MRC Technical Summary Report #2112

TWO-LEVEL MULTIFACTOR EXPERIMENT DESIGNS FOR DETECTING THE PRESENCE OF INTERACTIONS

Max D. Morris and Toby J. Mitchell

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

August 1980

(Received January 21, 1980)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

Department of Energy
Washington, D.C. 20545

(14) MRC-TSR-2112

(15) DAAG29-75-C-0024
DAAG29-80-C-0041

UNIVERSITY OF WISCONSIN-MADISON
MATHEMATICS RESEARCH CENTER

(6)

TWO-LEVEL MULTIFACTOR EXPERIMENT DESIGNS FOR DETECTING THE PRESENCE OF INTERACTIONS

(11) August 1980 (10) Max D. Morris* and Toby J. Mitchell**

(9) Technical Summary Report # 2112

(12) 41

August 1980

ABSTRACT

A design optimality criterion (tr(L) - optimality) is applied to the problem of designing two-level multifactor experiments to detect the presence of interactions among the controlled variables. Rules are given for constructing tr(L) - optimal foldover designs and tr(L) - optimal fractional factorial designs. Some results are given on the power of these designs for testing the hypothesis that there are no two-factor interactions. Modifications of the tr(L) - optimal designs to satisfy other experimental objectives (estimability of effects, detection of the presence of other nonlinear effects, estimation of the error variance) are suggested. Examples are given to demonstrate the application of these designs to (i) screening for interactions, and (ii) evaluating the first-order assumption in the sensitivity analysis of a computer code.

AMS (MOS) Subject Classifications: 62K05, 62K15.

Key Words: Design; experiment design; foldover design; fractional factorial design; interaction; main effects design; minimum aberration; optimal design; orthogonal array; Resolution IV design; screening design; sensitivity analysis; two-level design; tr(L) -optimality.

Work Unit Number 4 - Statistics and Probability.

*Department of Statistics, Mississippi State University.

**Mathematics Research Center, University of Wisconsin-Madison. Currently under assignment from the Computer Sciences Division, Union Carbide Corporation, Nuclear Division, Oak Ridge, Tennessee.

Sponsored by the United States Army under Contract Nos. DAAG29-75-C-0024 and DAAG29-80-C-0041 and the Applied Mathematical Sciences Research Program, Office of Energy Research, U.S. Department of Energy under contract W-7405-eng-26 with the Union Carbide Corporation.

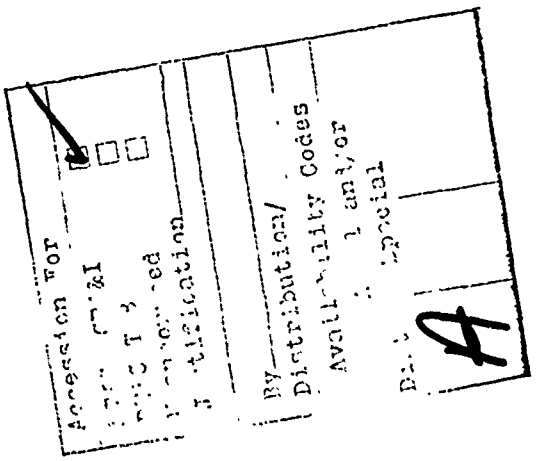
A 231100

SIGNIFICANCE AND EXPLANATION

Experiment designs (plans) in which each controlled variable (factor) can take one of two specified levels are frequently used in the initial stages of an experimental investigation, when the objective is to determine which factors are important and how they interact. This paper is concerned with the design of such two-level experiments to detect the presence of interactions. It is shown here how this can be done with very few experimental trials (runs), even when the number of factors under investigation is large.

Our approach is to define an "optimality" criterion which represents the ability of a design to detect the presence of interactions. This is based on the trace of a matrix that depends on the design. We then determine rules for constructing designs that maximize this quantity, over two broad classes of two-level experimental designs. The main drawback, from a practical point of view, is that the focus on a single criterion may result in the neglect of other design objectives. We therefore present some modifications which improve the designs with respect to other criteria.

The designs presented here have application to any experimental situation in which there is some doubt as to whether the controlled variables act independently on the response or whether they interact. A particular application, which is demonstrated in an example in the paper, is to the sensitivity analysis of computer codes which are used to model physical or economic systems.



The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the authors of this report.

TWO-LEVEL MULTIFACTOR EXPERIMENT DESIGNS FOR DETECTING THE PRESENCE OF INTERACTIONS

Max D. Morris* and Toby J. Mitchell**

1. Introduction

1.1. The Problem

We consider here the problem of designing two-level experiments to detect the presence of interactions among k experimentally controlled variables (factors) $X_1, X_2, \ldots, X_k$, with respect to their effect on the expectation η of a randomly distributed response variable y_x. There are no interactions over a specified region of interest χ if and only if η can be expressed as

$$\eta = \eta(X) = f_1(X_1) + f_2(X_2) + \ldots + f_k(X_k) \tag{1.1}$$

for suitably chosen functions $f_i(X_i)$ when $X = (X_1, X_2, \ldots, X_k) \in \chi$.

We shall restrict attention to just two levels of each factor: X_i^- and X_i^+, chosen so that all 2^k combinations of levels are in χ, and we shall define the "coded" factors $x_1, x_2, \ldots, x_k$ so that $x_i = -1$ when $X_i = X_i^-$ and $x_i = +1$ when $X_i = X_i^+$. We shall denote the 2^k corners of the cube $x_i = \pm 1$ by K. In terms of the coded factors, (1.1) becomes

$$\eta(x) = \beta_0 + \sum_{i=1}^{k} \beta_i x_i \tag{1.2}$$

---

*Department of Statistics, Mississippi State University.

**Mathematics Research Center, University of Wisconsin-Madison. Currently under assignment from the Computer Sciences Division, Union Carbide Corporation, Nuclear Division, Oak Ridge, Tennessee.

Sponsored by the United States Army under Contract Nos. DAAG29-75-C-0024 and DAAG29-80-C-0041 and the Applied Mathematical Sciences Research Program, Office of Energy Research, U.S. Department of Energy under contract W-7405-eng-26 with the Union Carbide Corporation.

for $x = (x_1, x_2, \ldots, x_k) \in K$, where $\beta_i = (f_i(X_i^+) - f_i(X_i^-))/2$, $i = 1,2,\ldots,k$, and $\beta_0 = \frac{1}{2} \sum_{i=1}^{k} (f_i(X_i^+) + f_i(X_i^-))$.

A simple departure from (1.1) allows pairwise interactions, i.e., η can be expressed as

$$\eta(x) = \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} f_{ij}(X_i, X_j) \quad . \tag{1.3}$$

In terms of the coded factors, (1.3) can be written

$$\eta(x) = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \beta_{ij} x_i x_j \tag{1.4}$$

for $x \in K$. This is a conventional model for the analysis of 2-level factorial experiments (possibly incomplete) in which interactions among three or more factors are assumed to be zero. We shall use it here in a somewhat different way, namely, as a device for the planning of experiments to indicate whether or not the factors affect the response independently, i.e., as in (1.1). Our approach will be to assume that (1.4) holds on K, then to use a design optimality criterion to construct designs which will be good for detecting the presence of non-zero β_{ij}'s. Even if higher order interactions are present, we would expect this approach to work, since (1.4) will be a better approximation to the true response over K than will (1.2). We should also emphasize that we do not expect nor require (1.4) to hold outside of K. (If η is a quadratic polynomial over some continuous region of interest which contains K, for example, (1.4) holds on K but not everywhere in the region.)

In this paper, we shall refer to the β_{ij}'s in (1.4) as "interactions" and the β_i's ($i \neq 0$) as "main effects". Except for a factor of 2, these are the same as the main effects and interactions conventionally defined for a two-level factorial experiment (Box and Hunter (1961)).

Good designs (e.g., the Resolution V fractional factorials) exist for estimating the main effects and interactions in the model (1.4). However, the number of runs required is at least $(k^2 + k + 2)/2$, and may be considerably greater than that if a regular fractional factorial design is used.

In this paper, we consider a less ambitious experimental goal, namely to determine whether or not significant interactions are present. Such information, obtained early in an investigation, can be useful in planning subsequent experiments. Initially, we shall ignore other considerations, such as estimability of the main effects and interactions and estimability of the error variance σ^2. These will be discussed in Section 5.

1.2. A Design Criterion: $\mathrm{tr}(\underset{\sim}{L})$ -optimality

In matrix notation, our model for a vector $\underset{\sim}{Y}$ of n observations, based on (1.4), is:

$$\underset{\sim}{Y} = \underset{\sim}{X}_1\underset{\sim}{\beta}_1 + \underset{\sim}{X}_2\underset{\sim}{\beta}_2 + \underset{\sim}{\varepsilon}, \quad E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}, \quad V(\underset{\sim}{\varepsilon}) = \underset{\sim}{I}\sigma^2 \tag{1.5}$$

where $\underset{\sim}{\beta}_1$ is the $(k+1)$-vector $(\beta_0,\beta_1,\ldots,\beta_k)'$ and $\underset{\sim}{\beta}_2$ is the k_2-vector $(\beta_{11},\beta_{12},\ldots,\beta_{k-1,k})'$ of interactions, with $k_2 = k(k-1)/2$. The matrices $\underset{\sim}{X}_1$ and $\underset{\sim}{X}_2$ depend through (1.4) on the $n \times k$ design matrix $\underset{\sim}{D}$, whose uth row is $(x_{1u},x_{2u},\ldots,x_{ku})$.

When the model (1.5) is fitted using the ordinary least squares criterion under the restriction that $\underset{\sim}{\beta}_2 = \underset{\sim}{0}$, the expected residual sum of squares is

$$\mathrm{ERSS} = (n-r(\underset{\sim}{X}_1))\sigma^2 + \underset{\sim}{\beta}_2'\underset{\sim}{L}\underset{\sim}{\beta}_2 \tag{1.6}$$

where $r(\underset{\sim}{X}_1)$ is the rank of $\underset{\sim}{X}_1$ and the lack-of-fit matrix $\underset{\sim}{L}$ is

$$\underset{\sim}{L} = \underset{\sim}{X}_2'(\underset{\sim}{I} - \underset{\sim}{X}_1(\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-}\underset{\sim}{X}_1')\underset{\sim}{X}_2 \quad . \tag{1.7}$$

We shall not require $\underset{\sim}{X}_1'\underset{\sim}{X}_1$ to be nonsingular, hence the use of the generalized inverse $(\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-}$ in (1.7).

Atkinson's (1972) general approach to the problem of detecting inadequacy of the model $E(\underset{\sim}{Y}) = \underset{\sim}{X}_1\underset{\sim}{\beta}_1$ was to select the design so as to maximize the determinant of $\underset{\sim}{L}$, or equivalently, to minimize the generalized variance of the least squares estimator of $\underset{\sim}{\beta}_2$.

This criterion can be applied only within the class of designs for which $\underset{\sim}{\beta}_2$ is estimable. Our approach here will be more closely related to the work of Atkinson and Fedorov (1975), whose T-optimality design criterion reduces to the maximization of $\lambda = \underset{\sim}{\beta}_2'\underset{\sim}{L}\underset{\sim}{\beta}_2$. This criterion, however, depends upon the value of $\underset{\sim}{\beta}_2$, which is unknown. Jones and Mitchell (1978) avoided this difficulty by utilizing the relationship between λ and the positive definite quadratic form $\tau = \underset{\sim}{\beta}_2'\underset{\sim}{T}\underset{\sim}{\beta}_2$, which, with proper choice of $\underset{\sim}{T}$, can be interpreted as a measure of the importance of the interaction terms. One of their criteria (A_2-optimality) requires maximizing the average value of λ (over $\underset{\sim}{\beta}_2$) for constant τ; this is equivalent to maximizing $\mathrm{tr}(\underset{\sim}{T}^{-1}\underset{\sim}{L})$. In the present setting, the Jones-Mitchell T-matrix can be shown to be the identity $\underset{\sim}{I}$, so A_2-optimization becomes maximization of the trace of $\underset{\sim}{L}$. In Appendix A, we show that the $\mathrm{tr}(\underset{\sim}{L})$ criterion can also be derived by maximizing the expectation of λ (no matter what the value of $\underset{\sim}{\beta}_2$ is) under random assignment of factor labels and factor level labels. This is the criterion we shall adopt in this paper.

1.3. Conventions and Notation.

Throughout this paper, the word "design" refers to a two-level design, except for a brief discussion of "center points" in Section 5.3. When we wish to indicate also the number of runs (n) and the number of factors (k), we shall write "(n,k)-design".

The following is a selected list of letters and symbols used in the text.

- k: Number of factors.
- k_2: Number of two-factor interactions = k(k-1)/2.
- K: Set of 2^k possible combinations of levels of the coded factors $x_1, x_2, \ldots, x_k$, where $x_i = \pm 1$.
- n: Number of runs in a design.
- $\tilde{n}$: Number of runs in the "half-design" used to construct a foldover design.
- a: Integer value of $k/\tilde{n}$.
- r: Remainder upon dividing k by $\tilde{n}$.

$\underset{\sim}{D}$: Design matrix.

$\underset{\sim}{\tilde{D}}$: Design matrix for the "half-design".

$\underset{\sim}{\beta}_1$: Vector of coefficients for the first-order model $(\beta_0,\beta_1,\ldots,\beta_k)$.

$\underset{\sim}{\beta}_2$: Vector of interaction coefficients $(\beta_{11},\beta_{12},\ldots,\beta_{k-1,k})$.

