2^{(n-m)-(l+1)}$ unconstrained ffd on factors $\{m+1, m+2, \dots, n\}$ and

$$\{k_1, k_2, \dots, k_{l+1}\}$$

is a set of generators of K. Denote by $k^{(i)}$ the interaction $k_i k_1 k_2 \cdots k_{l+1}$, $i = 1, 2, \ldots, l$. Let D be the $2^{n-(2l-1)}$ ffd with sdr H generated by

$$1k^{(1)}, 2k^{(2)} \cdots lk^{(l)}, (l+1)k_1, (l+2)k_2, \ldots, (2l+1)k_{l+1}.$$

Then D has resolution at least 4 and every interaction in H involves at least three factors from either $\{1, 2, \ldots, m\}$ or $\{m+1, m+2, \ldots, n\}$. Hence, every interaction is G-estimable under (D, G) by TFL, and therefore D is the G-best design.

4.2. *The case: $m < k$ and $k > 6$*

We first prove a lemma.

Lemma 4.2 (Resolution three lemma – RTL). *$R(n, k) \geqslant 3$ if and only if $2^{n-k} > n$, i.e., $n > \log_2(n) + k$.*

Proof. If $2^{n-k} > n$, then $2^{n-k} - (n-k) - 1 \geqslant k$. Therefore, there exist k distinct interactions $h_1, h_2, \ldots, h_k$ on the set of factors $\{k+1, k+2, \ldots, n\}$ with order at least 2. Let H be the sdr generated by

$$\{1h_1, 2h_2, \ldots, kh_k\}.$$

Then H has resolution at least 3. Therefore, $R(n, k) \geqslant 3$.

On the other hand, if $2^{n-k} \leqslant n$, then $2^{n-k} - (n-k) - 1 \leqslant k - 1$. Therefore, there are at most $k-1$ distinct interactions on the factors $\{k+1, k+2, \ldots, n\}$ with order at least 2. Let H_0 be the sdr of a 2^{n-k} design. Without loss of generality, we can assume that H_0 is generated by

$$\{1A_1, 2A_2, \ldots, kA_k\},$$

where A_i's are interactions on the set of factors $\{k+1, k+2, \ldots, n\}$. If any of the A_i's has order less than 2, then H has resolution less than 3. Assume, therefore, that all A_i's have order at least 2. Since there are at most $k-1$ distinct interactions on the factors $\{k+1, k+2, \ldots, n\}$ with order at least 2, at least two of A_i's are the same, say A_1 and A_2. Then 12 is contained in H_0. Hence, H_0 has resolution 2. We have, therefore, proved that $2^{n-k} \leqslant n$ implies that $R(n, k) \leqslant 2$. □

In the remainder of the paper there are many cases to examine and find G-best designs for. Formal *demonstrations are omited* in cases where the design is clearly G-best, either because every interaction is G-estimable, or because of the nature of construction.

4.2.1. *G-Best 2^{n-k} ($m = 1$) designs*

(a) If $R(n, k) \geqslant 4$, then the G-best 2^{n-k} ffd D has maximum resolution. In this design every interaction is G-estimable.

(b) If $R(n, k) = R(n-1, k) = 3$, then the G-best 2^{n-k} ffd D has sdr H which is the same as the sdr of a maximum resolution $2^{(n-1)-k}$ unconstrained ffd on factors $\{2, 3, \dots, n\}$. In this design every interaction is G-estimable.

(c) Assume $R(n, k) = 3$ and $R(n-1, k) = 2$. Let D_3 be any 2^{n-k} ffd with sdr H_3. Denote by $N_3(D_3)$, the number of 3-factor interactions of the form $1ij$ in H. Let D_2 be any $2^{(n-1)-k}$ unconstrained ffd on factors $\{2, 3, \dots, n\}$ with sdr H_2. Denote by $N_2(D_2)$, the number of i's such that i appears in a 2-factor interaction in H_2. Let $N_{3\mathrm{Min}}$ be the minimum number of $N_3(D_3)$ among all 2^{n-k} unconstrained resolution 3 ffds. Let $N_{2\mathrm{Min}}$ be the maximum number of $N_2(D_2)$ among all $2^{(n-1)-k}$ unconstrained resolution 2 ffds on factors $\{2, 3, \dots, n\}$.

