TECHNOMETRICS ©, VOL. 22, NO. 4, NOVEMBER 1980

Minimum Aberration 2^{k-p} Designs

Arthur Fries and William G. Hunter
Department of Statistics
and Engineering Experiment Station
University of Wisconsin–Madison
Madison, WI 53706

For studying k variables in N runs, all 2^{k-p} designs of maximum resolution are not equally good. In this paper the concept of aberration is proposed as a way of selecting the best designs from those with maximum resolution. Algorithms are presented for constructing these minimum aberration designs.

KEY WORDS

Fractional factorial designs
Resolution
Confounding
Aberration

1. INTRODUCTION

Fractional factorial designs—especially the two-level designs—are useful in a variety of experimental situations, for example, (i) screening studies in which only a subset of the variables is expected to be important, (ii) research investigations in which certain interactions are expected to be negligible and (iii) experimental programs in which groups of runs are to be performed sequentially, ambiguities being resolved as the investigation evolves (see Box, Hunter and Hunter, 1978). The literature on fractional factorial designs is extensive. For references before 1969, see the comprehensive bibliography of Herzberg and Cox (1969). For more recent references, see Daniel (1976) and Joiner (1975–79).

A useful concept associated with 2^{k-p} fractional factorial designs is that of resolution (Box and Hunter, 1961). A design is of resolution R if no c-factor effect is confounded with any other effect containing less than $R - c$ factors. For example, a design of resolution III does not confound main effects with one another but does confound main effects with two-factor interactions, and a design of resolution IV does not confound main effects with two-factor interactions but does confound two-factor interactions with one another. The resolution of a two-level fractional factorial design is the length of the shortest word in the defining relation. Usually an experimenter will prefer to use a design which has the highest possible resolution. But for studying variables in N runs, all 2^{k-p} designs which have maximum resolution are not equally good. (Note: $N = 2^{k-p}$.) The purpose of this paper is to provide a method for selecting a best subset of designs from the set of 2^{k-p} fractional factorial designs of highest resolution. "Best" is defined in terms of the concept of aberration.

Received March 1979; revised November 1979

2. AN EXAMPLE

To illustrate the main ideas, let us consider an example. Suppose with a two-level fractional factorial design in $N = 32$ runs a chemist wishes to study the joint effect of $k = 7$ variables such as reaction temperature, pH and concentration on the yield of a particular chemical reaction. If the chemist has prior knowledge concerning the possible importance of certain main effects and interactions, a 2^{7-2} design of resolution III (for example with generating relations **6** = **12** and **7** = **13**) might be better than one of resolution IV. Greenfield (1976, 1978) has solved this type of problem in which all the main effects and interactions to be estimated are specified. Usually such prior knowledge, however, is not sufficiently sharp to warrant such a choice. In this paper, therefore, we consider the more common situation in which prior knowledge is diffuse concerning the possible greater importance of certain specific main effects relative to others, certain specific two-factor interactions relative to others, and similarly for higher-order interactions. We will assume that the experimenter believes initially that main effects are more important than two-factor interactions, that two-factor interactions are more important than three-factor interactions, and so forth. The situation thus defined is one of great practical importance to experimenters. In these circumstances it is sensible to select a design that has maximum resolution.

TABLE 1—*Three choices for a* 2_{IV}^{7-2} *fractional factorial design for seven variables* **1, 2, 3, 4, 5, 6, 7** *in 32 runs. Entries underlined with a tilde (˜) are to be regarded as boldfaced characters.*

Design	(a)	(b)	(c)
Generators	6=123, 7=234	6=123, 7=145	6=1234, 7=1235
Defining relation	I=1236=2347=1467	I=1236=1457=234567	I=12346=12357=4567
Strings of aliased two-factor interactions (assuming three-factor and higher-order interactions are negligible)	12+36	12+36	45+67
	13+26	13+26	46+57
	14+67	14+57	47+56
	17+46	15+47	
	24+37	16+23	
	27+34	17+45	
	16+23+47		

It can be shown that $R = \text{IV}$ is the maximum attainable resolution for a 2^{7-2} design (we say more about this point in Section 4). The chemist, however, will probably regard some 2_{IV}^{7-2} designs as being better than others. For instance, consider designs (a), (b) and (c) in Table 1, all of which are resolution IV. For designs of resolution IV, unconfounded estimates are obtained for all main effects if one assumes that three-factor and higher-order interactions are negligible. If one makes this assumption, Table 1 provides a summary of these designs with regard to confounding among two-factor interactions. Unconfounded estimates are obtained for all two-factor interactions not shown in this table. That is, that part of the abbreviated confounding pattern is presented that is concerned with two-factor interactions for the situation in which it is assumed that three-factor and higher-order interactions are negligible. There is the greatest amount of confounding in design (a) and the least in (c). Therefore, on the basis of this analysis, design (c) is the best of the three.