$\underset{\sim}{X}_1,\underset{\sim}{X}_2$: Matrices of known constants in the model $E(\underset{\sim}{y}) = \underset{\sim}{X}_1\underset{\sim}{\beta}_1 + \underset{\sim}{X}_2\underset{\sim}{\beta}_2$.

$\underset{\sim}{L}$: Lack of fit matrix $\underset{\sim}{L} = \underset{\sim}{X}_2'\underset{\sim}{X}_2 - \underset{\sim}{X}_2'\underset{\sim}{X}_1(\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-}\underset{\sim}{X}_1'\underset{\sim}{X}_2$.

σ^2: Common variance of the individual observations $(y_i$'s).

n_i: Number of words of length i in the defining relation of a fractional factorial design.

q: Number of strings of aliased two-factor interactions (not counting the string that is confounded with the overall mean) in a fractional factorial design.

w: Likelihood-ratio statistic for testing the hypothesis that $\underset{\sim}{\beta}_2 = \underset{\sim}{0}$.

2. The Structure of $\mathrm{Tr}(\underset{\sim}{L})$-Optimal (n,k)-Designs

2.1. Orthogonal Arrays of Strength 3

It is clear from (1.7) that $\mathrm{tr}(\underset{\sim}{L})$ cannot exceed $\mathrm{tr}(\underset{\sim}{X}_2'\underset{\sim}{X}_2) = nk_2$, and that this upper bound is attained exactly if and only if $\underset{\sim}{X}_2'\underset{\sim}{X}_1 = 0$. This condition can occur if and only if all design moments of form [i], [ij], and [ijℓ] are zero, where

$$[i] = \frac{1}{n}\sum_{u=1}^{n} x_{iu}, \quad [ij] = \frac{1}{n}\sum_{u=1}^{n} x_{iu}x_{ju}, \quad [ij\ell] = \frac{1}{n}\sum_{u=1}^{n} x_{iu}x_{ju}x_{\ell u} \ . \tag{2.1}$$

This in turn can hold if and only if every subset of three columns of the design matrix $\underset{\sim}{D}$ forms a complete 2^3 factorial design (possibly replicated), i.e., $\underset{\sim}{D}$ is an orthogonal array of strength 3. We therefore have the following theorem:

Theorem 2.1. If n is a multiple of 8 and there exists an orthogonal array of strength 3 in n runs and k variables, then the set of all such orthogonal arrays is the set of all $\mathrm{tr}(\underset{\sim}{L})$-optimal (n,k)-designs. (An extended version of this theorem, applicable to a general factorial setup, appears in Morris and Mitchell (1977, pp. 45-55).)

The orthogonal arrays of the type declared optimal by the above theorem can easily be constructed by "folding over" an orthogonal main effects design, e.g. a Plackett-Burman (1946) design or a regular fractional factorial design of resolution III (Box and Hunter (1961)). The latter foldover designs are members of the familiar class of regular fractional factorial designs of resolution IV.

2.2. Foldover Designs and Tr($\underset{\sim}{L}$)-Optimality: A Conjecture

We now turn to values of n and k for which no orthogonal array of strength 3 exists. These include all cases in which $n < 2k$ or n is not a multiple of 8.

Our first attempt at the construction of tr($\underset{\sim}{L}$)-optimal designs in these situations was a limited computer search in which we used the design construction algorithm DETMAX (Mitchell (1974a)), modified for our purposes to find locally tr($\underset{\sim}{L}$)-optimal designs. Designs were generated for $k = 4$ with $n = 6,8,10$, and 12, and for $k = 5$ with $n = 8,12,16,20$, and 24. In every case, the design with maximum tr($\underset{\sim}{L}$) turned out to be a foldover design, i.e. the design matrix $\underset{\sim}{D}$ could be written as

$$\underset{\sim}{D} = \begin{pmatrix} \tilde{\underset{\sim}{D}} \\ -\tilde{\underset{\sim}{D}} \end{pmatrix} \qquad (2.2)$$

where the "half-design" matrix $\tilde{\underset{\sim}{D}}$ is an $\tilde{n} \times k$ matrix, $\tilde{n} = n/2$.

Foldover designs, introduced by Box and Wilson (1951), have proved to be extremely useful for estimating main effects free of bias from two-factor interactions. The results of our computer search indicated that this class of designs may also be "optimal" for detecting the presence of two-factor interactions. We express this specifically in the following conjecture, which we have not been able to prove.

Conjecture: For every n, a foldover design exists that is tr($\underset{\sim}{L}$)-optimal in the class of (n,k)-designs, where $\underset{\sim}{L}$ is defined as in (1.7) for the model (1.5).

Although $\mathrm{tr}(\underset{\sim}{L})$-optimal designs for even n are not <u>necessarily</u> foldovers (witness the resolution V fractional factorial designs), the conjecture implies that one need only search the class of foldovers to find a $\mathrm{tr}(\underset{\sim}{L})$-optimal design. This is what we shall do next.

2.3 Tr($\underset{\sim}{L}$)-Optimal Foldover Designs

Some simple matrix algebra shows that for a foldover design (2.2) and for the model defined by (1.5),

$$\mathrm{tr}(\underset{\sim}{L}) = \tilde{n}k^2 - \tilde{n}^{-1}\,\mathrm{tr}(\underset{\sim}{\tilde{D}}'\underset{\sim}{\tilde{D}})^2 \tag{2.3}$$

Thus the $\mathrm{tr}(\underset{\sim}{L})$ criterion for design selection is equivalent to minimizing $\mathrm{tr}(\underset{\sim}{\tilde{D}}'\underset{\sim}{\tilde{D}})^2$, which is Shah's (1960) criterion applied to a first order model with no constant term. (Also, see Kiefer (1974), Section 4H.) Thanks to some unpublished results of L. J. Gray and some helpful conversations with C. S. Cheng, an optimal $\underset{\sim}{\tilde{D}}$ can be constructed easily by referring to the following rules, derivations of which are given in Appendix B.

<u>Gray-Cheng Rules for Constructing</u> $\underset{\sim}{\tilde{D}}$ $(\tilde{n} \geq k)$

<u>Case 1:</u> $\tilde{n} \equiv 0 [\mathrm{mod}\ 4]$.

Choose $\underset{\sim}{\tilde{D}}$ to be a column-orthogonal $\tilde{n} \times k$ matrix. Examples most familiar to statisticians are the Resolution III two-level fractional factorials, and the Plackett and Burman (1946) designs.

<u>Case 2:</u> $\tilde{n} \equiv 1 [\mathrm{mod}\ 4]$.

Add any row with elements ± 1 to a column-orthogonal $(\tilde{n}-1) \times k$ matrix.

<u>Case 3:</u> $\tilde{n} \equiv 2 [\mathrm{mod}\ 4]$.

(a) If $k \leq \tilde{n}-2$, augment an $(\tilde{n}-2) \times k$ column-orthogonal matrix with two rows of $+1$'s and -1's, chosen so that the absolute value of their inner product is less than or equal to 1.

(b) If $k = \tilde{n}$ or $k = \tilde{n}-1$, remove from an $(\tilde{n}+2) \times k$ column-orthogonal matrix two rows whose inner product has absolute value less than or equal to 1. In Appendix 5, it is shown that two such rows exist.

Case 4: $\tilde{n} \equiv 3 \pmod 4$.

Remove any row from an $(\tilde{n}+1) \times k$ column-orthogonal matrix.

The Gray-Cheng rules can be applied in virtually all cases of practical interest, with the exception of the case $k = \tilde{n} = 1 \pmod 4$, where the column-orthogonal $(\tilde{n}-1) \times k$ matrix required by the rule for Case 2 does not exist. (See Raghavarao (1959) for special constructions when $\tilde{n} = 5$, 13, or 25.

Remark 1. Since $tr(\tilde{D}'\tilde{D})^2 = tr(\tilde{D}\tilde{D}')^2$, the same rules can be used when $\tilde{n} < k$: we simply transpose an optimal $k \times \tilde{n}$ matrix.

Remark 2. The $tr(L)$-optimal foldovers derived from these rules are not unique. Usually there are several ways to choose the basic column-orthogonal matrix and several ways to add or remove one or two rows according to the rules. These may yield different L (but the same $tr(L)$) when folded over.

Remark 3. In Cases 1,2, and 4, the class of foldover designs derived from the Gray-Cheng rules is the same as the class of designs obtained by folding over the X matrix (including the column of 1's) for designs constructed according to the rules given by Mitchell (1974b) to achieve D-optimality (in most cases) for the first-order model. In Case 3 there are some minor differences. We would therefore expect the $tr(L)$-optimal foldovers to be good for fitting the first-order model (when $n > 2k$) if interactions are found to be negligible.

Upper bounds on $tr(L)$ for foldover designs are easily obtained by substituting into (2.3) the minimum $tr(\tilde{D}'\tilde{D})^2$ given in Appendix 5. These bounds, which are given in Table 2.1, are attainable by all foldovers derived from the Gray-Cheng rules.

$\tilde{n} > k$		$\tilde{n} < k$	
$\tilde{n}$(mod 4)	tr($\underset{\sim}{L}$)	k(mod 4)	tr($\underset{\sim}{L}$)
0	$\tilde{n}k(k-1)$	0	$k^2(\tilde{n}-1)$
1	$\tilde{n}k(k-1)(1-\tilde{n}^{-2})$	1	$(k^2-1)(\tilde{n}-1)$
2,k even	$\tilde{n}k(k-1)(1-2(k-2)(k-1)^{-1}\tilde{n}^{-2})$	2,$\tilde{n}$ even	$k^2(\tilde{n}-1)-2(\tilde{n}-2)$
2,k odd	$\tilde{n}k(k-1)(1-2(k-1)k^{-1}\tilde{n}^{-2})$	2,$\tilde{n}$ odd	$k^2(\tilde{n}-1)-2(\tilde{n}-1)^2\tilde{n}^{-1})$
3	$\tilde{n}k(k-1)(1-\tilde{n}^{-2})$	3	$(k^2-1)(\tilde{n}-1)$

Table 2.1. Upper bounds on tr($\underset{\sim}{L}$) for foldover designs in k variables and $2\tilde{n}$ runs. These bounds are attained by all designs derived from the Gray-Cheng rules.

3. Tr($\underset{\sim}{L}$)-Optimal Fractional Factorial Designs

Since the fractional factorial designs are so well known and widely used, it is of interest to know which are the best with respect to the tr($\underset{\sim}{L}$) criterion, and how these compare with the optimal foldovers described in the previous section. We shall restrict our discussion to the regular 2^{k-p} fractional factorials. Every design in this class has a unique "defining relation" with 2^p-1 "words" which identify the effects that are completely confounded with the overall mean (Box and Hunter (1961)).

3.1. Characterization

From Theorem 2.1 and Theorem C2 in Appendix C, we can characterize tr($\underset{\sim}{L}$)-optimal fractional factorial designs as follows:

1. If $n > 2k$, the tr($\underset{\sim}{L}$)-optimal fractional factorials are the 2^{k-p} designs of resolution ≥ 4.
2. If $n < 2k$, the tr($\underset{\sim}{L}$)-optimal fractional factorials are the 2^{k-p} foldover designs of resolution 2 with the fewest 2-letter words in the defining relation.

3.2 Construction

The construction of designs of resolution ≥ 4 is well known, so there is no problem if $n > 2k$, unless one wants to use additional criteria to select from among the many designs available. For this purpose, we would recommend the "minimum aberration" criterion of Fries and Hunter (1979), which in the present case amounts to selecting designs that

have the fewest words of length 4 in the defining relation. (See Appendix D for a more detailed discussion.) Table 12.15 of Box, Hunter and Hunter (1978) gives a list of minimum aberration designs for $k \leq 11$, $n \leq 128$.

If $n < 2k$, we want to construct the 2^{k-p} foldover with the fewest words of length 2 in its defining relation, i.e., with the fewest pairs of completely confounded factors. This can be achieved only by distributing the factors as evenly as possible over the set of columns in $\underset{\sim}{D}^*$, the saturated design of resolution >4 in $\tilde{n} = n/2$ factors and n runs. (Examples of $\underset{\sim}{D}^*$ are the 2^2, 2^{4-1}_{IV}, 2^{8-4}_{IV} and 2^{16-11}_{IV} designs. See Box and Hunter (1961), Section 5.) The $tr(\underset{\sim}{L})$-optimal 2^{k-p} designs will therefore have the form $[\underset{\sim}{D}_1 : \underset{\sim}{D}_2]$, where $\underset{\sim}{D}_1$ consists of $a \geq 1$ copies of $\underset{\sim}{D}^*$ and $\underset{\sim}{D}_2$ consists of a subset of r distinct columns of $\underset{\sim}{D}^*$, and where a and r are the integer part of $k/\tilde{n}$ and the remainder, respectively:

$$a = \text{Int}(k/\tilde{n}) \tag{3.1}$$

$$r = k - a\tilde{n} \ . \tag{3.2}$$

For construction and analysis, it is convenient to write these designs in terms of "group-factors" $A_1, A_2, \ldots, A_{\tilde{n}}$ (Watson (1961)). An example, for $k = 12$ and $n = 8$, is given in Table 6.1 of Section 6. The aliasing relations can then be determined most easily by first writing down the n aliasing relations among the group factors in the usual way; then

(i) replace each group-factor main effect A_i by the sum of the main effects of the factors in Group A_i;

(ii) replace each two-factor interaction (A_iA_j) among group-factors by the sum of all two-factor interactions involving one factor from group A_i and one factor from group A_j, and

(iii) replace the overall mean (denoted by I in the notation of Box and Hunter (1961)) by $\mathcal{I}_0$ plus the sum of all two-factor interactions involving two factors from the same group.