If

$$2\,N_{3\mathrm{Min}} \leqslant N_{2\mathrm{Min}},$$

then the G-best 2^{n-k} ffd is the resolution 3 design D whose sdr H contains the minimum number of 3-factor interactions of the form $1ij$ among all 2^{n-k} resolution 3 ffds. The only interactions that are not G-estimable under (D, G) are the main effects $i_1, i_2, \dots, i_t, j_1, j_2, \dots, j_t$ and the 2-factor interactions $1i_1, 1i_2, \dots, 1i_t, 1j_1, 1j_2, \dots, 1j_t$, where $1i_1j_1, 1i_2j_2, \dots, 1i_tj_t$ are the 3-factor interactions of the form $1ij$ in H (note that $\{i_1j_1\}$ does not intersect $\{i_2j_2\}$ since the design D has resolution 3).

If

$$2\,N_{3\mathrm{Min}} > N_{2\mathrm{Min}},$$

then the G-best 2^{n-k} ffd D has sdr H which is the same as the unconstrained resolution two $2^{(n-1)-k}$ ffd on factors $\{2, 3, \dots, n\}$ which has minimum number of i's such that i appears in a 2-factor interaction in H among all unconstrained resolution two $2^{(n-1)-k}$ ffd on factors $\{2, 3, \dots, n\}$. The only interactions that are not G-estimable under (D, G) are the main effects $i_1, i_2, \dots, i_{n_2}$ and the 2-factor interactions $1i_1, 1i_2, \dots, 1i_{n_2}$, where $i_1, i_2, \dots, i_{n_2}$ are the factors appearing in the 2-factor interactions in H.

We omit the case of designs with $R(n, k) = 2$, since the low resolution eliminates their practical usefulness.

4.2.2. G-Best 2^{n-k} ($m = 2$) designs

(a) if $4 \geqslant R(n, k) = R(n-2, k) \geqslant 3$, then the G-best 2^{n-k} ffd is the design D with sdr the same as the sdr of a maximum resolution $2^{(n-2)-k}$ unconstrained ffd on factors $\{3, 4, \dots, n\}$. Under (D, G) every interaction is G-estimable.

(b) Assume $R(n, k) = 4$ and $R(n-2, k) = 3$. Lemma 2.1 yields the following:

If there exists a 2^{n-k} unconstrained resolution 4 ffd D with sdr H such that every interaction in H involves at least three factors from the set $\{3, 4, \dots, n\}$ then the design D is the G-best design. Under (D, G) every interaction is G-estimable.

If for every 2^{n-k} unconstrained resolution 4 ffd the sdr contains an interaction which involves only two factors from the set $\{3, 4, \dots, n\}$, then the G-best ffd D has the sdr the same as the sdr of a $2^{(n-2)-k}$ unconstrained resolution 3 ffd on the factors $\{3, 4, \dots, n\}$. Under (D, G) every interaction is G-estimable.

(c) Assume that $R(n, k) \leqslant 4$, and $R(n-2, k) = 2$. Let D_{n-2} be a $2^{(n-2)-k}$ unconstrained ffd on factors $\{3, 4, \dots, n\}$ with sdr H_{n-2}. Denote by $N(D_{n-2})$, the number of i's such that i appears in a 2-factor interaction in H_{n-2}.

Let D_{n-2} be the resolution two $2^{(n-2)-k}$ unconstrained ffd on factors $\{3, 4, \dots, n\}$, with sdr H_{n-2} such that $N(D_{n-2})$ is minimum among all $2^{(n-2)-k}$ unconstrained resolution two ffds on factors $\{3, 4, \dots, n\}$. The G-best 2^{n-k} ffd is the design D with sdr H generated by

$$\{12h_1, 12h_2, \dots, 12h_r, H_{n-2} \setminus \{h_1, h_2, \dots, h_r\}\},$$

where $h_1, h_2, \dots, h_r$ are the 2-factor interactions in H_{n-2}. Under (D, G) the only interactions which are not G-estimable are the 2-factor interactions $1i_1, 1i_2, \dots, 1i_l$, $2i_1, 2i_2, \dots, 2i_l$, where $i_1, i_2, \dots, i_l$ are the factors appearing in the 2-factor interactions in H_{n-2}.

Indeed, let D_0 be any other 2^{n-2} ffd with sdr H_0 and let $H^{(0)}_{n-2}$ be the projection of H_0 on $\{3, 4, \dots, n\}$. The projection is either injective or has kernel $\{I, 12\}$, since we do not allow a sdr to contain main effects.