The three word lengths in the defining relation for design (a) are all four, that is, the word length pattern is {4, 4, 4}. For design (b) it is {4, 4, 6}, and for design (c) it is {4, 5, 5}. Notice the defining relation for design (c) has only one word of length four, whereas (b) has two and (a) has three. Thus, (c) is the design which minimizes the number of words in the defining relation that are of minimum length. We call it a *minimum aberration design.* When comparing two designs using resolution as the criterion, one considers the lengths of the shortest word in each defining relation. If these lengths are equal, the two designs are regarded as being equivalent. With aberration as the criterion, however, one continues to examine the length of the next shortest word in each defining relation until one design is ranked superior to the other. With a resolution R design, main effects are confounded with interactions of order $R - 1$, two-factor interactions are confounded with interactions of order $R - 2$, and so forth. Given that resolution is maximized and equal to $R_{\max}$, minimizing aberration ensures that a design has the minimum number of words of length $R_{\max}$, which, in turn, means that the smallest number of main effects will be confounded with interactions of order $R_{\max} - 1$, the smallest number of two-factor interactions will be confounded with interactions of order $R_{\max} - 2$, and so forth. Hence, the concept of aberration is a natural extension of resolution.

In Section 3 we generalize the ideas presented above, giving a more precise definition of aberration and presenting a practical algorithm for construction of such designs. In Section 4 the problem of estimating bounds for the maximum attainable resolution is discussed.

3. AN ALGORITHM

To generalize the setup illustrated in Table 1, suppose that a 2_R^{k-p} design is constructed by first writing down a full two-level factorial design in $k - p$ factors and then defining the column vectors for p additional factors by associating them with certain interaction columns involving the first $k - p$ factors. Each such assignment results in a word (generator) equal to the identity **I**. For example, for design (a) the generators are **1236** and **2347**. Taking products of the p generators one at a time, two at a time, etc., gives the defining relation, which has $2^p - 1$ words plus **I**. For fixed N and k, and hence p, the problem is to select the best 2_R^{k-p} design. As always, we must carefully consider the question of what we mean by "best".

Suppose two 2^{k-p} designs (s) and (t) of maximum resolution $R_{\max}$ are to be compared, and their defining relations have these word-length patterns:

$$(s)\colon \{R_{\max}^{s_0}(R_{\max}+1)^{s_1}(R_{\max}+2)^{s_2}\cdots(R_{\max}+m)^{s_m}\}$$

$$(t)\colon \{R_{\max}^{t_0}(R_{\max}+1)^{t_1}(R_{\max}+2)^{t_2}\cdots(R_{\max}+n)^{t_n}\}.$$

Determine the first subscript i such that $s_i \neq t_i$. If $s_i < t_i$, then design (s) is the better design; otherwise (t) is the better design. We call designs for fixed N and

k that result from this procedure designs of *minimum aberration.* [Note: Possibly, though very rarely, two designs will have the same word length pattern even though one defining relation will not be a relabeling of the other (see Draper and Mitchell, 1968, 1970). In other words, the word length pattern does not uniquely define a design.]

We will now consider how this principle can be employed in practice to construct useful designs. The National Bureau of Standards tabulation of two-level fractional factorial designs (Connor and Zelen, 1959), which makes use of a similar criterion, indicates in this statement one of the kinds of problems that must be addressed: "Although considerable effort was made to find solutions which have the maximum number of two-factor interactions confounded [only] with three-factor and higher-order interactions, other solutions may exist having a larger number of measurable two-factor interactions." Criteria to be used in a related problem are suggested by Addelman (1969). In that paper it is implicit that consideration of resolution alone may not yield a unique design apart from relabeling.

Direct Approaches

The most direct approach is to write down all possible sets of p nonisomorphic generators (that is, sets of generators which are merely relabelings of others are not considered), calculate the associated word length patterns and select a choice of generators that will give a minimum aberration design. Actually, as shown in Appendix A, we only need to consider generators which contain all k variables. To establish the equivalence of two distinct choices of generators leading to the same word length pattern, it is possible to employ the "letter pattern comparison" test of Draper and Mitchell (1970). Instead of examining the p generators directly, it suffices to consider only the p higher-order interactions assigned to the last p variables. The word length pattern can be constructed by adding one to the length of each of these interactions, and by adding l to the length of each of their products taken l at a time.

An advantage to this direct approach is that it is straightforward and readily programmable. Disadvantages are first that one must check all the nonisomorphic choices of assignments before the best word length pattern can be established, and second that this purely computational procedure yields no insight into the problem. In particular, although counterexamples to conjectures can be discovered via this method, it is impossible to prove any general results.

Let us now consider an alternative direct approach that makes use of results of Brownlee, Kelly and Loraine (1948) and Burton and Connor (1957) (see Appendix B). One can examine all nonisomorphic choices of nonnegative integrals t's satisfying (B3), and then use (B4) to construct the corresponding word lengths and the associated word length patterns. Equations designated with B are in Appendix B. John (1966) has shown that when using (B4) all possible assignments of the t's to specific generators and their products must be considered. Different assignments may result in different word length patterns. This approach is unwieldy and its advantages are not apparent. Its great disadvantage, similar to the difficulty associated with the first approach, is that all nonisomorphic choices of the t's must be considered before the best word length pattern is determined.