For tr($\underset{\sim}{L}$) optimality, it doesn't matter which r of the $\tilde{n}$ columns of $\underset{\sim}{D}^*$ are chosen to form $\underset{\sim}{D}_2$. However, if we regard $\underset{\sim}{D}_2$ as an n-run design in r factors, and choose it to minimize the sum of squared lengths of the strings of two-factor interactions among those factors, then the design $\underset{\sim}{D} = [\underset{\sim}{D}_1 : \underset{\sim}{D}_2]$ will have minimum aberration among tr($\underset{\sim}{L}$)-optimal 2^{k-p} designs. (This is Theorem D1 in Appendix D.) The construction of such $\underset{\sim}{D}_2$ is easy when n = 4,8, or 16; any subset of r columns of D^* will do. When n = 32, proceed as follows:

(i) Write down the saturated 2^{16-11}_{IV} design with generators 1236, 1247, 1258, 1349, 135(10), 145(11), 234(12), 235(13), 245(14), 345(15), 12345(16).

(ii) Strike out the columns associated with the first (16-r) factors in the following list: 16, 15, 14, 13, 12, 1, 11, 10, 7, 6, 9, 5, 4, 3, 2.

This procedure was derived by writing down, for each k, all feasible integer vectors $(f_0, f_1, f_2, \ldots)$, where f_i is the number of strings of length i, finding the one which minimizes $\sum i^2 f_i$, and then finding the corresponding design. We have not attempted to derive similar procedures for n > 32.

3.3. Comparison of Tr($\underset{\sim}{L}$)-Optimal Fractional Factorials With Tr($\underset{\sim}{L}$)-Optimal Foldovers.

When n is a power of 2, one would generally prefer to use a fractional factorial design rather than the less familiar optimum foldovers of Section 2, mainly for reasons of simplicity of construction and analysis. As we shall now see, the optimal fractional factorials are either as good as or "almost" as good as the optimal foldovers with respect to tr($\underset{\sim}{L}$)-optimality.

If n, a power of 2, is greater than or equal to 2k, the tr($\underset{\sim}{L}$)-optimal foldovers and the tr($\underset{\sim}{L}$)-optimal fractional factorials are orthogonal arrays of strength 3, and so are optimal among (n,k)-designs by Theorem 2.1. In the more interesting case n < 2k, our main result (Theorem C2 of Appendix C) is as follows:

Given n (a power of 2) and $k > \tilde{n} = n/2$, a $\mathrm{tr}(\underset{\sim}{L})$-optimal 2^{k-p} design in n runs is $\mathrm{tr}(\underset{\sim}{L})$-optimal among all two-level foldover (n,k)-designs if and only if r (the remainder upon dividing k by $\tilde{n}$) is $0,1,2,\tilde{n}-1$, or $\tilde{n}-2$.

In the cases for which r does not satisfy these conditions, the $\mathrm{tr}(\underset{\sim}{L})$-optimal 2^{k-p} design is "almost" optimal among foldovers. For example, suppose $r = \tilde{n}/2$, where $\tilde{n} > 8$, which appears to be the "worst case" for the efficiency of the 2^{k-p} designs with $n < 2k$. Since $\tilde{n}$ is a power of 2 and $\tilde{n} > 8$, k is divisible by 4 (by (3.2)), so the upper bound on $\mathrm{tr}(\underset{\sim}{L})$ for foldover designs, given in Table 2.1, is $k^2(\tilde{n}-1)$. The efficiency of the optimal fractional factorial, relative to this upper bound, can be shown to equal $1 - \tilde{n}^2/(4k^2(\tilde{n}-1))$. For fixed $\tilde{n}$, this is minimized when $k = r + \tilde{n} = 3\tilde{n}/2$. (See (3.2).) Thus the efficiency of the optimal fractional factorial is at least $1 - (9(\tilde{n}-1))^{-1}$, which is at least .9841, since $\tilde{n} > 8$ here.

We conclude that if one is seeking a $\mathrm{tr}(\underset{\sim}{L})$-optimal (or nearly optimal) foldover in $2\tilde{n}$ runs where $\tilde{n}$ is a power of 2, one might as well restrict attention to the fractional factorials. These designs are easy to construct and the analysis of the data is easier to perform than for other types of $\mathrm{tr}(\underset{\sim}{L})$-optimal designs.

4. Power of the Likelihood-Ratio Test of the Hypothesis of No Interactions

In this section, we shall indicate roughly the ability of the designs of Sections 2 and 3 to detect the presence of interactions when a conventional statistical hypothesis test is used.

4.1. The LR Test for the Presence of Interactions.

We shall restrict attention here to two studies of the power of the likelihood-ratio (LR) test of the hypothesis that $\underset{\sim}{\beta}_2 = \underset{\sim}{0}$ in the model (1.5), where $\underset{\sim}{\varepsilon}$ is normally distributed and σ^2 is assumed "known". This is not intended to preclude the use of other formal or informal techniques of analyzing the data for the presence of interactions.

The LR test statistic w is $R(\underset{\sim}{\beta}_2|\underset{\sim}{\beta}_1)/\sigma^2$, where $R(\underset{\sim}{\beta}_2|\underset{\sim}{\beta}_1)$ is the increase in the residual sum of squares for the model (1.5) that results when $\underset{\sim}{\beta}_2$ is set to $\underset{\sim}{0}$. This statistic has a non-central chi-squared distribution with $r(\underset{\sim}{L})$ degrees of freedom and non-centrality parameter $\lambda/(2\sigma^2)$, where $r(\underset{\sim}{L})$ is the rank of $\underset{\sim}{L}$ and $\lambda = \underset{\sim}{\beta}_2'\underset{\sim}{L}\underset{\sim}{\beta}_2$.

The calculation of the LR statistic w is particularly easy for the foldover designs of Sections 2 and 3 when $n \leq 2k$. Let y_i^+ and y_i^- be one-half the sum and one-half the difference, respectively, of the two observations in the ith foldover pair. Then an equivalent form of the model (1.5) is

$$E(\underset{\sim}{y}^+) = \underset{\sim}{1}\beta_0 + \underset{\sim}{\tilde{X}}_2\underset{\sim}{\beta}_2 \qquad E(\underset{\sim}{y}^-) = \underset{\sim}{\tilde{D}}\underset{\sim}{\beta}_1, \tag{4.1}$$

where $\underset{\sim}{\tilde{D}}$ and $\underset{\sim}{\tilde{X}}_2$ are composed of the columns for main effects and interactions, respectively, in the half-design, and our notation has been changed temporarily so that $\underset{\sim}{\beta}_1$ now contains only main effects (not β_0). Note that the elements of $\underset{\sim}{y}^+$ and $\underset{\sim}{y}^-$ are all uncorrelated and have variance $\sigma^2/2$. When the half-design $\underset{\sim}{\tilde{D}}$ has full row rank $\tilde{n} = n/2$, as it does for the foldovers of Sections 2 and 3 with $\tilde{n} \leq k$, there is no contribution to the residual sum of squares from $\underset{\sim}{y}^-$, i.e., $\hat{\underset{\sim}{y}}^- = \underset{\sim}{y}^-$. It therefore follows that the residual sum of squares for the model (4.1) with $\underset{\sim}{\beta}_2 = \underset{\sim}{0}$ is just the sum of squared deviations of the y_i^+'s about their average. This residual sum of squares is in fact $R(\underset{\sim}{\beta}_2|\underset{\sim}{\beta}_1)$ since the row rank of $[\underset{\sim}{1}\vdots\underset{\sim}{\tilde{X}}_2]$ is also $\tilde{n}$, and the unrestricted model fits the data exactly. Thus

$$w = 2\sum(y_i^+ - \bar{y}^+)^2/\sigma^2 \tag{4.2}$$

for the designs of Sections 2 and 3 with $n \leq 2k$.

4.2. Power Study 1

This was a simulation study which was conducted for some $tr(L)$-optimal foldover designs constructed as indicated in Section 2. Cases examined were $k = 3$ through 8 with $n = 4,6,$ and 8.

For each design, the power of the LR test was investigated for two values of $\rho^2 = \underset{\sim}{\beta}_2'\underset{\sim}{\beta}_2/\sigma^2$, where ρ^2 may be viewed as a measure of the overall magnitude of the interactions. (We note that $\underset{\sim}{\beta}_2'\underset{\sim}{\beta}_2$ is the average squared residual per point which would occur if the first-order model were fitted to the "true" response (1.4) over K. Thus, for example, if the interactions are such that $\rho^2 = 4$, we would expect a "typical" deviation from the first-order model at a given combination of factor levels to be on the order of 2σ.)

In each simulation, $\underset{\sim}{\beta}_2$ was selected randomly 12,500 times from a uniform distribution on the sphere $\underset{\sim}{\beta}_2'\underset{\sim}{\beta}_2 = \rho^2\sigma^2$, ($\sigma = 1.0$ or 2.0), according to a method described by Marsaglia (1972). For each $\underset{\sim}{\beta}_2$, the non-centrality parameter was computed, then the corresponding power for the LR test at the $\alpha = .10$ level was calculated using an approximation to non-central chi-squared probabilities given by Severo and Zelen (1960). This procedure generated a distribution of power values, the quartiles of which are given in Table 4.1 for each case.

		ρ=1			ρ=2		
k	n	$P_{.25}$	$P_{.50}$	$P_{.75}$	$P_{.25}$	$P_{.50}$	$P_{.75}$
3	4	.217	.455	.652	.543	.949	.987
	6	.528	.675	.784	.986	.999	1.000
	8	.767	.767	.767	1.000	1.000	1.000
4	4	.189	.420	.661	.447	.924	.998
	6	.445	.644	.787	.955	.998	1.000
	8	.624	.766	.863	.998	1.000	1.000
5	4	.178	.382	.631	.409	.887	.996
	6	.387	.593	.751	.910	.995	1.000
	8	.549	.711	.831	.992	1.000	1.000
6	4	.176	.290	.483	.401	.736	.963
	6	.364	.576	.742	.884	.993	1.000
	8	.519	.694	.824	.987	1.000	1.000
7	4	.169	.367	.621	.376	.869	.995
	6	.359	.577	.740	.877	.993	1.000
	8	.510	.690	.821	.985	.999	1.000
8	4	.181	.369	.634	.421	.871	.996
	6	.362	.582	.752	.882	.994	1.000
	8	.517	.702	.834	.986	1.000	1.000

Table 4.1. Power study for small tr($\underset{\sim}{L}$)-optimal foldovers: quartiles of the distribution of the power of the LR test of the hypothesis $\underset{\sim}{\beta}_2 = \underset{\sim}{0}$ in (1.5) with normal $\underset{\sim}{\varepsilon}$ and σ^2 known, generated by selecting $\underset{\sim}{\beta}_2$ randomly from the sphere of radius $\rho\sigma$. The significance level the test is $\alpha = .10$. The results for each case are based on 12500 simulations.

4.3. Power Study 2

In this study we investigated the power of the LR test, again with σ^2 "known", under the assumption that the interactions are drawn independently from a normal distribution with mean 0 and variance σ_b^2, where $\sigma_b^2 = \rho^2\sigma^2/v_2$ i.e. $E(\underset{\sim}{\beta}_2'\underset{\sim}{\beta}_2) = \rho^2\sigma^2$. The designs considered were the tr($\underset{\sim}{L}$)-optimal fractional factorials of Section 3; for these designs the test statistic w and its mean and variance are easily calculated from the lengths of the strings of confounded two-factor interactions. (See Appendix E for details.)

The distribution of w was approximated by that of $w' = g_1\chi^2_{g_2}$, where g_1 and g_2 were chosen so that the mean and variance of w' matched those of w. We determined for each design the value of ρ for which the power is .90 for tests conducted at the $\alpha = .10$ level of significance. (Actually, we are discussing <u>expected</u> power here, where the expectation is taken over the assumed normal distribution of the interactions.) These "minimum detectable" values of ρ were calculated for (i) $n \geq 2k$, $3 \leq k \leq 10$, $n \leq 128$, for the designs in Table 12.15 of Box, Hunter, and Hunter (1978), which are minimum aberration designs of resolution ≥ 4, and for (ii) $n < 2k$, $3 \leq k \leq 10$, $n \geq 4$, for the minimum aberration resolution II foldovers presented in Section 3.2. Some results are shown in Table 4.2, for $k = 5$ and $k = 10$, as well as the limiting cases as $k \to \infty$. In the case $k = 5$, $n = 8$, for example, the interactions need to be big enough to cause a "typical" disturbance of magnitude 2.01σ at a randomly selected corner of the 5-cube in order to be detected with probability .90 by the LR test with $\alpha = .10$.

To obtain an approximation to the minimum detectable value of ρ, ($\rho^*(\alpha,P,k,n)$, say), for specified significance level α and power P, once can use the equation for the limiting value as $k \to \infty$, which can be shown to be

$$\rho^*(\alpha,P,\infty,n) = [(\chi^2_{q;\alpha}/\chi^2_{q;P} - 1)/2]^{1/2} \tag{4.3}$$

where $q = n/2-1$ and $\chi^2_{q;\alpha}$ is the upper $100\alpha\%$ percentage point of the χ^2_q distribution. (See Appendix E.) Since ρ^* does not change much with k, (4.3) can be used to approximate $\rho^*(\alpha,P,k,n)$.