Case 1: The projection is injective. Then the resolution of $H^{(0)}_{n-2}$ is at most 2, since $R(n-2, k) = 2$. If the resolution of $H^{(0)}_{n-2}$ is 1, without loss the main effect $3 \in H^{(0)}_{n-2}$. Then one of the interactions 13, 23 and 123 is contained in H_0. In any case, at least one of the main effects 1 or 2 is not G-estimable and therefore, the design D is G-better than D_0. If the resolution of $H^{(0)}_{n-2}$ is 2 and $ij \in H^{(0)}_{n-2}$, then at least one of the interactions ij, $1ij$, $2ij$ and $12ij$ is contained in H_0. If one of ij, $1ij$ and $2ij$ is contained in H_0, then either the main effect i or j is not G-estimable and so the design D is G-better than the design D_0. Therefore, we can assume that the interaction $12ij$ is contained in H_0 for any 2-factor interaction ij which is contained in $H^{(0)}_{n-2}$. Hence $1i$, $1j$, $2i$, $2j$ are not G-estimable for all $ij \in H^{(0)}_{n-2}$. Therefore D_0 is not G-better than D by the construction of D.

Case 2: The kernel of the projection is $\{I, 12\}$. Then the 2-factor interaction 12 is contained in H_0 and hence the main effects 1 and 2 are not G-estimable. Therefore the design D is G-better than the design D_0.

4.2.3. *G-Best 2^{n-k} ($m = 3$) designs*

If $n - 3 > \log_2(n-3) + (k-1)$, then $R(n-3, k-1)$, $\geqslant 3$ by RTL. Let H_0 be the sdr of a maximum resolution $2^{(n-3)+(k-1)}$ unconstrained ffd on $\{4, 5, \dots, n\}$. Let D be the 2^{n-k} design with sdr generated by $\{123, H_0\}$. Then under (D, G) every interaction is G-estimable, and the design is therefore G-best.

4.2.4. *G-Best 2^{n-k} ($m = 4$) designs*

If $n - 4 > \log_2(n-4) + (k-1)$, then $R(n-4, k-1) \geqslant 3$ by RTL. Let H_0 be the sdr of a maximum resolution $2^{(n-4)+(k-1)}$ unconstrained ffd on $\{5, 6, \dots, n\}$. Let D be the 2^{n-k} design with sdr generated by $\{1234, H_0\}$. Then under (D, G) every interaction is G-estimable.

4.2.5. G-Best 2^{n-k} ($m = 5$) designs

If $n - 5 > \log_2(n-5) + (k-2)$, then $R(n-5, k-2) \geq 3$ by RTL. Let H_0 be the sdr of a maximum resolution $2^{(n-5)+(k-2)}$ unconstrained ffd on $\{6, 7, \ldots, n\}$. Let D be the 2^{n-k} design with sdr generated by $\{123, 345, H_0\}$. Then under (D, G) every interaction is G-estimable.

For $6 \leq m < k$ the methods of constructing the G-best designs are essentially the same as the ways we constructed the designs above. For the simplicity of the paper we just briefly describe the idea.

(i) We first try to write $k = i + j$ such that we can construct a resolution 4 (or 3) 2^{m-i} design on $\{1, 2, \ldots, m\}$ and a resolution 4 (or 3) $2^{(n-m)-j}$ design on $\{m+1, m+2, \ldots, n\}$. We then combine them as a resolution 4 (or 3) 2^{n-k} design. The design constructed in this manner has all interactions G-estimable.

(ii) If we cannot construct a design as described in (i), we try to construct a 2^{m-i} design on $\{1, 2, \ldots, m\}$ and a $2^{(n-m)-j}$ design on $\{m+1, m+2, \ldots, n\}$ and then combine the two subgroups of defining relations into a sdr H of a 2^{n-k} design such that the number of interactions which involve at most two factors from $\{1, 2, \ldots, m\}$ and at most two factors from $\{m+1, m+2, \ldots, n\}$ is small.

5. G-Best 2^{n-k} ($k \leq 6$) designs

5.1. G-Best 2^{n-1} designs

The G-best 2^{n-1} ffd has defining relation $I = 12 \cdots n$ in which every interaction is G-estimable for $n > 3$.

5.2. G-Best 2^{n-2} designs

5.2.1. $n > 7$

The G-best 2^{n-2} ($n > 7$) ffd is the maximum resolution design in which every interaction is G-estimable by Section 3.