A Practical Algorithm

A more fruitful approach is to consider the word length patterns themselves and construct an algorithm as follows. Here we initially assume the maximum resolution $R_{\max}$ to be known and consider only words with lengths not less than $R_{\max}$. (We return to this point in Section 4.) Choose word length patterns satisfying conditions (B1) and (B2). Then use (B5) to solve for $\sum t^2$ and check if this is compatible with (B3). Any word length pattern satisfying all these necessary conditions is a candidate for an actual word length pattern corresponding to a 2_R^{k-p} design. For each set (possibly more than one) of t-values satisfying (B3) and (B5), the generators for the design are computed from (B4). As discussed in the last paragraph, the use of (B4) requires that all possible assignments of the t's to particular generators and their products be considered. Those assignments which result in a word length pattern other than the one with which we began must be discarded. For values of $k \leq 11$ this has presented no great difficulty. However for a general 2_R^{k-p} design, a specific subroutine serving this purpose as well as computer assistance is required.

The biggest advantage to this approach is that one need not check all possible word length patterns. Once an actual word length pattern has been established, by referring to Connor and Zelen (1959), say, or by writing out a set of generators and calculating the associated word length patterns, it suffices to consider only those word length patterns which by our criterion correspond to better designs. Elimination of all such possibilities then proves that the original design is best. If one begins with a design thought to be best, verification via this approach or discovery of a better design will not require a great deal of computational effort. This is, of course, equivalent to beginning at the theoretically best possible word length pattern and examining successively worse patterns until a minimum aberration design is found. The designs displayed in Table 12.15 of Box, Hunter and Hunter (1978) were obtained in this manner.

4. BOUNDS FOR MAXIMUM RESOLUTION

We now consider for given N and k (or equivalently k and p), the problem of determining $R_{\max}$, the maximum possible value of the resolution. This discussion will result in a modification of the last algorithm. For $p = 1$ it is obvious that $R_{\max} = k$. For $p = 2$, Robillard (1968) has shown that $R_{\max} = [2k/3]$, where $[x]$ is the greatest integer not exceeding x. More general bounds for s^{k-p} designs, where s is prime, have been obtained by Fujii (1976). Setting $s = 2$ and assimilating the above results, we have the following statement. The maximum resolution attainable for a 2^{k-p} design satisfies these equalities and inequalities, depending on the value of p:

$$\begin{aligned} R_{\max} &= k && \text{if } p = 1 \\ &= [2k/3] && \text{if } p = 2 \\ &= 2^{p-1}q && \text{if } r = 0, 1 \\ &\le 2^{p-1}q + [2^{p-2}(r-1)/(2^{p-1}-1)] && \\ & && \text{if } r = 2, 3, \ldots, 2^{p-1} - 1 \\ &\le 2^{p-1}q + [r/2] && \text{if } r = 2^{p-1}, \ldots, 2^{p-2} \end{aligned} \tag{1}$$

where $k = q(2^p - 1) + r$, $0 \le r \le 2^p - 2$, q is an integer. (Fujii states that the last inequality of (1) holds for $r = 2^{p-1} + 1, \ldots, 2^p - 2$, which for $p = 2$ becomes $r = 3, \ldots, 2$, which is nonsensical. Although it is known that $R_{\max} = [2k/3]$ when $p = 2$, the above reordering avoids any possible confusion.)

After reading an original draft of this paper, B. H. Margolin suggested that another bound for $R_{\max}$ could probably be obtained by considering the dual problem of the minimum number of observations required for a design to have resolution R. Following this suggestion and generalizing the work of Margolin (1969) and Webb (1968), we obtain the bound

$$R_{\max} \le 1 + 2H + I\left[N \ge \sum_{i=0}^{H} \binom{k}{i} + \binom{k-1}{H}\right], \tag{2}$$

where $N = 2^{k-p}$ is the number of observations, H is the largest integer such that $N \ge \sum_{i=0}^{H} \binom{k}{i}$ and I is the indicator function. A proof of (2) is given in Appendix C.

Table 2 displays a comparison of the $R_{\max}$ bounds from (1) and (2) for various values of k and p. It is seen that (1) is better for small values of p whereas (2) is better for large values of p. This suggests that we adopt the bound

$$R_{\max} \le \text{minimum}[R_{\max} \text{ bound (1)}, R_{\max} \text{ bound (2)}]. \tag{3}$$

Table 2 also compares $R_{\max}$ and the $R_{\max}$ bound obtained from (3). Here the asterisk (*) denotes those values of k and p for which the $R_{\max}$ bound exceeds $R_{\max}$. The table indicates that the bound given by (3) is relatively sharp.