Although the results of Tables 4.1 and 4.2 do not represent a very comprehensive study of power, they do serve to indicate roughly what the user can expect from the designs of Sections 2 and 3 with respect to their ability to detect the presence of interactions. In the next section, we shall consider other design objectives.

	n					
k	4	8	16	32	64	128
5	8.42	2.01	1.20	0.85	0.60	0.42
10	8.76	2.10	1.22	0.89	0.69	0.50
∞	9.23	2.20	1.27	0.90	0.68	0.54

Table 4.2. Power study 2: Minimum detectable values of ρ for the LR test of the hypothesis $\underset{\sim}{\beta}_2 = \underset{\sim}{0}$ at the $\alpha = .10$ significance level, σ^2 known, using tr($\underset{\sim}{L}$)-optimal fractional factorials described in Section 3. Here a value of ρ is "detectable" if the expected power of the test is at least .90 when the elements $\underset{\sim}{\beta}_2$ are drawn independently from a normal distribution with mean 0 and variance $\sigma_b^2 = \rho^2\sigma^2/k_2$, i.e., $E(\underset{\sim}{\beta}_2'\underset{\sim}{\beta}_2) = \rho^2\sigma^2$.

5. Modification of tr($\underset{\sim}{L}$)-Optimal Designs to Suit Additional Objectives

Seldom is an experiment planned in practice with just a single purpose in mind, so we shall now examine the designs of Sections 2 and 3 with respect to some other objectives and suggest some design modifications.

5.1. Fitting the First-Order Model

When $n \geq 2k$, the tr($\underset{\sim}{L}$)-optimal designs are orthogonal (or nearly so) for the first-order model: $E(\underset{\sim}{y}) = \underset{\sim}{X}_1\underset{\sim}{\beta}_1$, so they need no modification to estimate $\underset{\sim}{\beta}_1$ efficiently. When $n < 2k$, however, the tr($\underset{\sim}{L}$)-optimal foldovers presented in this paper do not permit estimation of $\underset{\sim}{\beta}_1$ in the first-order model. For these situations, we tried several approaches to the construction of "compromise designs" which would have relatively high values of tr($\underset{\sim}{L}$) and would also provide estimability of $\underset{\sim}{\beta}_1$ (Morris and Mitchell (1977)). Our most successful procedure was the following. The size of the final design, n, is specified as well as the size of a smaller foldover design, $2\tilde{n}$. A tr($\underset{\sim}{L}$)-optimal foldover design in $2\tilde{n}$ runs is then obtained and augmented with the $n - 2\tilde{n}$ runs which maximize the determinant of $\underset{\sim}{X}_1'\underset{\sim}{X}_1$ for the final design.

Compromise designs were constructed in this way for $k = 4$ through 8, with $n = k+2$ through $2k-1$ and varying $\tilde{n}$. The augmentation was done using the DETMAX algorithm (Mitchell (1974a)). The designs which have the minimum number of augmenting runs for fixed n are presented in Table 5.1. For these minimally augmented $tr(L)$-optimal designs, the augmentation does not affect $tr(L)$; in fact, the "extra" runs (those not marked with an asterisk in the table) are not used at all in the LR test of the hypothesis that $\beta_2 = 0$. The user of one of the compromise designs in Table 5.1 can therefore refer to the results of Section 4, particularly Table 4.1, for an indication of the ability of the design to detect the presence of interactions.

k	$2\tilde{n}$	n							
4	4	6	1101*	1110*	0111	1011			
4	6	7	1001*	1010*	1111*	1100			
5	4	7	01111*	01000*	00110	00011	00101		
5	6	8	01001*	01100*	00111*	10101	00000		
5	8	9	11000*	00001*	01010*	01101*	00100		
6	4	8	110100*	000101*	101110	111111	011100	010110	
6	6	9	111100*	011010*	010101*	111111	110110	001110	
6	8	10	010100*	010010*	111000*	110111*	011110	100110	
6	10	11	100111*	001101*	010001*	101001*	100000*	001010	
7	4	9	0100001*	0111111*	1110001	0111000	1101010	1001001	1101100
7	6	10	0100000*	0011100*	0000111*	1010001	0011011	0110111	1010010
7	8	11	0100010*	1010010*	0001000*	0000111*	1000001	1101010	1000100
7	10	12	0011001*	1011110*	1000101*	1111101*	1001011*	0010011	0001111
7	12	13	0110110*	1010000*	0011101*	0010011*	1111011*	1110101*	0101000
8	4	10	11011011*	00101011*	01000111	10000001	01001000	00011101	00010010	10001110
8	6	11	11100000*	10010010*	10101110*	11110111	00110100	01000110	10000101	01100011
8	8	12	10001001*	01100001*	11101101*	00010000*	00111101	01011011	11111000	01001100
8	10	13	10110000*	10101110*	00110001*	11111011*	11100101*	10011101	00110111	10000011
8	12	14	11111111*	11101000*	11110100*	11000111*	10110010*	10100101*	01110000	11010101
8	14	15	11011000*	10111110*	01010110*	11001111*	01111101*	00001100*	11100100*	10010101

Table 5.1. Compromise designs constructed by augmenting $tr(L)$-optimal designs to permit estimability of main effects. The notation "0" is used to refer to the level "-1". Each design also contains the foldover of runs marked with *.

Before going on to the consideration of other design objectives, we should remark that unaugmented tr($\underset{\sim}{L}$)-optimal designs with $n < 2k$ can be of practical use, even though they do not permit estimation of $\underset{\sim}{\beta}_1$. This is particularly true when one is dealing with a large number of factors and the number of runs is quite limited. Common practice is to use a first-order design in hopes that the main effects will override the interactions, and then perhaps to follow up with further runs to seek out interactions among the large main effects. When substantial interactions are present, however, inferences drawn from a main effects design, and subsequent experimental plans based on those inferences, may be misguided. What we are suggesting here is that in some cases it may be worth spending a few early runs (4 to 8, say) in order to find out, in a general way, how important the interactions are.

5.2. Identifying the Second-Order Interactions

Once the presence of interactions has been established, additional runs can be made to identify the larger ones. If one can afford it, one might wish to augment the initial tr($\underset{\sim}{L}$)-optimal design to provide estimates of all the interactions, e.g., Example 4 of Mitchell (1974a). In many situations, it will be more efficient to concentrate on a subset of interactions, as in the following example, condensed from Morris and Mitchell (1977).

Example 5.1. This is a hypothetical example with 7 factors, in which data were simulated according to the equation $y = 64 - 7x_1 - 19x_3 - 16x_1x_3 + \varepsilon$, where $\varepsilon \sim N(0,\sigma^2)$, and $\sigma = 5.5$. The initial design was a tr($\underset{\sim}{L}$)-optimal 2^{7-4}_{II} design, constructed as indicated in Section 3. The design points, data, and estimates of strings of confounded effects are shown in Table 5.1a. Assuming σ^2 is known, the LR test statistic for the hypothesis of no interactions is $2086.86/30.25 = 68.99$, which is highly significant when referred to the χ^2_3 distribution. Clearly, the most likely candidates for large interactions are those in the string $\beta_{13} + \beta_{14} + \beta_{23} + \beta_{24} + \beta_{57} + \beta_{67}$. (Since the estimate of the three strings of interactions are independent, the χ^2_3 statistic could have been partitioned to give a separate lack of fit test for each string. In the present

case, this would lead to rejection of the null hypothesis only for $\beta_{13}+\ldots+\beta_{67}$.) Eight additional runs were needed to estimate the main effects and the six suspected interactions. These runs were chosen using DETMAX to maximize the determinant of $\underset{\sim}{X}'\underset{\sim}{X}$ for all 16 runs, where the model is now

$$E(y) = \beta_0 + \sum_{i=1}^{7} x_i\beta_i + x_ix_3\beta_{13} + x_1x_4\beta_{14} + x_2x_3\beta_{23} + x_2x_4\beta_{24} + x_5x_7\beta_{57} + x_6x_7\beta_{67} \quad . \tag{5.1}$$

x_5, x_6	x_3, x_4	x_1, x_2	x_7	y
-1	-1	-1	-1	66.04
1	-1	-1	1	78.63
-1	1	-1	1	62.36
1	1	-1	-1	66.96
-1	-1	1	1	88.45
1	-1	1	-1	98.39
-1	1	1	-1	14.46
1	1	1	1	28.05

$$\hat{\beta}_0 + \hat{\beta}_{12} + \hat{\beta}_{34} + \hat{\beta}_{56} = 62.92$$

$$\hat{\beta}_1 + \hat{\beta}_2 = -5.58$$

$$\hat{\beta}_3 + \hat{\beta}_4 = -19.96$$

$$\hat{\beta}_5 + \hat{\beta}_6 = 5.09$$

$$\hat{\beta}_7 = 1.46$$

$$\hat{\beta}_{17} + \hat{\beta}_{27} + \hat{\beta}_{35} + \hat{\beta}_{36} + \hat{\beta}_{45} + \hat{\beta}_{46} = -0.54$$

$$\hat{\beta}_{15} + \hat{\beta}_{16} + \hat{\beta}_{25} + \hat{\beta}_{26} + \hat{\beta}_{37} + \hat{\beta}_{47} = 0.79$$

$$\hat{\beta}_{13} + \hat{\beta}_{14} + \hat{\beta}_{23} + \hat{\beta}_{24} + \hat{\beta}_{57} + \hat{\beta}_{67} = -16.12$$

Table 5.1a. Data and estimates of effects for the 2^{7-4}_{III} initial design in Example 5.1.

The data for the 8 new runs and the estimates of the parameters in the new model are given in Table 5.1b. The parameter estimates are not very precise, since we have added the fewest runs possible to achieve estimability. Even so, in 16 runs we have found that interactions are not negligible and have discovered the important one. This use of tr($\underline{L}$)-optimal designs to identify a few strings of potential interactions, which are then broken down by further runs, is very similar in spirit to Watson's (1961) approach to the problem of screening for main effects.

x_1	x_2	x_3	x_4	x_5	x_6	x_7	y
1	-1	1	-1	-1	-1	1	13.10
-1	-1	1	-1	-1	1	-1	70.08
1	-1	-1	1	-1	-1	1	91.69
-1	-1	1	-1	1	-1	-1	75.46
-1	1	1	-1	-1	1	1	62.95
1	-1	1	1	-1	-1	-1	22.07
1	-1	-1	-1	1	1	1	86.99
-1	1	1	-1	1	-1	1	74.80

$\hat{\beta}_0 = 62.92$	$\hat{\beta}_4 = 0.63$	$\hat{\beta}_{13} = -16.59$	$\hat{\beta}_{57} = 0.43$
$\hat{\beta}_1 = -6.61$	$\hat{\beta}_5 = 4.70$	$\hat{\beta}_{14} = 5.92$	$\hat{\beta}_{67} = -1.18$
$\hat{\beta}_2 = 1.03$	$\hat{\beta}_6 = 0.39$	$\hat{\beta}_{23} = -4.57$	
$\hat{\beta}_3 = -20.59$	$\hat{\beta}_7 = 1.46$	$\hat{\beta}_{24} = -0.14$	

Table 5.1b. Additional data and estimates of effects for the model (5.1) in Example 5.1.

Had the initial 8-run design in this example given no indication of the presence of interactions, we could have chosen our additional eight runs to give a good estimate of the parameters in the first-order model, as described in Section 5.1. The resulting 16-run design would then turn out to be the 2^{7-3}_{IV} design with generators 1234, 1256, and 1357 (Morris and Mitchell, 1977). The interactions should be examined again at this point.

5.3. Detecting the Presence of Other Non-Linear Effects

If some of the factors are continuous, then there may well be departures from the first-order model that do not involve interactions. The two-level designs considered in this paper will not be good for detecting such effects.

The most obvious augmentation in this case would involve adding one or more "center point" runs in which the quantitative factors are all set to a central value. Taking a formal design optimality approach, Jones and Mitchell (1975, Section 4.3.1) applied their A_2-optimality criterion (from which our $\mathrm{tr}(\underset{\sim}{L})$-optimality was derived) to the two-factor quadratic response surface model, and indeed found in all cases ($n = 4 - 10$) that the optimal designs for a rectangular region of interest were supported entirely on the corners and at the center of the region.

When $n < 2k$, the use of a center point also aids in identifying the interactions (by separating a string of interactions from β_0), but does not seem very efficient in terms of the $\mathrm{tr}(\underset{\sim}{L})$ criterion. It can be shown that the increase in $\mathrm{tr}(\underset{\sim}{L})$ resulting from the addition of a row of k 0's to a foldover (n,k)-design is $(nk_2 - \text{old } \mathrm{tr}(\underset{\sim}{L}))/(n+1)$, which is relatively small, especially when compared with the gain that can be made by adding a new foldover pair.

We have not considered the question of how many center points to add, nor the more interesting question of how to take center points when not all the variables are quantitative.

5.4. Estimation of σ^2

In our discussion of the LR test for the presence of interactions, and in Example 5.1, we assumed that σ^2 was "known". We shall now consider designs with the dual purpose of maximizing $\mathrm{tr}(\underset{\sim}{L})$ and obtaining an estimate of σ^2 through replication of some runs.

Consider the construction of a $\mathrm{tr}(\underset{\sim}{L})$-optimal foldover design under the restriction that $\tilde{n}_0$ rows of the $\tilde{n} \times k$ half-design $\underset{\sim}{\tilde{D}}$ are replicated once, where $\tilde{n} \leq k + \tilde{n}_0$. When $k \equiv 0,1$, or 3 (mod 4), this is achieved by replicating any $\tilde{n}_0$ foldover pairs of a

$tr(\underset{\sim}{L})$-optimal $(2\tilde{n}-2\tilde{n}_e,k)$-design. (See Appendix F.) If $k \equiv 2 \pmod 4$, we can use the following procedure. Partition a column orthogonal $(k-2) \times (\tilde{n}-\tilde{n}_e)$ matrix $\underset{\sim}{D}_0$ as $\underset{\sim}{D}_0 = [\underset{\sim}{A}:\underset{\sim}{B}]$, where $\underset{\sim}{A}$ has $\tilde{n}_e$ columns and $\underset{\sim}{B}$ has $\tilde{n} - 2\tilde{n}_e$ columns.