5.2.2. $n = 7$

The G-best 2^{7-2} ffd is the design D with defining relations $I = 12345 = 4567 = 12367$ in which every interaction is G-estimable.

5.2.3. $n = 6$

If $n = 6$, then $m = 1, 2, 3$.

The G-best 2^{6-2} ($m = 1$) ffd is the design D with defining relations $I = 1234 = 3456 = 1256$. Under (D, G) every interaction is G-estimable.

The G-best 2^{6-2} ($m = 2$) ffd is the design D with defining relations $I = 1234 = 1356 = 2456$. Under (D, G) the only interactions that are not G-estimable are the 2-factor interactions 13, 14, 23, 24.

The G-best 2^{6-2} ($m = 3$) ffd is the design D with defining relations $I = 123 = 456 = 123456$. Under (D, G) every interaction is G-estimable.

Observe that there are interactions which are not G-estimable in any resolution four 2^{6-2} ($m = 3$) design.

5.2.4. $n = 5$

The G-best 2^{5-2} ($m = 1$) ffd is the design D with defining relations $I = 12 = 1345 = 2345$. Under (D, G) the only interactions that are not G-estimable are the main effects 1 and 2 and the 2-factor interaction 12.

The G-best 2^{5-2} ($m = 2$) ffd is the design D with defining relations $I = 13 = 2345 = 1245$. Under (D, G) the G-estimable interactions are the main effects 2, 4, and 5 and the 2-factor interaction 23.

5.3. G-Best 2^{n-3} designs

5.3.1. $n > 9$

The G-best 2^{n-3} ($n > 9$) ffd is the maximum-resolution design in which every interaction is G-estimable by Section 3.

5.3.2. $n = 9$

The G-best 2^{9-3} ($m < 4$) ffd design D has sdr H generated by $\{14567, 2468, 3459\}$. Under (D, G) every interaction is G-estimable.

The G-best 2^{9-3} ($m = 4$) ffd is the design D with sdr

$$H = \{I, 12569, 13469, 2345, 124789, 135789, 23678, 45678\}.$$

Under (D, G) every interaction is G-estimable.

5.3.3. $n = 8$

The G-best 2^{8-3} ($m \leqslant 2$) ffd D has the sdr

$$H = \{I, 14568, 12457, 3467, 2678, 3578, 12356, 2348\}.$$

Under (D, G) every interaction is G-estimable.

The G-best 2^{8-3} ($m = 3$) ffd D has the sdr

$$H = \{I, 12356, 12378, 5678, 468, 1234568, 123467, 457\}.$$

Under (D, G) every interaction is G-estimable.

It can be verified that there are interactions which are not G-estimable in any resolution four 2^{8-3} ($m = 3$) design.

The G-best 2^{8-3} ($m = 4$) ffd is the design D with sdr

$$H = \{I, 1256, 1346, 2345, 12478, 13578, 23678, 45678\}.$$

Under (D, G) the only interactions that are not G-estimable are the 2-factor interactions 15, 16, 25, and 26.

5.3.4. $n = 7$

The G-best 2^{7-3} ffd is the design D with sdr

$$H = \{I, 1456, 2457, 3467, 1267, 1357, 2356, 1234\}.$$

For $m = 1$, every interaction is G-estimable. For $m = 2$ the only interactions which are not G-estimable are the 2-factors interactions ij, $i = 1, 2$; $j = 3, 4, 6, 7$. For $m = 3$, the only interactions which are not G-estimable are the 2-factors interactions ij, $i = 1, 2, 3$; $j = 5, 6, 7$.

5.3.5. $n = 6$

The G-best 2^{6-3} design D has sdr

$$H = \{I, 12, 14, 24, 1356, 2356, 3456, 123456\}.$$

For $m = 1$, the G-estimable interactions are the main effects 3, 5, 6 and the 2-factor interactions 13, 15, 16. For $m = 2$, the G-estimable interactions are the main effects 3, 5, 6. For $m = 3$, the G-estimable interactions are the main effects 3, 5, 6 and the 2-factor interaction 34.

Note that this G-best design has resolution 2. This is the best we can do. A resolution-3 design could be constructed, for example, the design D_0 with sdr

$$H_0 = \{I, 123, 145, 256, 346, 2345, 1356, 1246\}.$$

It is easy to check that for $m = 1$, the G-estimable interactions under (D_0, G) are 1, and 6; for $m = 2$, the only G-estimable interaction under (D_0, G) is 3; for $m = 3$, the G-estimable interactions under (D_0, G) are 1, 2, 3. Hence, D is G-better than D_0.

