The design of mixture experiments in the presence of covariates

Bryan Douglas Olin
Iowa State University
INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
The design of mixture experiments in the presence of covariates

Olin, Bryan Douglas, Ph.D.
Iowa State University, 1993
The design of mixture experiments in the presence of covariates

by

Bryan Douglas Olin

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of the
Requirements for the Degree of
DOCTOR OF PHILOSOPHY
Department: Statistics
Major: Statistics

Iowa State University
Ames, Iowa
1993
# TABLE OF CONTENTS

## CHAPTER 1. SOME PRELIMINARY REMARKS ABOUT MIXTURE EXPERIMENTS

1

## CHAPTER 2. DESIGN AND ANALYSIS METHODOLOGY FOR MIXTURE EXPERIMENTS

2.1 Introduction

3

2.2 Models for Mixture and Unrestricted Variables

7

2.2.1 General Models

8

2.2.2 Common Polynomial Models for Mixture Components

11

2.3 Standard Designs for Mixture Experiments

13

2.4 Orthogonality

15

## CHAPTER 3. TRANSFORMATION TO UNRESTRICTED VARIABLES

20

3.1 General Form of Transformations

21

3.2 Induced Matrices

28

3.3 Design Properties Preserved under Transformations

38

3.4 Comparison of Transformations

44

## CHAPTER 4. CONSTRUCTION OF BLOCK DESIGNS

47

4.1 Blocking in Mixture Experiments

48
5.6 Review of Construction Methods for Trend-Free Designs for Mixture Experiments ........................................ 111
5.7 The Construction of Trend-Free Designs for Mixture Experiments Through Transformation ........................................ 113
5.8 Construction of Trend-Free Designs Mixture Experiments in Constrained Regions .......................................... 117
5.9 Algorithmic Methods ................................................................. 124
  5.9.1 Random Sampling of Run Orders to Maximize $|Z'(I - P_U)Z|$ ......................................................... 126
  5.9.2 An Interchange Algorithm to Maximize $|Z'(I - P_U)Z|$ .............................................................. 128
5.10 Conclusions ........................................................................ 132

CHAPTER 6. CONCLUSIONS ................................................................. 134

BIBLIOGRAPHY __________________________________________________ 137

ACKNOWLEDGEMENTS .................................................................. 145

APPENDIX ..................................................................................... 146
CHAPTER 1. SOME PRELIMINARY REMARKS ABOUT MIXTURE EXPERIMENTS

Mixture experiments have been successfully applied in the chemical, food, pharmaceutical, and materials (steel, aluminum, etc.) industries for almost a half-century. They provide scientists with an efficient, systematic methodology for collecting data on product characteristics dependent upon the proportions of the product's constituents. Cornell (1990) discusses how the taste of a punch varies with the proportions of the constituent fruit juices. Williams and Amidon (1984) model the solubility of phenobarbital as a function of the proportion of several chemicals in solution. Cornell and Piepel (1991) document the extensive use of mixture experiments in many areas of applications. In addition, they catalog much of the research on the statistical design and analysis of mixture experiments.

Many designs for completely randomized experiments with mixtures have been developed and used. However, as in the evolution of factorial experiments, needs arise for designs where randomization is restricted, allowing the experimenter to control for extraneous sources of variation. To meet these needs in factorial experiments, researchers devised confounding and other methods allowing for the introduction of blocks; others constructed run orders of designs for factorial experiments allowing the estimators of the factorial effects to be uncorrelated with estimators of deterministic,
polynomial trend parameters. Blocking variables and deterministic trends comprise
the covariates in this paper's title.

Designs for which covariate parameters and factor effects are orthogonally esti-
mated have become important in the design of factorial experiments. Such designs
are also important in the design of mixture experiments; doubly so, perhaps, due to
the heavy use of mixture experiments in the process industries where deterministic
trends are common. Until the last decade, little substantial research has been aimed
at finding orthogonal block designs for mixture experiments or trend-free mixture
orders of mixture experiments.

This paper suggests some new methods for constructing block designs and trend-
free designs for mixture experiments. Chapter 2 introduces design and analysis issues
in mixture experiments. Chapter 3 explores the role of transformations. Some meth-
ods for constructing block designs are provided in Chapter 4. Finally, Chapter 5
discusses the design of mixture experiments in the presence of deterministic, polyno-
mial trends
CHAPTER 2. DESIGN AND ANALYSIS METHODOLOGY FOR MIXTURE EXPERIMENTS

In this chapter, the foundation is laid for discussing the design and analysis of mixture experiments. Most importantly, the model used throughout the paper is introduced in Section 2.2. The remainder of the chapter is devoted to explaining the notion of a mixture experiment and some standard mixture models and designs, as well as the idea of orthogonality.

2.1 Introduction

This section introduces mixture experiments and delineates the differences between the various types of explanatory variables that play significant roles throughout the rest of the paper.

To facilitate discussion of experimental design issues for mixture components, we begin with a typical mixture experiment scenario. In order to achieve a degree of generality, the notion of a process variable is introduced and discussed. In practice, process variables are important. However, they are not dealt with in the remainder of the paper, although the block designs of Chapter 4 provide a method of including them in designs.

Suppose a steel company’s objective is to formulate a steel composed of iron,
nickel, chromium, and molybdenum for a particular application. Explanatory variables might include the proportions of the elements. Each combination of proportions of the four different elements is called a mixture. The proportions of elements in each mixture must sum to one. To be useful, the steel should possess desirable properties such as high tensile and/or yield strength. This is a typical response surface problem: find a mixture of iron, nickel, chromium, and molybdenum producing a steel with maximum strength, or model strength as a function of these proportions.

Factorial experiments have been considered for such investigation, but require that the level (proportion) of each factor can be chosen independently of the levels of the other factors. The levels cannot be chosen independently since the proportions of the ingredients must sum to one—if all but one of the proportions are known, the last can be obtained by subtraction. Thus, traditional designs for factorial experiments and most other designs for response surface investigations are eliminated from consideration.

Because the proportions of a mixture are constrained to sum to one, the space from which to choose mixtures is no longer the usual hypercube. For our example, plausible values for the four mixture components are restricted to lie in a three-dimensional regular simplex. A $(p - 1)$-dimensional simplex is a geometric figure defined by the convex hull of $p$ vertices in $p$-dimensions that may not lie in a $(p - 2)$-dimensional subspace of $\mathbb{R}^p$. "Regular" implies that the \( \binom{p}{2} \) edges connecting each pair of the $p$ vertices have equal lengths; in general, the simplex for mixture experiments is defined by the convex hull of the vertices $(1,0,\ldots,0)$, $(0,1,\ldots,0)$, $\ldots$, and $(0,0,\ldots,1)$, called pure mixtures. In our case, the vertices defining the simplex will consist of 100% of one of the elements, that is the points $(1,0,0,0)$, $(0,1,0,0)$, $(0,0,1,0)$,
and \((0, 0, 0, 1)\). Equivalently, the simplex is defined by the constraints that the proportions in the mixture must lie between zero and one and the sum of the proportions must be one, i.e.

\[
0 \leq z_j \leq 1 \quad \text{for all} \quad j = 1, \ldots, p \quad \text{and} \quad \sum_{j=1}^{p} z_j = 1, \tag{2.1}
\]

where \(z_j\) is the proportion of component \(j\) in the mixture and \(p\) is the number of mixture components. For \(p = 2\), the region is a line segment; for \(p = 3\), an equilateral triangle; for \(p = 4\), a tetrahedron. Given constraint (2.1) and statistical inferences of interest, the design problem is to select mixtures from the simplex allowing the inferences of interest to be made. The steel example fits this framework.

Extending the steel example, suppose that the engineer, suspects that other factors affect steel strength, namely quenching time and temperature. Levels of these factors can be selected independently of one another—any temperature level may appear in combination with any amount of quenching time. Variables such as quenching time and temperature are called process variables to distinguish them from the mixture components. To reiterate, the primary difference between the two types of variables is subtle: mixture components levels are proportions constrained to add to one and cannot be changed independently of one another; process variables have no such restrictions. To study process variables effects, standard factorial designs combined with mixture designs may be used. Certain level combinations may produce undesirable results, but one of the objectives of experimentation may be to find level combinations producing high strength and level combinations which should be avoided. Designs allowing us to determine the joint effects of processing conditions and alloy (mixture) on strength need to be found, increasing the complexity of our example.
As an example, suppose a $2^{r-q}$ fractional factorial design is used to study $r$ process variables effects. Assume that the chosen design has $N_{mix}$ mixtures and at each mixture a $2^{r-q}$ fractional factorial design is run, requiring a total of $2^{r-q} \times N_{mix}$ mixtures. Alternatively, at each point of a $2^{r-q}$ fractional factorial design a mixture design consisting of $N_{mix}$ points can be run; the mixture designs need not be the same. This experiment is potentially large, possibly too large to run as a completely randomized experiment. With such a prodigious number of mixtures, there is a danger that the experimental material used to make the mixtures may not be homogeneous. Also, conducting such a large experiment may take an inordinate amount of time. Degradation of the experimental material may occur over time, inducing deterministic trends. Finally, the cost of a large experiment may be too great, calling for smaller, more parsimonious designs. Good designs account for these and other potential sources of variation and economic considerations.

Lack of homogeneity may be dealt with by breaking the experimental material into homogeneous sets called blocks; the associated variables added to the model are called blocking variables. A block may correspond to a raw material lot, to a time period such as a day, to a particular set of environmental conditions, or some group of variables which are expensive to control (c.f. Moen, Nolan, Provost, 1991).

While blocks are necessary to account for lack of homogeneity in the experimental material, they may affect the estimates of the mixture component and process variable parameters. Level combinations of the mixture components and process variables should be allocated to the blocks so that the presence of block effects has a minimal impact on the estimation of the mixture component and process variable parameters. Section 2.4 defines the term "minimal." Block designs are discussed in Chapter 4.
As with blocking variables, estimation of the mixture and process variable parameters is dependent on the absence or presence of trend parameters in the model. When designing experiments in the presence of deterministic trends a run order is selected which minimizes the impact of trend parameters on estimating the mixture component and process variable parameters. Only deterministic trends are considered; some of the proposed methods can easily be modified to include stochastic trends. Designs for mixture experiments in the presence of deterministic trends are discussed in Chapter 5.

In general, the term covariate describes both blocking and trend variables, or any variable other than a mixture component or a process variable.

The $p$ mixture components may be transformed to a set of $p-1$ unrestricted variables. Methods for designing and analyzing experiments with unrestricted variables are well-known and easy to work with, in contrast to methods for mixture components. Chapter 3 discusses transformations of mixture components into unrestricted variables.

### 2.2 Models for Mixture and Unrestricted Variables

In this section, a family of models including covariates is introduced. These models are written in terms of both mixture components and unrestricted variables. Initially, a generic model notation is given. Then, specific models are discussed, providing the reader with a flavor of typical models used in practice. Foremost among these are the Scheffé canonical polynomial models.
2.2.1 General Models

Throughout the paper it will sometimes be convenient to present the model for individual responses, $y$. Other times, particularly when discussing designs and their properties, matrix notation for the vector of responses, $y$, proves more convenient. This section presents both notations. The notation is general and accommodates process variables. Chapter 3 discusses the use of unrestricted variables, so models for unrestricted variables are also introduced.

The model for mixture components is expressed in terms of an individual response $y$:

$$ y = u'\gamma + F_d(z, w)'\tau + \epsilon, $$

(2.2)

where

- $u' = (u_1, \ldots, u_b)$ is a vector of covariate values
- $\gamma$ is a $b \times 1$ vector of covariate parameters
- $z' = (z_1, \ldots, z_p)$ is a vector of proportions (a mixture)
- $w$ is a $r \times 1$ vector of level combinations of process variables
- $F_d(z, w)$ is a $s \times 1$ vector-valued function relating the response to changes in mixture variable proportions and process variable levels; $d$ is the degree of $F$ if $F$ is a polynomial
- $\tau$ is a $s \times 1$ vector of regression parameters corresponding to $F_d(z, w)$
- $\epsilon$ is a random error term.
In matrix notation, we have

\[ y = U\gamma + Z\tau + e, \]  

(2.3)

where

- \( y \) is a \( N \times 1 \) vector of responses

- \( U \) is a \( N \times b \) matrix of covariate values, with rows \( u_i' \)

- \( Z \) is a \( N \times s \) matrix with rows \( F_d(z, w)' \)

- \( e \) is a vector of random errors.

The model in terms of the unrestricted variables is written as

\[ y = u'\gamma + G_d(x, w)'\beta + e, \]  

(2.4)

where \( u, w, \gamma, \) and \( e \) have the same interpretations as before and

- \( x' = (1, x_1, \ldots, x_{(p-1)})' \), where \( x_j \) is the level of the \( j \)-th unrestricted variable

- \( G_d(x, w) \) is a \( s \times 1 \) vector-valued function relating the response to changes in the unrestricted and process variables; \( d \) is the degree of \( G \) if it is a polynomial

- \( \beta \) is the \( s \times 1 \) vector of regression parameters corresponding to \( G_d(x, w) \).

The matrix counterpart is

\[ y = U\gamma + X\beta + e \]  

(2.5)

where \( y, U, \gamma, \beta, \) and \( e \) have the same interpretations as above and \( X \) is a \( N \times s \) matrix with rows \( G_d(x, w)' \).
Including process variables allows for their presence when needed. However, for the remainder, the \( w \) will be suppressed, indicating the absence of process variables.

The following assumptions about the error terms are made as appropriate:

- the errors are independent

- \( E(e) = 0 \)

- \( E(ee') = \sigma^2 I. \)

Additional assumptions about \( e \) will be stated when necessary.

Finally, let \( S \) denote the design space for the mixture experiment, either the \( p \) dimensional simplex or some subregion of the simplex.

Inference centers on \( \tau \) or \( \beta \). We consider \( \gamma \) to be a vector of nuisance parameters for two reasons. Modeling the response as a function of \((z, w)\) or \((x, w)\) is of primary interest. Also, in the case of blocking, there is usually a recognition that block contrasts may be important due to inhomogeneity of the experimental material. Inferences about the block effects to determine their magnitudes are primarily to check if the suspected block effects are indeed significant. A situation where inferences about the block contrasts may be of importance is if the blocks represent level combinations of "noise" variables. It is desirable that the estimators of estimable functions of \( \tau \) and \( \gamma \) be uncorrelated (see Section 2.4). The same arguments may be applied to deterministic trends. Deterministic time trends may be induced by lurking noise variables that are highly correlated with time, position, or space.
2.2.2 Common Polynomial Models for Mixture Components

As mentioned earlier, mixture components have a redundancy which has implications with regards to the model matrix rank.

Consider a small example for $p = 3$ and $d = 2$, i.e.

$\mathbf{F}_2(\mathbf{z})' = (1, z_1, z_2, z_3, z_1z_2, z_1z_3, z_2z_3, z_1^2, z_2^2, z_3^2)$.

Because of constraint (2.1), the following equalities hold:

\[ z_1 + z_2 + z_3 = 1 \]
\[ z_1^2 = z_1z_2 - z_1z_3. \]  

(2.6)

Similar expressions hold for $z_2^2$ and $z_3^2$. The model is overparametrized; the largest attainable rank of $Z_2$ is six. In the overparametrized model there are 10 parameters, so $\tau$ is non-estimable.

In general, a $d$-th order polynomial in $p$ mixture components contains $\binom{p+d}{d}$ terms; the maximum possible rank of $Z$ is only $\binom{p+d-1}{d}$.

Scheffé (1958,1963) proposed reparametrizing the $d$-th-order polynomial so that it includes only $\binom{p+d-1}{d}$ terms, reflecting (2.1). These polynomial models are collectively known as the Scheffé or canonical polynomial models.

Rather than expressing a general form for the canonical polynomials, only the form of the first- and second-order models are given.

The first-order canonical polynomial in $p$ components is written as

\[ \mathbf{F}_1(\mathbf{z})'\mathbf{\tau} = \sum_{j=1}^{p} \tau_j z_j. \]  

(2.7)

The second-order Scheffé polynomial is

\[ \mathbf{F}_2(\mathbf{z})'\mathbf{\tau} = \sum_{j=1}^{p} \tau_j z_j + \sum_{j=1}^{p-1} \sum_{k=j+1}^{p} \tau_{jk} z_j z_k. \]  

(2.8)
Parameter interpretation for canonical polynomials is relatively straightforward. For both first- and second-order models, $\tau_j$ represents the expected response at the vertex where $z_j = 1$. The $\tau_{jk}$ correspond to non-linear blending effects, called synergistic or antagonistic blending of $z_j$ and $z_k$. Synergism and antagonism indicate the joint effect that both mixture components exert on the response when both are present in the mixture, e.g. if a large response is desirable and $\tau_{jk}$ is positive (negative), $z_j$ and $z_k$ are said to be synergistic (antagonistic) i.e. their joint presence has a desirable (undesirable) effect on the response. Terms of order greater than 2, in higher order models are also called synergisms or antagonisms (Scheffé 1963), but interpretations are more complex.

When $S$ is a subregion of the simplex defined by the constraints

$$0 \leq L_j \leq z_j \leq U_j \leq 1, \quad \text{for } j = 1, \ldots, p$$

or, more generally by constraints of the form

$$C \leq \sum_{j=1}^{p} A_j z_j \leq D,$$

such interpretations of the parameters may be invalid or not meaningful, depending upon the constraints. Often the original mixture components are transformed into pseudocomponents (see Cornell, 1990, ch. 4). Interpretations similar to those in the previous paragraph may be appropriate for the model fitted to the pseudocomponents. When transformed back to the original mixture components, the original model parameters are interpreted as functions of the parameters of the model fitted to the pseudocomponents.
2.3 Standard Designs for Mixture Experiments

Given a hypothesized model such as those in the previous section, a design must be formulated that allows model parameters or functions of the model parameters to be estimable. This section gives some "standard" designs for mixture experiments: simplex lattice, simplex centroid designs, and extreme vertices designs. The first two are generally appropriate only when $S$ is the entire simplex. These designs may often be modified for constrained regions, see Cornell (1990, ch. 5).

Scheffé (1958, 1963) formulated two types of designs: the $\{p, d\}$ simplex lattice designs and the $p$-component simplex-centroid design.

The most widely used family of designs for mixture experiments are the $\{p, d\}$-simplex lattice designs. In this notation, $p$ refers to the number of mixture components and $d + 1$ is the number of levels that the mixture components may assume. Each mixture component assumes levels from the set: $\{0, \frac{1}{d}, \frac{2}{d}, \ldots, 1\}$. A $\{p, d\}$ simplex lattice design has $\binom{p+d-1}{d}$ mixtures. Under a $\{p, d\}$ simplex lattice design, all $\binom{p+d-1}{d}$ parameters of the $d$-th order canonical polynomial are individually estimable. However, when fitting a $d$-th order canonical polynomial using responses generated from mixtures in a $\{p, d\}$ simplex lattice, there are no degrees of freedom available for estimating $\sigma^2$ or testing lack of fit, without additional runs. Kiefer (1961) proved that the $\{p, 1\}$ and $\{p, 2\}$ simplex lattice designs are D-optimal (see the Appendix) for first- and second-order canonical polynomials, respectively.

As an example, consider the case of a $\{3, 3\}$ simplex lattice design. The design points are:

$$(1, 0, 0), \left(\frac{2}{3}, \frac{1}{3}, 0\right), \left(\frac{1}{3}, \frac{2}{3}, 0\right), (0, 1, 0), \left(0, \frac{2}{3}, \frac{1}{3}\right), \left(0, \frac{1}{3}, \frac{2}{3}\right), (0, 0, 1), \left(\frac{1}{3}, 0, \frac{2}{3}\right), \left(\frac{2}{3}, 0, \frac{1}{3}\right), \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) .$$
Simplex centroid designs are another popular class of design which support fitting a subset of the $d$-th order canonical polynomials. These designs have a total of $2^p - 1$ mixtures: the $p$ vertices $(1,0,0,...,0)$, the $\binom{p}{2}$ binary blends where $z_j = \frac{1}{2}, z_k = \frac{1}{2}$, the $\binom{p}{3}$ ternary blends where three of the mixture components are $\frac{1}{3}$, ..., and the overall centroid, $(\frac{1}{p},\ldots,\frac{1}{p})$.

Under a $p$ component simplex centroid design, the $2^p - 1$ parameters of the model

$$y = \sum_{j=1}^{p} \tau_j z_j + \sum_{j=1}^{p} \sum_{k=j+1}^{p} \tau_{jk} z_j z_k + \ldots + \tau_{123\ldots p} z_1 z_2 z_3 \cdots z_p$$

are all estimable.

For example, take $p = 3$. Then the simplex centroid design consists of the mixtures $(1,0,0), (0,1,0), (0,0,1), (\frac{1}{2}, \frac{1}{2}, 0), (\frac{1}{2}, 0, \frac{1}{2}), (0, \frac{1}{2}, \frac{1}{2}),$ and $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$.

Constrained regions require different strategies for design construction. Often, simplex lattice or simplex centroid designs will not suffice because the mixtures in these standard designs are outside of the constrained region. Instead, designs typically consist of mixtures corresponding to the vertices of the region and the centroids of $c$-dimensional faces, $c = 2,3,\ldots,p-1$. A subset of these mixtures may be selected, depending on the hypothesized model, perhaps using an optimality criterion such as those discussed in the Appendix. Such designs are called extreme vertices designs. Snee and Marquardt (1974,1976) discuss methods of designing and analyzing screening designs (for first-order models) in constrained regions, as does Piepel (1990,1991). Snee (1975) gives strategies for choosing mixtures from a constrained region so that the parameters of the second-order polynomial are estimable. Cornell (1990, ch.4) summarizes the recommendations of these authors.

While both the simplex lattice and simplex centroid designs are versatile and widely used, they do not meet the needs of every situation, as the previous para-
graph shows. Using design algorithms that select runs that maximize optimality criteria provides a more flexible approach to the construction of designs for mixture experiments.

2.4 Orthogonality

Orthogonality is a notion that is useful in areas other than mixture experiments. To introduce it in its full generality, we assume that the generic model

$$y = S\alpha + T\delta + e, \quad (2.9)$$

holds, where $S$ is $N \times t$ and $T$ is $N \times q - t$. Suppose that $\text{rank}(S) = t$ and $\text{rank} [T'(I - P_S)T] = q - t$, are attainable for some designs where $P_S = S(S'S)^{-1}S'$. Let $\mathcal{R}$ be the design space. Suppose that $1$ is a column of $S$ so that the overall mean $\mu$ is contained in $\alpha$.

**Definition 2.4.1** A design is said to be orthogonal under model (2.9) for two parameters $\alpha$ and $\delta$ if one of the following equivalent conditions holds:

1. $SS(\delta|\mu) = SS(\delta|\alpha),$

2. $P_1T = P_S T$

3. The best linear unbiased estimator of any estimable linear function of $\delta$, say $\lambda'\delta$, is the same under the model including all of $\alpha$ as it is under a model with only $\mu$.

When trying to find designs orthogonal for parameters $\alpha$ and $\delta$, it is useful to look at the efficiencies of parameter estimators for comparative purposes.
**Definition 2.4.2** Suppose the model (2.9) holds for a specified $S$ and $T$. Then the efficiency of a vector of linearly independent, estimable functions of $\delta$, say $\Lambda \delta$, is the ratio of the determinant of the covariance matrix of $\Lambda \delta$ under a model with only $\mu$ to the determinant of the covariance matrix of $\Lambda \delta$ with $\alpha$ in the model, i.e.

$$
\text{eff}(\Lambda \delta) = \left( \frac{\left| \Lambda [T'(I - P_S)T]^{-1} \Lambda' \right|^\frac{1}{q-t}}{\left| \Lambda [T'(I - P)T]^{-1} \Lambda' \right|} \right)
$$

(2.10)

Note that when $\Lambda$ is $1 \times (q-t)$, this reduces to a ratio of variances.