Now let the $k \times \tilde{n}$ matrix $\underset{\sim}{D}_1$ have the form

$$\underset{\sim}{D}_1 = \begin{bmatrix} \underset{\sim}{A} & \underset{\sim}{A} & \underset{\sim}{B} \\ \underset{\sim}{a}_1' & \underset{\sim}{a}_1' & \underset{\sim}{b}_1' \\ \underset{\sim}{a}_2' & \underset{\sim}{a}_2' & \underset{\sim}{b}_2' \end{bmatrix}$$

where $|2\underset{\sim}{a}_1'\underset{\sim}{a}_2 + \underset{\sim}{b}_1'\underset{\sim}{b}_2| \leq 1$. (Note: If $\tilde{n}-\tilde{n}_e = k$ or $\tilde{n}-\tilde{n}_e = k-1$, choose $\underset{\sim}{D}_0$ instead to be a column-orthogonal $(k+2) \times (\tilde{n}-\tilde{n}_e)$ matrix. $\underset{\sim}{D}_1$ is then formed by removing from $\underset{\sim}{D}_0$ two rows $(\underset{\sim}{a}_1' \; \underset{\sim}{a}_1' \; \underset{\sim}{b}_1')$ and $(\underset{\sim}{a}_2' \; \underset{\sim}{a}_2' \; \underset{\sim}{b}_2')$ that satisfy the above property.) If we now transpose $\underset{\sim}{D}_1$ and fold it over, the result will be a $2\tilde{n} \times k$ foldover design which is $tr(\underset{\sim}{L})$-optimal subject to the restriction that $\tilde{n}_e$ foldover pairs are replicated once. A short proof is given in Appendix F.

When $\tilde{n} > k + \tilde{n}_e$, we have not found a general procedure for constructing $tr(\underset{\sim}{L})$-optimal designs subject to replication of $\tilde{n}_e$ foldover pairs. However, the rules of Section 2.3 are not very restrictive, and it is often possible to construct designs that satisfy these rules and also replicate some runs. For example, the 6×4 matrix with rows $(1,-1,-1,1)$, $(1,1,-1,-1)$, $(1,-1,1,-1)^*$, $(1,1,1,1)^*$, where the asterisks indicate replication, yields a $tr(\underset{\sim}{L})$-optimal design (Case 3 of Section 2.3) when folded over.

6. Example: Sensitivity Analysis for a Computer Code

The Oak Ridge Inverse Code (ORINC), (Ott and Hedrick (1977)), is used to calculate temperature and heat flux at the surface of the electric heater rods in a simulated nuclear reactor, given the heat generation rate, the geometry, thermophysical parameters, and the thermocouple temperature at an axial position of one of the rods.

To determine the sensitivity of ORINC's results to variations in key parameters, a computational experiment was conducted. The experimental design was a 32-run 2^{12-7}_{IV} fractional factorial design in the 12 factors (parameters): (1) MgO radius, (2) inconel thickness, (3) Bn thickness, (4) inner sheath thickness, (5) outer sheath thickness, (6) gap size, (7) thermocouple temperature, (8) power peaking factor, (9) voltage, (10) amperage, (11) MgO conductivity, and (12) Bn conductivity. The two levels of each parameter were at one standard deviation above and below the nominal value of that parameter, where the standard deviations were based on given "uncertainty distributions". Sensitivities were defined in terms of main effects, calculated in the usual way. Strings of two-factor interactions were also estimated and found to be negligible. Assessments of importance of effects were based on relative magnitude; there is no statistical error involved.

In the following, we shall use some of the data from this computer experiment to demonstrate how a small preliminary $tr(\underset{\sim}{L})$-optimal design might have been used to provide an early assessment of the importance of interactions. The chosen 8-run $tr(\underset{\sim}{L})$-optimal 2^{12-9} design, augmented by the center point, is shown in Table 6.1, with the heat flux results of the ORINC runs and the calculated effects. (Table 6.1 shows only the heat flux $y(t)$ at time 0; however, each ORINC run gives the values of heat flux as a function of time, and the effects may be plotted in this way.)

Group A_1 2,11,12	Group A_2 1,3,4	Group A_3 6,7,10	Group A_4 5,8,9	Response (y(0)-500000)/100
-	-	-	-	401
+	-	-	+	392
-	+	-	+	261
+	+	-	-	239
-	-	+	+	422*
+	-	+	-	400*
-	+	+	-	267*
+	+	+	+	261*
0	0	0	0	329

*Also used in the followup 2^{12-8}_{III} design

Group Aliasing	Factor Aliasing	Effects
A_1	2 + (11) + (12)	- 7.4
A_2	1 + 3 + 4	-73.4
A_3	6 + 7 + (10)	7.1
A_4	5 + 8 + 9	3.6
$A_1A_2 + A_3A_4$	12 + 23 + 24 + 1(11)+...+5(10) + 8(10) + 9(10)	0.4
$A_1A_3 + A_2A_4$	26 + 27 + 2(10) + 6(11)+...+45 + 48 + 49	0.4
$A_1A_4 + A_2A_3$	25 + 28 + 29 + 5(11)+...+46 + 47 + 4(10)	- 0.1
	2(11) + 2(12) + (11)(12) + 13+...+56 + 59 + 89	1.4**

**obtained by subtracting the center point response from the average of the other points

Table 6.1. A tr($\underset{\sim}{L}$)-optimal 2^{12-9} design plus center point, with data from Example 6.1. The numbers in the factor aliasing relations stand for subscripts on the coefficients (β's) in the model.

On the basis of these results, we would tentatively infer that interactions are negligible, although we still need to be aware of possible "cancellations" within interaction strings. We can then proceed with a first-order design with some confidence that the larger main effects will correctly identify the parameters to which the ORIV

results are most sensitive. In the present case, 12 additional runs, combined with the four marked with an asterisk in Table 6.1, yield a 2^{12-8}_{III} design with generators 125, 136, 147, 238, 249, 34(10), 123(11), and 234(12). The main effects are given in Table 6.2. (For simplicity we calculated these effects using only the 16 runs of the 2^{12-8} design.)

Factor	Effect	Factor	Effect	Factor	Effect
1	-16.7	5	-34.2	9	37.8
2	-15.9	6	0	10	7.2
3	-26.8	7	0	11	8.6
4	-29.8	8	0	12	0

Table 6.2. Main effects from 2^{12-8}_{III} design in Example 6.1

Had this investigation involved a very large number of factors, augmentation to estimate all main effects might not be feasible. A reasonable approach in this case might be to estimate individual effects only within main effect strings that appear to be large in the initial design (Watson (1961)). For these factor screening applications, one should attempt to assign "+" and "-" to each factor in such a way that a "+" corresponds to an anticipated increase in response. If one's guesses of the direction of effects are correct, this will eliminate the possibility of "cancellations" within strings of main effects.

7. Summary and Conclusions.

We have given here the results of the application of a design optimality criterion (maximization of $\mathrm{tr}(\underset{\sim}{L})$ where $\underset{\sim}{L} = \underset{\sim}{X}_2'\underset{\sim}{X}_2 - \underset{\sim}{X}_2'\underset{\sim}{X}_1(\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-}\underset{\sim}{X}_1'\underset{\sim}{X}_2$) to the problem of designing two-level n-run experiments to detect the presence of two-factor interactions ($\underset{\sim}{\beta}_2$) among k-factors in the model $E(\underset{\sim}{y}) = \underset{\sim}{X}_1\underset{\sim}{\beta}_1 + \underset{\sim}{X}_2\underset{\sim}{\beta}_2$, where $\underset{\sim}{\beta}_1$ consists of a constant term β_0 and main effects.

When n is a multiple of 8, the $tr(\underset{\sim}{L})$-optimal designs are orthogonal arrays of strength 3 (e.g., Resolution IV fractional factorials), if such an array exists (Section 2.1). In other cases, it appears that we can restrict attention to the class of foldover designs (Section 2.2). A simple set of rules can be used to construct $tr(\underset{\sim}{L})$-optimal foldovers (Section 2.3) for nearly all n and k of practical interest.

Within the class of regular fractional factorial designs, the $tr(\underset{\sim}{L})$-optimal designs are the resolution IV designs if $n \geq 2k$. If $n < 2k$, the optimal fractional factorials are foldovers with the fewest words of length two in the defining relation (Section 3.1). These can be easily constructed through the use of "group-factors" (Section 3.2). A comparison of $tr(\underset{\sim}{L})$-optimal fractional factorials with the $tr(\underset{\sim}{L})$-optimal foldovers, when n is a power of two, indicates that the former are either equally good or nearly as good as the latter with respect to $tr(\underset{\sim}{L})$ (Section 3.3). To choose among the optimal fractional factorial designs, we recommend the Fries-Hunter minimum aberration criterion.

The results of two different studies of power (Section 4) give a rough indication of the ability of the $tr(\underset{\sim}{L})$-optimal designs to detect the presence of interactions when a likelihood-ratio (χ^2) test of the hypothesis $\underset{\sim}{\beta}_2 = \underset{\sim}{0}$ is used, with σ^2 "known".

Designs presented in this paper have some weaknesses with respect to other design objectives. These can be overcome through augmentation of various kinds. To achieve estimability of $\underset{\sim}{\beta}_1$ when $k + 2 \leq n \leq 2k-1$, we present some "compromise" designs which have a $tr(\underset{\sim}{L})$-optimal design as a nucleus (Section 5.1). Augmentation to identify important individual interactions is illustrated by means of an example (Section 5.2). If the factors are continuous, the addition of a center point is an aid to detection of the presence of other non-linear effects, particularly quadratic terms (Section 5.3). Estimation of σ^2 can be achieved by replicating some foldover pairs, and some simple rules are given in Section 5.4 for constructing $tr(\underset{\sim}{L})$-optimal foldovers subject to the specified replication requirements.

When k is large and the number of runs in limited, some of the designs presented here are effective as preliminary designs for detecting in relatively few runs whether it is reasonable to proceed with an experimental strategy based on a first-order model. An example of this type of application, to a sensitivity analysis of a computer code, is given in Section 6.

8. Acknowledgments

The rules in Section 2.3 and Appendix B for constructing designs to minimize $\mathrm{tr}(\underset{\sim}{D}'\underset{\sim}{D})^2$ arose as a result of discussions with L. J. Gray of the Mathematics and Statistics Research Department, Computer Sciences Division, Union Carbide Corporation Nuclear Division, and with C. S. Cheng of the Department of Statistics, University of California at Berkeley. (If we have misapplied their results here, the fault is entirely ours.)

We are also grateful to R. E. Textor and K. W. Childs, of the Computer Sciences Division and L. J. Ott of the Engineering Technology Division, Union Carbide Corporation Nuclear Division, for their collaboration on the sensitivity analysis example of Section 6.

Appendix A: Maximization of $\underset{\sim}{\beta}_2'\underset{\sim}{L}\underset{\sim}{\beta}_2$ Under a Design Randomization Scheme

Consider an (n,k)-design D, with corresponding matrices $\underset{\sim}{X}_1$, $\underset{\sim}{X}_2$, and $\underset{\sim}{L}$ as defined in (1.5) and (1.7). We further define $\underset{\sim}{H} = \underset{\sim}{X}_1(\underset{\sim}{X}_1'\underset{\sim}{X}_1)^{-}\underset{\sim}{X}_1'$; thus $\underset{\sim}{L} = \underset{\sim}{X}_2'(\underset{\sim}{I}-\underset{\sim}{H})\underset{\sim}{X}_2$. For any $\underset{\sim}{x}$, $\underset{\sim}{H}\underset{\sim}{x}$ is the projection of $\underset{\sim}{x}$ onto the space spanned by the columns of $\underset{\sim}{X}_1$ and $\underset{\sim}{x}'(\underset{\sim}{I}-\underset{\sim}{H})\underset{\sim}{x}$ is the distance from $\underset{\sim}{x}$ to that space.

We propose to select the design D_R for the experiment by the following two-stage randomization scheme $R = R_1R_2$:

R_1: Randomly relabel the factors in D so that each one of the k! possible labelings has the same probability of realization.

R_2: With probability 0.5, reverse the levels of factor i in D, independently for each i = 1,2,...,k.

The matrices $\underset{\sim}{X}_{1R}, \underset{\sim}{X}_{2R}, \underset{\sim}{H}_R$, and $\underset{\sim}{L}_R$ are obtained from D_R in the same way that $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \underset{\sim}{H}$, and $\underset{\sim}{L}$ are obtained from D.

The expectation of $\lambda_R = \underset{\sim}{\beta}_2' \underset{\sim}{L}_R \underset{\sim}{\beta}_2$ under the randomization R is

$$\begin{aligned} E_R(\lambda_R) &= E_{R_1} E_{R_2|R_1} (\underset{\sim}{\beta}_2' \underset{\sim}{L}_R \underset{\sim}{\beta}_2) \\ &= E_{R_1} E_{R_2|R_1} (\underset{\sim}{\beta}_2' \underset{\sim}{X}_{2R}' (\underset{\sim}{I} - \underset{\sim}{H}) \underset{\sim}{X}_{2R} \underset{\sim}{\beta}_2) \quad . \end{aligned} \tag{A.1}$$

(The substitution of $\underset{\sim}{H}$ for $\underset{\sim}{H}_R$ is justified by the fact that the columns of $\underset{\sim}{X}_{1R}$ span the same space as the columns of $\underset{\sim}{X}_1$, so the distance from any vector to that space is invariant under the randomization.