The discussion of the algorithm in Section 3 assumed $R_{\max}$ to be known. This was necessary because a minimal word length was required before possible word length patterns could be investigated. Results (1), (2) and (3) permit the removal of this restriction. Given only k and p, one can determine the best word length pattern as follows. Let B denote the $R_{\max}$ bound. Let $R = B$ and use the algorithm in Section 3 to find, if possible, the best word length pattern. If there is no solution, set $R = B - 1$ and proceed as before. Continue in this manner until the best word length pattern is found. Figure 1 summarizes this procedure.

To illustrate the above procedure we return to the 2_{IV}^{7-2} example. In this case the values of k and p are small so only a few calculations will be required, but in general the computations become lengthy and computer assistance is required. With $k = 7$ and $p = 2$, (1) gives $B = R_{\max} = 4$ and (B1) yields $\sum w = 14$. The only word length patterns satisfying both of these conditions are $\{4, 4, 6\}$ and $\{4, 5, 5\}$. We first examine $\{4, 5, 5\}$ because it is the "best" available word length pattern (see Section 3). Exactly two words have odd length so that (B2) holds. From (B3) and (B5) we calculate $\sum t = 7$ and $\sum t^2 = 17$ from which it follows that the t's must take on the values 2, 2 and 3. Consider the assignment $t(1) = t(2) = 2$, $t(12) = 3$. Employing (B4) we see that the generators have lengths $w(1) = t(1) + t(12) = 5$, $w(2) = t(2) + t(12) = 5$, and their product has length $w(12) = t(1) + t(2) = 4$. The generators, up to a relabeling of factors, are $W(1) = \mathbf{12346}$ and $W(2) = \mathbf{12357}$. These choices satisfy the conditions imposed on their lengths and on the length of their product. The determination of the generators also follows from Figure 2, appropriately modified for the present example. The same generators are obtained using $t(1) = t(12) = 2$ and $t(2) = 3$, the only other possible assignment of the t's.

5. SUMMARY

The statistical literature gives the curious impression that one should only worry about confounding for *fractional* factorials. However, if a *full* 2^k design is performed and the true, but unknown, response function is quadratic, the corresponding quadratic coefficients are confounded with the mean. Whether it is prudent to run a full or fractional design depends on such factors as the object of the investigation, the size of the budget, and plausible forms of the true response function. For a 2^{k-p} design, it is not completely satisfactory to consider the confounding pattern by itself because it ignores second-order

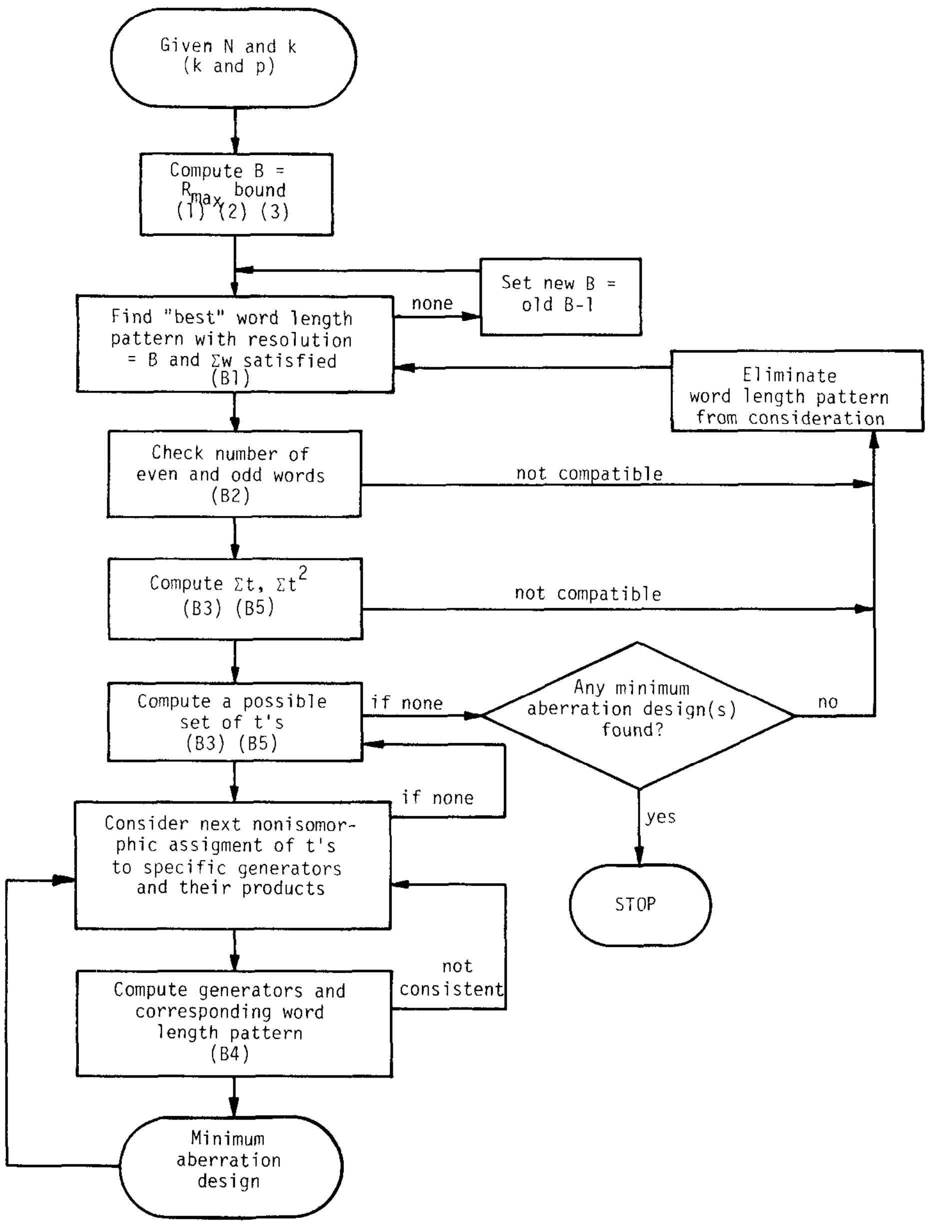
TABLE 2—*A comparison of R_{max} and the R_{max} bounds obtained from equations (1), (2) and (3). The asterisk (*) indicates that R_{max} bound (3) > R_{max}.*