5.4. G-Best 2^{n-4} designs

Let H be the sdr generated by

$$\{134579, 12467t_0, 2567t_1, 34568\},$$

where t_0, t_1 denote 10, 11. Then it is easy to check that H has resolution 5 and so $R(n, 4) \geqslant 5$ for $n \geqslant 11$. G-best designs are thus given by Section 3.

5.4.1. $n = 10$

Let D be the 2^{10-4} ffd with sdr

$$H = \{I, 1234\} \oplus \{I, 4567, 459t_0, 468t_0, 5689, 4789, 578t_0, 679t_0\}.$$

Then D is the G-best 2^{10-4} $(m = 1, 3, 4)$ design. Under (D, G) every interaction is G-estimable.

The G-best 2^{10-4} ($m = 2, 5$) ffd D has the sdr H generated by

$$\{1356, 2347, 14589, 2359t_0\}.$$

Under (D, G) every interaction is G-estimable.

5.4.2. $n = 9$

The G-best 2^{9-4} ($m = 1, 2$) design D has sdr H generated by

$$\{12346, 12357, 2458, 3459\}.$$

Under (D, G) every interaction is G-estimable.

The G-best 2^{9-4} ($m = 3$) design D has sdr H generated by

$$\{1456, 2678, 3479, 589\}.$$

Under (D, G) every interaction is G-estimable.

The G-best 2^{9-4} ($m = 4$) design D has sdr

$$H = \{I, 123\} \otimes \{I, 3456, 3489, 3579, 4578, 3678, 4679, 5689\}.$$

Under (D, G) the only interactions which are not G-estimable are the 2-factors interactions ij, $i = 3, 4$; $j = 5, 6, 8, 9$.

Note that it can be checked that every resolution 4 design has at least nine 2- factors intersections which are not G-estimable.

5.4.3. $n = 8$

The G-best 2^{8-4} design D has sdr H generated by

$$\{1567, 2568, 3578, 4678\}.$$

For $m = 1$, every interaction is G-estimable. For $m = 2$, the only interactions which are not G-estimable are the 2-factors intersections ij, $i = 1, 2$; $j = 7, 8$. For $m = 3$, the only interactions which are not G-estimable are the 2-factors interactions ij, $i = 1, 2, 3$; $j = 6, 7, 8$. For $m = 4$, the G-estimable non-zero interactions are all the main effects.

5.4.4. $n = 7$

The G-best 2^{7-4} ($m = 1, 2$) design D has sdr H generated by

$$\{12, 13, 14, 567\}.$$

For $m = 1$, the G-estimable interactions are the main effects 5, 6, and 7 and the 2-factor interactions 15, 16, and 17. For $m = 2$, the G-estimable interactions are the main effects 5, 6, and 7.

The G-best 2^{7-4} ($m = 3$) design D has sdr H generated by

$$\{123, 46, 56, 67\}$$

in which the G-estimable interactions are the main effects 1, 2, and 3.

5.5. G-Best 2^{n-5} designs

Let H be the sdr generated by

$$\{12359, 1246t_0, 1347t_1, 2348t_2, 12345678t_3\},$$

where t_0, t_1, t_2, t_3 denote 10, 11, 12, 13. Then it is easy to check that H has resolution 5. Therefore $R(n, 5) \geqslant 5$ for $n \geqslant 13$. Section 3 yields G-best designs in these cases.

5.5.1. $n = 12$

Let H be the sdr generated by

$$\{12359, 1246t_0, 1347t_1, 2348t_2, 12345678\}.$$

Then, it is straightforward to check that H has resolution 4 and every interaction in H involves at least three factors from either $\{1, 2, \ldots, m\}$ or $\{m-1, m+2, \ldots, n\}$ for $m = 1, 2, 3, 4, 5, 6$. Hence, the G-best 2^{12-5} ffd is the design with sdr H in which every interaction is G-estimable.

5.5.2. $n = 11$

(a) $m = 1, 2$. Let H be an unconstrained 2^{10-5} ffd of resolution 4 on the factors $\{2, 3, 4, \ldots, 11\}$. Then the 2^{11-5} design D with the sdr H is G-best for $m = 1, 2$. Under (D, G) every interaction is G-estimable.