Efficiency has the property that

$$
\text{eff}(\Lambda \delta) \leq 1.
$$

The motivation for efficiency is that if a design is orthogonal for $\alpha$ and $\Lambda \delta$, then the efficiency of $\Lambda \delta$ is one. Thus, when attempting to find an orthogonal or nearly-orthogonal design for $\alpha$ and $\Lambda \delta$, we might consider orderings of the mixtures maximizing $| \Lambda [T'(I - P_S)T]^{-1} \Lambda' |$, or equivalently maximizing the efficiency. Such a criterion is considered in Sections 4.6.2 and 5.9.1. Efficiency is used elsewhere to compare designs or elucidate properties of a given design.

Now recall model (2.3). Let,

$$
V = [1, \tilde{V}] = U [1_N, M_2] = UM,
$$

where $\text{rank}(\tilde{V}) = b - 1$, $\tilde{V}'1 = 0$, and $M$ is non-singular. Then,

$$
P_U = P_V,
$$

leading to the following theorem.
Theorem 2.4.1 A design for mixture components is orthogonal for \( \tau \) and \( \gamma \) if and only if

\[ \hat{V}'Z = 0, \]

for \( \hat{V} \) as in the previous paragraph.

Proof:

\[
P_U = P_V
= [1, \hat{V}] \begin{pmatrix} \frac{1}{N} & 0 \\ 0 & (\hat{V}'\hat{V})^{-1} \end{pmatrix} \begin{bmatrix} 1' \\ \hat{V}' \end{bmatrix}
= P_1 + P_{\hat{V}}
\]

Thus,

\[ P_UZ = P_1Z \]

if and only if

\[ P_{\hat{V}}Z = 0. \]

This is equivalent to

\[ \hat{V}'Z = 0. \]

Theorem 2.4.1 gives a very easy condition to check for orthogonality. It implies that a design is orthogonally blocked if and only if,

\[ | Z'Z | = | Z'(I - P_{\hat{V}})Z |, \]

or equivalently

\[ Z'Z = Z'(I - P_{\hat{V}})Z, \]
providing the impetus for arranging the mixtures so that $|Z'(I-P_{\gamma})Z|$ is maximized; see Section 4.6.

Note that when $U$ consists of orthogonal polynomial coefficients as in Chapter 5, we can analogously consider a design to be trend-free if

$$|Z'Z| = |Z'(I-P_U)Z|,$$

or equivalently

$$Z'Z = Z'(I - P_U)Z.$$

Now consider interpretations of orthogonality regarding the correlations of ordinary least squares estimators of estimable functions $\Lambda\gamma$ and $\Delta\tau$, where $\Lambda = AU$ for $A1_N = 0_b$ and $\Delta = BZ$. In the case that $U$ is a matrix of blocking variables, the class of estimable functions $\Lambda\gamma$ includes block contrasts and all other estimable functions of $\gamma$. If $U$ is a matrix of orthogonal polynomial coefficients, $\Lambda\gamma$ can be the entire vector $\gamma$ since every column of $U$ is orthogonal to $1_N$. Using Theorem 2.4.1, we show that if a design is orthogonal for $\gamma$ and $\tau$, then the ordinary least squares estimators $\Lambda\hat{\gamma}$ and $\hat{\Delta}\tau$ are uncorrelated.

Suppose that model (2.3) holds. Then by the definition of orthogonality we have the following

$$cov(\Lambda\hat{\gamma}, \hat{\Delta}\tau) = -\sigma^2A(U'U)^{-1}U'Z[Z'(I-P_U)Z]^{-}\Delta'$$

$$= -\sigma^2AP_UZ[Z'(I-P_U)Z]^{-}\Delta'$$

$$= -\sigma^2AP_UZ[Z'(I-P_1)Z]^{-}\Delta'$$

$$= 0.$$

Thus, in either case the estimators of $\Delta\tau$ and $\Lambda\gamma$ are uncorrelated. When $\Lambda\hat{\gamma}$ and $\Delta\hat{\tau}$ are uncorrelated, there will be no confounding of $\Delta\tau$ and $\Lambda\gamma$, making it easy
to assess the influence they exert on the response. In the extreme case when $\Lambda \hat{\gamma}$ and $\Delta \hat{\tau}$ are completely correlated, it will be impossible to determine to what significance should be attributed, $\Delta \tau$ or $\Lambda \gamma$. This is certainly a less than ideal situation since the primary purpose is to determine the effect that the mixture components have on the response.

Orthogonality should not be a primary criterion. Orthogonality ignores the magnitude of the estimated variances of parameter estimators, so it may be most appropriate as a secondary criterion, albeit an important one. Information may still be gained if the correlation between $\Lambda \hat{\gamma}$ and $\Delta \hat{\tau}$ is not too severe. Given a set of mixtures selected by some criterion, one might consider orderings or allocations of these mixtures which minimize the correlation between $\Lambda \hat{\gamma}$ and $\Delta \hat{\tau}$. This strategy is one principal topic of future sections concerned with selecting designs when blocking variables and deterministic trends are the covariates.
CHAPTER 3. TRANSFORMATION TO UNRESTRICTED VARIABLES

Because mixture components are restricted by constraint (2.1), traditional methods for designing and analyzing experiments with several factors cannot be applied. As a result, interest developed in devising transformations mapping the $p$ mixture components into $p - 1$ unrestricted variables. This simplifies analysis and allows the rich class of designs for unrestricted variables to be tapped—with caution.

Initially, transformations allowed the use of standard regression methodology for data subject to constraint (2.1). Fitting an overparametrized model in the mixture components, i.e. one including an intercept and other terms rendered redundant by the restriction was computationally intensive because it required the computation of a generalized inverse. Canonical polynomials had just been introduced and were still being assimilated into mainstream statistical practice. Both Claringbold (1955) and Becker (1969) formulated transformations for analysis purposes.

Transformations can also play a role in designing mixture experiments. A design in the $p - 1$ unrestricted variables can be constructed according to some criterion and then mapped back into the mixture components. The resulting design for the mixture experiment may have many of the same design optimality properties of the design in the unrestricted variables, as will be evident in Section 3.3. Among the first to exploit
transformations for design purposes were Draper and Lawrence (1965), Murty (1966), and Thompson and Myers (1969). Draper and Lawrence found minimum integrated bias and minimum integrated mean squared error designs for unrestricted variables, noting that the two optimality properties are invariant under non-singular linear transformations. Thompson and Myers and Murty explored the possibilities of using fractional factorial and response surface designs in the unrestricted variables as the basis for obtaining designs for mixture experiments. Becker (1969) briefly discussed this usage of transformations.

Runs in designs for unrestricted variables must be selected from the region in \( p - 1 \) dimensional space that is the image of the simplex (or some subregion) under the transformation used. Otherwise, when the runs are transformed into mixtures, the mixtures will lie outside the simplex.

This chapter examines several common transformations regarding their suitability for creating designs and develops a general framework encompassing all the transformations that are discussed. Section 3.3 demonstrates the invariance of several of the optimality criteria discussed in the Appendix under the transformations given in this chapter. Chapters 4 and 5 exploit this invariance, using transformations as a means of constructing orthogonal block designs and trend-free designs for mixture experiments.

### 3.1 General Form of Transformations

Transformations map a mixture, \( z \), into a \( p \times 1 \) vector, \( x \), containing the levels of the \( p - 1 \) unrestricted variables and a one in the remaining position of the vector.
This is a simple linear transformation of the form

\[ x' = z'Q^{-1}, \]

where \( Q^{-1} \) is a non-singular \( p \times p \) matrix with \( 1_p \) as a column.

As an example, consider the following transformation for \( p = 4 \). Let,

\[ (1, x_1, x_2, x_3) = (z_1, z_2, z_3, z_4)Q^{-1}, \]

with

\[ Q^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}. \]

Under this transformation, the centroid \((0.25, 0.25, 0.25, 0.25)\) is mapped to \((1, 0.25, 0.25, 0.25)\), the vertex \((1, 0, 0, 0)\) goes to \((1, 0, 0, 0)\), and the other vertices are mapped to \((1, 1, 0, 0)\), \((1, 0, 1, 0)\), and \((1, 0, 0, 1)\), respectively.

The same type of transformation as given above exists for any \( p \), i.e.

\[ Q^{-1} = \begin{pmatrix} 1 & 0'_{p-1} \\ 1_{p-1} & I_{p-1} \end{pmatrix}. \]

Finding suitable (e.g. factorial, etc...) designs in the unrestricted variables requires an understanding of the mechanics of each transformation.

Let \( z' \) be an arbitrary point in the simplex and \( x = (1, x_1, \ldots, x_{p-1})' \). In general, each transformation has the form

\[ z' = x'Q, \]

(3.1)
where $Q = (q_{ij})$ ($i = 0, \ldots, p - 1$ and $j = 1, \ldots, p$) is a non-singular, $p \times p$ matrix. In particular, for each of the transformations examined in this chapter, $Q$ may be expressed as

$$Q = \begin{pmatrix} q_0 \\ Q_1 \end{pmatrix}$$

(3.2)

where

$$\sum_{j=1}^{p} q_{0j} = 1.$$  

(3.3)

Expressing $z_j$ as a function of the $x'_s$ and $q'_s$,

$$z_j = q_{0j} + \sum_{k=1}^{p-1} q_{kj} x_k$$

(3.4)

and

$$-q_{0j} \leq \sum_{k=1}^{p-1} q_{kj} x_k \leq 1 - q_{0j}.$$  

(3.5)

Combining equations (3.3) and (3.4) with constraint (2.1) gives

$$\sum_{j=1}^{p} \sum_{k=1}^{p-1} q_{kj} x_k = 0.$$  

(3.6)

The relations developed above can be used to find conditions on $x$ ensuring that the corresponding transformed vector $z$ meets constraints 2.1. Equation (3.6) is true regardless of the values of the $x'_s$. When $\sum_{j=1}^{p} q_{kj} = 0$, equation (3.6) simplifies, allowing any choice to be made for $x$. Unfortunately, this is not the case for any of the transformations discussed in this chapter. Without such simplifications, additional conditions must be placed on the $x'_s$. A general necessary and sufficient condition on the $x'_s$ is unknown, making design construction for mixture experiments using unrestricted variables difficult. Some statements may be made about specific transformation. As we shall see, for Thompson and Myers's or Claringbold's
transformations, a weak sufficient for constraints 2.1 to be satisfied is that the level combinations of the unrestricted variables must lie in the region defined by

\[ \hat{x}'\hat{x} \leq 1, \]

where \( \hat{x} = (x_1, \ldots, x_{p-1}) \) is the vector of unrestricted variables. Nigam (1973) gave another sufficient condition on the unrestricted variables for Claringbold's transformation which utilized the relations given above. Thus, relationships (3.4) and (3.6) can aid the construction of mixture designs, determining the range of permissible values that the \( x_i' \)s may assume—but seemingly only on a case by case basis.

Given a design for a first-order model in the unrestricted variables and equation (3.1), we have the relationship

\[ F_1(z)' = G_1(x)'Q \quad (3.7) \]

for non-singular \( Q \), where \( F_1(z)' \) is a row in the model matrix for a first-order model in the mixture components and \( G_1(x)' \) is a row in the model matrix for a first-order model in the unrestricted variables. As a consequence \( C(Z) = C(X) \), so

\[ Z1 = 1 = XQ1, \quad (3.8) \]

where \( Z \) and \( X \) are the model matrices with rows \( F_1 \) and \( G_1 \), respectively. Essentially, \( Q \) reparametrizes the first-order canonical polynomial into a model which includes an intercept and \( p - 1 \) unrestricted variables.

An important property of some transformations is that the centroid of the design space in the mixture variables is the origin in the space of the unrestricted variables. This is a desirable property when the centroid represents usual operating conditions.
or when the design space is a constrained subspace of the simplex. In both cases, primary interest lies in assessing the effect of moving away from the centroid.

Section 3.2 shows that equation (3.7) implies the existence of a non-singular linear relationship between a \( d \)-th order polynomial in the mixture components and a \( d \)-th order polynomial in the unrestricted variables. Section 3.3 uses this result to prove the invariance of various design optimality criteria under transformations of same form as equation (3.7).

Becker (1969), Thompson and Myers (1968), and Draper and Lawrence (1965) propose various forms of \( Q \).

Becker (1969) used transformations primarily for analysis purposes, enabling him to use standard regression models rather than the overparametrized polynomials or canonical polynomials. He notes that any transformation that rotates the simplex so that the \( j \)-th mixture component axis is orthogonal to the rotated simplex is acceptable. However, some transformations are easier to implement than others. Becker chooses \( Q_1 \) as

\[
Q_1 = \left( I_{p-1} - \frac{1}{p+\sqrt{p}} 1_{p-1}1_{p-1}' + \frac{1}{\sqrt{p}} 1_{p-1} \right),
\]

where \( z_0 \in S \) is some reference point. This expression actually contains two transformations; one chooses either \( \frac{1}{p+\sqrt{p}} \) or \( \frac{1}{p-\sqrt{p}} \), requiring the use of \( -\frac{1}{\sqrt{p}} 1_{p-1} \) or \( \frac{1}{\sqrt{p}} 1_{p-1} \), respectively.

Thompson and Myers (1968) consider a different scenario. They assume that \( S \) is an ellipsoidal subregion of the simplex having the form

\[
\sum_{j=1}^{p} \left( \frac{z_j - z_{0j}}{h_j} \right)^2 \leq 1,
\]
where \( h_j \) is the half-length of the ellipsoid's axis corresponding to the \( j \)-th mixture component and \( z_0' = (z_{01}, \ldots, z_{0p}) \) represents the center of the region of interest. Note that this region may be expressed as

\[
S = \{ z : [z - z_0]' H^{-2} [z - z_0] \leq 1, \ 0 \leq z_j \leq 1, \ \sum_{j=1}^{p} z_j = 1 \},
\]

where \( H \) is a diagonal matrix with elements \( h_j \). This is equivalent to the spherical region in \( \mathbf{x} = (x_1, \ldots, x_{p-1}) \),

\[
\mathbf{x}'\mathbf{x} \leq 1,
\]

where

\[
\mathbf{x}' P = (z - z_0)' H^{-1}
\]

and \( P \) is such that \( PP' = I_{p-1} \).

The transformation suggested by Thompson and Myers fits within the general framework; their choice of \( Q_1 \) reflects the region of interest and has elements which are functions of the half-lengths. In this case,

\[
Q_1 = PH
\]

and \( q_0 = z_0 \). The choice of the \((p-1) \times p\) matrix \( P \) is not unique except that it must be orthogonal to the vector \((h_1, \ldots, h_p)'\) and \( PP' = I_{p-1} \). Thompson and Myers gave the general form of a matrix described as follows. The first row is always chosen to be the normalized version of \((-h_2, h_1, 0, \ldots, 0)\). Succeeding rows have elements that are functions of the \( h_j \)'s, chosen so that each row is orthogonal to previous rows. In the \( j \)-th row, entries beyond the \((j + 1)\)-st column are zero. As an example, consider
the case of \( p = 5 \). \( P \) is obtained by normalizing the rows of the orthogonal matrix

\[
\begin{pmatrix}
-h_2 & h_1 & 0 & 0 & 0 \\
 h_1 h_3 & h_2 h_3 & -(h_1^2 + h_2^2) & 0 & 0 \\
 h_1 h_3 h_4 & h_2 h_3 h_4 & h_3^2 h_4 & -h_3(h_1^2 + h_2^2 + h_3^2) & 0 \\
 h_1 h_3 h_4 h_5 & h_2 h_3 h_4 h_5 & h_3^2 h_4 h_5 & h_3 h_4^2 h_5 & -h_3 h_4(h_1^2 + h_2^2 + h_3^2 + h_4^2)
\end{pmatrix}.
\]

Claringbold's (1955) transformation may be arrived at as a special case of Thompson and Myers by setting \( h_j = 1 \) for \( j = 1, \ldots, p \) and pre- and post-multiplying \( P \) by permutation matrices which reverse the ordering of rows and columns.

The two previous transformations allow \( z_0 \) to be any point in the simplex or subregion thereof. In the next transformation, the only allowable value of \( z_0 \) is the centroid (i.e. \( \frac{1}{p} 1_p \)), rendering the transformation unsuitable when the design space is anything other than the entire simplex or a subregion centered at the centroid. The transformation due to Draper and Lawrence (1965) involves Hadamard matrices.

**Definition 3.1.1** A \( p \times p \) matrix \( H \) with elements 1 and \(-1\) where

\[
H' H = p I_p
\]

is called a Hadamard matrix of order \( p \).

A necessary condition for the existence of a \( p \times p \) Hadamard matrix is that \( p = 2 \) or \( p \equiv 0 \) (mod 4). Seberry and Yamada (1992) discuss the construction and properties of Hadamard matrices in detail.

Draper and Lawrence (1965b) use

\[
Q = \frac{1}{p} H,
\]
where $H$ is a Hadamard matrix of order $p$ with a first row of ones. As an example, consider the case of $p = 4$.

$$H = \begin{pmatrix}
1 & 1 & 1 & 1 \\
1 & 1 & -1 & -1 \\
1 & -1 & 1 & -1 \\
-1 & 1 & 1 & -1
\end{pmatrix}.$$  

This transformation conforms to the others—the first row is the reference point, and the lower sub-matrix is a $(p - 1) \times p$ matrix. In principle, any Hadamard matrix with all ones in the first row may be used. Draper and Lawrence chose a Hadamard matrix having a special interpretation. The last three columns of $H'$ form the half fraction $I = -x_1x_2x_3$ of a $2^3$ factorial design. Because Hadamard matrices exist only when the number of mixture components is a multiple of four ($p = 2$ excepting), this transformation suffers in comparison to the other transformations which exist for any number of mixture components.

### 3.2 Induced Matrices

Before showing how the various optimality criteria may be transmitted via the transformations of the previous sections, the tools to prove such assertions need to be developed. This section does so, introducing the necessary concept of induced vectors and matrices. Also, this section details the relationship between $d$-th order polynomials in the mixture components and $d$-th order polynomials in the unrestricted variables, establishing the existence of a matrix $Q_{(d)}$ such that

$$F_d(z)' = G_d(x)'Q_{(d)}$$
when a non-singular matrix $Q$ exists such that $z' = x'Q$. Two key theorems on induced matrices are given. These theorems provide the engine that drives all of the results on the invariance of the various optimality criteria.

The following notation proves to be useful in the definition of induced matrices. Let $C = (c_{ij})$ be a $m \times n$ matrix. $C(i \mid j)$ is the sub-matrix of $C$ obtained by deleting row $i$ and column $j$. Let $S_d(m)$ be the set of all non-decreasing sequences of length $d$ chosen from the integers $1, \ldots, m$. Suppose that $i \in S_d(m)$ and $\kappa \in S_d(n)$. Define $C[i \mid \kappa]$ to be the matrix constructed from $C$ using rows numbered by $i$ and columns numbered by $\kappa$. Finally, let $\nu(i)$ be the product of the factorials of the multiplicities of the distinct integers in the sequence $i$.

The definition of an induced matrix requires the notion of a permanent.

**Definition 3.2.1** Let $C$ be a $m \times n$ ($m \leq n$) matrix. Then the **permanent** of $C$, denoted by $\text{per}(C)$, is defined as

$$\text{per}(C) = \sum_{\sigma} c_{\sigma(1)}c_{\sigma(2)}c_{\sigma(3)} \cdots c_{\sigma(m)},$$

where $\sigma$ is a one-to-one function from $\{1, 2, \ldots, m\}$ to a subset of $\{1, 2, \ldots, n\}$ of size $m$ and the summation extends over all such functions $\sigma$.

One property of the that will prove useful is that $\text{per}(C) = \text{per}(C')$, see Minc (1978) or Marcus and Minc (1992). For instance, this allows us to talk about induced vectors as indicated after Definition 3.2.2.

A permanent is a "positive" determinant—suitably generalized to non-square matrices. A type of Laplace development similar to that for determinants exists for permanents, e.g. for any $i$, $1 \leq i \leq m$,

$$\text{per}(C) = \sum_{j=1}^{n} c_{ij} \text{per}(C(i \mid j)).$$
For more information on permanents, see Minc (1978).

Consider the following example. Suppose that

\[
C = \begin{pmatrix}
0 & 1 & 1 & 2 \\
1 & 1 & 0 & 1 \\
0 & 0 & 1 & 1
\end{pmatrix}.
\]

Then, expanding by the first row,

\[
\text{per}(C) = 1 \times \text{per} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} + 1 \times \text{per} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} + 2 \times \text{per} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\]

\[
= 1 \times (1 \times 2 + 1 \times 1) + 1 \times (1 \times 1 + 1 \times 1) + 2 \times (1 \times 1 + 1 \times 1)
\]

\[
= 9
\]

**Definition 3.2.2** Let \( C \) be a \( m \times n \) matrix. The \( d \)-th order induced matrix of \( C \), denoted by \( C_{(d)} \), is the \( \binom{m+d-1}{d} \times \binom{n+d-1}{d} \) matrix with entries obtained by labeling the rows by elements of \( S_d(m) \) in lexicographical order and labeling the columns by elements of \( S_d(n) \) in lexicographical order. The \((\iota,\kappa)\) entry of \( C_{(d)} \) is given by

\[
C_{(d)}[\iota,\kappa] = \frac{\text{per}(C[\iota | \kappa])}{\sqrt{\nu(\iota)\nu(\kappa)}}.
\]  

(3.9)

If \( C \) in Definition 3.2.2 is \( m \times 1 \), we will call \( C_{(d)} \) the \( d \)-th order induced vector of \( C \).

The next three theorems provide useful results on induced matrices that will be used to prove the main result of this section.

**Theorem 3.2.1** Let \( a \) be a \( m \times 1 \) vector and \( a_{(d)} \) the \( \binom{m+d-1}{d} \times 1 \) \( d \)-th order induced vector of \( a \). Then,

\[
a_{(d)} a_{(d)} = (a^t a)^d.
\]
Proof:

Let \( \mathbf{a}' = (a_1, a_2, \ldots, a_m) \), \( \kappa = (1, 1, \ldots, 1) \) be a sequence of \( d \) ones, and \( \iota = (i_1, \ldots, i_d) \). Then,

\[
\mathbf{a}[^{\iota} | \kappa] = \begin{pmatrix}
  a_{i_1} & a_{i_1} & \cdots & a_{i_1} \\
  a_{i_2} & a_{i_2} & \cdots & a_{i_2} \\
  \vdots & \vdots & \ddots & \vdots \\
  a_{i_d} & a_{i_d} & \cdots & a_{i_d}
\end{pmatrix},
\]

and

\[
\text{per}(\mathbf{a}[^{\iota} | \kappa]) = d! \prod_{i=1}^{d} a_{i_i}.
\]

Thus, the \((i, 1)\) entry of \( \mathbf{a}[^d] \) is

\[
\sqrt{d!} \prod_{i=1}^{d} a_{i_i}.
\]

This gives

\[
\mathbf{a}[^d] \mathbf{a}[^d] = \sum_{\iota \in S_d(m)} \frac{d!}{\nu(\iota)} (\prod_{i=1}^{d} a_{i_i})^2.
\]

Now, notice that for a \( m \times 1 \) vector \( \mathbf{a} \)

\[
\mathbf{a}' \mathbf{a} = \sum_{j=1}^{m} a_j^2
\]

and by the multinomial theorem we have that

\[
(a' a)^d = \sum_{u_1=0}^{d} \sum_{u_2=0}^{d} \cdots \sum_{u_{m-1}=0}^{d} \frac{d!}{u_1! u_2! \cdots u_{m-1}!} \left( \prod_{i=1}^{m} a_i^{u_i} \right)^2,
\]

(3.11)

where the \( u_i \) is the multiplicity of \( i \) in a sequence of length \( d \), \( v_j = \sum_{k=1}^{j} u_k \), and \( u_m \) is defined as \( v_m \). This sum is taken over all possible sequences of \( \{1, \ldots, m\} \) having
length \( d \), i.e. over all \( \iota \in S_d(m) \) allowing us to conclude that

\[
a'_{(d)}a_{(d)} = (a'a)^d,
\]

the desired result.