k	p	R_{max} bound (1)	R_{max} bound (2)	R_{max} bound (3)	R_{max}	k	p	R_{max} bound (1)	R_{max} bound (2)	R_{max} bound (3)	R_{max}
5	2	3	3	3	3	12	8	5	3	3	3
6	3	3	3	3	3	12	7	5	4	4	4
6	2	4	4	4	4	12	6	5	4	4	4
7	4	3	3	3	3	12	5	5	5	5	5
7	3	4	4	4	4	12	4	6	6	6	6
7	2	4	5	4	4	12	3	6	8	6	6
8	4	4	4	4	4	12	2	8	9	8	8
8	3	4	4	4	4	13	9	6	3	3	3
8	2	5	5	5	5	13	8	6	4	4	4
9	5	4	3	3	3	13	7	6	4	4	4
9	4	4	4	4	4	13	6	6	5	5	4*
9	3	4	5	4	4	13	5	6	6	6	4*
9	2	6	6	6	6	13	4	6	7	6	6
10	6	4	3	3	3	13	3	6	8	6	6
10	5	4	4	4	4	13	2	8	9	8	8
10	4	5	5	5	4*	14	10	6	3	3	3
10	3	5	5	5	5	14	9	6	4	4	4
10	2	6	6	6	6	14	8	6	4	4	4
11	7	5	3	3	3	14	7	6	5	5	4*
11	6	5	4	4	4	14	6	6	6	6	5*
11	5	5	4	4	4	14	5	6	7	6	5*
11	4	5	6	5	5	14	4	7	8	7	6*
11	3	6	7	6	6	14	3	7	10	7	7
11	2	7	9	7	7	14	2	9	11	9	9

quadratic and higher-order effects. In practice one must view the confounding pattern against the backdrop of many other factors. What is usually done is to concentrate on a careful examination of the confounding pattern itself, while in some fashion keeping in mind these other factors. But it can be tedious to work out the complete confounding patterns for different candidate designs and difficult to compare them in order to choose the "best" one. In many circumstances it is reasonable to consider abbreviated rather than complete confounding patterns because this saves time. Moreover it may be more sensible anyway because it is not logical in general to be worrying about the possible influence of kth, $(k-1)$th, ... order interactions while ignoring second-order quadratic effects. By examining abbreviated confounding patterns, one essentially eliminates from consideration the high-order interactions. Taking this process of simplifying the selections of 2^{k-p} designs to a useful limit, one arrives at a concept like resolution, which does not require the writing out of even the abbreviated confounding patterns but essentially only the defining relations themselves.

The point raised in the present paper is that using resolution alone may short circuit the selection process too severely in some situations because all designs of the same resolution are not equally good. The concept of aberration allows one to more closely approximate what should ideally be done in comparing 2^{k-p} designs, with not much more work than considering resolution alone. As discussed in the body of this paper, the concept of aberration is a natural extension of resolution. We have presented three algorithms for generating minimum aberration designs. It is not clear which, if any, can be extended to s^{k-p} designs, where s is a prime number greater than two.