(b) $m = 3, 4$. Let H_0 be an unconstrained 2^{8-4} ffd of resolution 4 on the factors $\{4, 5, \ldots, 11\}$. Then the 2^{11-5} design D with the sdr $H = \{I, 1234\} \oplus H_0$ is G-best for $m = 3, 4$. Under (D, G) every interaction is G-estimable.

(c) $m = 5$. The G-best 2^{11-5} $(m = 5)$ ffd has the sdr H generated by

$$\{6789, 89t_0t_1, 168t_0, 2368t_1, 4568t_1\},$$

Under (D, G) every interaction is G-estimable.

5.5.3. $n = 10$

(a) $m = 1, 2$. The G-best 2^{10-5} design D has sdr H generated by

$$\{1678, 27890, 36890, 4679, 5670\}$$

in which every interaction is G-estimable.

(b) $m = 3$. The G-best 2^{10-5} $(m = 3)$ design D has sdr H generated by

$$\{123, 4890, 589, 680, 790\}.$$

Under (D, G) every interaction is G-estimable.

(c) $m = 4$. The G-best 2^{10-5} $(m = 4)$ design D has sdr H generated by

$$\{1234, 3789, 4780, 5790, 6890\}.$$

Under (D, G) the only interactions which are not G-estimable are the 2-factor interactions ij, $i = 1, 2, 3, 4$; $j = 9, 0$.

(d) $m = 5$. The G-best 2^{10-5} ($m = 5$) design D has sdr H generated by

$\{16890, 27890, 3670, 4678, 5679\}$

in which the only interactions which are not G-estimable are the 2-factor interactions ij, $i = 1, 2$, $j = 6, 7$ and ij, $i = 3, 4, 5$; $j = 8, 9, 0$.

5.5.4. $n = 9$

(a) $m = 1$. The G-best 2^{9-5} design D has sdr H generated by

$\{12, 3789, 478, 579, 689\}$

in which the interactions which are not G-estimable are the main effects 1 and 2 and the 2-factor interaction 12.

Note that any resolution 3 design contains at least one 3-factor interaction of form $1ij$, where $i, j > 1$, since $R(8, 5) = 2$ by RTL. Therefore, i, j, $1i$, and $1j$ are not G-estimable.

(b) $m = 2$. The G-best 2^{9-5} design D has sdr H generated by

$\{123, 1245, 1267, 1289, 468\}$.

Under (D, G) the main effects $3, 4, \ldots, 9$ are G-estimable.

(c) $m = 3$. The G-best 2^{9-5} design D has sdr H generated by

$\{12, 13, 345, 567, 789\}$.

Under (D, G) the main effects $4, 5, \ldots, 9$ are G-estimable.

Note that for the resolution 3 design in (b), the only G-estimable main effects are 1, 2 and 3.

(d) $m = 4$. The G-best 2^{9-5} design D has sdr H generated by

$\{12, 13, 14, 567, 789\}$

Under (D, G) the main effects 5, 6, 7, 8, 9 are G-estimable.

Note that for the resolution 3 design in (b), the only G-estimable main effects are 1, 2, and 5.

Since $R(8, 5) = 2$ we omit the cases for $n \leqslant 8$.

5.6. G-Best 2^{n-6} designs

Let H be the sdr generated by

$\{1789t_0, 278t_1t_2, 378t_3t_4, 478t_0t_2t_4, 59t_1t_3t_4, 679t_0t_2t_3\}$,

where t_0, t_1, t_2, t_3, t_4 denote 10, 11, 12, 13, 14. Then it is straightforward to check that H has resolution 5. Therefore, $R(n, 6) \geqslant 5$ for $n \geqslant 14$. Section 3 yields G-best designs in these cases.

5.6.1. $n = 13$

The G-best 2^{13-6} ($m = 1, 3, 4, 5, 6$) ffd is the design with sdr H generated by

$$\{1234, 12579, 1268t_0, 1458t_1, 24578t_2, 1478t_3\}$$

in which every interaction is G-estimable.

The G-best 2^{13-6} ($m = 2$) ffd is the design with sdr H generated by

$$\{1346, 15679, 1236t_0, 1345t_1, 23457t_2, 1347t_3\}$$

in which every interaction is G-estimable.