**Theorem 3.2.2** Let \( C \) be a \( m \times n \) matrix and \( b \) and \( a \) be \( n \times 1 \) and \( m \times 1 \) vectors such that

\[
a = Cb.
\]

Then

\[
a_{(d)} = C_{(d)}b_{(d)}
\]

for any positive integer \( d \).

**Proof:**

Let \( \iota \in S_d(m) \) and \( \kappa = (1, \ldots, 1) \) be a sequence of \( d \) ones. Suppose that \( C = (c_{ij}) \) and \( b = (b_j) \). Need to show that

\[
(Cb)_{(d)} = C_{(d)}b_{(d)}.
\]

From equation (3.10), in the \( \iota \) position of \((Cb)_{(d)}\) we have

\[
\sqrt{d!} \prod_{i=1}^{d} \left( \sum_{j=1}^{n} c_{\iota,i}b_j \right).
\]

Note the following:

\[
\prod_{i=1}^{d} \sum_{j=1}^{n} c_{\iota,i}b_j = \sum_{\sigma} \prod_{i=1}^{d} c_{\iota,\sigma(i)}b_{\sigma(i)}
\]

(3.12)

\[
= \sum_{\kappa} \prod_{i=1}^{d} c_{\iota,\kappa(i)}b_{\kappa(i)}
\]

(3.13)

\[
= \sum_{\kappa} \sum_{\sigma} \prod_{i=1}^{d} \frac{c_{\iota,\sigma(i)}b_{\kappa(i)}}{\nu(\kappa)}
\]

(3.14)
Here $\sigma^* (i)$ maps $\{1, \ldots, d\}$ into $\{1, \ldots, n\}$, $\sigma$ is a bijection from $\{1, \ldots, d\}$ to $\{1, \ldots, n\}$, and $\kappa \in S_d(n)$.

Now, in position $i$ of $C_{(d)} b_{(d)}$, from Definition 3.2.2, we have the entry

$$
\sum_{\kappa} \frac{\text{per}(C[i \mid \kappa])}{\nu(i) \nu(\kappa)} \sqrt{\prod_{i=1}^{d} b_{\kappa_i}} = \sqrt{\frac{d!}{\nu(i)} \sum_{\kappa} \frac{\text{per}(C[i \mid \kappa])}{\nu(\kappa)} \prod_{i=1}^{d} b_{\kappa_i}} 
$$

$$
= \sqrt{\frac{d!}{\nu(i)} \sum_{\kappa} \prod_{i=1}^{d} \frac{c_{\sigma(i) \kappa_i} b_{\kappa_i}}{\nu(\kappa)}}. \quad (3.15)
$$

Thus, $(C b)_{(d)} = C_{(d)} b_{(d)}$.

Finally, a theorem on the uniqueness of induced matrices is proven.

**Theorem 3.2.3** Let $C$ be an $m \times n$ matrix and for each $n \times 1$ vector $b$ define $a = C b$.

Then there is a unique $(m+d-1) \times (n+d-1)$ matrix $C_0$ such that $a_{(d)} = C_0 b_{(d)}$ for all $b$ and by Theorem 3.2.2, $C_0 = C_{(d)}$.

**Proof:**

The existence of $C_0$ with $a_{(d)} = C_0 b_{(d)}$ follows from Theorem 3.2.2. To show the uniqueness of $C_0$, suppose that there exist $C_0$ and $C_1$ such that $C_0 \neq C_1$ and $C_0 b_{(d)} = C_1 b_{(d)}$ for all $b \in \mathcal{R}^n$. Then $[C_0 - C_1] b_{(d)} = 0_m$ for all $b \in \mathcal{R}^n$. However, the elements of $[C_0 - C_1] b_{(d)}$ are linear combinations of distinct polynomial terms in $b_1, \ldots, b_n$, for all $b \in \mathcal{R}^n$. The only way that $[C_0 - C_1] b_{(d)} = 0_m$ for all $b \in \mathcal{R}^n$ is for $C_0 - C_1 = 0$.

Now that the mechanics of induced matrices have been developed, two theorems are given that have a direct bearing on the current problem—demonstrating the invariance of the various design optimality criteria.

**Theorem 3.2.4** Let $C$ and $D$ be $m \times n$ and $n \times q$ matrices, respectively. Then

$$
C_{(d)} D_{(d)} = (C D)_{(d)}.
$$
Proof:

Let \( \mathbf{b} \) be an arbitrary vector in \( q \) indeterminates and \( C \) and \( D \) be arbitrary, fixed \( m \times n \) and \( n \times q \) matrices. Define \( \mathbf{a} \) as

\[
\mathbf{a} = C \mathbf{D} \mathbf{b}
\]

and \( \mathbf{c} \) as

\[
\mathbf{c} = D \mathbf{b}.
\]

By Theorem 3.2.2,

\[
\mathbf{a}_{(d)} = C_{(d)} \mathbf{c}_{(d)}
\]

and

\[
\mathbf{c}_{(d)} = D_{(d)} \mathbf{b}_{(d)},
\]

giving

\[
\mathbf{a}_{(d)} = C_{(d)} D_{(d)} \mathbf{b}_{(d)}.
\]

(3.16)

But, we also have that

\[
\mathbf{a}_{(d)} = (C \mathbf{D})_{(d)} \mathbf{b}_{(d)}.
\]

(3.17)

Applying Theorem 3.2.3 yields the desired result.

Theorem 3.2.4 implies the following useful corollary.

**Corollary 3.2.1** Let \( C \) be a \( m \times m \) matrix, \( C_{(d)} \) be the corresponding \( d \)-th order induced matrix, and let \( I_m \) be the \( m \times m \) identity matrix. Then,

1. \( I_{(d)} \) is the \( \binom{m+d-1}{d} \times \binom{m+d-1}{d} \) identity matrix.

2. \( (C')_{(d)} = C'_{(d)} \).
3. \( C \) is an orthogonal matrix if and only if \( C_{(d)} \) is an orthogonal matrix.

4. \( C_{(d)} \) is non-singular if and only if \( C \) is non-singular, in which case \( C_{(d)}^{-1} = (C^{-1})_{(d)} \)

**Proof:**

1. Let \( a \) be an \( m \times 1 \) vector. Then, \( a = I_m a \). Applying Theorem 3.2.2 gives \( a_{(d)} = I_{(d)} a_{(d)} \). By Theorem 3.2.3, we have the desired result.

2. As noted earlier, \( \text{per}(C) = \text{per}(C') \). The result follows by applying Definition 3.2.2 and Theorem 3.2.3 to \( C' \) and noticing that the rows and columns are then labeled by elements of \( S_d(m) \).

3. Let \( C \) in Theorem 3.2.4 be orthogonal and \( D = C' \). Then \( (CC')(d) = I(d) = (C'C)(d) \), by part 1. Theorem 3.2.4 and part 2 give \( C_{(d)} C_{(d)} = C_{(d)} C'_{(d)} = I(d) \). Thus, \( C_{(d)} \) is orthogonal. Let \( C_{(d)} \) be orthogonal. Then \( C_{(d)} C_{(d)} = C_{(d)} C'_{(d)} = I(d) \). By Theorem 3.2.4 and part 2 we have \((CC')(d) = (C'C)(d) = I(d) \). Theorem 3.2.3 and part 1 show that \( CC' = C'C = I_m \), i.e. \( C \) is orthogonal.

4. Suppose that \( C \) is non-singular. Then \( C^{-1} \) exists such that \( CC^{-1} = C^{-1}C = I_m \). Applying Theorem 3.2.4 and then part 1, we have that \( C_{(d)}(C^{-1})_{(d)} = (C^{-1})_{(d)} C_{(d)} = I(d) \). Thus, \( C_{(d)} \) is non-singular and by uniqueness, \( (C^{-1})_{(d)} = C_{(d)}^{-1} \). Suppose that \( C_{(d)} \) is non-singular. Then \( C_{(d)}^{-1} \) exists such that \( C_{(d)}^{-1} C_{(d)} = C_{(d)} C_{(d)}^{-1} = I(d) \). By Theorem 3.2.4 we have that \( (CC')(d) = (C'C)(d) = (I)(d) \) for some matrix \( 
\hat{C} \). By part 1, \( C\hat{C} = \hat{C}C = I \), i.e. \( C \) is non-singular, and
by the uniqueness of inverses of matrices, \( \bar{C} = C^{-1} \). Thus, by Theorem 3.2.3, 
\[
(C^{-1})_{(d)} = C_{(d)}^{-1}.
\]

Marcus and Minc (1992), Ryser (1961), and Aitken (1948) provide more details on induced vectors and induced matrices.

To illustrate some of the concepts presented in this section, the following are pertinent examples of induced vectors. These examples provide a parametrization of \( d \)-th order mixture models that is different from that of Section 2.2.2.

The first-order model in the mixture components is the same for both the induced vector and the canonical polynomial parametrizations, i.e.
\[
F_1(z)' = (z_1, \ldots, z_p).
\]
However, \( F_d(z) \) may be found using the induced vector parametrization, generating the \( d \)-th order induced vector of \( z \), i.e.
\[
F_d(z)' = (z_1, \ldots, z_p)_{(d)}.
\]

For example, the 2-nd order induced vector for two mixture components is given by
\[
F_2(z)' = (z_1, z_2)_{(2)}
= (z_1^2, z_1z_2\sqrt{2}, z_2^2).
\]

In contrast, the 2-nd order canonical polynomial is
\[
F_2(z)' = (z_1, z_2, z_1z_2).
\]

If we just have two mixture components, the fourth-order model is
\[
F_4(z)' = (z_1, z_2)_{(4)}
= (z_1^4, 2z_1^3z_2, \sqrt{6}z_1^2z_2^2, 2z_1z_2^3, z_2^4).
\]
The corresponding quartic canonical polynomial is given by

\[ F_4(z) = (z_1, z_2, z_1z_2, z_1z_2(z_1 - z_2), z_1z_2(z_1 - z_2)^2). \]

Recall that a \( d \)-th order canonical polynomial has \( \binom{p+d-1}{d} \) terms, the same number of elements as the \( d \)-th order induced vector of the \( p \times 1 \) vector \( z \). It can be shown that induced vectors of order \( d \) based on \( (z_1, \ldots, z_p) \) are full-rank reparametrizations of the \( d \)-th order canonical polynomials, given that the design is such that \( \text{rank}(Z) = \binom{p+d-1}{d} \). For instance, consider the quadratic model with \( p = 2 \).

\[
(z_1^2, z_1z_2\sqrt{2}, z_2^2) = (z_1, z_2, z_1z_2) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ -1 & \sqrt{2} & -1 \end{pmatrix}.
\]

Finding canonical polynomials for \( d \geq 3 \) is laborious; to this author's knowledge, polynomials of order greater than four have not appeared in print. While there appears to be no general rule for deriving \( d \)-th order canonical polynomials, one can begin by examining terms in the \( d \)-th order polynomial, \( z_1^{\xi_1} \cdots z_p^{\xi_p} (\sum_{j=1}^p \xi_j = d) \), and eliminating terms for which \( \max\{\xi_1, \xi_p\} - \min\{\xi_1, \xi_p\} \geq 1 \). After this step, it is unclear how to systematically combine terms to produce a \( d \)-th order canonical polynomial. In contrast, finding a \( d \)-th order induced vector is a relatively simple task, requiring a straightforward application of Definition 3.2.2.

The drawback of using the induced vector representation is that the terms in these models are less familiar. For example, the quadratic model above contains pure quadratic and interaction terms, but no linear terms. Fitting regression models is usually an iterative sequence of design and analysis, starting with polynomials of lower degree and progressively fitting higher order polynomials as the need arises.
However, such a sequential strategy cannot be employed when using the induced vector parametrization. Suppose that at stage 1 a \((d - 1)\)-st order induced vector model has been fit. Then at stage 2 it is desirable to fit a \(d\)-th order induced vector model. Although the column space of the model matrix for \(d\)-th order induced vector at stage 2 contains the column space of the model matrix at stage 1, there is no hierarchical structure, confusing interpretations. Induced vectors are introduced here primarily as a mechanism to prove the results of the next section. The results of the next section are applicable to both the induced vector parametrization and the canonical polynomial parametrization.

Induced matrices also play an important role in the formulation of rotatability (see the Appendix). Draper (1984) developed a notion called Schlӓflian rotatability, named after another name for induced matrices—Schlӓflian matrices. He worked with \(d\)-th order polynomials in \(p\) unrestricted variables having \(\binom{p+d}{d}\) terms, using the induced vector representation of the polynomials. He claimed that Schlӓflian rotatability was a generalization of the usual notion of rotatibility.

Sa and Edwards (1993), without acknowledging so, used 2-nd order induced vectors for convenience when constructing simultaneous confidence intervals for the quantity \(E(\hat{y}(x)) - E(\hat{y}(0))\) under the assumption that the response is adequately modeled by a quadratic polynomial. Using induced vectors may yield additional results for higher order models.

### 3.3 Design Properties Preserved under Transformations

As stated earlier, arguably the most important use of transformations is in the design stage. This section demonstrates the invariance of several desirable design
properties under linear transformations. These properties include:

- Orthogonal blocking
- Orthogonality to deterministic time trends
- D-, G-, and V-optimality
- Minimum integrated bias and integrated mean-squared error properties
- Rotatability (For orthogonal transformations)

Orthogonality was defined in Section 2.4. The other optimality criteria discussed in this section are defined in Appendix.

Throughout this section, assume that a design has been chosen, specifying $Z$ and $X$.

**Theorem 3.3.1** Suppose that the models

$$y = \tilde{V}\zeta + Z\tau + e$$

and

$$y = \tilde{V}\zeta + X\beta + e,$$

hold, where the rows of $Z$ and $X$ are given by $F_d(z) = z(d)$ and $G_d(x) = x(d)$, respectively. Let $Z$ have full column rank, $\tilde{V}'1_N = 0_{d-1}$ and $z' = x'Q$ for a non-singular matrix $Q$. Then, $P_{\tilde{V}}Z = 0$ if and only if $P_{\tilde{V}}X = 0$.

**Proof:**

By Theorems 3.2.3 and 3.2.1, there exists a non-singular matrix $Q(d)$ for all positive integers $d$ such that $z'(d) = x'(d)Q(d)$. Since

$$P_{\tilde{V}}ZQ^{-1}(d) = P_{\tilde{V}}XQ(d)Q^{-1}(d) = P_{\tilde{V}}X,$$
The result follows.

Theorem 3.3.1 is a powerful tool to construct design for mixture components for which $z$ and $\tau$ may be estimated orthogonally, where $z$ might be a vector of block contrasts or trend parameters. This implies that orthogonal block designs or trend-free designs in the unrestricted variables may be mapped in the mixture component space, $S$, and the resulting designs will also be orthogonally blocked or trend-free. These topics will be dealt with in Chapters 4 and 5.

Murty (1966) and Nigam (1970) both realized that transformations could be used to construct orthogonal block designs for mixture components. In fact, Nigam erroneously concluded that the use of transformations was the only way in which orthogonal block designs could be constructed. As we shall see in the next chapter, there are other methods of constructing orthogonal block designs.

Because $X$ and $Z$ have the same column space, provided $Q$ is non-singular, we have that

$$(I - P_U)Z [Z'(I - P_U)Z]^\top Z'(I - P_U) = (I - P_U)X [X'(I - P_U)X]^\top X'(I - P_U).$$

Also, if a design in the unrestricted variables is D-optimal for the maximal set of rank$[(I - P_U)Z]$ linearly independent estimable functions, then so is the corresponding design in the mixture components from because D-optimality is invariant under full-rank reparametrizations [see Box and Draper (1971), Hedayat (1981)]. It is important to note that the design region over which the unrestricted variable design is optimal must be the image of the mixture component design space. This is somewhat of a drawback because of the complex region in the unrestricted variables. While the invariance of D-optimality is important conceptually, a more efficient method of searching for a D-optimal design searches in the mixture region rather than the region
for the unrestricted variables. Invariance is mentioned primarily for completeness.

Folks (1958) showed that the bias and variance functions, and thus the integrated bias, integrated variance (V-optimality criteria), and the integrated mean-squared error are invariant under certain non-singular linear transformations. Using the more general notion of bias introduced in the Appendix, we extend his conclusions to any non-singular, linear transformation. As before, we assume that the induced vector parametrization is being used.

**Theorem 3.3.2** Assume (realistically) that the specified models

\[ y = Z_1 \tau_1 + e \quad (3.18) \]
\[ y = X_1 \beta_1 + e, \quad (3.19) \]

with rows \( F_{d_1}(z)' \) and \( G_{d_1}(x)' \), respectively, are incorrect. Suppose (unrealistically) that the true models are

\[ y = Z_2 \tau_2 + e \quad (3.20) \]
\[ y = X_2 \beta_2 + e, \quad (3.21) \]

respectively, having rows \( F_{d_2}(z)' \) and \( G_{d_2}(x)' \) such that \( d_1 \leq d_2 \). Then bias is invariant under non-singular, linear transformations of the form \( z' = x'Q \).

**Proof:** Let \( z \) be a mixture and \( x \) be such that \( z' = x'Q \), for non-singular \( Q \).

Then, defining bias for a mean response as in equation (A.9) gives

\[
B(z) = F_{d_1}(z)'(Z_{d_1}', Z_{d_1})^{-1}Z_{d_1}'Z_{d_2} \tau_2 - F_{d_2}(z)' \tau_2 \\
= G_{d_1}(x)'Q_{(d_1)}(Q_{(d_1)}'X_{d_1}'X_{d_1}Q_{(d_1)})^{-1}Q_{(d_1)}'X_{d_1}'X_{d_2}Q_{(d_2)}Q_{(d_2)}^{-1} \beta_{d_2} - G_{d_2}(x)'Q_{(d_2)}^{-1}Q_{(d_2)}^{'(d_2)} \beta_{d_2}
\]
Next, we show that $V$-optimality (the variance function) (see the Appendix) is invariant under such transformations. If so, then the integrated mean-square error is also invariant since the mean-square error is a simple sum of squared-bias and variance. Suppose that models 3.18 and 3.19 hold with the induced vector parametrization.

**Theorem 3.3.3** The variance function, $V(z)$ is invariant under non-singular, linear transformations of the form $z' = x'Q$.

**Proof:** Let $z$ and $x$ be arbitrary points in $\mathcal{S}$ and the image of $\mathcal{S}$ under the transformation $z' = x'Q$. By Theorem 3.2.4, we have

$$F_{d_1}(z)' = G_{d_1}(x)'Q_{(d_1)},$$

for a non-singular $Q_{(d_1)}$. Then, the variance is

$$V(z) = \frac{\text{Nvar}(\hat{\beta}(z))}{\sigma^2} = \left[ F_{d_1}(z)' \left[ Z_{d_1}'Z_{d_1} \right]^{-1} F_{d_1}(z) \right]$$

$$= \left[ G_{d_1}(x)'Q_{(d_1)} \left[ Q_{(d_1)}'X'XQ_{(d_1)} \right]^{-1} Q_{(d_1)}'G_{d_1}(x) \right]$$

$$= \left[ G_{d_1}(x)'[X'X]^{-1} G_{d_1}(x) \right]$$

$$= V(x). \quad (3.22)$$

Theorem 3.3.3 implies that $G$-optimality is also invariant under non-singular linear transformations.
Once again, we assume that models 3.18 and 3.19 are incorrect as specified and that the true models are given as models 3.20 and 3.21, respectively.

Theorem 3.3.4 Suppose that \( z' = x'Q \) where \( Q \) is a non-singular matrix. Then the following statements are true.

1. A design in the mixture components is a minimum integrated bias design if and only if the design in the unrestricted variables is a minimum integrated bias design.

2. A design in the mixture components is a minimum integrated variance design if and only if the design in the unrestricted variables is a minimum integrated variance design.

3. A design in the mixture components is a minimum integrated mean squared error design if and only if the design in the unrestricted variables is a minimum integrated mean squared error design.

Proof: The first two statements follow directly from Theorems 3.3.2 and 3.3.3, respectively. The last statement follows from parts 1 and 2.

Finally, we prove a theorem about the invariance of rotatability, where rotatability is defined in the Appendix.

Theorem 3.3.5 Let \( z' = x'Q \) for a non-singular \( Q \) and suppose that the models

\[
y = Z\tau + e
\]

and

\[
y = X\beta + e
\]

holds. If \( Q \) is orthogonal, then rotatability is invariant under this transformation.
Proof:

Suppose that we have a \( d \)-th order rotatable design in the unrestricted variables centered at \( x_0 \). Define \( \rho[a, b] = (a - b)'(a - b) \). Then,

\[
\text{var}(\hat{y}(z)) = \text{var}(\hat{y}(x)) = v(\rho[x, x_0]) = v(\rho[z, z_0])
\]

since by Theorem 3.3.3

\[
\text{var}(\hat{y}(x)) = \text{var}(\hat{y}(z))
\]

and \( x - x_0 = Q'(z - z_0) \), with \( Q \) orthogonal, so

\[
\rho[x, x_0] = [x - x_0]'[x - x_0] = (z - z_0)'(z - z_0) = \rho[z, z_0]
\]

(3.23)

3.4 Comparison of Transformations

Although all of the transformations assume the same form, some are simpler to work with than others, particularly when the aim is to select a design for a mixture experiment. The important fact to consider about a transformation is into what subspace of \((p - 1)\) dimensional space is \( S \) being mapped. Important points to consider when selecting a transformation include

- Can “standard” designs in the unrestricted variables be used without much alteration?
• Is there any intuition for choosing the transformation?

Regarding the first point, it seems that only Thompson and Myers (and Claringbold) provide much guidance on the translation of a design from unrestricted variables to a design in mixture components. Standard designs may be used in conjunction with Thompson and Myers with little modification. Neither the transformations by Draper and Lawrence or that of Becker are amenable to easy rules for selecting the runs from \((p - 1)\)-dimensional space.

Each transformation is sensitive to the choice of reference point. The only transformation restricting the choice of \(z_0\) is that of Draper and Lawrence, rendering it unsuitable for use with constrained regions with centroids other than the simplex centroid. Other transformations tacitly allow \(z_0\) to be different—as long as \(z_0\) remains in \(S\).

Only by using Thompson and Myers’ method is choosing the design easy. Choosing a design is easy because we have restricted ourselves to an ellipsoid in \(p - 1\) dimensional space, a region common for choosing response surface designs. However, the design region in the whole simplex is limited to an ellipsoidal region. As the number of mixture components increases, the volume of the region covered by the largest sphere drops off rapidly—a rather unpleasing consequence.

One should not be entirely dissuaded from this approach, though. In typical response surface investigations, the design region is generally not taken to be the entire region of operability; in fact the operability region is often unknown. Usually, a small subregion is chosen and using a sequence of designs in combination with steepest ascent and canonical analysis, one explores increasingly larger portions of the operability region. Merely because the (absolute) operability region is known for
mixtures, we are not beholden to design an experiment providing uniform coverage of the whole simplex. Often it is sensible to restrict oneself to a certain region. If this is so, than the approach of Thompson and Myers is sensible and one has all response surface designs at his disposal.