FIGURE 1. *Summary of algorithm for construction of minimum aberration 2^{k-p} fractional factorial designs (appropriate equation numbers are given in parentheses).*



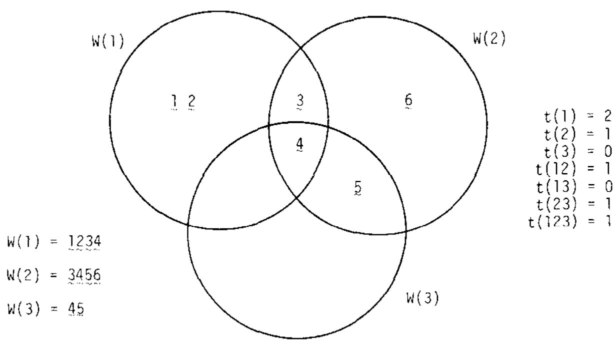


FIGURE 2. An illustration of the definition of $t(i_1 \cdots i_s)$. Entries underlined with a tilde (˜) are to be regarded as boldface characters.

6. ACKNOWLEDGMENTS

We are grateful to the referees for their careful reading of this manuscript and especially for their help in clarifying the explanation of certain detailed points concerning the algorithm for the construction of minimum aberration designs. We are also indebted to B. H. Margolin for his suggestion, mentioned in Section 4, which resulted in an improved bound for R_{max}.

APPENDIX A

Lemma: The best word length pattern corresponds to a defining relation in which all k factors are present, that is, all of the first $k - p$ factors are used in the assignment of higher-order interactions to the last p factors.

Proof: Given k and p (we need not be given R), we want to show that any word length pattern that does not use all of the first $k - p$ factors in the assignment of higher-order interactions to the last p factors is inferior to some word length pattern that uses all of the factors. Let the first word length pattern under consideration be denoted by $R^{a_0}(R + 1)^{a_1}(R + 2)^{a_2} \cdots (R + m)^{a_m}$, where $a_0 > 0$, $a_i \geq 0$ for $1 \leq i \leq m$. Add exactly one of the unassigned factors to a generator and examine the resulting word length pattern. The length of that generator will increase by one. Also, since this added factor appears only in one generator, cancellation because of multiplication is not possible and all products including that generator will have the word length increased by one. Thus the resulting word length pattern is of the form $R^{b_0}(R + 1)^{b_1} \cdots (R + m)^{b_m}(R + m + 1)^{b_{m+1}}$ where $b_0 \leq a_0$, $b_i \geq 0$ for $1 \leq i \leq m + 1$. If $b_0 < a_0$ the second pattern is better. If $b_0 = a_0$ we have $b_0 = a_0, b_1 = a_1, \ldots, b_j = a_j, b_{j+1} < a_{j+1}$ for some j satisfying $1 \leq j \leq m$, from which we conclude that the second pattern is better. Continuing to add the unassigned factors one at a time to the generators completes the proof.

APPENDIX B

The following theorems hold if all k variables appear in the defining relation.

Theorem B1: Let $w_1, \ldots, w_m$ denote the lengths of $m = 2^p - 1$ words. Then the following are two necessary conditions that these words correspond to a defining relation of a 2^{k-p} design:

$$\sum_{i=1}^{m} w_i = 2^{p-1}k, \tag{B1}$$

Either the w's all are even or exactly 2^{p-1} of them are odd. (B2)

The conditions of Theorem B1 do not require that the lengths of particular generators and their products be specified. Before we state necessary and sufficient conditions for the existence of a defining relation for which the lengths of the generators and their products are given we must introduce some notation.

Let $i_1, \ldots, i_s$ be s integers such that $0 < i_1 < \cdots < i_s < p + 1$. Denote the product of the i_1th, i_2th, ... and i_sth generator by $W(i_1 \cdots i_s)$ and let the length of the word $W(i_1 \cdots i_s)$ be denoted by $w(i_1 \cdots i_s)$. Then $W(i)$ is the ith generator and there are exactly $2^p - 1$ words corresponding to the collection of $2^p - 1$ symbols $(i_1, \ldots, i_s)$ which we will denote by S. Also introduce the symbols $0 = 0(i_1, \ldots, i_s)$ and $E = E(i_1, \ldots, i_s)$ to denote the collection of symbols which contain an odd number of indices from $(i_1, \ldots, i_s)$, and the collection of symbols which contain none or an even number of indices from $(i_1, \ldots, i_s)$, respectively. Letting $n(0)$ and $n(E)$ denote the number of symbols in 0 and E it is easily seen that $n(0) = 2^{p-1}$ and $n(E) = 2^{p-1} - 1$.

Finally let $t = t(i_1 \cdots i_s)$ denote the number of letters which occur in all of the s generators $W(i_1), \ldots, W(i_s)$, but not in the remaining generators. Figure 2 illustrates this definition.

The t's are the number of letters in the basic disjoint sets from which it follows that

$$\sum_{S} t(i_1 \cdots i_s) = k. \tag{B3}$$

Furthermore any set of t's which are positive integers or zero and satisfy (B3) corresponds to a constructible defining relation involving k factors. These results are also stated within the following theorems from Burton and Connor (1957).