We can express $\underset{\sim}{X}_{2R}$ as

$$\underset{\sim}{X}_{2R} = \underset{\sim}{X}_2 \underset{\sim}{P}_{R_1} \underset{\sim}{Q}_{R_2} \tag{A.2}$$

where $\underset{\sim}{P}_{R_1}$ is a permutation matrix which permutes the columns of $\underset{\sim}{X}_2$ according to R_1 and $\underset{\sim}{Q}_{R_2}$ is a diagonal matrix with diagonal elements +1 or -1 reflecting the effect of R_2 on the columns of $\underset{\sim}{X}_2 \underset{\sim}{P}_{R_1}$. A specific diagonal element of Q_{R_2} has the form $q_i q_j$ where q_i and q_j are (independently) +1 or -1 with probability 0.5. Given R_1,

$$\begin{aligned} E_{R_2|R_1} (\underset{\sim}{\beta}_2' \underset{\sim}{X}_{2R}' (\underset{\sim}{I} - \underset{\sim}{H}) \underset{\sim}{X}_{2R} \underset{\sim}{\beta}_2) &= E_{R_2|R_1} (\underset{\sim}{\beta}_2' \underset{\sim}{Q}_{R_2}' \underset{\sim}{P}_{R_1}' \underset{\sim}{L} \underset{\sim}{P}_{R_1} \underset{\sim}{Q}_{R_2} \underset{\sim}{\beta}_2) \\ &= \mathrm{tr}[\underset{\sim}{P}_{R_1}' \underset{\sim}{L} \underset{\sim}{P}_{R_1} E_{R_2|R_1} (\underset{\sim}{Q}_{R_2} \underset{\sim}{\beta}_2 \underset{\sim}{\beta}_2' \underset{\sim}{Q}_{R_2}')] \quad . \end{aligned} \tag{A.3}$$

A typical diagonal element of $E_{R_2|R_1}(\underset{\sim}{Q}_{R_2}\underset{\sim}{\beta}\underset{\sim}{\beta}'\underset{\sim}{Q}'_{R_2})$ is $E_{R_2|R_1}(q_i^2 q_j^2 \beta_{ij}^2) = \beta_{ij}^2$ and typical off-diagonal elements are $E_{R_2|R_1}(q_i q_j q_k q_\ell \beta_{ij}\beta_{k\ell}) = 0$ $(i \neq j \neq k \neq \ell)$ and $E_{R_2|R_1}(q_i^2 q_j q_k \beta_{ij}\beta_{ik}) = 0$ $(i \neq j \neq k)$. Hence (A.3) can be simplified and substituted into (A.1) to yield

$$E_R(\lambda_R) = E_{R_1}\, tr[\underset{\sim}{P}'_{R_1}\underset{\sim}{L}\underset{\sim}{P}_{R_1}\underset{\sim}{B}] = tr[\underset{\sim}{L}E_{R_1}(\underset{\sim}{P}_{R_1}\underset{\sim}{B}\underset{\sim}{P}'_{R_1})] \qquad (A.4)$$

where $\underset{\sim}{B}$ is a diagonal matrix with diagonal elements β_{ij}^2. Since $\underset{\sim}{P}_{R_1}$ is a permutation matrix, $\underset{\sim}{P}_{R_1}\underset{\sim}{B}\underset{\sim}{P}'_{R_1}$ is a diagonal matrix obtained by permuting the diagonal elements of $\underset{\sim}{B}$. Over all such permutations generated by the randomization procedure R_1, the expectation $E_{R_1}(\underset{\sim}{P}_{R_1}\underset{\sim}{B}\underset{\sim}{P}'_{R_1})$ is just $\bar{b}\,\underset{\sim}{I}$ where $\bar{b} = (\sum_{i<j}\sum \beta_{ij}^2)/k_2$. Substituting into (A.4), we finally obtain

$$E_R(\lambda_R) = \bar{b}\, tr(\underset{\sim}{L}) \quad . \qquad (A.5)$$

This result implies that $E_R(\lambda_R)$ is maximized by choosing D to maximize $tr(\underset{\sim}{L})$, regardless of the value of $\underset{\sim}{\beta}$. (The subsequent data analysis should, of course, be made conditional on the design that was actually selected.)

Appendix B: Minimization of $Tr(\underset{\sim}{D}'\underset{\sim}{D})^2$, Where $|d_{ij}| = 1$.

The following results justify the Gray-Cheng rules for constructing the $\tilde{n} \times k$ matrix $\tilde{\underset{\sim}{D}}$ in Section 2.3. For simplicity of notation, we use $\underset{\sim}{D}$ and n here instead of $\tilde{\underset{\sim}{D}}$ and $\tilde{n}$.

Let $\underset{\sim}{D}$ be an $n \times k$ matrix whose elements d_{ij} must be $+1$ or -1. We want to minimize $tr(\underset{\sim}{D}'\underset{\sim}{D})^2$, which is the sum of squares of the elements of $\underset{\sim}{D}'\underset{\sim}{D}$. Since the diagonal elements of $\underset{\sim}{D}'\underset{\sim}{D}$ are equal to n for all $\underset{\sim}{D}$, we can restrict attention to the off-diagonal elements.

We shall assume here that $k \leq n$. The results for $k > n$ follow directly from the fact that $\mathrm{tr}(\underset{\sim}{D}'\underset{\sim}{D})^2 = \mathrm{tr}(\underset{\sim}{D}\underset{\sim}{D}')^2$.

Case 1: $n \equiv 0, \mod 4$.

If $\underset{\sim}{D}$ is column-orthogonal, it is optimal, since the off-diagonal elements achieve their minimum in absolute value, 0. We then have $\mathrm{tr}(\underset{\sim}{D}'\underset{\sim}{D})^2 = kn^2$. This construction can be used whenever a Hadamard matrix of order n exists. (As of 1977, the smallest order for which a Hadamard matrix had not been constructed was 268, according to Hedayat and Wallis (1978). We are not aware of any changes in this list since then.)

Case 2: $n \equiv 1, \mod 4$.

Since the off-diagonal elements cannot be 0 in this case, it is evident that if all the off-diagonal elements of $\underset{\sim}{D}'\underset{\sim}{D}$ are +1 or -1, then $\underset{\sim}{D}$ is optimal. We can construct such a $\underset{\sim}{D}$ by augmenting an $(n-1) \times k$ column-orthogonal matrix with any row of +1's and -1's. We then have $\mathrm{tr}(\underset{\sim}{D}'\underset{\sim}{D})^2 = kn^2 + k(k-1) = k(n^2 + k-1)$. The only subcases in which this construction cannot be used (assuming a Hadamard matrix of order $(n-1)$ exists) are those in which $k = n$. Solutions for $n = 5$, 13, and 25 are given by Raghavarao (1959); we are not aware of solutions for other cases with $k = n$.

Case 3: $n \equiv 2, \mod 4$.

By Ehlich's (1964) Lemma 3.4, the maximum possible number of zeros in $\underset{\sim}{D}'\underset{\sim}{D}$ is $k^2/2$ if k is even and $(k^2-1)/2$ if k is odd. Suppose $\underset{\sim}{D}$ is formed by augmenting an $(n-2) \times k$ column-orthogonal matrix with two rows of +1's and -1's, chosen so that their inner product is 0 if k is even and +1 or -1 if k is odd. Then $\underset{\sim}{D}'\underset{\sim}{D}$ will contain the maximum number of zeros possible, and all the non-zero off-diagonal elements of $\underset{\sim}{D}'\underset{\sim}{D}$ will attain their lower bound in absolute value, 2; hence $\underset{\sim}{D}$ is optimal.

The construction above suffices when $k \leq n-2$. If $k = n$ or $k = n-1$, we can resort to a different method. By a similar argument to the one above, it can be shown that the removal of two rows from an $(n+2) \times k$ column-orthogonal matrix $\underset{\sim}{A}$, again chosen to have

inner product with absolute value 0 or 1, yields an optimal $\underset{\sim}{D}$. The only question is whether two such rows can be found in $\underset{\sim}{A}$. We first treat the case $k = n$, and assume there is no orthogonal pair of rows in $\underset{\sim}{A}$. Then the inner product of any two rows of $\underset{\sim}{A}$ has absolute value at least 2, so

$$\mathrm{tr}(\underset{\sim}{A}\underset{\sim}{A}')^2 \geq (n+2)n^2 + 4(n+2)(n+1) \qquad (B.1)$$

where we use the fact that the left hand side is equal to the sum of squares of the elements of $\underset{\sim}{A}\underset{\sim}{A}'$. But

$$\mathrm{tr}(\underset{\sim}{A}\underset{\sim}{A}')^2 = \mathrm{tr}(\underset{\sim}{A}'\underset{\sim}{A})^2 = n(n+2)^2 \qquad (B.2)$$

since $\underset{\sim}{A}$ is column-orthogonal, and it is easily shown that (B.2) and (B.1) are incompatible. $\underset{\sim}{A}$ must therefore have at least one pair of orthogonal rows. An analogous argument can be used to prove the same proposition for the case $k = n-1$.

The optimum values of $\mathrm{tr}(\underset{\sim}{D}'\underset{\sim}{D})^2$ for Case 3 can easily be shown to be $k[n^2 + 2(k-2)]$ when k is even and $kn^2 + 2(k-1)^2$ when k is odd.

Case 4: $n \equiv 3$, mod 4.

If we remove any row from an $(n+1) \times k$ column-orthogonal matrix, the resulting matrix $\underset{\sim}{D}$ will be optimal, by the same argument used for Case 2 above. As in Case 2, the optimum $\mathrm{tr}(\underset{\sim}{D}'\underset{\sim}{D})^2$ is $k(n^2 + k-1)$.

Remark: The above arguments establish lower bounds for $\mathrm{tr}(\underset{\sim}{D}'\underset{\sim}{D})^2$ even for the (sparse) pairs (n,k) for which the suggested construction is not possible.

Appendix C - On $Tr(\underset{\sim}{L})$-Optimality in the Class of Regular Fractional Factorials

Theorem C1. No 2^{k-p} fractional factorial design having words of length 1 or 3 in its defining relation can be $tr(\underset{\sim}{L})$-optimal among regular fractional factorials.

The proof is by construction of the superior 2^{k-p} design $F_2(F_1,i)$ where F_2 is obtained from a given 2^{k-p} design F_1 by folding over, for suitably chosen i, the half of F_1 in which $x_i = 1$. We shall use the notation [i], [ij], [ijℓ] to refer to first-, second-, and third-order design moments, respectively.

We note that $tr(\underset{\sim}{L})$ is the sum over all pairs $(i < j)$ of the squared distance from $\underset{\sim}{x}_{ij}$ (the column of $\underset{\sim}{X}_2$ corresponding to β_{ij}) to the space spanned by the columns of $\underset{\sim}{X}_1$. In a regular fractional factorial design this squared distance is either n (if $\underset{\sim}{x}_{ij}$ is orthogonal to $\underset{\sim}{X}_1$) or 0, so $tr(\underset{\sim}{L})$ is just n times the number of columns in $\underset{\sim}{X}_2$ that are orthogonal to $\underset{\sim}{X}_1$.

Given a 2^{k-p} design F_1 with words of length 1 or 3 in its defining relation, the construction of $F_2(F_1,i)$ with larger $tr(\underset{\sim}{L})$ is based on the following lemmas.

Lemma C1. For any i such that $[i] = 0$, any column of $\underset{\sim}{X}_2$ that is orthogonal to $\underset{\sim}{X}_1$ in F_1 is also orthogonal to $\underset{\sim}{X}_1$ in $F_2(F_1,i)$.

Proof of Lemma C1. Suppose $\underset{\sim}{x}'_{ij}\,\underset{\sim}{X}_1 = \underset{\sim}{0}$ in F_1. This implies in particular $[i] = [j] = [ij] = 0$ in F_1, so $\underset{\sim}{x}_i$ and $\underset{\sim}{x}_j$ form a 2^2 factorial design (possibly replicated), a property we shall hereafter refer to as **Property A**. It is easily seen that $\underset{\sim}{x}_i$ and $\underset{\sim}{x}_j$ also have this property in $F_2(F_1,i)$. Because $F_2(F_1,i)$ is a foldover, all its odd-order design moments are 0, so $\underset{\sim}{x}'_{ij}\,\underset{\sim}{X}_1 = \underset{\sim}{0}$ in $F_2(F_1,i)$ iff $[ij] = 0$ in $F_2(F_1,i)$. We have already established that $[ij] = 0$, so the lemma is proved for columns of the form $\underset{\sim}{x}_{ij}$. We still need to consider columns of form $\underset{\sim}{x}_{j\ell}\,(j\neq i, \ell\neq i)$, with $\underset{\sim}{x}'_{j\ell}\,\underset{\sim}{X}_1 = \underset{\sim}{0}$ in F_1. We then have $[j\ell] = 0$ and $[ij\ell] = 0$, and we recall that i was chosen such that $[i] = 0$. Thus $\underset{\sim}{x}_i$ and $\underset{\sim}{x}_{j\ell}$ have Property A in F_1, and also in $F_2(F_1,i)$, which leads as above to the result that $\underset{\sim}{x}'_{j\ell}\,\underset{\sim}{X}_1 = \underset{\sim}{0}$ in F_2.

Remark on Lemma C1. The lemma shows that for any given 2^{k-p} fractional factorial design, there exists a 2^{k-p} fractional factorial foldover design with $tr(\underset{\sim}{L})$ at least as great as that of the given design. The lemmas which follow establish the $tr(\underset{\sim}{L})$- superiority of the foldover when the given design has moments of order 1 or 3.