If the entire simplex is of interest, it is possible to select several points around the simplex. These points will act as the ellipsoid centers, around which designs may be constructed. Then, the overall design may consist of the union of all of the smaller designs. This approach may also be used with other transformations.
CHAPTER 4. CONSTRUCTION OF BLOCK DESIGNS

This chapter discusses aspects of designing mixture experiments when blocking variables are present. We begin by examining problems of estimability resulting from the inclusion of blocking variables, giving conditions for orthogonal blocking, and summarizing current methods for constructing block designs. A combinatorial structure called a *mixture mate* is introduced in Section 4.3. Mixture mates form the elements from which orthogonal block designs for mixture experiments are built. In particular, we construct mixture mates by borrowing techniques from the theory of trade-off in $t$-designs. Several new methods of constructing block designs using mates are proposed and evaluated. Section 4.4 shows how to find orthogonal or nearly-orthogonal block designs when $S$ is a subregion of the simplex. When orthogonality is too stringent for a particular situation, non-orthogonal block designs are needed; some easy ways to obtained designs are discussed in Section 4.5. Finally, block designs are also constructed using computer algorithms seeking to optimize various objective functions.

One option of constructing orthogonal block designs for mixture experiments is to use a randomized complete block arrangement. A randomized complete block would select a set of mixtures and repeat all of these mixtures in each of the blocks. However, as we will show, constructing orthogonal block designs that do not have the
entire set of mixtures in each block is not too difficult. Eliminating the restriction that the entire set of mixtures must appear in every block facilitates the consideration of more distinct combinations of mixtures. This can allow a more complete coverage of the region of interest and possibly enables lack-of-fit tests to be performed if the mixtures are chosen judiciously.

4.1 Blocking in Mixture Experiments

4.1.1 Blocking and Estimability

Throughout this chapter we will suppose that the model

$$y = U\gamma + Z\tau + e$$

holds, where $U$ is a $N \times b$ matrix corresponding to blocking variables with columns $u_{il}, l = 1, \ldots, b$ and $\gamma$ is the vector of block effects. For $U = (u_{il}), l = 1, \ldots, b$ and $i = 1, \ldots, N$, we will adopt the convention that

$$u_{il} = \begin{cases} 1 & \text{if unit } i \text{ is in block } l \\ 0 & \text{otherwise.} \end{cases}$$

$Z$ is the $N \times s$ matrix corresponding to a $d$-th order canonical polynomial model, with rows $F_d(z)'$. Let $z_k$, be the $k$-th column of $Z$, $k = 1, \ldots, s$ and $\tau$ is the vector of regression parameters. The first $p$ columns of $Z$ may be taken to correspond to a first-order canonical polynomial. Then, the columns of $[U, Z]$ satisfy

$$\sum_{l=1}^{b} u_{il} = 1_N = \sum_{k=1}^{p} z_k,$$

implying that

$$\text{rank } [U, Z] \leq s + b - 1.$$
Thus, neither the parameters corresponding to the first-order portion of the $d$-th order canonical polynomial model nor the elements of $\gamma$ are estimable in model (4.1).

Suppose that $\text{rank } [U, Z] = s + b - 1$. Partition $\tau$ as $\tau' = (\tau_1', \tau_2')$, where $\tau_1$ is $p \times 1$ and corresponds to the parameters in the first-order portion of the model and $\tau_2$ is a $(s - p) \times 1$ vector of parameters corresponding to the higher order portion of the model based on the mixture components. Then all estimable functions are characterized by

$$\lambda' \gamma + \phi_1' \tau_1 + \phi_2' \tau_2,$$

where $\lambda' 1_b = \phi_1' 1_p$ (4.4)

Block contrasts result by taking each of $\phi_1$ and $\phi_2$ to be zero vectors of appropriate lengths. Elements of $\tau_2$ are individually estimable, seen by taking $\lambda$ and $\phi_1$ to be zero vectors of appropriate lengths.

4.1.2 Conditions for Orthogonal Blocking

Specific conditions on the mixtures allocated to the blocks are required to achieve orthogonal blocking. These conditions are derived from Definition 2.4.1.

John (1984) gave necessary and sufficient conditions on the mixtures within a block for orthogonal blocking of designs for second-order canonical polynomial models having equal block sizes. We derive a generalized version of John's result, which applies to all models satisfying equation (4.4) and allows for unequal block sizes. Let $B_l$ be the $b$ blocks, $l = 1, \ldots, b$, and $n_l$ be the number of mixtures in block $l$.

**Theorem 4.1.1**: Under model (4.1) a necessary and sufficient condition for

$$P_1 Z = P_0 Z$$
is that

\[ \frac{u'_{l}z_{k}}{1_{N}z_{k}} = \frac{n_{l}}{N} \text{ for all } l = 1, \ldots, b \text{ and } k = 1, \ldots, s. \]

Proof:

Under the convention of equation (4.2), \( u_{l} \) is a \( N \times 1 \) vector with \( n_{l} \) 1's in positions corresponding to mixtures in block \( l \) and zeros elsewhere. With observations ordered by blocks, \( P_{U} = U(U'U)^{-1}U' \) is a block diagonal matrix of the form

\[ P_{U} = \sum_{l=1}^{b} \frac{1}{n_{l}} u_{l}u'_{l}. \]

So,

\[ P_{U}Z = \begin{pmatrix} \frac{1}{n_{1}} 1_{n_{1}}u'_{1}z_{1} & \cdots & \frac{1}{n_{1}} 1_{n_{1}}u'_{1}z_{s} \\ \vdots & \ddots & \vdots \\ \frac{1}{n_{b}} 1_{n_{b}}u'_{b}z_{1} & \cdots & \frac{1}{n_{b}} 1_{n_{b}}u'_{b}z_{s} \end{pmatrix} \]

and

\[ P_{Y}Z = \frac{1}{N} 1_{N} (1'_{N}z_{1}, \ldots, 1'_{N}z_{s}). \]

These two matrices are equal if and only if their entries are equal, giving the desired result.

When a quadratic canonical polynomial is considered and block sizes are unequal, the conditions for orthogonal blocking become:

1. For each mixture component, \( z_{j} \),

\[ \frac{\sum_{i \in B_{l}} z_{ij}}{N} = \frac{n_{l}}{N} \text{ for all } l = 1, \ldots, b. \]  \hfill (4.5)
2. For each pair of mixture components, $z_j$ and $z_k$,

$$\sum_{i \in B_l} \frac{z_{ij} z_{ik}}{N} = \frac{n_l}{N} \quad \text{for all } l = 1, \ldots, b. \quad (4.6)$$

The results of John (1984) follow from Theorem 4.1.1 by taking $d = 2$ and $n_l = n$ for all $l$.

### 4.2 Review of Methods for Constructing Orthogonal Block Designs for Mixture Experiments

The techniques for constructing orthogonal block designs for mixture experiments may be grouped into three general approaches:

- constructions using transformation techniques,
- constructions using combinatorial tools, and
- constructions exploiting geometric methods.

Techniques of the first type draw from Chapter 3. Murty (1966) obtained orthogonal block designs for mixture components using the result of Theorem 3.3.1. Given an orthogonal block design for the unrestricted variables, transform the runs in the design back into mixtures. Theorem 3.3.1 guarantees the preservation of orthogonality under this construction.

Many attempts to devise orthogonal block designs for mixture experiments were thwarted by the constraint (2.1). Nigam (1970,1976) described several methods for constructing orthogonal block designs for second-order canonical polynomial models.
However, the conditions he imposed were unnecessarily restrictive and rendered the quadratic terms non-estimable.

Nigam (1970) initiated the use of combinatorial tools, the second category, constructing orthogonal block designs using perpendicular arrays, a special case of mixture mates. Later, John (1984) constructed orthogonal block designs for $p$ mixture components using $p \times p$ Latin squares. However, without augmentation, John’s designs produce model matrices for the second-order canonical polynomial models for which the individual quadratic parameters are non-estimable. The singularity problems may be ameliorated by inclusion of additional points, such as centroids. Recently, Draper, Prescott, Lewis, Dean, John, and Tuck (1993), Prescott, Draper, Dean, and Lewis (1993), and Lewis, Dean, Draper, and Prescott (1993) have extensively studied and extended John’s work allowing the construction of multiple blocks and an arbitrary number of mixture components.

The third method for constructing orthogonal block designs includes several geometric techniques. Crosier (1991) exploits properties of constrained regions, while Saxena and Nigam (1974) use symmetric-simplex designs; see these papers for details.

**4.3 Combinatorial Methods for Constructing Block Designs for Mixture Experiments**

This section introduces the principal combinatorial structure for constructing orthogonal block designs for mixtures, mixture mates. Two types of mixture mates will be defined and ways to construct each will be given. Finally, Section 4.3.4 introduces some ways to construct large mixture mates by patching together smaller mixture mates.
4.3.1 Methods for Constructing b Mixture Mates of Strength t

Mixture mates have been used by Nigam (1970, 1976), but were not really defined until John (1984). Recently, papers by Draper, Prescott, Lewis, Dean, John, and Tuck (1993), Prescott, Draper, Dean, and Lewis (1993) and Lewis, Dean, Draper, and Prescott (1993) have more fully developed the notion of mixture mates. The construction of mixture mates has been limited to equal block sizes and construction of pairs of mates; methods to construct more than two mates were given in Draper, et. al. and Prescott, et al. Also, the block designs constructed using mates are useful for 2nd order canonical polynomials, with a few extensions to higher order canonical polynomials. This section provides a general framework for the discussion of mixture mates, allowing for unequal block sizes, an arbitrary number of blocks and higher-order canonical polynomials.

Let \( a \) and \( b \) be \( q \times 1 \) vectors and \( a \times b \) denote the \( q \times 1 \) vector created by taking the Schur (element-wise) product of \( a \) and \( b \). In Definition 4.3.1, the Schur product of the vectors are polynomials since the elements of the vectors are chosen from a symbol set.

**Definition 4.3.1** Let \( R_l = (r_{ij}), l = 1, \ldots, b, \) be distinct \( n_l \times p \) arrays where \( N = \sum_{l=1}^{b} n_l \) and \( r_{ij} \) are chosen from some set of symbols, \( \mathcal{K} \). Let \( t \) be fixed and \( r_{ji} \) be the \( j \)-th column of array \( l \) and \( r_{j.} \) be the \( N \times 1 \) vector containing the \( j \)-th columns of all the arrays. Then \( R_1, \ldots, R_b \) are said to be \( b \) symbolic mixture mates of strength \( t \) if the following conditions hold

1. \( N(r_{j_{1}i} \times r_{j_{2}i} \times \cdots \times r_{j_{l}i})' \mathbf{1}_{n_l} = n_l(r_{j_{1}i} \times r_{j_{2}i} \times \cdots \times r_{j_{l}i})' \mathbf{1}_N \) for all sets of \( t \) columns \( (j_1, \ldots, j_t) \in S_t(p) \), (see Chapter 3), for all \( l \).
2. \( \sum_{j=1}^{p} r_{ijl} = r \) for all \( i \) and for all \( l \), where \( r \) is an element of \( \mathbb{Z}[\mathbb{H}] \), where \( \mathbb{Z}[\mathbb{H}] \) is the ring of polynomials in elements of \( \mathbb{H} \) with integer coefficients selected from \( \mathbb{Z} \).

In Definition 4.3.1, \( t \) is called strength of the symbolic mixture mates, that is, the largest number of columns for which part 1 is true. The strength \( t \) can be related to the degree of the polynomial being fitted. Given \( b \) mixture mates of strength \( t \), if the \( \mathbb{H} \) has at least \( t + 1 \) symbols that can be associated with \( t + 1 \) levels, the mates provide a blocking arrangement for a design fitting a \( t \)-order canonical polynomial. This can be seen by equating parts 1 and 2 to the conditions required in Theorem 4.1.1. Also, part 2 will be termed row consistency.

Two specific types of mixture mates have already been formulated and used to construct orthogonal block designs, complementable perpendicular arrays [a generalization of Nigam's (1970) method] and combinations of Latin squares [John (1984), Draper, et. al. (1993), Prescott, et. al. (1993), and Lewis, et. al. (1993)]. Nigam’s method produces a pair of symbolic mixture mates of strength two and the Latin square method can produce multiple symbolic mixture mates of strength two (and sometimes higher). Using either method yields a set of symbolic mixture mates with each mate having the same number of rows.

**Definition 4.3.2** A perpendicular array is a \( \binom{v}{2} \times f \) array \( C \) with entries from a set \( \mathbb{H} \) of \( v \) symbols having the property that any \( \binom{v}{2} \times 2 \) subarray contains each of the \( \binom{v}{2} \) unordered pairs of symbols in \( \mathbb{H} \) exactly once as a row.

To construct a pair of symbolic mixture mates of strength two, complementable perpendicular arrays are needed.
Definition 4.3.3 Let $C$ denote a $\binom{v}{2} \times f$ perpendicular array. Let the entries of the perpendicular array $\bar{C}$ be defined on the set $\mathbb{R}$. If the $2 \binom{v}{2} \times f$ array
\[
\begin{bmatrix}
C \\
\bar{C}
\end{bmatrix}
\]
is such that each $2 \binom{v}{2} \times 2$ subarray contains each ordered pair of elements from $\mathbb{R}$ occurs exactly once as, then $C$ is said to be a complementable perpendicular array. $\bar{C}$ is called the complement of $C$.

To exploit complementable perpendicular arrays for blocking purposes, take $f = v = p$.

It has been shown [c.f. Rao (1963) and Schellenberg, van Rees, and Vanstone (1978)] that if $p$ is an odd prime power then there is a $\binom{v}{2} \times p$ complementable perpendicular array. By definition, both $C$ and $\bar{C}$ contain each ordered pair of symbols with equal frequency, one property of a pair of symbolic mixture mates of strength two.

Suppose that $p$ is an odd prime power. Let the perpendicular array and its complement be denoted by $C$ and $\bar{C}$. Schellenberg, van Rees, and Vanstone (1978, p.142) note that the number of times a symbol occurs in each column of a $\binom{p}{2} \times p$ perpendicular array is $\frac{p(p-1)}{2}$, also true for the complement, satisfying part 1 of Definition 4.3.1 for $t = 2$ for the case that $j_1 = j_2$. By Definitions 4.3.2 and 4.3.3, part 1 of Definition 4.3.1 is satisfied for $t = 2$ for the case $j_1 \neq j_2$. $C$ and $\bar{C}$ satisfy part 2 because each row contains each of the $p$ symbols exactly once. Then $B_1$ and $B_2$ are symbolic mixture mates of strength two.

Table 4.1 contains a perpendicular array and its complement, constituting a pair of mixture mates of strength two. All that remains to construct two orthogonal
blocks for a mixture experiment is to assign proportions to each of the five symbols. Nigam chose to set all but two or three symbols to zero and then divide the non-zero entries by the resulting row sum. For example, suppose that symbols 0, 2, 3 are set to 0; the first run in Block 1 of Table 4.1 would be (0, 0.2, 0, 0.8). A more flexible way is to assign the symbols to any five proportions, so long as they add to one. In the spirit of Czitrom (1988, 1989), these proportions may be chosen according to a design optimality criterion such as those discussed in the Appendix.

<table>
<thead>
<tr>
<th></th>
<th>$B_1$</th>
<th></th>
<th>$B_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 4.1: A pair of symbolic mixture mates of strength two for $p=5$

Nigam's method has its failings. Most importantly, one can construct only two large blocks; more blocks may be obtained by patchwork (see Section 4.3.1). The block size increases with the number of mixture components; each block has size $\binom{p}{2}$.

John (1984), Draper, Prescott, Lewis, Dean, John, and Tuck (1993), Prescott, Draper, Dean, and Lewis (1993), and Lewis, Dean, Draper, and Prescott (1993) construct symbolic mixture mates of strength two (and in some cases greater than two) using Latin squares. John and Draper, et. al. covers the case of $p = 4$; Prescott, et. al. deals with $p = 5$ and Lewis, et. al. discusses methods for arbitrary $p$. 
Definition 4.3.4 A $p \times p$ matrix is said to be a Latin square if each cell contains exactly one symbol, selected from a set $\mathcal{R}$, such that each of the $p$ symbols of $\mathcal{R}$ appears exactly once in each of the rows and exactly once in each of the columns.

Observe that the first and second sets of five rows for each of the perpendicular arrays $B_1$ and $B_2$ in Table 4.1 are Latin squares.

By Definition 4.3.4, a pair of Latin squares clearly satisfies parts 1 (for $t = 1$) and 2 of Definition 4.3.1. However, only special pairs of Latin squares satisfy part (1) for $t = 2$. The references above construct pairs of Latin squares meeting part (1) for $t = 2$, noting that in certain cases part (1) is met for $t \geq 2$. Draper, et. al. and Prescott, et. al. observe that the Latin squares of order 4 and 5 meet condition (1) of Definition 4.3.1 for $t = 4, 5$.

Lewis, Dean, Draper, and Prescott (1993) give general rules for finding symbolic mixture mates of order two using Latin squares of arbitrary order $p$. Methods of obtaining mixture mates are provided for the special cases that a Latin square is cyclic, cyclic equivalent, or when $p$ is even but the Latin square is not cyclic or cyclic equivalent. Table 4.2 shows a pair of symbolic mixture mates of strength two constructed by the method of Lewis, Dean, Draper, and Prescott for cyclic equivalent Latin squares.

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>3</th>
<th>4</th>
<th>1</th>
<th>5</th>
<th>2</th>
<th>0</th>
<th>3</th>
<th>2</th>
<th>5</th>
<th>1</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>3</td>
<td></td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td></td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td></td>
<td>3</td>
<td>0</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td></td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>4</td>
<td>1</td>
<td></td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 4.2: A pair of mixture mates of strength two, $p=6$
Essentially, when using these methods, we only have \( p - 1 \) "degrees of freedom" when choosing the design. Once we choose the proportions assigned to \( p - 1 \) of the symbols, the design is wholly determined. Because of this, complementable perpendicular arrays or Latin squares limit the variety of mixtures that can appear in the corresponding designs. This is restrictive when finding an optimal designs according to the criteria in the Appendix. Finding optimal designs boils down to choosing an optimal set of \( p \) proportions to which the symbols are assigned. In an attempt to loosen these restrictions, we introduce a more general structure than symbolic mixture mates.

**Definition 4.3.5** Let \( R_l = (r_{ij}) \), \( l = 1, \ldots, b \), be distinct \( n_l \times p \) arrays where \( N = \sum_{l=1}^{b} n_l \) and \( r_{ij} \) are chosen from a set of nonnegative integers, \( \mathcal{I} \). Let \( t \geq 1 \) and \( r_{ji} \) be the \( j \)-th column of array \( l \) and \( r_{j.} \) be the vector containing the \( j \)-th columns of all the arrays. Then \( R_1, \ldots, R_b \) are said to be \( b \) integral mixture mates of strength \( t \) if the following conditions hold:

1. \( N(r_{j_1} \ast r_{j_2} \ast \cdots \ast r_{j_t})^T 1_{n_l} = n_l(r_{j_1} \ast r_{j_2} \ast \cdots \ast r_{j_t})^T 1_N \) for a set of \( t \) columns \((j_1, \ldots, j_t) \in S_t(p)\), (where \( S_t(p) \) was defined in Chapter 3), for all \( l \).

2. \( \sum_{j=1}^{p} r_{ij} = r \) for some integer \( r \), for all \( i \) and for all \( l \).

Integral mixture mates are more general structures because the preservation of numerical sums is a much weaker condition than the preservation of symbolic sums. Thus, symbolic mixture mates may be seen as a special case of integral mixture mates. By providing methods to construct integral mixture mates, the class of orthogonal block designs for mixture experiments can be enriched, enabling the practitioner to patch together both integral and symbolic mixture mates (see Section 4.3.4) to
produce orthogonal block designs having the desired number of blocks and block sizes.

The crucial recognition is that, given a set of $b$ mixture mates of strength $t$, symbolic or integral, we automatically have an orthogonal block design for a mixture experiment having $b$ blocks.

A caveat on the use of mixture mates, symbolic or integral, is that they will often need to be augmented with one or more distinct runs to prevent the corresponding model matrix for a $t$-th order canonical polynomial from being singular. In some cases, such as Latin squares and perpendicular arrays, there is an obvious reason for this; for more general structures, the need and reason for augmentation are not as clear. In general, dependencies among the columns of the model matrices exist whenever all symbols occur the same number of times in each row.

Suppose that we consider fitting a $t$-th order canonical polynomial. Both perpendicular arrays and Latin squares are row consistent, so the first $p$ columns in the corresponding model matrix add to a constant. For a given mixture, the all of the symbols will appear exactly once in the first $p$ columns of the model matrix. Because of this, all possible products of pairs of elements in the $\binom{p}{2}$ columns corresponding to the quadratic portion of the model. Thus, for each mixture, these columns also add to a constant, the sum of the $\binom{p}{2}$ products of pairs of symbols. This produces a dependence between the columns for the linear portion of the model and the columns corresponding to the quadratic portion, decreasing the model matrix's rank by one. This same event occurs when cubic and higher order terms are included in the model.

To rectify the situation, the design must be augmented. The added mixtures may not be mixtures already in the design or other mixtures where each symbol
appears an equal number of times.

The next two sections provide some theory for constructing \( b \) integral mixture mates of strength \( t \) through the use of trades on \( m \)-ary block designs for treatments.

### 4.3.2 Results from Trade-off Theory

This section discusses how to construct several equivalent block designs for treatments. Section 4.3.3 converts the block-treatment incidences matrices from the equivalent block designs for treatments into \( b \) integral mixture mates of strength \( t \).

Let \( V = \{1, \ldots, v\} \) be a set of \( v \) treatments and \( v \Sigma^m_k \) be the set of \( k \)-collections of the elements of \( V \) where each element may appear 0, 1, \ldots, or \( m - 1 \) times in each collection.

**Definition 4.3.6** The block-treatment incidence matrix for a block design, \( M \), is the \( n \times v \) matrix whose \((i,j)\)-th entry, \( m_{ij} \), corresponds to the number of replications of treatment \( j \) in block \( i \).

Associated with a block-treatment incidence matrix is a block design for treatments.

**Definition 4.3.7** A proper, \( m \)-ary block design for treatments is a block design for which \( v \) treatments are arranged in blocks of equal size.

Let \( M \) a block-treatment incidence matrix for a proper \( m \)-ary block design for treatments. Then \( M \) has the property \( M1_v = k1_n \) for some block size \( k \) and whose entries \( m_{ij} \in \{0, \ldots, m - 1\} \). This implies that the block-treatment incidence matrix for a proper, \( m \)-ary block design is row-consistent in the sense of Definition 4.3.5. Note that \( m \)-ary block designs are often called \( n \)- or \( p \)-ary block designs in the literature.
In addition, a proper, \( m \)-ary block design for treatments is \( balanced \) if \( C = \text{diag}(r_1, \ldots, r_v) - (1/k)MM' \), where \( r_j \) is the number of times treatment \( j \) is replicated, is completely symmetric, i.e. all off-diagonal elements are equal. If in Definition 4.3.7, \( c = 2 \) and the design is balanced, we have a \( balanced \) \( incomplete \) \( block \) or \( BIB \) design. For more information on \( BIB \) designs, see Raghavarao (1988).

To illustrate the concepts discussed in this section, consider the pair of Latin squares in Table 4.2. View each Latin square as a block-treatment incidence matrix of a proper, \( 6 \)-ary block design for treatments where column \( j \) corresponds to treatment \( j \), row \( i \) to block \( i \), and the \((i,j)\)-th entry \( m_{ij} \) corresponds to the number of times treatment \( j \) appears in block \( i \). These block-treatment incidence matrices are equivalent in the following sense: the number of replications (column sum) for each treatment is the same for both designs, and the number of times two treatments appear together in a block (the sum of the entries of the Schur product of the corresponding pair of columns) is the same for both designs. Thus, we have a pair of integral mixture mates of order two.

"Equivalent" block designs for treatments can be constructed using \( trades \).