Theorem B2: If the w's satisfy the $2^p - 1$ equations

$$\sum_{0} w(j_1 \cdots j_p) - \sum_{E} w(j_1 \cdots j_p) = 2^{p-1}t(i_1 \cdots i_s),$$

where $\sum_0 t + \sum_E t = k$ in the sense of implying t's which are nonnegative integers, then the defining relation exists. Furthermore, from the $2^p - 1$ equations

$$\sum_{0(i_1 \cdots i_s)} t(j_1 \cdots j_p) = w(i_1 \cdots i_s) \tag{B4}$$

it follows that the t's are sufficient to construct the defining relation, and the defining relation corresponding to a set of t's is unique (apart from relabeling).

Theorem B3: A necessary condition for the defining relation to exist is that there are k or fewer positive integers whose sum is k and whose squares add to $2^{-p+2} \sum w^2 - k^2$, that is

$$\sum w^2 = 2^{p-2}(\sum t^2 + k^2). \tag{B5}$$

APPENDIX C

Lemma: A 2^{k-p} design of even resolution $R = 2l$ must contain at least $\sum_{i=0}^{l-1} \binom{k}{i} + \binom{k-1}{l-1}$ runs.

(Margolin (1969) and Webb (1968) independently established this result for $l = 2$. Below we modify Webb's proof only slightly to obtain the above generalization.)

Proof: Let $\mathbf{X}_1^{(1)}, \ldots, \mathbf{X}_k^{(1)}$ be the column vectors associated with the k main effects, let $\mathbf{X}_1^{(2)}, \ldots, \mathbf{X}_{\binom{k}{2}}^{(2)}$ be the column vectors associated with the $\binom{k}{2}$ two-factor interactions, ..., let $\mathbf{X}_1^{(l-1)}, \ldots, \mathbf{X}_{\binom{k}{l-1}}^{(l-1)}$ be the column vectors associated with the $\binom{k}{l-1}(l - 1)$th-factor interactions, and define the matrix $\mathbf{X}$ by

$$\mathbf{X} = \left[\mathbf{X}_1^{(1)}, \ldots, \mathbf{X}_k^{(1)}, \mathbf{X}_1^{(2)}, \ldots, \mathbf{X}_{\binom{k}{2}}^{(2)}, \ldots, \mathbf{X}_1^{(l-1)}, \ldots, \mathbf{X}_{\binom{k}{l-1}}^{(l-1)}\right].$$

Let $\mathbf{Z}_0, \mathbf{Z}_1, \ldots, \mathbf{Z}_{\binom{k-1}{l-1}}$ represent the column vectors associated with the grand mean and the interactions of the first factor with the $\binom{k-1}{l-1}$ $(l-1)$th-factor interactions which do not include the first factor. Define the matrix $\mathbf{Z}$ by $\mathbf{Z} = [\mathbf{Z}_0, \mathbf{Z}_1, \ldots, \mathbf{Z}_{\binom{k-1}{l-1}}]$. The columns of the matrix $[\mathbf{X}, \mathbf{Z}]$ are linearly independent by the orthogonality of the vectors.

Now suppose the design contains $N < \sum_{i=0}^{l-1} \binom{k}{i} + \binom{k-1}{l-1}$ runs. Select any vectors $\mathbf{W}_1, \ldots, \mathbf{W}_{N - \sum_{i=1}^{l-1} \binom{k}{i}}$ such that the matrix $[\mathbf{X}, \mathbf{W}]$ is of rank N. Here we have defined $\mathbf{W} = [\mathbf{W}_1, \ldots, \mathbf{W}_{N - \sum_{i=1}^{l-1} \binom{k}{i}}]$ and implicitly used $N \geq \sum_{i=0}^{l-1} \binom{k}{i}$. Since $[\mathbf{X}, \mathbf{W}]$ is of full rank there exist matrices $\mathbf{H}_1$ and $\mathbf{H}_2$ such that $\mathbf{Z} = \mathbf{X}\mathbf{H}_1 + \mathbf{W}\mathbf{H}_2$. We next show $\mathbf{H}_1 = \mathbf{0}$. Write the model for the design as $E\mathbf{Y} = \mathbf{X}\mathbf{B}_1 + \mathbf{W}\mathbf{B}_2 + \mathbf{Z}\mathbf{B}_3$ where $\mathbf{Y}$ is the $N \times 1$ vector of responses, $\mathbf{B}_1$ is the $(\sum_{i=1}^{l-1} \binom{k}{i}) \times 1$ vector of parameters to be estimated, and $\mathbf{B}_2$ and $\mathbf{B}_3$ are respectively $(N - \sum_{i=0}^{l-1} \binom{k}{i}) \times 1$ and $(\binom{k-1}{l-1} + 1) \times 1$ vectors. Since $E\hat{\mathbf{B}}_1 = \mathbf{B}_1$ and $\hat{\mathbf{B}}_1 = \mathbf{J}\mathbf{Y}$ for some matrix $\mathbf{J}$ we have $E\mathbf{J}\mathbf{Y} = \mathbf{J}(\mathbf{X}\mathbf{B}_1 + \mathbf{W}\mathbf{B}_2 + \mathbf{Z}\mathbf{B}_3) = \mathbf{B}_1$. It follows that $\mathbf{J}\mathbf{X} = \mathbf{I}$, $\mathbf{J}\mathbf{W} = \mathbf{0}$ and $\mathbf{J}\mathbf{Z} = \mathbf{0}$. But