Lemma C2. If $k-p > 2$ and there exists j such that $[j] \neq 0$ in F_1, then there exists i such that $F_2(F_1,i)$ has greater $tr(\underset{\sim}{L})$ than does F_1.

Proof of Lemma C2. Let $\underset{\sim}{x}_i$ and $\underset{\sim}{x}_\ell$ be two columns having Property A in F_1. (Two such columns always exist when $k-p > 2$.) Since $[j] \neq 0$ in F_1, $\underset{\sim}{x}'_{j\ell}\,\underset{\sim}{X}_1 \neq \underset{\sim}{0}$ in F_1. But $[j\ell] = 0$ in F_1 and also in $F_2(F_1,i)$, which implies that $\underset{\sim}{x}'_{j\ell}\,\underset{\sim}{X}_1 = \underset{\sim}{0}$ in $F_2(F_1,i)$. The set of columns in $\underset{\sim}{X}_2$ that are orthogonal to $\underset{\sim}{X}_1$ in $F_2(F_1,i)$ therefore includes $\underset{\sim}{x}_{j\ell}$ as well as all the columns of $\underset{\sim}{X}_2$ that were orthogonal to $\underset{\sim}{X}_1$ in F_1 (by Lemma C1) so Lemma C2 is proved.

Remark on Lemma C2. In the case $k-p = 1$, which is not covered by the lemma, there are only 2 runs, and $tr(\underset{\sim}{L})$ is always 0.

Lemma C3. Let all first-order design moments in the 2^{k-p} design F_1 be 0, and suppose that F_1 has at least one non-zero third-order moment $[ijk]$. Then $F_2(F_1,i)$ has greater $tr(\underset{\sim}{L})$ than does F_1.

Proof of Lemma C3. In F_1, $[ij] = 0$ (otherwise $[k] \neq 0$); hence $\underset{\sim}{x}_i$ and $\underset{\sim}{x}_j$ have Property A in F_1 and in $F_2(F_1,i)$. Thus $[ij] = 0$ in $F_2(F_1,i)$ so $\underset{\sim}{x}'_{ij}\,\underset{\sim}{X}_1 = \underset{\sim}{0}$ in $F_2(F_1,i)$. Recall that $\underset{\sim}{x}'_{ij}\,\underset{\sim}{X}_1 \neq \underset{\sim}{0}$ in F_1 (because $[ijk] \neq 0$ there), so by the same argument used in the proof of Lemma C2, we conclude that there are more columns of $\underset{\sim}{X}_2$ orthogonal to $\underset{\sim}{X}_1$ in $F_2(F_1,i)$ than in F_1.

Proof of Theorem C1. Every word of length 1 or 3 in the defining relation corresponds to a non-zero first or third order design moment. Lemmas C2 and C3 imply that such designs always have lower $tr(\underset{\sim}{L})$ than some 2^{k-p} fractional factorial foldover, which has no words of odd length because it is a foldover.

Theorem C2. If $n < 2k$, a necessary condition for a 2^{k-p} design to be $tr(\underset{\sim}{L})$-optimal in the class of regular fractional factorials is that it be a foldover design.

Proof: With appropriate relabeling of the variables, p generators of a 2^{k-p} design can be chosen so that they have the form: $W_1(k-p+1), W_2(k-p+2),\ldots,W_p(k)$, where each W_i is a word composed of letters (variables) in the set $1,2,\ldots,k-p$. There are, in general, $n = 2^{k-p}$ candidates for each W_i (including the "word" with no letters), which we denote by C_j, $j = 1,2,\ldots,n$. If we denote by $m(C_j)$ the number of times C_j is selected as one of the W_i's, then we see that the vector $\underset{\sim}{m} = (m(C_1),\ldots,m(C_n))'$ determines the resulting 2^{k-p} design.

We first consider a design F_1 with words of odd length in the defining relation, and show that it cannot be $tr(\underset{\sim}{L})$-optimal. If there are any words of length 1 or 3 present, the result follows immediately from Theorem C1, so we need consider only designs whose first- and third-order design moments are 0. For such designs, $tr(\underset{\sim}{L}) = n(k-n_2)$, where n_2 is the number of words of length 2 in the defining relation, so $tr(\underset{\sim}{L})$- optimality is equivalent to minimization of n_2. Words of length 2 occur in two ways: (i) as generators in which W_i has length 1 and (ii) as the product of two generators having identical W_i. Thus,

$$n_2 = \sum_j m_j(m_j-1)/2 + \sum_{j\in J_1} m_j \qquad (C.1)$$

where $J_1 = \{j|C_j$ has length $1\}$. Since F_1 has words of odd length (>5) in its defining relation, any set of generators for F_1 must contain at least one word of odd length. Without loss of generality, we choose one of these odd generators and denote it by $C_{j^*}k$ where C_{j^*} has even length >4. It follows that $m_j = 0$ if C_j has length one less than C_{j^*} and C_j is contained in C_{j^*} (i.e., the length of $C_jC_{j^*}$ is 1); otherwise, there would be a word of length 3 in the defining relation. Denote one of these j's for which $m_j = 0$ by j'. Now we find j" such that (i) $m_{j"} > 2$ or (ii) $j" \in J_1$ and $m_{j"} > 1$. (Such a j" must exist; otherwise, $n_2 = 0$ by equation (C.1) and F_1 would be a resolution IV design in $n < 2k$ runs, which is impossible (Webb (1968), Margolin (1969).) Define a new 2^{k-p} design F_2 by adding 1 to $m_{j'}$ (making

it 1) and subtracting 1 from m_{j*}. From (C.1), we see that this reduces the contribution of m_{j*} to n_2 but leaves the contribution of $m_{j\cdot}$ at 0. Thus F_2 has fewer two-letter words in its defining realtion than does F_1, and so has larger $tr(\underset{\sim}{L})$.

We have thus established that in order for a 2^{k-p} design with $n < 2k$ to be $tr(\underset{\sim}{L})$-optimal, its defining relation must consist entirely of words of even length. But this is the same as requiring it to be a foldover design. (For example, it is easy to see that if a design with no odd words in its defining relation is split into two parts, according to whether $x_1 = +1$ or $x_1 = -1$, each part is the negative of the other.) Theorem C2 is therefore proved.

Remark: Theorem C2 can be extended to the case $n = 2k$. The $tr(\underset{\sim}{L})$- optimal fractional factorial design in this case is the "minimal" or "saturated" resolution IV design and must therefore be a foldover (Margolin (1969)).

Theorem C3: Given n (a power of 2) and $k > \tilde{n} = n/2$, a $tr(\underset{\sim}{L})$-optimal 2^{k-p} design in n runs is $tr(\underset{\sim}{L})$-optimal among all two-level foldover (n,k)-designs if and only if r (the remainder upon dividing k by $\tilde{n}$) is 0, 1, 2, $\tilde{n} - 1$, or $\tilde{n} - 2$.

Proof: We shall consider only the case $k \equiv 2 \pmod 4$ in detail. The argument for the other cases is similar. As noted in the proof of Theorem C2, the value of $tr(\underset{\sim}{L})$ for optimal fractional factorials is:

$$tr(\underset{\sim}{L}) = n[k(k-1)/2 - n_2] \qquad (C.2)$$

where n_2 is the number of two-letter words in the defining relation. Since n_2 is the same as the number of pairs of completely confounded factors, we can refer to the construction of Section 3.2 to obtain

$$\begin{aligned} n_2 &= r(a-1)a/2 + (\tilde{n}-r)a(a-1)/2 \\ &= (k-r)(r+k-\tilde{n})/2\tilde{n} \end{aligned} \qquad (C.3)$$

where a and r are defined by (3.1) and (3.2). Substituting (C.3) into (C.2) we have

$$tr(\underset{\sim}{L}) = k^2(\tilde{n}-1) - r(\tilde{n}-r) \qquad (C.4)$$

for the optimal fractional factorials. These designs are necessarily foldovers, since $n < 2k$ (Section 3.1). We now refer to Table 2.1, which gives, for the case $k \equiv 2 \pmod 4$ and $\tilde{n}$ even:

$$tr(\underset{\sim}{L}^*) = k^2(\tilde{n}-1) - 2(\tilde{n}-2) \qquad (C.5)$$

where $\underset{\sim}{L}^*$ is the lack-of-fit matrix for an optimal foldover. Equations (C.4) and (C.5) are the same iff $r = 2$ or $r = \tilde{n}-2$. Similar arguments, applied to the cases $k \equiv 0$, 1, or 3 (mod 4) yield Theorem C3.

Appendix D. "Minimum Aberration" As A Supplementary Criterion for Choosing Among Tr($\underset{\sim}{L}$)- Optimal Fractional Factorial Designs.

Fries and Hunter (1979) introduced the concept of aberration as an extension of resolution in classifying 2^{k-p} fractional factorial designs according to their confounding properties. Let $(n_1,n_2,\ldots,n_k)$ be the word-length pattern for the 2^{k-p} design D_1, i.e., n_i is the number of words of length i in the defining relation. Similarly, let $(n_1',n_2',\ldots,n_k')$ be the word-length pattern for another 2^{k-p} design D_2. Then D_1 has lower aberration than D_2 (which we express by $D_1 < D_2$) if and only if there exists J such that $n_j = n_j'$, $j = 1,2,\ldots,J-1$ and $n_J < n_J'$. Clearly, if $D_1 < D_2$ and $D_2 < D_3$, then $D_1 < D_3$, so the concept of aberration may be used to rank designs. The best designs under this criterion are the minimum aberration designs: D_1 is a minimum aberration design in a given class if there is no design D in such that $D < D_1$. This criterion is consistent with the more familiar "maximum resolution" criterion but is much more sensitive to differences in the structure of the aliasing (confounding) relationships.

We recommend the use of the minimum aberration criterion to supplement $tr(\underset{\sim}{L})$- optimality when choosing a 2^{k-p} design to detect the presence of interactions. Minimum aberration designs appear to be good with respect to:

(i) maximizing the number of degrees of freedom q for two-factor interactions, and

(ii) distributing the $k_2 = k(k-1)/2$ interactions evenly over the q strings of completely confounded interactions.

We shall consider the cases $n \geq 2k$ and $n < 2k$ separately.

<u>$n \geq 2k$.</u>

The $tr(\underset{\sim}{L})$-optimal 2^{k-p} designs are precisely those of resolution ≥ 4. In this case, we are unable to prove a direct relationship between minimum aberration and (i) and (ii) above, but the following results may be useful for those who wish to explore the matter further. Let h_i be the number of two-factor interactions that appear in exactly i of the n_4 4-letter words in the defining relation, $i = 0,1,\ldots,n_4$. Then

$$\sum h_i = k_2 = k(k-1)/2 \tag{D.1}$$

$$(\sum i h_i)/6 = n_4 \tag{D.2}$$

$$\sum h_i/(i+1) = q \tag{D.3}$$

$$\text{Average (string length)} = k_2/q \tag{D.4}$$

$$\text{Average [(string length)}^2\text{]} = (k_2+6n_4)/q \quad . \tag{D.5}$$

If we note that the number of strings of length $i+1$ is $h_i/(i+1)$, these results are all straightforward. One consequence of them is that, among designs with the same degrees of freedom for interactions (q), the minimum aberration design distributes the interactions "evenly" among the q strings by minimizing the dispersion of the string lengths, where we define <u>dispersion</u> to be the sum of squares deviations from the average.

$n < 2k$.

The $tr(\underset{\sim}{L})$-optimal 2^{k-p} designs in this case are foldover designs of resolution 2, having the fewest possible words of length 2 in their defining relations. (Appendix C.) The aliasing relationships for such a design include $\tilde{n} = n/2$ strings of two-factor interactions, counting the one which is completely confounded with the overall mean. If we define the length of each such string to be the number of two-factor interactions in it, then the average string length is obviously $k_2/\tilde{n}$. The following lemma provides a formula for the average squared string length.

Lemma D1. In a 2^{k-p} foldover design of resolution 2, the average squared length of the $\tilde{n}$ strings of two-factor interactions is $[k(k-1)/2 + 6n_4 + 2(k-2)n_2]/\tilde{n}$, where n_4 and n_2 are the number of 4- and 2-letter words, respectively, in the defining relation.

Proof. Let n_{ij} be the number of four-letter words that include i and j, and let n'_{ij} be the number of two-letter words that include i or j (but not both). Then β_{ij} will be confounded with exactly $n_{ij} + n'_{ij}$ other two-factor interactions, and the length of the string that includes β_{ij} is $n_{ij} + n'_{ij} + 1$. Now let v_ℓ be the number of interactions which are in a string of length ℓ. Clearly, v_ℓ/ℓ is the number of strings of length ℓ, and the sum of squared string lengths is $\sum_\ell \ell^2(v_\ell/\ell) = \sum_{ij} (n_{ij} + n'_{ij} + 1)$, where the second summation is over all pairs $1 \le i < j \le k$. Since each 4-letter word in the defining relation contributes one unit to each of 6 different n_{ij}'s, $\sum_{ij} n_{ij} = 6n_4$. Similarly, $\sum_{ij} n'_{ij} = 2(k-2)n_2$, and the lemma then follows directly.

Since k, n_2, and $\tilde{n}$ are all fixed in the class of $tr(\underset{\sim}{L})$-optimal 2^{k-p} designs with $n < 2k$, this lemma shows that the minimum aberration criterion here is equivalent to minimizing the dispersion of the lengths of the strings of two-factor interactions. (This holds true even if we omit the string which is confounded with the overall mean, since the length of that string is fixed at n_2.)