**Definition 4.3.8** Let \( T_1 \) and \( T_2 \) be two disjoint collections of blocks from \( v\Sigma^m_k \) such that \( |T_1| = |T_2| = n \). Then \((T_1,T_2)\) is called a \((v,k,t)\)-trade if for a fixed \( t \) and all \((j_1, \ldots, j_t) \in S_t(v)\)

\[
\sum_{i=1}^n m_{ij_1} \cdots m_{ij_t}
\]  

is independent of \( T_1 \) and \( T_2 \).

Definition 4.3.8 provides the appropriate generalization of Hedayat’s definition of a trade to \( m \)-ary block designs. When \( m = 2 \) it reduces to Hedayat’s definition.
Heuristically, \((v, k, t)\)-trades produce proper \(m\)-ary block designs for treatments whose matrices \(C = \text{diag}(r_1, \ldots, r_v) - (1/k)MM'\) are identical. More salient to the construction of mixture designs, the block-treatment incidence matrices have identical column sums and sums of Schur products of \(t\)-tuples of columns for any group of \(t\). Thus, the block-treatment incidence matrices for a group of proper, \(m\)-ary block designs for treatments obtained by \((v, k, t)\)-trades satisfy the conditions of Definition 4.3.5.

Clearly, the preservation of symbolic products of \(t\)-tuples of columns is a special case of the preservation of numerical sums of Schur products of \(t\)-tuples of columns. Thus, as mentioned in Section 4.3.1, the construction of mixture mates using perpendicular arrays and Latin squares is a special case of the trade-off. When symbolic products and sums are preserved, there is more freedom in assigning the entries of \(M\) to proportions.

One major drawback to trade-off is that trade-off theory for binary designs is in its infancy and results for \(m\)-ary block designs are non-existent. As a result, for purposes of illustration, future examples are constructed from binary designs. Trades on BIB designs may be performed through simple permutations of the incidence matrix columns. Not all possible trades may be found through permutations. Hedayat (1990) and Khosrovshahi and Ajoodani-Namini (1988) provide more sophisticated ways of devising trades on BIB designs. Restricting attention to \(m\)-ary designs constructed from binary designs is unnecessary, but finding trades for \(m\)-ary designs would almost certainly require the use of a computer algorithm.

Consider an example of trade-off in a BIB (binary) design for treatments with \(v = n = 7\), \(k = r = 3\), and \(\lambda = 1\), denote this by a BIB(7,7,3,3,1); \(\lambda\) denotes the
number of blocks in which two treatments appear together. Four block-treatment incidence matrices for a BIB(7,7,3,3,1) design are given in Table 4.3. Designs 1-3 can be obtained from one another via (7, 3, 2)-trades. Design 2 by interchanging columns 1 and 6 of Design 1’s block-treatment incidence matrix: Design 3 may be obtained by interchanging the columns 3 and 4 of Design 1’s block-treatment incidence matrix. Design 4 is presented for future reference.

Table 4.3: Four equivalent BIB Designs constructed via (7,3,2)-trades

<table>
<thead>
<tr>
<th>Design 1</th>
<th>Design 2</th>
<th>Design 3</th>
<th>Design 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 0 0 0 0 1 0</td>
<td>0 0 0 0 1 1 1</td>
<td>1 0 0 0 1 0 1</td>
<td>1 1 0 0 0 1 0</td>
</tr>
<tr>
<td>1 1 0 0 0 1 0</td>
<td>1 1 0 0 0 1 0</td>
<td>1 1 0 0 0 1 0</td>
<td>1 0 0 0 1 0 1</td>
</tr>
<tr>
<td>0 1 1 0 0 0 1</td>
<td>0 1 1 0 0 0 1</td>
<td>0 1 0 1 0 0 1</td>
<td>1 0 1 0 1 0 0</td>
</tr>
<tr>
<td>1 0 1 1 0 0 0</td>
<td>0 1 1 0 1 0 0</td>
<td>1 0 1 1 0 0 0</td>
<td>0 1 0 1 1 0 0</td>
</tr>
<tr>
<td>0 1 0 1 1 0 0</td>
<td>0 1 0 1 1 0 0</td>
<td>0 1 1 0 1 0 0</td>
<td>0 1 1 0 0 1 0</td>
</tr>
<tr>
<td>0 0 1 0 1 1 0</td>
<td>1 0 1 0 1 0 0</td>
<td>0 0 1 1 1 0 0</td>
<td>0 0 0 1 1 1 0</td>
</tr>
<tr>
<td>0 0 0 1 0 1 1</td>
<td>1 0 0 1 0 1 1</td>
<td>0 0 1 0 0 1 1</td>
<td>0 1 1 0 0 0 1</td>
</tr>
</tbody>
</table>

4.3.3 The Method of Trade-Off Applied to Constructing Mixture Mates

Trades on m-ary block designs provide the following method for constructing integral mixture mates of strength $t$.

Select a $m$-ary block design for treatments having $n$ blocks and $p$ treatments. Choosing $t$ to be larger than $d$ since at least $d + 1$ levels is a sufficient condition for $d$-th order parameters to be estimated orthogonally to the block contrasts. Selecting $m$ to be at least $d + 1$ is sufficient for the $d$-th order canonical polynomial parameters to be estimable. Next, generate $b - 1$ $(p, k, t)$-trades on the $m$-ary block design for treatments. One must ensure that the trades are unique; otherwise some of the block designs for treatments will be identical. Using the trades generated, the
incidence matrices for \( b-1 \) additional block designs for treatments can be constructed, producing \( b \) integral mixture mates of strength \( t \).

Once the \( b \) integral mixture mates of strength \( t \) have been constructed, assign proportions to the \( m_{ij} \)'s in the incidence matrices. The \( m_{ij} \)'s that appear in more than one incidence matrix must be assigned the same proportion to preserve orthogonality. This assignment may be conducted in several ways. The simplest is to divide each integer in the incidence matrix by \( k \), the row sum. Let \( 0, \ldots, m-1 \) be the integers that appear in the incidence matrices and associate these with indeterminates \( x_0, \ldots, x_{m-1} \). Let the matrix \( A \) have \( m \) columns and have the number of rows correspond to the number of unique patterns of symbols in the \( b \) integral mixture mates. Take the entries of \( A \) to be the multiplicities of \( 0, \ldots, m-1 \) in these unique patterns. Another method of assigning proportions to the symbols is to solve the system of equations

\[
A\chi = 1
\]

subject to the constraint that the solutions for the \( \chi_i \)'s must lie in the interval \([0, 1]\). The examples presented use only \( m \)-ary block designs constructed using BIB designs. This is unnecessary, but given the computational issues surrounding trades for even binary designs, the designs constructed are adequate to illustrate the techniques involved. The methods used to construct \( m \)-ary block designs which use BIB designs include Murty and Das (1967), Das and Rao (1968), Nigam (1974), Sinha and Saha (1979), and Sujatha and Surendran (1987). For each of these methods of construction, a trade on a BIB design used in the construction produces a trade on the resulting \( m \)-ary block design.

One drawback of restricting attention to \( m \)-ary designs constructed using BIB
designs is that BIB's typically have a large number of blocks due to the identity $n \geq v$, so the block size of the mixture experiment is large. Exacerbating the problem is that some of the $m$-ary designs are constructed by taking Kronecker products of the block-treatment incidence matrices of BIB designs and other methods that enlarge the number of rows in the incidence matrix. A partial remedy of the large block size is to delete rows occurring in every block, decreasing the block size without disturbing orthogonality.

Another problem with the trade-off method for constructing block designs for mixture experiments is that the designs produced are not readily adaptable to situations in which the mixture components are constrained to lie in a subspace of the simplex.

Consider the following method of obtaining three integral mixture mates of strength two. Let $M_i, i = 1, 2, 3, 4$ be the incidence matrices from Table 4.3. The following method of Das and Rao (1968) can be used to construct three equivalent 4-ary designs with incidence matrices, $\hat{M}_i$, in Table 4.4 given by

$$\hat{M}_i = M_i M_4$$

for $i = 1, 2, 3$. Notice that these are integral mixture mates:

$$m'_{j,i}l_7 = 9, \text{ for all } j \text{ and for all } l,$$

$$m'_{j,i}m_{j,i} = 11, \text{ for nonidentical all pairs } (j_1, j_2) \in S_2(7) \text{ and } l = 1, 2, 3, 4,$$

and

$$m'_{j,i}m_{j,i} = 15 \text{ for identical all pairs in } S_2(7).$$

To be used in mixture experiments, the numbers 0, 1, 2, 3 must be assigned to proportions. One possible assignment is $0 \rightarrow 0, 1 \rightarrow \frac{1}{5}, 2 \rightarrow \frac{1}{5}, \text{ and } 3 \rightarrow \frac{1}{3}$. Another
method to assign the integers 0,1,2 to proportions is to associate the integers with indeterminates $\chi_0, \chi_1, \chi_2, \chi_3$ whose solutions are restricted to $\chi_i \in [0,1]$, and solve the following system of equations:

$$\chi_0 + 3\chi_1 + 3\chi_2 = 1$$
$$6\chi_1 + \chi_3 = 1$$

This produces the solutions $0 \leq \chi_1 \leq 1/6$, $0 \leq \chi_2 \leq 1/3$, $\chi_0 = 1 - 6\chi_1$, and $\chi_3 = 1 - 6\chi_1$. Either assignment forms a blocking arrangement for 7 mixture components in 3 orthogonal blocks. Notice that one run, $(2 1 0 1 1 2 2)$, is common to all three designs and may be eliminated without disturbing the orthogonality conditions.

Suppose that we wish to construct $b = 4$ mixture mates for $p = 6$ mixture components and fit a quadratic canonical polynomial. One (not unique) way of doing this is based on a BIB design for treatments having 13 treatments and blocks, block size and treatment replication of 4, and $\lambda = 1$. Table 4.5 gives an incidence matrix for a $BIB(13,13,4,4,1)$. To construct a 3-ary block design, add 5 pairs of columns and one triplet of columns together. To get trades, simply add different pairs and triplets of columns together. The four block designs for treatments in Table 4.6 were obtained as follows:

### Table 4.4: A triplet of integral mixture mates of strength two for $p=7$

<table>
<thead>
<tr>
<th></th>
<th>$M_1$</th>
<th>$M_2$</th>
<th>$M_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

This produces the solutions $0 \leq \chi_1 \leq 1/6$, $0 \leq \chi_2 \leq 1/3$, $\chi_0 = 1 - 6\chi_1$, and $\chi_3 = 1 - 6\chi_1$. Either assignment forms a blocking arrangement for 7 mixture components in 3 orthogonal blocks. Notice that one run, $(2 1 0 1 1 2 2)$, is common to all three designs and may be eliminated without disturbing the orthogonality conditions.

Suppose that we wish to construct $b = 4$ mixture mates for $p = 6$ mixture components and fit a quadratic canonical polynomial. One (not unique) way of doing this is based on a BIB design for treatments having 13 treatments and blocks, block size and treatment replication of 4, and $\lambda = 1$. Table 4.5 gives an incidence matrix for a $BIB(13,13,4,4,1)$. To construct a 3-ary block design, add 5 pairs of columns and one triplet of columns together. To get trades, simply add different pairs and triplets of columns together. The four block designs for treatments in Table 4.6 were obtained as follows:
Design 1—add columns \((1,2,13),(3,4),(5,6),(7,8),(9,10),(11,12)\)

Design 2—add columns \((1,2,12),(3,4),(5,6),(7,8),(9,10),(11,13)\)

Design 3—add columns \((1,11,13),(2,3),(5,6),(7,8),(9,10),(4,12)\)

Design 4—add columns \((1,2,13),(4,12),(3,6),(7,8),(9,10),(5,11)\).

Table 4.6 gives the resulting 3-ary block designs, eliminating some of the runs common to all four block designs.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>0</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

All that remains is to assign proportions to \(0, 1,\) and \(2\). Let \(\chi_0, \chi_1,\) and \(\chi_2\) be in \([0, 1]\). Then, the equations \(A\chi = 1\) are

\[
2\chi_0 + 4\chi_1 = 1
\]

\[
3\chi_0 + 2\chi_1 + \chi_2 = 1.
\]

The solutions to these equations are

\[
\chi_0 = \frac{1}{2} - 2\chi_1
\]
Table 4.6: Four mixture mates of strength two derived from 
(13,3,2)-trades on a BIB(13,13,4,4,1) design

<table>
<thead>
<tr>
<th></th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>210010</td>
<td>210010</td>
<td>120010</td>
<td>201010</td>
<td></td>
</tr>
<tr>
<td>012010</td>
<td>012010</td>
<td>002011</td>
<td>011011</td>
<td></td>
</tr>
<tr>
<td>110101</td>
<td>110101</td>
<td>200101</td>
<td>110101</td>
<td></td>
</tr>
<tr>
<td>011011</td>
<td>011011</td>
<td>111010</td>
<td>002011</td>
<td></td>
</tr>
<tr>
<td>100012</td>
<td>100012</td>
<td>200011</td>
<td>002011</td>
<td></td>
</tr>
<tr>
<td>101011</td>
<td>201010</td>
<td>101011</td>
<td>110011</td>
<td></td>
</tr>
<tr>
<td>111100</td>
<td>011101</td>
<td>111100</td>
<td>101101</td>
<td></td>
</tr>
<tr>
<td>020101</td>
<td>120100</td>
<td>010102</td>
<td>021100</td>
<td></td>
</tr>
<tr>
<td>201100</td>
<td>101101</td>
<td>201100</td>
<td>201100</td>
<td></td>
</tr>
<tr>
<td>210010</td>
<td>110011</td>
<td>110011</td>
<td>210010</td>
<td></td>
</tr>
<tr>
<td>101101</td>
<td>201100</td>
<td>011101</td>
<td>111100</td>
<td></td>
</tr>
</tbody>
</table>

\[
\frac{1}{8} \leq \chi_1 \leq \frac{1}{4}
\]

\[
\chi_2 = 4\chi_1 - \frac{1}{2}. 
\]

This design has 23 distinct runs spread among the four blocks.

4.3.4 Constructing Mixture Mates by Patchwork

This section discusses methods of constructing a large group of mixture mates (integral or symbolic or a combination) by patching together several smaller sets of mixture mates, each set having a common strength. The methods contained in this section are useful if it is easier to produce mixture mates of common strength for small \( b \) than to construct mixture mates of common strength \( a \) for larger \( b \). Using these methods, one can produce virtually any number of mixture mates with a common strength. However, mates constructed by patchwork may have a large number of mixtures per mate.
Suppose that we can easily construct $q$ sets of mixture mates of common strength $t_i$ with the $i$-th set containing $b_i$ mates. Each set of $b_i$ mixture mates might be constructed using the techniques of the previous section. Interest lies in constructing $b'$ mixture mates, where $\max\{b_1, b_2, \ldots, b_q\} < b' \leq b_1 \times b_2 \times \cdots \times b_q$.

Construct a $b_1 \times b_2 \times \cdots \times b_q$ factorial design. If $b'$ is much smaller than the product, a fractional factorial may be considered. Associate the $b_i$-level factor with the $b_i$ mixture mates of strength $t_i$. Arbitrarily assign each of these mates to one of the $b_i$ factor levels. Repeat this for each of the $q$ sets of mixture mates. Then each run in the (fractional) factorial design represents a "new" mate built from the $q$ sets of "old" mates, one mate from each set. Finally, select $b'$ of the $b_1 \times b_2 \times \cdots \times b_q$ (or fraction thereof) mates formed in this manner to make up the required $b'$ orthogonal blocks.

The $b'$ mates chosen depend on the application. For instance if the blocks represent process variable level combinations, the $b'$ mates chosen should form a (fractional) factorial design where the process variable main effects and interactions of interest are estimable. These main effects and interactions will be estimated by estimators of block contrasts. Because $\hat{t}$ is orthogonal to the block contrast estimators, it is uncorrelated with the estimators of process variable main effects and interactions.

The method given in Draper, Prescott, Lewis, Dean, John, and Tuck (1993) and Prescott, Draper, Dean, John and Lewis (1993) is a special case of the method given above, taking $b_i = 2$ for all $i$.

This method allows us to patch together mates constructed using the Latin square method and the trade-off method. This allows for greater flexibility in design size, than does the exclusive use of Latin squares. In general, using a larger number
of mates per set reduces the number of mates needed in a block, potentially reducing the block size. For example, consider, the case of finding ten blocks. Using pairs of mixture mates requires 4 mates per block; using a set of 4 mixture mates requires only two mates per block. Unlike randomized complete blocks, using mates produces a design having a larger number of distinct runs, allowing lack of fit tests to be performed.

One drawback is that as the number of blocks $b'$ increases, then so does the block size. This is evident since the number of mates per block increases with $b'$. Section 4.5 provides techniques of nonorthogonal blocking which produces smaller block sizes, but at the expense of having a design for which the estimators of the mixture parameters are not orthogonal to all block contrasts.

To illustrate the patchwork technique, consider the case of finding more than two blocks using pairs of mates of strength 2. Each pair of mates forms a design having two blocks. Once a pair of mates have been found, they can be combined with other pairs to form designs having more than two blocks. Let $M_i^{(0)}$ and $M_i^{(1)}$, $i = 1, 2, 3$ be three pairs of mixture mates. Form a $2^3$ factorial design and assign factor $i$ to the $i$-th pair of mixture mates. Then for example associate the high levels with $M_i^{(1)}$ and the low levels with $M_i^{(0)}$ for each $i$. This gives the design in Table 4.7, taken from Draper, Prescott, Lewis, Dean, John, and Tuck (1993).

<table>
<thead>
<tr>
<th>Block 1</th>
<th>Block 2</th>
<th>Block 3</th>
<th>Block 4</th>
<th>Block 5</th>
<th>Block 6</th>
<th>Block 7</th>
<th>Block 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1^{(0)}$</td>
<td>$M_1^{(1)}$</td>
<td>$M_1^{(0)}$</td>
<td>$M_1^{(1)}$</td>
<td>$M_1^{(0)}$</td>
<td>$M_1^{(1)}$</td>
<td>$M_1^{(0)}$</td>
<td>$M_1^{(1)}$</td>
</tr>
<tr>
<td>$M_2^{(0)}$</td>
<td>$M_2^{(0)}$</td>
<td>$M_2^{(1)}$</td>
<td>$M_2^{(0)}$</td>
<td>$M_2^{(0)}$</td>
<td>$M_2^{(0)}$</td>
<td>$M_2^{(0)}$</td>
<td>$M_2^{(0)}$</td>
</tr>
<tr>
<td>$M_3^{(0)}$</td>
<td>$M_3^{(0)}$</td>
<td>$M_3^{(0)}$</td>
<td>$M_3^{(0)}$</td>
<td>$M_3^{(0)}$</td>
<td>$M_3^{(0)}$</td>
<td>$M_3^{(0)}$</td>
<td>$M_3^{(0)}$</td>
</tr>
</tbody>
</table>

Table 4.7: Eight mixture mates formed by the patchwork method
As is, the table is a blocking arrangement for $b = 8$ blocks. If fewer than eight blocks are required, any combination of the eight blocks may be used to form the desired number of blocks. For fewer than four blocks, it is unnecessary to have three mates in a block.

Consider finding $b' = 10$ blocks using sets of 4 mixture mates. Using sets of 4 mixture mates allows the use of only two mates per block, possibly giving a reduction in block size over choosing pairs of mates which would require 4 mates per block or triplets of mates, requiring three mates per block. For this scenario, select $q = 2$. Construct two sets of 4 mixture mates of strength $t$, say $M_1^{(0)}, M_1^{(1)}, M_1^{(2)}, M_1^{(3)}$ and $M_2^{(0)}, M_2^{(1)}, M_2^{(2)}, M_2^{(3)}$. Form a $4^2$ factorial design with runs

$$(00, 10, 20, 30, 01, 11, 21, 31, 02, 12, 22, 32, 03, 13, 23, 33).$$

Then, make the following assignments: $M_1^{(1)} \rightarrow 0, M_1^{(2)} \rightarrow 1, M_1^{(3)} \rightarrow 2, M_1^{(4)} \rightarrow 3$. Similarly for the second set of 4 mixture mates. This produces the blocking arrangement in Table 4.8. Then choose any $b' = 10$ of the blocks.

Finally, consider an alternative to the design of Table 4.7. Suppose that we have $q = 2$ sets of mixture mates—4 mixture mates of strength $t$ and 2 mixture mates also of strength $t$. Then, construct a $4 \times 2$ factorial with level combinations $(00, 10, 20, 30, 01, 11, 21, 31)$. Make the following associations: $M_1^{(0)} \rightarrow 0, M_1^{(1)} \rightarrow 1, M_1^{(2)} \rightarrow 2, M_1^{(3)} \rightarrow 3$ and $M_2^{(0)} \rightarrow 0$ and $M_2^{(1)} \rightarrow 1$. This gives the block design in Table 4.9.
### 4.4 Construction of Block Designs in Constrained Regions Using Orthogonal Arrays

When $S$ is a constrained region, a non-mate based method for constructing orthogonal block designs for mixture experiments borrows from confounding in $d^p$ (fractional) factorial designs and asymmetrical (mixed-level) orthogonal arrays.

Piepel (1990, 1991) described a technique to construct screening designs for mixture experiments using screening designs for $p - 1$ process variables as a template. Piepel only considers the case of $d = 2$ levels, but the method works for arbitrary $d$.

Throughout this section assume that $S$ is a constrained subregion of the simplex of the form

$$0 \leq L_j \leq z_j \leq U_j \leq 1 \quad j = 1, \ldots, p.$$

The constraints may arise from physical or economic considerations.

**Definition 4.4.1** An asymmetrical (mixed-level) orthogonal array of strength $t$ with

---

**Table 4.8: A Set of 16 orthogonal blocks**

<table>
<thead>
<tr>
<th>Block 1</th>
<th>Block 2</th>
<th>Block 3</th>
<th>Block 4</th>
<th>Block 5</th>
<th>Block 6</th>
<th>Block 7</th>
<th>Block 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1^{(0)}$</td>
<td>$M_1^{(1)}$</td>
<td>$M_1^{(2)}$</td>
<td>$M_1^{(3)}$</td>
<td>$M_1^{(0)}$</td>
<td>$M_1^{(1)}$</td>
<td>$M_1^{(2)}$</td>
<td>$M_1^{(3)}$</td>
</tr>
<tr>
<td>$M_2^{(0)}$</td>
<td>$M_2^{(1)}$</td>
<td>$M_2^{(2)}$</td>
<td>$M_2^{(3)}$</td>
<td>$M_2^{(0)}$</td>
<td>$M_2^{(1)}$</td>
<td>$M_2^{(2)}$</td>
<td>$M_2^{(3)}$</td>
</tr>
</tbody>
</table>

**Table 4.9: A Set of 8 orthogonal blocks**

<table>
<thead>
<tr>
<th>Block 1</th>
<th>Block 2</th>
<th>Block 3</th>
<th>Block 4</th>
<th>Block 5</th>
<th>Block 6</th>
<th>Block 7</th>
<th>Block 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_1^{(0)}$</td>
<td>$M_1^{(1)}$</td>
<td>$M_1^{(2)}$</td>
<td>$M_1^{(3)}$</td>
<td>$M_1^{(0)}$</td>
<td>$M_1^{(1)}$</td>
<td>$M_1^{(2)}$</td>
<td>$M_1^{(3)}$</td>
</tr>
<tr>
<td>$M_2^{(0)}$</td>
<td>$M_2^{(1)}$</td>
<td>$M_2^{(2)}$</td>
<td>$M_2^{(3)}$</td>
<td>$M_2^{(0)}$</td>
<td>$M_2^{(1)}$</td>
<td>$M_2^{(2)}$</td>
<td>$M_2^{(3)}$</td>
</tr>
</tbody>
</table>
N runs and p factors, $p_i$ of which have $d_i$ levels, $i = 1, \ldots, e$, denoted by $OA(N, \prod_{i=1}^{e} d_i^{p_i}, t)$, is a $N \times p$ array in which $p_i$ of the columns are based on symbol set $D_i$ with cardinality $d_i$ where for every distinct $t$-tuple of columns, say $(i_1, \ldots, i_t)$, every element from $D_{i_1} \times D_{i_2} \times \cdots \times D_{i_t}$ appears equally often as a row of the $N \times t$ subarray. If $d_i = d$ for all $i$, then we have an orthogonal array.