We remark that in this case, as in several other cases, the number of parameters is less than the number of runs. In such a situation one can consider other criteria for selection of design, such as D-optimality or various other optimality criteria, in addition to our criterion of G-estimability. The case we just considered involves 128 runs and $1 + 13 + 2 \times 11 = 36$ parameters, for example. Clearly, some other designs may be preferable in this situation. Additional issues related to choice of design may lead to a ffd in which all interactions are G-estimable and some other desirable features may be present as well.

5.6.2. $n = 12$

The G-best 2^{12-6} ($m = 1, 2$) ffd is the design with sdr H generated by

$$\{135t_2, 148t_2, 13479, 1346t_0, 167t_1t_2, 23467t_2\}$$

in which every interaction is G-estimable.

The G-best 2^{12-6} ($m = 3, 4$) ffd is the design with sdr H generated by

$$\{12567, 12789, 129t_0t_1, 34567, 349t_0t_1, 68t_0t_2\}$$

in which every interaction is G-estimable.

The G-best 2^{12-6} ($m = 5, 6$) ffd is the design with sdr H generated by

$$\{1235, 1248, 13479, 1346t_0, 1267t_1, 23467t_2\}$$

in which every interaction is G-estimable.

5.6.3. $n = 11$

The G-best 2^{11-6} design D has sdr H generated by

$$\{1236, 1247, 1348, 1259, 135t_0, 145t_1\}$$

For $m = 1$, every interaction is G-estimable. For $m = 2$, all main effects and the 2-factors intersections ij, $i = 1, 2$, $j = 8$, t_0, t_1 are G-estimable. For $m = 3$, all main effects and the 2-factors intersections ij, $i = 1, 2, 3$, $j = 6, t_2$ are G-estimable. For $m = 4$, all main effects and the 2-factors intersections 16, 17, and 18 are G-estimable. For $m = 5$, all main effects and the 2-factors intersections $1j$, $j = 6, 7, \ldots, t_1$ are G-estimable.

5.6.4. $n = 10$

(a) $m = 1, 2, 3$. The G-best 2^{10-6} design D has sdr H generated by

$$\{12, 13, 489t_0, 589, 68t_0, 79t_0\}.$$

For $m = 1$, the only interactions which are not G-estimable are the main effects 1, 2, and 3, and the 2-factors interactions 12 and 13. For $m = 2$ or 3, the main effects $4, 5, \ldots, t_0$ are G-estimable.

Note that for any resolution-3 design there are at least four main effects which are not G-estimable in case of $m = 1$ and there are at least five main effects which are not G-estimable in case of $m = 2$ or 3.

(b) $m = 4$. The G-best 2^{10-6} design D has sdr H generated by

$$\{12, 13, 14, 567, 789, 68t_0\}$$

in which the main effects 5, 6, 7, 8, 9, t_0 are G-estimable.

(c) $m = 5$. The G-best 2^{10-6} design D has sdr H generated by

$$\{12, 13, 14, 15, 678, 89t_0\}$$

in which the main effects 6, 7, 8, 9, t_0 are G-estimable.

Since $R(9, 6) = 2$, a resolution that is too small, we omit the cases for $n \leqslant 9$.

Acknowledgements

The authors thank the referee for a careful reading of the manuscript, and the many helpful comments and suggestions that led to a more focussed presentation, especially for the sporadic cases that appear in Section 5.

References

Addelman, S., 1963. Techniques for constructing fractional replicate plans. J. Amer. Statist. Assoc. 58, 45–71.

Box, G.E.P., Hunter, J.S., 1961a. The 2^{k-p} Fractional factorial designs. Part I. Technometrics 3 (3), 311–351.

Box, G.E.P., Hunter, J.S., 1961b. The 2^{k-p} Fractional factorial designs. Part II. Technometrics 3 (4), 449–458.

Daniel, C., 1962. Sequential fractional replicates in the 2^{p-q} series. J. Amer. Statist. Assoc. 57, 403–429.

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

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

Raktoe, B.L., Hedayat, A., Federer, W.T., 1981. Factorial Designs. Wiley, New York.

Srivastava, J.N., 1990. Modern factorial design theory for experimenters and statisticians. In: Subir Ghosh (Ed.), Statistical Design and Analysis of Experiments. Marcel Dekker, New York, Basel.

Wu, C.F.J., Chen, J., 1991. Some results on s^{n-k} fractional factorial designs with minimum aberration or optimal moments. The Annals of Statistics 2, 1028–1041.