The construction of minimum aberration $tr(\underset{\sim}{L})$-optimal 2^{k-p} designs when $n < 2k$ is facilitated by the following theorem.

Theorem D1. Let $\underset{\sim}{D} = (\underset{\sim}{D}_1 : \underset{\sim}{D}_2)$ be the $n \times k$ design matrix of a 2^{k-p} fractional factorial design ($n < 2k$) such that $\underset{\sim}{D}_1$ consists of $a \geq 1$ copies of the columns of the saturated design $\underset{\sim}{D}^*$ of resolution ≥ 4 in $\tilde{n} = n/2$ factors and n runs and $\underset{\sim}{D}_2$ consists of a subset of r distinct columns of $\underset{\sim}{D}^*$. If $\underset{\sim}{D}_2$ minimizes the dispersion of the lengths of the two-factor interaction strings in the class of regular fractional factorial foldover designs in n runs and r factors, then $\underset{\sim}{D}$ has minimum aberration in the class of $\text{tr}(\underset{\sim}{L})$-optimal 2^{k-p} designs.

Proof: By the results of Section 3.2, $\underset{\sim}{D}$ is $\text{tr}(\underset{\sim}{L})$-optimal no matter which r columns of $\underset{\sim}{D}^*$ are chosen to form $\underset{\sim}{D}_2$. Now consider the strings of two-factor interactions in the aliasing relationships of $\underset{\sim}{D}$. If we ignore all interactions involving the factors in $\underset{\sim}{D}_2$, the string that is confounded with β_0 will have length $\tilde{n}a(a-1)/2$, and each of the remaining $(\tilde{n}-1)$ strings will have length $a^2\tilde{n}/2$. Now consider all interactions of the form β_{ij}, where x_i is a factor in $\underset{\sim}{D}_1$ and x_j is a factor in $\underset{\sim}{D}_2$. For fixed j, there are $\tilde{n}a$ such interactions, a of them in each string. If we do the same for all r i's, and include these interactions in the aliasing relations for $\underset{\sim}{D}$, the β_0- string will have length $ra + \tilde{n}a(a-1)/2$ and each of the remaining strings will have length $c = ra + a^2\tilde{n}/2$. The only interactions we have ignored so far are those that involve two factors from $\underset{\sim}{D}_2$. If we finally include these in the aliasing relations for $\underset{\sim}{D}$, we will add c_i, say, to each string length, $i = 1,2,\ldots,\tilde{n}$, where the c_i's are the string lengths for the (n,r)-design $\underset{\sim}{D}_2$ alone. The dispersion of the string lengths (excluding the β_0-string) is therefore the same in $\underset{\sim}{D}$ as it is in $\underset{\sim}{D}_2$. If the choice of $\underset{\sim}{D}_2$ minimizes this dispersion, then, by the remark after Lemma D1, $\underset{\sim}{D}$ has minimum aberration in the class of $\text{tr}(\underset{\sim}{L})$-optimal 2^{k-p} designs and Theorem D1 is proved.

Appendix E. The Approximate Distribution of the Likelihood Ratio Statistic for Testing the Hypothesis of No Interaction, When the Interactions are Drawn From a Normal Distribution, and a Tr($\underset{\sim}{L}$)-Optimal Fractional Factorial Design is Used.

This Appendix provides back-up material for Section 4.3, which gives the results of a power study conducted under the assumption that the "true" iteractions (β_{ij}'s) are drawn independently from a normal distribution with mean 0 and variance σ_b^2.

We restrict this discussion to tr($\underset{\sim}{L}$)-optimal 2^{k-p} designs in n runs. If w_i is the value of the contrast in the observations which estimates the $i^{\underline{th}}$ string of confounded two-factor interactions, the likelihood ratio statistic w for testing $\underset{\sim}{\ell}_2 = \underset{\sim}{0}$ in the model (1.4) is, for σ^2 "known",

$$w = n \sum_{i=1}^{q} w_i^2/\sigma^2 \qquad (E.1)$$

where q is the number of such strings. (The string of two-factor interactions that is confounded with β_0 is not included.)

Under the assumed normal distribution of the interactions and the assumption of normally distributed errors in (1.5), the w_i's are independently $N(0, \ell_i\sigma_b^2 + \sigma^2/n)$ where ℓ_i is the length of the string associated with w_i. The mean and variance of w are therefore:

$$E(w) = nS_1\sigma_b^2/\sigma^2 + q = n\rho^2 S_1/k_2 + q \qquad (E.2)$$

$$V(w) = 2\sum_i (n\ell_i\sigma_b^2/\sigma^2 + 1)^2 = 2(n^2\rho^4 S_2/k_2^2 + 2n\rho^2 S_1/k_2 + q) \qquad (E.3)$$

where $S_1 = \sum_{i=1}^{q} \ell_i$, $S_2 = \sum_{i=1}^{q} \ell_i^2$, and $\rho^2 = k_2\sigma_b^2/\sigma^2$. When $\ell_i = \ell$ for all i, w is distributed as $(n\ell\rho^2/k_2 + 1)\chi_q^2$. Although the ℓ_i's are seldom identical, they are

generally quite close, especially when the choice among $tr(\underset{\sim}{L})$-optimal designs is made using the minimum aberration criterion as recommended in Section 3.2. We would therefore expect the distribution of w to be well approximated by that of $w' = g_1\chi^2_{g_2}$, where g_1 and g_2 are chosen so that the mean and variance of w' match those of w given in (E.2) and (E.3). This approximation was the basis for the power study discussed in Section 4.3.

If k is a multiple of $\tilde{n}$, i.e., $k = m\tilde{n}$, then $\tilde{n}m(m-1)/2$ interactions will be confounded with β_o and the remaining ones will be distributed equally among $q = \tilde{n}-1$ strings. Thus $\ell_i = \ell = k^2/n$, so w is distributed as $(k^2\rho^2/k_2 + 1)\chi^2_{n-1}$, which approaches $(2\rho^2+1)\chi^2_{n-1}$ as $k \to \infty$. Using this as an approximation for the distribution of w when k is large, we find that the power will be P when $\rho^2 = (\chi^2_{n-1;\alpha}/\chi^2_{n-1;P} - 1)/2$. This result was used to find the "minimum detectable" value of ρ for large k in (4.3).

Appendix F. Tr($\underset{\sim}{L}$)-Optimality Under a Replication Restriction

Consider the construction of a $tr(\underset{\sim}{L})$-optimal foldover design under the restriction that $\tilde{n}_e$ rows of the $\tilde{n} \times k$ half-design $\tilde{D}$ are replicated once, where $\tilde{n} \le k + \tilde{n}_e$. From the discussion in Section 2, we see that the problem is the same as that of choosing the $k \times \tilde{n}_e$ matrix $\underset{\sim}{A}$ and the $k \times (\tilde{n}-2\tilde{n}_e)$ matrix $\underset{\sim}{B}$ so as to minimize $tr(\tilde{\underset{\sim}{D}}\tilde{\underset{\sim}{D}}')^2$ where $\tilde{\underset{\sim}{D}}' = [\underset{\sim}{A}:\underset{\sim}{A}:\underset{\sim}{B}]$. If $k \equiv 0$, 1, or 3 (mod 4) this problem is solved simply by choosing $[\underset{\sim}{A}:\underset{\sim}{B}]$ to be the design matrix of a $tr(\underset{\sim}{L})$-optimal half-design in k runs and $\tilde{n} - \tilde{n}_e$ factors. This can be verified by noting that the inner product of any pair of columns of $\tilde{\underset{\sim}{D}}'$ will then have its lowest possible magnitude (0 or 1), except for the $\tilde{n}_e$ pairs corresponding to the required replicated columns. If $k \equiv 2$ (mod 4), the solution is slightly more complicated. We construct $[\underset{\sim}{A}:\underset{\sim}{B}]$ by augmenting a column-orthogonal matrix with two rows $(\underset{\sim}{a}_1':\underset{\sim}{b}_1')$ and $(\underset{\sim}{a}_2':\underset{\sim}{b}_2')$, chosen so that $|2\underset{\sim}{a}_1'\underset{\sim}{a}_2 + \underset{\sim}{b}_1'\underset{\sim}{b}_2| \le 1$. (If $\tilde{n}-\tilde{n}_e = k$ or $\tilde{n}-\tilde{n}_e = k-1$, two such rows will have be removed instead. See Case 3 of Appendix B.) It can be verified that $\tilde{\underset{\sim}{D}}'$ will then have the maximum possible number of

pairs of orthogonal columns, by Ehlich's (1964) Lemma 3.4. The inner product of any of the remaining pairs of columns of $\tilde{D}'$ will achieve its lowest possible absolute value (2), except for the $\tilde{n}_e$ pairs corresponding to the required replicated columns.

Once the matrix $\tilde{D}'$ has been constructed, we simply transpose it and fold it over to obtain a tr(L)-optimal design subject to the restriction that $\tilde{n}_e$ of the foldover pairs be replicated. In practice, care is required only when $k \equiv 2 \pmod 4$; for all other values of k, the procedure is equivalent to replicating any $\tilde{n}_e$ foldover pairs of a tr(L)-optimal design.

REFERENCES

Atkinson, A. C. (1972). Planning experiments to detect inadequate regression models. Biometrika 59, 275-293.

Atkinson, A. C. and Fedorov, V. V. (1975). The design of experiments for discriminating between rival models. Biometrika 62, 57-70.

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

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

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

*Ehlich, H. (1964). Determinantenabschätzungen für binäre Matrizen. Math. Zeitschr. 83, 123-132.

Fries, A. and Hunter, W. G. (1979). Minimum aberration 2^{k-p} designs. Technical Report #551, Department of Statistics, University of Wisconsin-Madison.

*Hedayat, A. and Wallis, W. D. (1978). Hadamard matrices and their applications. Ann. Statist. 6, 1184-1238.

Jones, E. R. and Mitchell, T. J. (1975). Response surface designs for the detection of model inadequacy. UCCND-CSD-21, available from National Technical Information Service, 5285 Port Royal Rd., Springfield, VA 22161.

Jones, E. R. and Mitchell, T. J. (1978). Design criteria for detecting model inadequacy. Biometrika 65, 541-551.

Kiefer, J. (1974). General equivalence theory for optimum designs (approximate theory). Ann. Math. Statist. 2, 849-879.

*Margolin, B. H. (1969). Results on factorial designs of resolution IV for the 2^m and $2^n 3^m$ series. Technometrics 11, 431-444.

Marsaglia, G. (1972). Choosing a point from the surface of a sphere. Ann. Math. Statist. 43, 645-646.

Mitchell, T. J. (1974a). An algorithm for the construction of "D-optimal" experimental designs. Technometrics 16, 203-210.

Mitchell, T. J. (1974b). Computer construction of "D-optimal" first-order designs. Technometrics 16, 211-220.

Morris, M. D. and Mitchell, T. J. (1977). Designs for the detection of inadequacy in factorial models. ORNL/CSD/TM-30, available from National Technical Information Service, 5285 Port Royal Rd., Springfield, VA 22161.

Ott, L. J. and Hedrick, R. A. (1977). ORINC - a one-dimensional implicit approach to the inverse heat conduction problem. ORNL/NUREG-23, Oak Ridge National Laboratory, Oak Ridge, Tenn.

Plackett, R. L. and Burman, J. P. (1946). The design of optimal multifactorial experiments. Biometrika 33, 305-325.

Raghavarao, D. (1959). Some optimum weighing designs. Ann. Math. Statist. 30, 295-303.

Severo, N. C. and Zelen, M. (1960). Normal approximation to the chi-square and non-central F probability functions. Biometrika 47, 411-416.

Shah, K. R. (1960). Optimality criteria for incomplete block designs. Ann. Math. Statist. 31, 791-794.

Watson, G. S. (1961). A study of group-screening method. Technometrics 3, 371-388.

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

*referred to only in Appendices

MDM/TJM/jvs

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER #2112	2. GOVT ACCESSION NO. AD-A093 573	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Two-Level Multifactor Experiment Designs for Detecting the Presence of Interactions		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Max D. Morris and Toby J. Mitchell		8. CONTRACT OR GRANT NUMBER(s) W-7405-eng-26 DAAG29-75-C-0024 DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 4 - Statistics & Probability
11. CONTROLLING OFFICE NAME AND ADDRESS See Item 18 below		12. REPORT DATE August 1980 13. NUMBER OF PAGES 44
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

Department of Energy
Washington, D.C.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Design; experiment design; foldover design; fractional factorial design; interaction; main effects design; minimum aberration; optimal design; orthogonal array; Resolution IV design; screening design; sensitivity analysis; two-level design; $tr(\underset{\sim}{L})$-optimality.

20 ABSTRACT (Continue on reverse side if necessary and identify by block number)

A design optimality criterion "$tr(\underset{\sim}{L})$ - optimality" is applied to the problem of designing two-level multifactor experiments to detect the presence of interactions among the controlled variables. Rules are given for constructing $tr(\underset{\sim}{L})$-optimal foldover designs and $tr(\underset{\sim}{L})$ - optimal fractional factorial designs. Some results are given on the power of these designs for testing the hypothesis that there are no two-factor interactions. Modifications of the $tr(\underset{\sim}{L})$ - optimal designs to satisfy other experimental objectives (estimability of effects, detection of the presence of other nonlinear effects, estimation of the error

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ABSTRACT (Continued)

variance) are suggested. Examples are given to demonstrate the application of these designs to (i) screening for interactions, and (ii) evaluating the first-order assumption in the sensitivity analysis of a computer code.