An orthogonal fraction of a factorial design is a fraction which allows unbiased estimation of the main effects and certain interactions deemed important. The corresponding estimators of these main effects and interactions are uncorrelated, i.e. if $\hat{\beta}$ is the estimator of these main effects and interactions, $\text{var}(\hat{\beta}) = D$, where $D$ is some diagonal matrix.

This method uses a $b \times d^{(p-1)}$ (fractional) factorial experiment. One can think of a $d^{(p-1)}$ factorial experiment confounded into $b$ ($b$ a power of $d$) blocks as a $b \times d^{(p-1)}$ fractional factorial experiment, the $b$-level factor corresponding to the blocking variable. Dey (1985), Raghavarao (1988), Wang and Wu (1991), Hedayat, Pu, and Stufken (1992), and Wu, Zhang, and Zhang (1992), among others, provide methods of constructing asymmetric orthogonal arrays. More generally, we can consider $b \times d^{(p-1)}$ factorial experiments where $b$ is not necessarily a power of $d$. Addelman (1962) shows how to construct orthogonal fractions of asymmetrical experiments.

Construct a $b \times d^{(p-1)}$ experiment using a method such as one of those listed above. Associate the $b$-level factor with the blocking variable and select $p - 1$ of the $p$ mixture components to associate with the $d$-level factors. The $p - 1$ mixture components used will depend on the relative importance of each of the components, an issue discussed later. For each of the $p - 1$ mixture components assign a proportion to each of the $d$ factor levels. This gives a $p$ column array with one column associated
with a \( b \)-level factor and the remainder of the array filled out by proportions.

Create a \((p + 1)\)-st column in the array. This additional column corresponds to the \( p \)-th mixture component. For each row (excepting the blocking factor) find 
\[
z_p = 1 - \sum_{j=1}^{p-1} z_j
\] and enter it into the additional column. Apply the following rule: if \( z_p < L_p \) or \( z_p > U_p \) in a row, set \( z_p \) to be the closer of \( L_p \) or \( U_p \). Then try to adjust \( z_{p-1} \) for that row so that \( \sum_{j=1}^{p} z_j = 1 \) while still keeping \( z_{p-1} \) within its constraints. If the adjustment to \( z_{p-1} \) fails to bring the row sum to one, fix \( z_{p-1} \) at the value bringing \( \sum_{j=1}^{p} z_j \) the closest to one and continue to work backwards toward \( z_1 \), making adjustments to each component which keep all of the components within their constraints and make the row sum one. Repeat this adjustment process until each row sums to one and each entry in the array (not including the blocking factor) is within its constraints. To create a block design, assign row \( i \) to block \( l \) if the \( b \)-level factor is at level \( l \) in row \( i \).

Piepel suggests to assign the \( p - 1 \) components with the smallest values of \( R_j = U_j - L_j \) to the \( p - 1 \) columns of the \( d \)-level factors in the \( b \times d^{p-1} \) factorial. Piepel bases this recommendation on the ability of the components to absorb adjustments. The larger the range of \( z_j \) is, the more able \( z_j \) is to absorb a change without forcing an adjustment to \( z_{j-1} \). Other choices are possible. The best way might be to assign the first \( p - 1 \) columns to the most important \( p - 1 \) mixture components from the experimenter's point of view. The components that are assigned to the first \( p - 1 \) columns are likely to go through less adjustment than are the components in the last few columns.

The first example illustrating this procedure involves a confounded two-level fractional factorial design; the second uses an asymmetrical orthogonal array.
Saxena and Nigam (1977) consider a study of several different constituents of a lubricant oil and use a symmetric-simplex design [c.f. Murty and Das (1968)]. Blocking is artificially introduced here for the sake of example. The constituents are: pour point depressant \((z_1)\), detergent dispersant \((z_2)\), anti-oxidants \((z_3)\), viscosity index improvers \((z_4)\), and SAE-10 oil \((z_5)\). These constituents have the constraints

\[
0.0010 \leq z_1 \leq 0.0050 \\
0.0005 \leq z_2 \leq 0.0100 \\
0.0005 \leq z_3 \leq 0.0100 \\
0.0400 \leq z_4 \leq 0.1000 \\
0.8750 \leq z_5 \leq 0.9580.
\]

A design capable of fitting a first-order canonical polynomial may be formed using a 2^5 fractional factorial design with defining relation \(I=12345\) that has been split into two blocks. The first column of the array corresponds to a blocking factor and the five two level factors correspond to columns 2-6. After assigning the extreme proportions to levels \(-1\) and \(1\), and applying the method outlined above, the design is for all intensive purposes orthogonally blocked.

The next example looks at a design for fitting a second-order model. To do so, we use a orthogonal fraction of a 5 \(\times\) 3^3 factorial design as a template.

Anik and Sukumar (1984) modeled solubility of butoconazole nitrate in a five component solution: polyethylene glycol 400 \((z_1)\), glycerin \((z_2)\), polysorbate 60 \((z_3)\), water \((z_4)\), and poloxamer 407. Poloxamer 407 was held constant, presumably at some optimum level found by some preliminary investigations. Anik and Sukumar used an extreme vertices design and did not consider blocking. The following constraints
defined the experimental region $S$ for the other four variables:

$$z_1 + z_2 + z_3 + z_4 = 0.9$$

$$0.1 \leq z_1 \leq 0.4$$

$$0.1 \leq z_2 \leq 0.4$$

$$0.0 \leq z_3 \leq 0.08$$

$$0.3 \leq z_4 \leq 0.7$$

The researchers’ objective was to model solubility as a function of the mixture components. They assumed that a second-order model would be adequate. For purposes of illustration, assume that a design with 5 blocks is desirable. We choose 5 blocks to show how flexible the method is regarding the number of blocks.

In this case take $d = 3$ since we are attempting to fit a second-order canonical polynomial. Construct a $5 \times 3^3$ fractional factorial design using Addelman’s (1962)
method. The design is given in Table 4.11. The 5-level factor in column one is the blocking factor; all runs at a common level of the blocking factor are grouped into a block. Then, for each mixture variable, assign the levels 0, 1, 2 as follows:

\[
\begin{array}{ccc}
 z_1 & z_2 & z_3 \\
 0 & 0.10 & 0.00 \\
 1 & 0.25 & 0.04 \\
 2 & 0.40 & 0.08 \\
\end{array}
\]

The estimator of \( \tau \) is not orthogonal to the block contrast estimators. The global measure of design efficiency with respect to the unblocked design, defined in Section 2.4, is

\[
\left( \frac{|Z_2'(I-P_\gamma)Z_2|}{|Z_2'Z_2|} \right)^{1/10} = 0.905.
\]

Recall that \( |Z_2'Z_2| \) is a possibly unattainable upper bound for \( |Z_2'(I-P_\gamma)Z_2| \), but provides a rough check on how our procedure performed. In this case it appears that the block contrast columns in the model matrix and the columns of the model matrix for the second-order canonical polynomial terms are nearly orthogonal. Estimators of individual elements of \( \tau \) are quite efficient.

\[
\begin{align*}
\text{eff}(\tau_1) &= 0.999 & \text{eff}(\tau_{13}) &= 0.824 \\
\text{eff}(\tau_2) &= 0.999 & \text{eff}(\tau_{14}) &= 0.923 \\
\text{eff}(\tau_3) &= 0.984 & \text{eff}(\tau_{23}) &= 0.647 \\
\text{eff}(\tau_4) &= 0.999 & \text{eff}(\tau_{24}) &= 0.932 \\
\text{eff}(\tau_{12}) &= 0.874 & \text{eff}(\tau_{34}) &= 0.934. \\
\end{align*}
\]

Aside from \( \tau_{23} \), all of the other parameter estimators have reasonably high efficiencies. This design may be reasonable if the goal is parameter estimation. The global
measure becomes salient when considering the model for estimating mean responses; for this purpose the design fares reasonably well. This is surprising since the $5 \times 3^3$ used did not permit orthogonal estimation of the linear by linear two factor interactions and the main effects, both of which are involved in the quadratic canonical polynomial parameter estimators.

In this $5 \times 3^3$ experiment for independent factors, the estimators of the main effects of the four factors in the array are uncorrelated with one another. However, in the array used estimators of two-factor interaction among the four factors are correlated. This is reflected in the high efficiencies of the estimators of the first-order canonical polynomial parameters, relative to the low efficiencies of the estimators of the second-order canonical polynomial parameters. Correlation of the two-factor interaction contrast estimators in the original four array factors translates into lower efficiency of the estimators of the parameters corresponding to the two-factor interactions in the canonical polynomial. In particular, correlation of the two-factor interaction contrast estimators and estimators of the main effects of the blocking variable are the culprit because of equation (2.6). Gosset [see Hardin and Sloane (1992)], an optimal design program found a $5 \times 3^3$ design that is better suited for the procedure than the design in Table 4.11 and the individual efficiencies of the parameter estimators are comparable. The global measure of efficiency is 0.937, quite a bit better than the design in Table 4.11.

Another contributor to low overall or individual efficiency is the adjustment step. The design for the original $p$ factors allows main effects and certain interactions to be estimated orthogonally to one another. This is true even when the symbols of the $p - 1$ $d$-level factors are initially assigned to proportions. However, in the adjustment
Table 4.11: Five blocks for the Anik and Sukumar example

<table>
<thead>
<tr>
<th>$u_1$</th>
<th>$z_1$</th>
<th>$z_2$</th>
<th>$z_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>
step, when too many of these proportions are changed from their initial proportions, the condition of zero correlation between \( \hat{\tau} \) and the block contrast estimators may be destroyed. Piepel (1991) indicates that when adjustments must be made to 40% or more of the first \( p - 1 \) columns, the orthogonality between columns of the model matrix is significantly altered, causing the design to deviate from an orthogonal block design.

This method works well when the structure of the initial design for the \( p \) factors remains intact after the levels have been assigned to proportions. This happens when there are relatively few adjustments that need to be made to the first \( p - 1 \) mixture components, as noted previously. However, even when few adjustments are made, there are cases when the \( p \)-th mixture component assumes a very limited number of levels, or is forced to assume values which are all near its upper or lower bound, producing a design which is not very attractive from a practical standpoint. This happens when \( z_p \) has a very small lower bound, usually near zero and when \( \sum_{j=1}^{p-1} z_j \) is frequently near one for many level combinations in the original array. Another case in which this occurs is when most of the mixture components have ranges \( R_j \) that are large, i.e. say over 0.5.

4.5 Nonorthogonal Block Designs for Mixture Experiments

Producing uncorrected estimators of \( \tau \) and \( \Lambda \gamma \) is a desirable property of a block design. However, conditions are not always amenable for orthogonal blocking. In particular, uncorrected estimators of \( \tau \) and all block contrast estimators is stronger than what is required. This section explores some options and formulates some methods of constructing block designs for mixture experiments when some of the block
contrasts are negligible. Assume that the experimenter has some knowledge about which block contrasts are important and which are negligible. For example, blocks may represent spatial or temporal units. In this case, it may be reasonable to assume that the effects of blocking are adequately represented by a smooth trend consisting of low-order orthogonal polynomial block contrasts. In this case, it is crucial that \( f \) be uncorrelated with estimators of these low-order polynomial block contrasts, but not necessarily estimators of the negligible higher-order contrasts. Another example arises when the blocks represent level combinations of process variables and it is assumed that certain process variable interactions are negligible. In this case, it is appropriate to find a block design for which \( f \) and the estimators of the block contrasts identified with main effects of the process variables are uncorrelated, but not those identified with the negligible interactions.

This scenario is likened to that of aliasing in factorial experiments where it is assumed \textit{a priori} that certain of the effects are likely negligible. If true, the effects of interest are estimated unbiasedly; if not the effects of interest may be biased by the effects thought to be negligible. The same is true for the approach of this section. If block contrasts thought to be negligible turn out to be nonnegligible, estimators of the mixture parameters may be highly correlated with the block contrasts.

Recall that the matrix \( U \) represents the blocking variables. \( U \) may be reparameterized as

\[
U = \begin{bmatrix} 1, \hat{V} \end{bmatrix} M,
\]

where \( \hat{V} \) represents the \( b - 1 \) orthogonal block contrasts and \( M \) is non-singular. Analogously, \( \gamma = M^{-1} \zeta \). Partition \( \hat{V} \) as

\[
\hat{V} = \begin{bmatrix} \hat{V}_2, \hat{V}_3 \end{bmatrix},
\]
where \( \hat{V}_2 \) represents the model matrix for the set of \( r \) contrasts deemed important and \( \hat{V}_3 \) the set of \( b-r-1 \) negligible block contrasts. Let \( \xi_2 \), \( r \times 1 \), and \( \xi_3 \), \( (b-r-1) \times 1 \), represent the corresponding vectors of regression parameters. The gives the model

\[ y = \hat{V}_2 \xi_2 + Z \tau + e \]  

(4.9)

**Definition 4.5.1** A block design is defined to be partially orthogonal if the following (equivalent) statements are true.

1. \( SS(\tau | \xi_2, \mu) = SS(\tau | \mu) \)

2. \( P_{\xi_2} Z = 0 \)

3. The BLUE of any estimable function of \( \tau \) under model (4.9) is the same regardless of whether \( \xi_2 \) is in the model or not.

When \( r = b - 1 \), Definition 4.5.1 corresponds to orthogonal blocking (Definition 2.4.1).

The first method exploits the structure of orthogonal polynomials for a given number of blocks \( b \). Fisher and Yates (1963) provide an extensive list of the coefficients of the orthogonal polynomials up to degree five for a variety of values of \( b \), where the blocking variable has equally spaced ordinal intervals. Taken as a vector, the coefficients of the odd degree polynomials are orthogonal to \( 6 \times 1 \) vectors whose last \( \frac{b}{2} \) entries are the first \( \frac{b}{2} \) entries repeated in the opposite order, for \( b \) even. For odd \( b \), the coefficient vectors are orthogonal to \( 6 \times 1 \) vectors whose \( \frac{b-1}{2} \)-th element is 0 and the last \( \frac{b-1}{2} \) entries are the first \( \frac{b-1}{2} \) repeated in opposite order. For example, the vector \((1,2,3,3,2,1)\) is orthogonal to the linear, cubic and quintic coefficients for \( b = 6 \).
Applying the result above, partially orthogonal block designs for mixture experiments can be constructed by the following method. For $b$ even, select $\frac{b}{2}$ blocks of equal size according to some criterion and order them, giving the first $\frac{b}{2}$ blocks. Repeat these blocks in the opposite order to obtain the remaining $\frac{b}{2}$ blocks. For $b$ odd, select $\frac{b-1}{2} + 1$ blocks and arbitrarily order these; repeat the first $\frac{b-1}{2}$ in the opposite order after the $\frac{b-1}{2} + 1$-st block. This procedure produces a block design which is partially orthogonally blocked with respect to all orthogonal polynomial contrasts of odd degree for $b$ blocks.

Note that this method is very flexible in that it does not require the use of mixture mates. Also, by judicious choice of mixtures in the blocks, it may be possible to achieve orthogonality to some of the orthogonal polynomials of even degree for a given $b$. However, this must be evaluated on a case by case basis. We reiterate that orthogonal or partially orthogonal blocking is somewhat of a secondary criteria and that mixtures should not be chosen solely to achieve orthogonal blocking.

To illustrate this technique, consider a small example. Suppose that we wish to construct a block design for five blocks. As discussed above, we select three blocks, say $B_1, B_2,$ and $B_3$. Then the block design will consist of the following five blocks, in order, $B_1, B_3, B_2, B_3,$ and $B_1$. To show that these are orthogonal to the odd order orthogonal polynomials for $b = 5$, consider the following. Recall Definition 4.5.1 that a block design for mixtures is partially orthogonal if $\hat{V}_2 Z = 0$. Let $z_{ki}$ be the entries in the $k$-th column of the model matrix corresponding to block $l$. Then, letting $\hat{V}_2$ contain the coefficients of the linear and cubic polynomials, we have that

$$\hat{V}_2 z_{k} = \begin{pmatrix} -21'_{n_1} z_{k1} - 1'_{n_2} z_{k3} + 0'_{n_2} z_{k2} + 1'_{n_3} z_{k3} + 21'_{n_1} z_{k1} \\ -1'_{n_1} z_{k1} + 21'_{n_2} z_{k3} + 0'_{n_2} z_{k2} - 21'_{n_3} z_{k3} + 1'_{n_1} z_{k1} \end{pmatrix}$$
The next method uses asymmetrical (mixed-level) orthogonal arrays (see Section 4.4) in combination with mixture mates to find partially orthogonal block designs.

Assume that model (4.9) holds and a design in \( b = kq \) blocks is desired. Further suppose that it is possible to construct \( k \) different sets of \( q \) mixture mates of strength \( t \) and that an \( OA(b, k \times \prod_{i=1}^{\xi} f_i^{t_i}, 2) \) exists for \( \sum_{i=1}^{\xi} b_i(f_i - 1) + k - 1 \leq b - 1 \) and \( f_i \geq 2 \) for all \( i = 1, \ldots, \xi \).

For \( i = 1, \ldots, \xi \), associate each of the \( b_i \) factors with \( f_i \) orthogonal block contrasts. That is, the main effects of the \( b_i \) factors having \( f_i \) levels will represent \( f_i - 1 \) of the \( b - 1 \) orthogonal block contrasts.

Assign each set of \( q \) mixture mates to one level of the \( k \)-level factor. Let \( M_1, \ldots, M_q \) be a set of mixture mates. Each time level \( j \) appears in the asymmetrical orthogonal array, substitute one of \( M_1, \ldots, M_q \). Repeat this until all of \( M_1, \ldots, M_q \) have been assigned to exactly one occurrence of level \( j \).

A sufficient condition for producing uncorrelated estimators of block contrasts and \( \tau \) is that each level of the \( f_i \)-level factors, \( i = 1, \ldots, \xi \), comes equally often with each level of the \( k \)-level factor, and thus with the \( q \) mixture mates. Because of this we need to use asymmetrical orthogonal arrays. Another important reason to use asymmetrical orthogonal arrays is so that the

\[
\sum_{i=1}^{\xi} b_i(f_i - 1) \leq b - k
\]

contrasts corresponding to the main effects of the \( f_i, \ i = 1, \ldots, \xi \) level factors will be orthogonal to one another. These contrasts represent the orthogonal block contrasts
that are uncorrelated with $\hat{\tau}$ and a $b - k$ of the possible $b - 1$ orthogonal contrasts that comprise the sum of squares due to blocks.

As an example, reconsider the problem of constructing a mixture experiment having eight blocks. The blocks may correspond to level combinations of 4 process variables under study. A $2^{4-1}$ fractional factorial design is used for the process variable. The template for this design is an orthogonal main effects plan given in Addelman (1962), an $OA(8, 4 \times 2^4, 2)$. Four pairs of mixture mates are constructed using the techniques of Section 4.3. Let $M_i^{(0)}$ and $M_i^{(1)}$ represent a pair of mates of strength $t$, for $i = 0, 1, 2, 3$. Associate each pair of mixture mates with one of the levels of the four level factor. Here $M_0^{(0)}$ and $M_0^{(1)}$ are associated with level 0, $M_1^{(0)}$ and $M_1^{(1)}$ are associated with level 1, $M_2^{(1)}$ and $M_2^{(1)}$ with level 2, and $M_3^{(0)}$ and $M_3^{(1)}$ with level 3. In this design the mixture parameter estimators are orthogonal to the four first-order terms in the process variables, but not to the interactions. If the blocks did not correspond to process variable settings, the mixture parameter estimators would be orthogonal to four of the seven block contrasts.

Table 4.12: A design for process variables and mixture components

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>$M_i^{(0)}$</th>
<th>$M_i^{(1)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>$M_0^{(0)}$</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>3</td>
<td>$M_1^{(0)}$</td>
</tr>
<tr>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>2</td>
<td>$M_2^{(0)}$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>$M_3^{(0)}$</td>
</tr>
<tr>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>$M_0^{(1)}$</td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>2</td>
<td>$M_1^{(1)}$</td>
</tr>
<tr>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>3</td>
<td>$M_2^{(1)}$</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$M_3^{(1)}$</td>
</tr>
</tbody>
</table>
Notice that the previous design takes advantage of foldover pairs in the process variables. Notice that for each pair of mixture mates all of the process variables come at their high level once and their low level in the other member. Because of this the pairs of mixture mates will be orthogonal to each pair of runs and all mixture mates taken together will be orthogonal to the four process variable main effects. This procedure has the potential to have much smaller block sizes that the procedure outlined by Draper, et. al. (1993) and Prescott, et. al. (1993), depending on the method used to construct the mixture mates.

4.6 Algorithmic Methods for Blocking Existing Designs

This section considers blocking a group of \( N \) mixtures that have been previously selected according to physical or economic considerations or an optimality criterion, possibly one discussed in the Appendix. The techniques proposed in previous sections may be unsuitable for mixtures selected according to such specifications. Unfortunately, there is no single "optimum" answer to this problem.

To solve this dilemma, we adopt the following tack. Block contrasts are viewed as nuisance parameters and the objective is to find an allocation of mixtures to blocks that minimizes the correlation between the block contrast estimators and \( \hat{\tau} \). Recall from Theorem 3.3.1 that the correlations are null if and only if \( P_{\hat{\tau}} Z = 0 \). Since the mixtures have been pre-selected \( Z \) is fixed up to row permutations, we propose maximizing \( | Z'(I - P_{\hat{\tau}})Z | \) over all possible row permutations, i.e. over all possible allocations to blocks. For a fixed set of mixtures, an upper bound (possibly a poor one) for \( | Z'(I - P_{\hat{\tau}})Z | \) is \( Z'Z \), the value attainable if orthogonal blocking is possible for the given set of mixtures. A measure of closeness to orthogonal blocking
is provided by the efficiency measure proposed in Section 2.4.

One method looks at all possible partitionings of the mixtures into \( b \) blocks and selects those (or the one) for which \( \left| Z'(I - P_Y)Z \right| \) is maximized. However, for \( N \) mixtures and \( b \) blocks of sizes \( n_l, \ l = 1, \ldots, b \), this approach is computationally intensive and probably unnecessary. For example, when \( N = 20 \) with two blocks of size \( n_1 = n_2 = 10 \), a complete search examines 92,378 possibilities.

Of course, not all of these possibilities will produce different designs in the eyes of these criteria; there will be different equivalence classes producing designs having the same value of the objective function. In general, finding these classes is also difficult and time consuming.

An alternate way to find the best design by this criterion is to sample from all possible allocations. This approach is flexible; it can handle deterministic and/or stochastic time trends and even situations for which there are trends within the blocks. The next section discusses the method of Cook and Nachtsheim (1989) and Section 4.6.2 introduces a methodology based on sampling from all possible run orders. These methods are applied to blocking the minimum bias designs of Draper and Lawrence (1965a,b) for unequal block sizes to demonstrate the methods' versatility. These mixtures cannot be blocked using any of the methods discussed earlier in this chapter.

4.6.1 Maximization of \( |Z'(I - P_Y)Z| \) Using an Interchange Algorithm

The method of Cook and Nachtsheim (1989) is actually more general than described here; it constructs response surface designs from a non-singular starting design and a set of candidate points and then blocks the design if needed. A non-singular
starting design is a design which permits unbiased estimation of all parameters of interest. However, the pertinent portion of their algorithm deals with the blocking of response surface designs which have been previously generated. Given a non-singular starting design for a specified model, the *interchange* part of the algorithm is concerned with switching runs across blocks to improve in $| Z'(I - P_\mathbf{y})Z |$.