$$\begin{aligned} \mathbf{JZ} &= \mathbf{J}(\mathbf{XH}_1 + \mathbf{WH}_2) \\ &= \mathbf{JXH}_1 + \mathbf{JWH}_2 = \mathbf{IH}_1 + \mathbf{OH}_2 = \mathbf{H}_1 \end{aligned}$$

so that we must have $\mathbf{H}_1 = \mathbf{0}$ and $\mathbf{Z} = \mathbf{W}\mathbf{H}_2$. This contradicts the assumption $N < \sum_{i=0}^{l-1} \binom{k}{i} + \binom{k-1}{l-1}$ since there are more linearly independent $\mathbf{Z}$'s than there are linearly independent $\mathbf{W}$'s.

Theorem: Let $N = 2^{k-p}$, let H be the largest integer such that $N \geq \sum_{i=0}^{H} \binom{k}{i}$ and let I be the indicator function. Then

$$R_{\max} \leq 1 + 2H + I\left[N \geq \sum_{i=0}^{H} \binom{k}{i} + \binom{k-1}{H}\right]. \quad \text{(C1)}$$

Proof: Suppose $R_{\max} = 1 + 2l$ is odd. Then the parameters estimated include the grand mean, the k main effects, the $\binom{k}{2}$ two-factor interaction, ..., and the $\binom{k}{l}$ l-factor interactions. This requires a minimum of $1 + \binom{k}{1} + \binom{k}{2} + \cdots + \binom{k}{l} = \sum_{i=0}^{l} \binom{k}{i} \leq N$ runs. Then $l \leq H$ and $R_{\max} = 1 + 2l \leq 1 + 2H$ so that (C1) holds. Now suppose $R_{\max} = 2l$ is even. From the definition of H we have $H \geq l - 1$. If $H = l - 1$, then $R_{\max} = 2l = 2 + 2H$ and (C1) is satisfied since $N \geq \sum_{i=0}^{l-1} \binom{k}{i} + \binom{k-1}{l-1}$ by the preceding lemma. If $H \geq l$ then $R_{\max} = 2l \leq 2H$ and (C1) holds.

REFERENCES

ADDELMAN, S. (1969). Sequences of two-level fractional factorial designs. *Technometrics, 11*, 477–509.

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

BOX, G. E. P., HUNTER, W. G. and HUNTER, J. S. (1978). *Statistics for Experimenters.* New York: John Wiley and Sons, Inc.

BROWNLEE, K. A., KELLY, B. K. and LORAINE, P. K. (1948). Fractional replication arrangements for factorial experiments with factors at two levels. *Biometrika, 35*, 268–276.

BURTON, R. C. and CONNOR, W. S. (1957). On the identity relationship for fractional replicates of the 2^n series. *Ann. Math. Statist., 28*, 762–767.

CONNOR, W. S. and ZELEN, M. (1959). Fractional factorial experimental designs for factors at two levels. *National Bureau of Standards Applied Mathematics Series, 48.*

DANIEL, C. (1976). *Applications of Statistics to Industrial Experimentation.* New York: John Wiley and Sons, Inc.

DRAPER, N. R. and MITCHELL, T. J. (1968). Construction of the set of 256-run designs of resolution ≥ 5 and the set of even 512-run designs of resolution ≥ 6 with special reference to the unique saturated designs. *Ann. Math. Statist., 39*, 246–255.

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

FUJII, Y. (1976). An upper bound of resolution in symmetrical fractional factorial designs. *Ann. Statist., 4*, 662–667.

GREENFIELD, A. A. (1976). Selection of defining contrasts in two-level experiments. *Applied Statistics, 25*, 64–67.

GREENFIELD, A. A. (1978). Selection of defining contrasts in two-level experiments—a modification. *Applied Statistics, 27*, 78.

HERZBERG, A. M. and COX, D. R. (1969). Recent work on the design of experiments. *J. Roy. Statist. Soc., A-132*, 29–67.

JOHN, P. W. M. (1966). On identity relationships for 2^{n-r} designs having words of equal length. *Ann. Stat., 37*, 1842–1843.

JOINER, B. L. (ed.) (1975–79). *Current Index to Statistics.* Washington, D.C.: American Statistical Association.

MARGOLIN, B. H. (1969). Resolution IV fractional factorial designs. *J. Roy. Statist. Soc., B-31*, 514–523.

ROBILLARD, P. (1968). Combinatorial problems in the theory of factorial designs and error correcting codes. *Instit. Statist. Mimeo Ser. 594.*, University of North Carolina, Chapel Hill.

WEBB, S. R. (1968). Non-orthogonal designs of even resolution. *Technometrics, 10*, 291–299.