Cook and Nachtsheim's algorithm must be modified to handle mixture components since they parametrize the blocking variables by indicator variables. In the formulation of their model, they assume that the column space of $Z$ does not contain the vector 1, which cannot be avoided with mixture components. Rather than parametrizing blocks with indicator variables, the matrix of $b - 1$ orthogonal block contrasts, $\hat{V}$ is used. Parametrization has little effect on the idea behind the algorithm.

Essentially, the Cook and Nachtsheim interchange algorithm sequentially considers all interchanges of a point in one block with points in other blocks. The algorithm requires a non-singular starting design $\text{rank}(\hat{V}, Z) = s + b - 1$ and a convergence criteria, $\zeta$, as the initial inputs. Then, starting with mixture 1 in block 1, the algorithm sequentially considers possible interchanges with all mixtures in other blocks, continuing until it reaches observation $n_i$ in block $b$, at which time it starts over with observation 2 of block 1. This pattern continues until the last mixture is reached.

After making an interchange, $| Z'(I - P_\mathbf{y})Z |$ is evaluated. If it increases, the interchange is permanent and $| Z'(I - P_\mathbf{y})Z |$ and $Z'(I - P_\mathbf{y})Z$ are immediately updated; otherwise, the interchange is reversed.

Let the model matrix at the previous step be denoted by $Z^{(0)}$. After the last
pair of mixtures has been considered for interchange, a check is made:
is
\[
\frac{|Z'(I - P_{t})Z| - |Z^{(0)}(I - P_{t})Z^{(0)}|}{|Z^{(0)}(I - P_{t})Z^{(0)}|} < \zeta
\]
If so, the algorithm stops; otherwise another round of interchanges begins, starting with the first observation in the first block.

As it stands, the algorithm is completely deterministic in its outcome, that is, given a starting design and convergence criterion, it will always converge to the same final design. Using different starting designs, such as the designs generated by the sampling method (see Section 4.6.2), as starting designs is one way to introduce a degree of randomness into the process.

For example, consider blocking the minimum integrated bias design for \(p = 3\) constructed by Draper and Lawrence (1965a) into three blocks with sizes \(n_1 = n_2 = 3\) and \(n_3 = 5\). Draper and Lawrence's design protects against biases incurred by the possible significance of cubic terms when a quadratic canonical polynomial is to be fit. The interchange algorithm with a convergence criterion of \(0.0000001\) gives the design in Table 4.13.

For this example \(|Z'(I - P_{t})Z| = 1.3 \times 10^{-6}\), and the overall efficiency is 87.6%. Both are vast improvements over the values using the run order as given in the original paper, \(8 \times 10^{-8}\) and 55.1%, respectively.

Consider finding three blocks of sizes six, six, and eight for a minimum bias design for \(p = 4\). This design was chosen from the group of minimum bias designs that fit second-order models and protected against the possible significance of third-order parameters. This design is formed using point sets \((1,1,4)\) of Draper and Lawrence (1965b). The arrangement in Table 4.14 gives a determinant of \(4.30 \times 10^{-11}\), whereas the unblocked value is \(1.08 \times 10^{-10}\), giving an overall efficiency of 91.2%, once again
Table 4.13: Draper and Lawrence example (p=3) run in three blocks constructed by the Interchange Method

<table>
<thead>
<tr>
<th>Block</th>
<th>$z_1$</th>
<th>$z_2$</th>
<th>$z_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1844</td>
<td>0.1844</td>
<td>0.6320</td>
</tr>
<tr>
<td>1</td>
<td>0.0753</td>
<td>0.5913</td>
<td>0.3334</td>
</tr>
<tr>
<td>1</td>
<td>0.6776</td>
<td>0.2411</td>
<td>0.0813</td>
</tr>
<tr>
<td>2</td>
<td>0.0813</td>
<td>0.0813</td>
<td>0.8374</td>
</tr>
<tr>
<td>2</td>
<td>0.6220</td>
<td>0.2560</td>
<td>0.1220</td>
</tr>
<tr>
<td>2</td>
<td>0.2560</td>
<td>0.6220</td>
<td>0.1220</td>
</tr>
<tr>
<td>3</td>
<td>0.4107</td>
<td>0.0447</td>
<td>0.5446</td>
</tr>
<tr>
<td>3</td>
<td>0.5913</td>
<td>0.0753</td>
<td>0.3334</td>
</tr>
<tr>
<td>3</td>
<td>0.2411</td>
<td>0.6776</td>
<td>0.0813</td>
</tr>
<tr>
<td>3</td>
<td>0.4823</td>
<td>0.4822</td>
<td>0.0354</td>
</tr>
<tr>
<td>3</td>
<td>0.0447</td>
<td>0.4107</td>
<td>0.5446</td>
</tr>
</tbody>
</table>

an improvement over the ordering given by Draper and Lawrence.

4.6.2 A Sampling Approach to Maximizing $|Z'(I - P_\gamma)Z|$  

This approach randomly samples from the discrete space of all possible run orders. The algorithm is described in full generality to allow for other criteria to be used in place of the determinant criteria advanced in previous sections. As with the interchange method, a non-singular starting design is required.

Number the mixtures $1, 2, \ldots, N$ in order so that each has a number associated with it. Adopt the convention that the first $n_1$ places comprise the first block, the next $n_2$ places comprise the second block, and the last $n_6$ comprise the 6-th block.

Generate $J$ random orderings of the numbers $1, 2, \ldots, N$. Rearrange the mixtures according to this ordering, based on the number each mixture was initially assigned. Allocate the mixtures to blocks by the places that the mixtures now hold, i.e. the mixtures in the first $n_1$ places in the ordering are assigned to the first block, etc. Find
Table 4.14: Draper and Lawrence example \((p=4)\) run in three blocks constructed by the Interchange Method

<table>
<thead>
<tr>
<th>Block</th>
<th>(z_1)</th>
<th>(z_2)</th>
<th>(z_3)</th>
<th>(z_4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.081750</td>
<td>0.754750</td>
<td>0.081750</td>
<td>0.081750</td>
</tr>
<tr>
<td>1</td>
<td>0.499100</td>
<td>0.027100</td>
<td>0.342900</td>
<td>0.130900</td>
</tr>
<tr>
<td>1</td>
<td>0.081750</td>
<td>0.081750</td>
<td>0.754750</td>
<td>0.081750</td>
</tr>
<tr>
<td>1</td>
<td>0.081750</td>
<td>0.081750</td>
<td>0.081750</td>
<td>0.754750</td>
</tr>
<tr>
<td>1</td>
<td>0.320875</td>
<td>0.226375</td>
<td>0.226375</td>
<td>0.226375</td>
</tr>
<tr>
<td>1</td>
<td>0.226375</td>
<td>0.320875</td>
<td>0.226375</td>
<td>0.226375</td>
</tr>
<tr>
<td>2</td>
<td>0.027100</td>
<td>0.342900</td>
<td>0.499100</td>
<td>0.130900</td>
</tr>
<tr>
<td>2</td>
<td>0.226375</td>
<td>0.226375</td>
<td>0.226375</td>
<td>0.320875</td>
</tr>
<tr>
<td>2</td>
<td>0.754750</td>
<td>0.081750</td>
<td>0.081750</td>
<td>0.081750</td>
</tr>
<tr>
<td>2</td>
<td>0.130900</td>
<td>0.027100</td>
<td>0.499100</td>
<td>0.342900</td>
</tr>
<tr>
<td>2</td>
<td>0.226375</td>
<td>0.027100</td>
<td>0.499100</td>
<td>0.342900</td>
</tr>
<tr>
<td>2</td>
<td>0.226375</td>
<td>0.226375</td>
<td>0.320875</td>
<td>0.226375</td>
</tr>
<tr>
<td>2</td>
<td>0.342900</td>
<td>0.499100</td>
<td>0.027100</td>
<td>0.130900</td>
</tr>
<tr>
<td>3</td>
<td>0.027100</td>
<td>0.130900</td>
<td>0.342900</td>
<td>0.499100</td>
</tr>
<tr>
<td>3</td>
<td>0.342900</td>
<td>0.130900</td>
<td>0.499100</td>
<td>0.027100</td>
</tr>
<tr>
<td>3</td>
<td>0.499100</td>
<td>0.342900</td>
<td>0.130900</td>
<td>0.027100</td>
</tr>
<tr>
<td>3</td>
<td>0.130900</td>
<td>0.342900</td>
<td>0.027100</td>
<td>0.499100</td>
</tr>
<tr>
<td>3</td>
<td>0.027100</td>
<td>0.499100</td>
<td>0.130900</td>
<td>0.342900</td>
</tr>
<tr>
<td>3</td>
<td>0.130900</td>
<td>0.499100</td>
<td>0.342900</td>
<td>0.027100</td>
</tr>
<tr>
<td>3</td>
<td>0.499100</td>
<td>0.130900</td>
<td>0.027100</td>
<td>0.342900</td>
</tr>
<tr>
<td>3</td>
<td>0.342900</td>
<td>0.027100</td>
<td>0.130900</td>
<td>0.499100</td>
</tr>
</tbody>
</table>
| $Z'(I - P_\psi)Z$ | for each of the $J$ random orderings. Choose the ordering with the maximum value of $|Z'(I - P_\psi)Z|$. Hopefully, this ordering will have an acceptable value and efficiency; this is not guaranteed since this is only a random sample. As with any random sample, the larger the sample, the more precise inference we can make about the maximum of $|Z'(I - P_\psi)Z|$ for the entire population of orderings.

As a way to compare the relative performances of the sampling method and the interchange method, we apply the sampling method to same examples of the previous section.

Table 4.15: Draper and Lawrence example ($p=3$) run in three blocks constructed by the Sampling method with $J=60$

<table>
<thead>
<tr>
<th>Block</th>
<th>$z_1$</th>
<th>$z_2$</th>
<th>$z_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.2411</td>
<td>0.6776</td>
<td>0.0813</td>
</tr>
<tr>
<td>1</td>
<td>0.6220</td>
<td>0.2560</td>
<td>0.1220</td>
</tr>
<tr>
<td>1</td>
<td>0.1844</td>
<td>0.1844</td>
<td>0.6320</td>
</tr>
<tr>
<td>2</td>
<td>0.0753</td>
<td>0.5913</td>
<td>0.3334</td>
</tr>
<tr>
<td>2</td>
<td>0.5913</td>
<td>0.0753</td>
<td>0.3334</td>
</tr>
<tr>
<td>2</td>
<td>0.4823</td>
<td>0.4822</td>
<td>0.0354</td>
</tr>
<tr>
<td>3</td>
<td>0.4107</td>
<td>0.0447</td>
<td>0.5446</td>
</tr>
<tr>
<td>3</td>
<td>0.2560</td>
<td>0.6220</td>
<td>0.1220</td>
</tr>
<tr>
<td>3</td>
<td>0.0913</td>
<td>0.0813</td>
<td>0.8374</td>
</tr>
<tr>
<td>3</td>
<td>0.6776</td>
<td>0.2411</td>
<td>0.0813</td>
</tr>
<tr>
<td>3</td>
<td>0.0447</td>
<td>0.4107</td>
<td>0.5446</td>
</tr>
</tbody>
</table>

Using the sampling method, sampling $J = 60$ times we get $|Z'(I - P_\psi)Z| = 0.000001603$, slightly larger than that obtained using the interchange method of the previous section. This gives an overall efficiency of 90.6%.

Consider the second Draper and Lawrence example for $p = 4$, given in Table 4.16.

The sampling procedure found a design which is slightly better than the design
Table 4.16: Draper and Lawrence example (p=4) run in three blocks constructed by the Sampling Method with J=60

<table>
<thead>
<tr>
<th>Block</th>
<th>( z_1 )</th>
<th>( z_2 )</th>
<th>( z_3 )</th>
<th>( z_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.081750</td>
<td>0.754750</td>
<td>0.081750</td>
<td>0.081750</td>
</tr>
<tr>
<td>1</td>
<td>0.226375</td>
<td>0.226375</td>
<td>0.320875</td>
<td>0.226375</td>
</tr>
<tr>
<td>1</td>
<td>0.226375</td>
<td>0.320875</td>
<td>0.226375</td>
<td>0.226375</td>
</tr>
<tr>
<td>1</td>
<td>0.342900</td>
<td>0.130900</td>
<td>0.499100</td>
<td>0.027100</td>
</tr>
<tr>
<td>1</td>
<td>0.499100</td>
<td>0.130900</td>
<td>0.027100</td>
<td>0.342900</td>
</tr>
<tr>
<td>1</td>
<td>0.081750</td>
<td>0.081750</td>
<td>0.081750</td>
<td>0.754750</td>
</tr>
<tr>
<td>2</td>
<td>0.027100</td>
<td>0.499100</td>
<td>0.130900</td>
<td>0.342900</td>
</tr>
<tr>
<td>2</td>
<td>0.499100</td>
<td>0.342900</td>
<td>0.130900</td>
<td>0.027100</td>
</tr>
<tr>
<td>2</td>
<td>0.320875</td>
<td>0.226375</td>
<td>0.226375</td>
<td>0.226375</td>
</tr>
<tr>
<td>2</td>
<td>0.342900</td>
<td>0.027100</td>
<td>0.130900</td>
<td>0.499100</td>
</tr>
<tr>
<td>2</td>
<td>0.081750</td>
<td>0.081750</td>
<td>0.754750</td>
<td>0.081750</td>
</tr>
<tr>
<td>2</td>
<td>0.130900</td>
<td>0.499100</td>
<td>0.342900</td>
<td>0.027100</td>
</tr>
<tr>
<td>3</td>
<td>0.226375</td>
<td>0.226375</td>
<td>0.226375</td>
<td>0.320875</td>
</tr>
<tr>
<td>3</td>
<td>0.754750</td>
<td>0.081750</td>
<td>0.081750</td>
<td>0.081750</td>
</tr>
<tr>
<td>3</td>
<td>0.027100</td>
<td>0.342900</td>
<td>0.499100</td>
<td>0.130900</td>
</tr>
<tr>
<td>3</td>
<td>0.130900</td>
<td>0.342900</td>
<td>0.027100</td>
<td>0.499100</td>
</tr>
<tr>
<td>3</td>
<td>0.027100</td>
<td>0.130900</td>
<td>0.342900</td>
<td>0.499100</td>
</tr>
<tr>
<td>3</td>
<td>0.342900</td>
<td>0.499100</td>
<td>0.027100</td>
<td>0.130900</td>
</tr>
<tr>
<td>3</td>
<td>0.130900</td>
<td>0.027100</td>
<td>0.499100</td>
<td>0.342900</td>
</tr>
<tr>
<td>3</td>
<td>0.499100</td>
<td>0.027100</td>
<td>0.342900</td>
<td>0.130900</td>
</tr>
</tbody>
</table>
found by interchange in the previous section. The determinant was $4.42 \times 10^{-11}$, giving an efficiency of 91.4%—not all that great of an improvement over the interchange method.

The key point to realize is that the sampling procedure has the potential to find better designs than the interchange because it randomly samples from all possible run orders; sometimes it may find very good designs, sometimes it may find very bad designs. In contrast, the interchange algorithm is deterministic, handicapped by the starting design. If the starting design is poor, it may be unable to substantially improve the design. Generating a starting design for the interchange method using the sampling design may substantially improve the interchange method’s ability to find a good block design.

### 4.7 Conclusions

Each of the methods given in this chapter have their appropriate time and place.

The most flexible in terms of application are the computer-intensive methods, which may be applied to any design scenario. They are useful for non-standard situations such as unequal block sizes. Even if an orthogonal design exists, neither algorithm is guaranteed to find it. The interchange method is dependent upon the starting design it is fed. The sampling method, dependent upon what sample it gets—a sample of designs which are poor with respect to the criterion of choice is always possible.

The best recommendation for the interchange method to overcome its limitation of dependency is to try several different starting designs. The interchange method may be efficiently used in tandem with the sampling method; the best designs found
by the sampling method may serve as starting designs. When using the sampling method, it is best to try several samples, perhaps with varying sample sizes. Taking larger samples is also beneficial to get a better handle on the maximum value of the determinant.

Symbolic and integral mixture mates of strength $t$ are versatile tools for constructing orthogonal block designs. Both types of mixture mates may be used in a lego-like manner to assemble block designs for virtually any number of mixture components or model since the idea of a pair of mixture mates of strength $t$ has been extended to $b$ mates of strength $t$. Using trade-off methods to construct mixture mates greatly increases the viability of mixture mates as a design tool because a wider variety of blocking arrangements may be formulated—trade-off releases us from limitations imposed by only considering pairs of mates as with Latin squares and perpendicular array methods.

Mates are the primary tool to assemble orthogonal block designs—for equal block sizes. Requiring equal block sizes limits the scope of the use of mixture mates. While the definition of both symbolic and integral mixture mates of strength $t$ allows for unequal block sizes, it is unclear how any of the present construction methods can be adapted to find mixture mates having different numbers of rows. Mates may also be used to construct non-orthogonal block designs when used in concert with asymmetrical orthogonal arrays, useful when patchwork techniques yield block sizes that are too large and when not all block contrasts are deemed to be important.

Using orthogonal arrays and fractional factorials as in Section 4.4 is useful only when $S$ is a constrained region inside the simplex. When first-order models provide an adequate fit, this method performs well, producing orthogonal or very nearly
orthogonal blocks. However, they often place severe limitations upon the values the mixture component with the largest range may assume. For example, consider the Anik and Sukumar example. The largest proportion of the mixture that water may assume is 70%. This happens only twice. For most of the runs, water is unnecessarily restricted to 0.5 and below. Such restrictions seem to be arbitrarily placed upon the component with the largest range. This arbitrariness can be lessened by assigning the least important mixture component to the $p$-th column, the column getting adjusted most often and most likely the one to have its values concentrated at or near either the low or high boundary, rather than being spread at both boundaries and in between.
CHAPTER 5. CONSTRUCTION OF TREND-FREE AND NEARLY TREND-FREE DESIGNS FOR MIXTURE EXPERIMENTS

This chapter discusses the construction of mixture experiments when deterministic trends are present. Deterministic trends include trends across observations taken at equally spaced time intervals and spatial position of units. Stochastic trends governed by time series or spatial processes will be considered at another time in another place. Section 5.3 gives a short background and development of trend-free designs. For readers still not convinced of the need for or validity of trend-free designs, Section 5.4 discusses randomization issues. Among the most useful trend-free designs in the literature are trend-free (fractional) factorial designs, reviewed in Section 5.5. Section 5.6 discusses the lone paper on the construction of trend-free designs for mixture components. Finally, Sections 5.7, 5.8, and 5.9 provide several ways of designing experiments for mixture components that are trend-free or nearly trend-free.

5.1 Introduction

The models used throughout this chapter will be models

\[ y = U\gamma + Z\tau + e \]  \hspace{1cm} (5.1)

and

\[ y = U\gamma + X\beta + e \]  \hspace{1cm} (5.2)
of Section 2.2.1.

In this chapter the columns of the matrix $U$ correspond to $b$ orthogonal polynomials across positions $1, \ldots, N$. For example, consider a trend across four points taken at equally spaced time intervals. Then the linear portion of the trend, the first column of $U$, is given by $(-3, -1, 1, 3)'$, the quadratic portion of the trend, the second column of $U$ is $(1, -1, -1, 1)'$, and the cubic portion of the trend, the third column of $U$ is $(-1, 3, -3, 1)$.

Coefficients of orthogonal polynomials for equally spaced levels for up to $N = 75$ have been tabulated in Fisher and Yates (1963), among other places. In the present context, the orthogonal polynomials represent time effects that may be induced by a variety of factors: tool wear, process drift, aging of the experimental material, etc. Orthogonal polynomials also have a variety of other uses.

**Definition 5.1.1** Let $U$ be a $N \times b$ matrix of orthogonal polynomial coefficients associated with deterministic trends. A design for a mixture experiment is said to be trend-free under model 5.1 if

$$U'Z = 0.$$ 

An ordering of mixtures is said to be $b$-trend-free if $U'Z = 0$ where the columns of $U$ contain the coefficients for all of the orthogonal polynomials of at most degree $b$.

By definition 5.1.1, a design is trend-free if for functions $\Lambda \gamma$ and $\Lambda \tau$, estimable under model 5.1,

$$\text{cov}(\Lambda \hat{\gamma}, \Lambda \hat{\tau}) = 0.$$
5.2 An Example of the Use of Trend-Free Designs

To illustrate the need for designs accounting for the possible existence of time trends, consider the results of an experiment performed by Williams and Amidon (1984) to determine the solubility of phenobarbital in a solution containing ethanol, propylene glycol, and water. Solubility is defined in terms of total vapor pressure. They empirically modeled solubility as a function of the proportion of these three variables in the solution. The data and design are given in Table 5.2.

Table 5.1: Data and design for Williams and Amidon (1984) example

<table>
<thead>
<tr>
<th>Ethanol ($z_1$)</th>
<th>Water ($z_2$)</th>
<th>Propylene Glycol ($z_3$)</th>
<th>Vapor Pressure ($y$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1191</td>
<td>0.3033</td>
<td>0.5776</td>
<td>18.44</td>
</tr>
<tr>
<td>0.0817</td>
<td>0.6219</td>
<td>0.2964</td>
<td>25.81</td>
</tr>
<tr>
<td>0.1857</td>
<td>0.4764</td>
<td>0.3379</td>
<td>30.71</td>
</tr>
<tr>
<td>0.0463</td>
<td>0.9265</td>
<td>0.0272</td>
<td>30.73</td>
</tr>
<tr>
<td>0.3246</td>
<td>0.2804</td>
<td>0.3950</td>
<td>33.27</td>
</tr>
<tr>
<td>0.1359</td>
<td>0.6994</td>
<td>0.1647</td>
<td>35.03</td>
</tr>
<tr>
<td>0.1096</td>
<td>0.8259</td>
<td>0.0645</td>
<td>37.03</td>
</tr>
<tr>
<td>0.3475</td>
<td>0.4415</td>
<td>0.2110</td>
<td>42.00</td>
</tr>
<tr>
<td>0.5003</td>
<td>0.2555</td>
<td>0.2442</td>
<td>43.70</td>
</tr>
<tr>
<td>0.2588</td>
<td>0.6638</td>
<td>0.0774</td>
<td>48.42</td>
</tr>
<tr>
<td>0.4874</td>
<td>0.4143</td>
<td>0.0983</td>
<td>48.79</td>
</tr>
<tr>
<td>0.7188</td>
<td>0.2272</td>
<td>0.0540</td>
<td>53.70</td>
</tr>
</tbody>
</table>

Fitting a linear trend across mixtures to vapor pressure produces the following result (estimated standard errors of the parameter estimates are listed below parameter estimates):

$$\hat{y} = 37.30 + 1.42 u_1$$

(0.51) (0.07)  \hspace{1cm} (5.3)

where $u' = (-11, -9, -7, -5, -3, -1, 1, 3, 5, 7, 9, 11)$. 
Fitting a second-order canonical polynomial gives

\[ \hat{y} = 43.13z_1 + 28.35z_2 + 0.1789z_3 \]

\[ (5.91) \quad (2.28) \quad (10.64) \]

\[ + 88.99z_1z_2 + 23.09z_1z_3 - 0.86z_2z_3 \]

\[ (19.39) \quad (26.18) \quad (23.78) \]

From examining the parameter estimates of this model and their standard errors, it appears that ethanol and water and their interaction are important, but perhaps propylene glycol may be deleted from the model altogether. A second-order model in ethanol and water could then be fit.

However, when a linear time trend is included in the model, we get a different story. Then,

\[ \hat{y} = 0.75u + 40.99z_1 + 32.53z_2 + 9.63z_3 \]

\[ (0.39) \quad (5.06) \quad (2.91) \quad (10.1) \]

\[ + 37.99z_1z_2 + 23.26z_1z_3 + 21.48z_2z_3 \]

\[ (31.35) \quad (21.86) \quad (23.07) \]

Now it appears that none of the second-order parameters appear to be significant. The ethanol-water interaction no longer seems important, meaning that ethanol and water have an additive effect. A first-order model in ethanol and water now seems to be adequate, simplifying the situation. The correlations between estimators of the linear trend and the second-order model parameters are

\[ \rho(\hat{\gamma}_1, \hat{\gamma}_1) = -0.222 \]
\begin{align*}
\rho(\hat{\tau}_1, \hat{\tau}_2) &= 0.755 \\
\rho(\hat{\tau}_1, \hat{\tau}_3) &= 0.488 \\
\rho(\hat{\tau}_1, \hat{\tau}_{12}) &= -0.856 \\
\rho(\hat{\tau}_1, \hat{\tau}_{13}) &= 0.040 \\
\rho(\hat{\tau}_1, \hat{\tau}_{23}) &= 0.506.
\end{align*}

These correlations indicate the degree of confounding between the linear time trend and the estimated model parameters. The linear trend is highly correlated with the two-factor interaction between ethanol and water, as can be seen in equation (5.4). When the ethanol-water interaction is removed from the model, the linear trend is significant. Because of the confounding between \( \hat{\tau} \) and \( \hat{\gamma} \), it is unclear whether the ethanol-water interaction or the linear time trend is important or neither. Had a linear trend been anticipated at the design stage, a more judicious choice of design or ordering of the mixtures could have been chosen. This would allow these two effects to be separated, clarifying how this linear time trend and the ethanol-water interaction contribute to vapor pressure. Ignoring the time trend produces incorrect measures of the uncertainty in the parameter estimates.

Another pertinent example comes from Hill (1960). A chemist may wish to study a chemical reaction as a function of the constituents of the chemical, under constant activity of a catalyst. However, the catalyst’s effectiveness may degrade over time, giving the chemist two costly choices: replace or regenerate the catalyst. Without doing either, degradation may inflate the variance of the estimators of the model parameters of interest. A third choice could be made prior to conducting the experiment—design the experiment so that the estimators of the parameters of
interest are uncorrelated or nearly uncorrelated with any trend induced by catalyst degradation.

5.3 Background on Trend-Free Experimental Designs

Study of systematic designs and run orders originated in the early 1950's with the work of Cox (1951,1952) and Box (1952).

Cox (1951,1952) was concerned with the influence that a smooth polynomial time trend would have on the estimators of treatment differences. He attempted to find treatment sequences that provided estimators of the treatment differences and time trend parameters that were uncorrelated. He found several exact solutions for small values of $N$, $p$, the number of treatments, and $b$, the degree of the polynomial trend. The present author would like to emphasize that in the absence of trend-free orders, Cox found many good nearly trend-free orders that minimized the correlation between estimators of treatment differences and trend parameters.

Box (1952) developed a flexible method for obtaining $b$-trend-free designs for $p - 1$ process variables. Factor levels were generated for each factor using a linear combination of the orthogonal polynomials of degree $b+1$ to $N-b-1$. The coefficients of the linear combinations were randomly chosen in such a way that the $p-1$ process variables were mutually orthogonal; these coefficients were generated by a process Box called angular randomization.

Hill (1960) blended the approaches of both Box and Cox to find a wide variety of trend-free designs for quantitative variables. He was the first to find trend-free designs for two-level factorials, constructing a trend-free $2^4$ factorial design. Hill was also a pioneer in providing a systematic method to construct trend-free block designs.
Trend-free block designs are designs in which the (mixture, treatment, or factorial) parameters are orthogonal to common trends within blocks.

Since Hill, several authors have investigated methods of constructing trend-free factorial designs. Most notable are Daniel and Wilcoxon (1966), Dickinson (1974), Joiner and Campbell (1976), Cheng (1985), Coster and Cheng (1985), Cheng and Jacroux (1988), Steinberg (1988), Jacroux and Saha Ray (1989), Cheng (1990), Jacroux (1990), and John (1990). All primarily concentrate on the construction of trend-free factorial designs. The results of Cheng and Jacroux (1988), Coster and Cheng (1988), and John (1990) will be given in greater detail in section 5.5. Cheng (1990) and Jacroux (1990) summarize many results on trend-free factorial designs in a palatable form. Joiner and Campbell's method is similar to the sampling method proposed in Section 5.9.1 and will be discussed in that section.

5.4 Randomization and Trend-Free Designs

Only limited randomization of run order is possible in trend-free designs—a point of contention for many statisticians. This section briefly tackles this issue, attempting to persuade the reader that in certain situations, trend-free designs are as natural and beneficial as the restricted randomization that occurs in block designs. Daniel and Wilcoxon (1966) provide an eloquent defense of trend-free designs—one that this author cannot and will not try to duplicate, only draw from.

Daniel and Wilcoxon (1966,p.260) stated “Randomization is used, then after we have exhausted our knowledge of the behavior of the system under study and have taken serious steps to control what can be controlled.” Deterministic trends, along with blocks, exemplify “knowledge” of the system under observation and should be
considered when designing an experiment. Logically, those who accept the use of blocks should also accept the use of time trends.

Trend-free designs exploit the experimenter's knowledge of the possible deterministic trends just as blocking exploits knowledge of the size and disposition of homogeneous groups of experimental material. Using a completely randomized design when significant block differences may exist is an invitation for disaster. Unrecognized block effects may dominate the factor effects of interest, making it completely impossible to detect any significant differences between factor effects. Also, because the estimated standard errors will be incorrect, the rate of type I errors increases; recall the Williams and Amidon (1984) example of Section 5.2. When block designs are used, the randomization is restricted to within the blocks. Kempthorne (1983) states the case for blocking using an agricultural example of comparing specific treatments applied to cows, across breeds of cows. Kempthorne (1983, p. 165) stated “...It is sufficient to know that the contribution [of breed to \( \sigma^2 \)] may be large, relative to the variance between cows of the same breed. The experimenter would then be ignoring valuable a priori information by using complete randomization. Furthermore, he would allow the possibility arising of particular treatment comparisons he wishes to make being affected by breed differences which he does not know precisely.”

Extending Kempthorne's logic, it is possible that an unknown, but highly significant deterministic trend may swamp any effects of interest by inflating the amount of uncertainty in parameter estimates. Certain run orders produced by randomization procedures may be very undesirable, in some cases giving designs where the effects of interest are completely confounded with time trends. In the Williams and Amidon (1984) example of Section 5.2 it is almost impossible to determine whether changes
in the response were due to the ethanol-water interaction or to the linear time trend. By considering the time trend when designing the experiment, a design may be constructed or the run order can be arranged so that the estimators of the mixture parameters and time trend parameters are nearly or completely uncorrelated.

Essentially, one can view trend-free designs as a sort of non-orthogonal block design. The blocks have size one, and the effects of interest are only orthogonal to selected block contrasts. This concept is very similar to that of Section 4.5 where we found designs in which the estimated mixture parameters were orthogonal only to a subset of the $b - 1$ estimated block contrasts.

Some randomization is possible in trend-free designs. The run order is fixed, but factor labels may be permuted.

5.5 The Construction of Trend-Free Factorial Designs

The methods of Sections 5.7 and 5.8 exploit trend-free (fractional) factorial designs. This section summarizes selected methods that select run orders such that the inner product of the contrast (main effects, interactions) coefficients and the coefficients of the orthogonal polynomial trends is zero.

Before embarking on a discussion of trend-free run orders, it is worthwhile to examine the type of factorial designs in the unrestricted variables that will be required to fit $d$-th order polynomials in the mixture components.

In general, if a $d$-th order canonical polynomial is adequate for representing the response in the region of interest, then the class of designs to be considered must permit $d$-th order main-effects and $d$-th order interactions to be estimated unbiasedly. A sufficient condition is that the unrestricted variables have $d + 1$ levels. The reso-
olution of the design is another important consideration; for estimating the terms of a $d$-th order polynomial in the unrestricted variables, the design must be a complete factorial or at least of resolution $2d + 1$ [see Raktoe, Hedayat, and Federer (1981) or Dey (1985) for a discussion of resolution]. Alternatively, if orthogonal arrays are to be used [c.f. Section 4.5] the strength of the orthogonal array should be $2d$.

Daniel and Wilcoxon (1966) inspired the results presented here. Cheng (1985), Cheng and Jacroux (1988), Coster and Cheng (1988), and John (1990) all examined techniques which extended those given by Daniel and Wilcoxon. Daniel and Wilcoxon noticed that when the runs in a factorial design are arranged in standard run order [c.f. Box, Hunter, and Hunter (1978)], coefficients of the $c$-factor interactions are orthogonal to coefficients of orthogonal polynomial trends up to order $c - 1$.

Coster and Cheng (1988) proposed constructing run orders for trend-free factorial designs using the *generalized foldover* method. Cheng and Jacroux (1988) expanded upon Daniel and Wilcoxon's ideas, showing how to construct trend-free designs by selecting run orders generated by identifying contrasts of interest with higher order interactions. Cheng (1990) notes that the foldover technique can duplicate any trend-free design constructed using the methods of Cheng and Jacroux or Daniel and Wilcoxon by careful choice of *generators*. Generalized foldover is applicable to factors having more than two levels [Coster and Cheng (1988), John (1990)]. Because of its versatility, only the generalized foldover method will be further pursued; all other methods may be reached as special cases.

Before introducing generalized foldover, it is necessary to introduce terminology pertaining to designs for (fractional) factorial experiments. Throughout this section, $p - 1$ will denote the number of factors and $A_j$, $j = 1, \ldots, p - 1$, will denote the $j$-th
factor. For each \( j = 1, \ldots, p - 1 \), let \( \zeta_j \in \{0, 1, \ldots, d - 1\} \). Denote a main effect of factor \( j \) by \( A_j \) and a \( c \)-factor interaction by \( A_{j_1} \cdots A_{j_c} \). A run in the design is denoted by a string of lowercase letters called a *word*. The exponents of the letters of a word, say \( a_j^{\zeta_j} \cdots a_{p-1}^{\zeta_{p-1}} \), indicate the level for each factor in the run. The product of two words is formed by taking the letter-wise product, according to the rules

\[
a_j^{\zeta_j} a_j^{\zeta_j'} = a_j^{\zeta_j + \zeta_j'}
\]

and

\[
a_j^d = 1.
\]

**Definition 5.5.1** A set of words is said to be independent if no word in the set may be expressed as the product of other words in the set.

Independence of a set of factorial effects is defined similarly. A set of \( f \) independent words can generate the runs of a \( d^f \) (fractional) factorial design. Sets of \( f \) independent words are called *generators*.

To illustrate these concepts, consider a small example. Suppose that \( p - 1 = 3 \) and \( d = 2 \). Table 5.2 gives the runs of a complete \( 2^3 \) factorial design, where -1 and 1 represent low and high factor levels, respectively. Each column gives the coefficients for a particular \( c \)-factor contrast \((c = 1, 2, 3)\), e.g. the fourth columns gives the coefficients for the two-factor interaction \( A_1 A_2 \). The words in the eighth column give the run order. For instance, in the fourth row, \( a_1 a_2 \) indicates that factors 1 and 2 are at their high levels and factor 3 is at its low level.

One possible group of generators for this example is \( a_1 a_2, a_1 a_3, \) and \( a_1 a_2 a_3 \).
Forming all possible products gives

\[ a_1a_2a_1a_3 = a_1^2a_2a_3 = a_2a_3 \]
\[ a_1a_2a_1a_2a_3 = a_1^2a_2^2a_3 = a_3 \]
\[ a_1a_3a_1a_2a_3 = a_1^2a_2a_3^2 = a_2 \]
\[ a_1a_2a_1a_3a_1a_2a_3 = a_1^3a_2^2a_3^2 = a_1. \]

Adding the element (1) where all factors are at the low level as the first run gives all eight runs of the complete \(2^3\) factorial design. The generators used to obtain the run order in Table 5.2 are \(a_1, a_2,\) and \(a_3.\)

| Table 5.2: A complete \(2^3\) factorial design, listed in standard order |
|---|---|---|---|---|---|---|---|---|
| \(A_1\) | \(A_2\) | \(A_3\) | \(A_1A_2\) | \(A_1A_3\) | \(A_2A_3\) | \(A_1A_2A_3\) |
| -1 | -1 | -1 | 1 | 1 | 1 | -1 | (1) |
| 1 | -1 | -1 | -1 | -1 | 1 | 1 | \(a_1\) |
| -1 | 1 | -1 | -1 | 1 | -1 | 1 | \(a_2\) |
| 1 | 1 | -1 | 1 | -1 | -1 | -1 | \(a_1a_2\) |
| -1 | -1 | 1 | -1 | 1 | -1 | 1 | \(a_3\) |
| 1 | -1 | 1 | -1 | 1 | -1 | -1 | \(a_1a_3\) |
| -1 | 1 | -1 | -1 | -1 | 1 | 1 | \(a_2a_3\) |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | \(a_1a_2a_3\) |

As alluded to earlier the foldover method of constructing run orders plays a central role in the construction of trend-free designs. Let \(B = (b_{\xi+1}, b_{\xi+2}, \ldots, b_{\xi+\nu})\), be a sequence of \(d\nu\) consecutive words corresponding to runs \(\xi + 1, \ldots, \xi + d\nu\). Take \(\nu\) to be a power of \(d\).

**Definition 5.5.2** Let \(B\) be given as above and \(e\) be a word. Then, the generalized foldover of \(B\) is the sequence of \(d^2\nu\) words

\[ B^*(e) = (B, B(e), B(2e), \ldots, B((d-1)e)), \]
where

\[ B(ke) = (b_{t+1}e^k, b_{t+2}e^k, \ldots, b_{t+d}e^k), \]

for some \( k \in \{0, 1, 2, \ldots, d - 1\} \).

In simpler terms, given a sequence of runs, to find the generalized foldover of the sequence, simply multiply each word in the sequence \( B \) by each power of the word \( e \), i.e. \( e^0, e^1, e^2, e^3, \ldots, e^{d-1} \), reducing the exponents by modulo \( d \) when necessary.

As an example of generalized foldover, consider an example of the \( 2^3 \) experiment whose runs are given in Table 5.2. Suppose we have the sequence of words \( B = ((1), a_1a_2a_3, a_1a_2, a_3) \). Then, since \( d = 2 \),

\[ B^*(a_1a_3) = (B, B(a_1a_3)) = ((1), a_1a_2a_3, a_1a_2, a_3, a_1a_3, a_2, a_2a_3, a_1), \]

an alternative to the run order given in Table 5.2.

This gives rise to the following method for constructing a run order of a \( s^f \) (fractional) factorial design using generalized foldover.

**Method 5.5.1** Let \( g_1, g_2, \ldots, g_f \) be \( f \) generators. Let \( (1) \) be the run where all of the factors are at the low level. Generate the run order as follows:

1. Set \( B_0 = (1) \).

2. Let \( B_i = B_{i-1}^*(g_i) \), and for \( i = 1, 2, \ldots, f \) for \( B^* \) as in Definition 5.5.2.

Then \( B_f \) is a run order of a \( s^f \) (fractional) factorial experiment.

To illustrate Method 5.5.1 consider constructing a run order for a \( 2^3 \) factorial design using the generators \( a_1, a_2, \) and \( a_3 \).
1. \( B_0 = (1) \).

2. \( B_1 = B_0^* = (B_0, B_0(a_1)) = ((1), a_1) \).

3. \( B_2 = B_1^* = (B_1, B_1(a_2)) = ((1), a_1, a_2, a_1a_2) \).

4. \( B_3 = B_2^* = (B_2, B_2(a_3)) = ((1), a_1, a_2, a_1a_2, a_3, a_1a_3, a_2a_3, a_1a_2a_3) \),

precisely the run order given in Table 5.2.

Cheng (1990) gives the following result about run orders for (fractional) factorial designs constructed using the generalized foldover method. It is powerful because it states that the generators completely determine the degree to which a main effect or interaction is trend-free. This result is an extension of Theorem 3 of Coster and Cheng (1988, p. 1194).

**Theorem 5.5.1** Suppose that a \( d^f \) \((f \leq p-1)\) design has been constructed using the generalized foldover method (Method 5.5.1). Let \( m + 1 \leq f \).

1. If a factor appears at nonzero levels in \((m + 1)\) generators, the corresponding main effect is \(m\)-trend-free.

2. A \(c\)-factor interaction, \(A_{j_1}A_{j_2} \cdots A_{j_c}\), is \(m\)-trend-free if and only if there are at least \((m + 1)\) generators, each of which contains an odd number of letters of the corresponding word \(a_{j_1}^c \cdots a_{j_c}^c\).

John (1990) also used generalized foldover to devise rules for constructing trend-free designs for two-level factors. He also examined the construction of trend-free designs for variables having three levels. However, the rules that he gives are (for the most part) special cases of Theorem 5.5.1. An exception is John’s discussion of kept factors. For more details on the role of kept factors, see John (1990).
5.6 Review of Construction Methods for Trend-Free Designs for Mixture Experiments

Although mixture experiments are used heavily by the process industries (chemical and oil companies, food industry, etc.) where significant deterministic and stochastic time trends often exist, little research has been conducted on finding trend-free designs for mixture experiments.

The exception is Goel (1980). Goel attempts to find design points in the simplex in such a way that parameter estimators associated with the time trends are uncorrelated with estimators of the second-order polynomial parameters. It would seem as though his method is more general than he says, not restricted to only second-order models. He notes that construction of such designs is dependent on the number of mixtures that it is feasible to use, say \( N_i \). Goel assumes that it is possible to select a design with \( N = N_i/2 \) or \( N = (N_i + 1)/2 \) (as \( N_i \) is even or odd) distinct mixtures which allows estimation of all of the \( d \)-th order canonical polynomial parameters. If this is not possible, Goel’s method is not applicable.

If \( N_i \) is even, randomize the run order of the \( N \) mixtures and repeat the mixtures in reverse order giving a design in \( N_1 = 2N \) mixtures. If \( N_i \) is odd, randomize the run order of the \( N \) mixtures and repeat the first \( (N - 1) \) mixtures in reverse order after the \( N \)-th mixture, giving a design in \( N_1 = 2N - 1 \) mixtures. For the odd case, every mixture will be replicated twice except the \( N \)-th mixture.

The designs constructed by Goel’s scheme produce mixture parameter estimators which are uncorrelated with the estimators of the deterministic trend parameters for odd order orthogonal polynomials only.

For examples of both cases, consider the \( \{3,2\} \)-simplex lattice design and the
three variable simplex centroid design. One ordering for the simplex lattice design is (1,0,0), (0,0,1), (0.5,0,0.5), (0.5,0.5,0), (0.5,0.5,0.5), (0,0,1), (0,1,0), (0,0.5,0.5), (0.5,0.5,0), (0.5,0.5,0), (0,0,1), and (1,0,0).

For the simplex centroid design, fix (0,0.5,0.5) as the unreplicated point, giving the ordering (0,0,1), (0.5,0.5,0), (1,0,0), (0.33,0.33,0.34), (0,1,0), (0.5,0.5,0.5), (0.5,0,0.5), (0,1,0), (0.33,0.33,0.34), (1,0,0), (0.5,0.5,0), and (0,0,1).

Goel’s method is unattractive if the experimenter is unwilling or unable to replicate the entire experiment. Complete replication may not be practical or smart in situations where cost of a mixture is important. For small $N$ complete replication may be within the realm of possibility. In the two examples given this only involved an additional 6 and 7 runs. What should one do when faced with 7 variables? Running a $\{7,2\}$ simplex lattice or 7 variable simplex centroid require 21 and 127 runs, respectively. Completely replicating experiments of these sizes may be costly and inefficient. Incomplete replication, taking additional runs at strategic locations to check lack of fit and to estimate pure error may be more prudent. Goel’s method cannot handle such situations.

In the present author’s opinion it is a gross misallocation of resources to completely replicate an experiment solely to produce mixture orders that are trend-free only for orthogonal polynomial trends of odd order. Goel’s unsatisfactory and inflexible method motivated additional research on the problem of constructing trend-free mixture designs.
5.7 The Construction of Trend-Free Designs for Mixture Experiments Through Transformation

In this section we construct $b$-trend-free designs for mixture experiments utilizing the transformations of Chapter 3 in conjunction with the $b$-trend-free designs for factorial experiments given in Section 5.5.

The central result of this section, Theorem 5.7.1, is a direct consequence of Theorem 3.3.1. Theorem 5.7.1 suggests that if we select a $b$-trend-free design for the unrestricted variables and map this design into the simplex (or subregion) according to the transformation $z' = x'Q$ ($Q$ non-singular), then the resulting design for the mixture components will also be $b$-trend-free. Theorem 5.7.1 assumes that $F_d(z)$ is the induced vector parametrization or some reparametrization.

**Theorem 5.7.1** Let $z' = x'Q$ for $Q$ non-singular. A design for a mixture experiment is $b$-trend-free if and only if the corresponding design in the unrestricted variables is $b$-trend-free.

The proof follows directly from Theorem 3.3.1.

Theorem 5.7.1 places the results of Section 5.5 on factorial designs at our disposal and provides a general rule by which trend-free designs for mixture experiments may be constructed.

As an example, consider fitting a first-order canonical polynomial in 4 mixture variables. Suppose that it is desirable for the estimators of the first-order model parameters to be uncorrelated with a linear trend. Let $x_1, x_2, \text{ and } x_3$ be three unrestricted variables (see Chapter 3). For this example, consider using the run order generated from Table 5.4. This run order was constructed applying the foldover

...
method with generators $x_1x_2$, $x_1x_3$, and $x_1x_2x_3$. The design for the unrestricted variables $x_1$, $x_2$ and $x_3$ is 1-trend-free. In order to ensure that these points are mapped back into the mixture space, we use 0's and -1's and use Claringbold's transformation utilizing a result of Nigam (1973). This gives the 1-trend-free design in Table 5.3.

As is evident from Table 5.4, this design does not explore the simplex very well, which may or may not be a problem, depending on the scenario. One of the mixture components, $z_1$, never assumes values larger than 0.25. Even the other variables do not encompass a large portion of the simplex. The largest range of any one variable is for $z_4$, with a range of 0.4053. This problem was mentioned in Chapter 3 as one of the major flaws of the current understanding of transformations as methods for generating designs.

Table 5.3: A $2^3$ factorial design

<table>
<thead>
<tr>
<th>$x_1$</th>
<th>$x_2$</th>
<th>$x_3$</th>
<th>$x_1x_2$</th>
<th>$x_1x_3$</th>
<th>$x_2x_3$</th>
<th>$x_1x_2x_3$</th>
<th>$x_3$</th>
<th>$x_2$</th>
<th>$x_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>-1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>0</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When the design space has an ellipsoidal shape, the transformation due to Thompson and Myers may be used to provide designs that have other useful properties in addition to being trend-free.

Consider the following example taken from Cornell (1990, p. 125), where the
Table 5.4: A 1-trend-free design constructed using the transformation method

<table>
<thead>
<tr>
<th>$z_1$</th>
<th>$z_2$</th>
<th>$z_3$</th>
<th>$z_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>0.25</td>
<td>0.0450</td>
<td>0.4541</td>
</tr>
<tr>
<td>0.25</td>
<td>0.0143</td>
<td>0.3679</td>
<td>0.3679</td>
</tr>
<tr>
<td>0.00</td>
<td>0.3334</td>
<td>0.3333</td>
<td>0.3333</td>
</tr>
<tr>
<td>0.00</td>
<td>0.0976</td>
<td>0.2471</td>
<td>0.6553</td>
</tr>
<tr>
<td>0.00</td>
<td>0.0976</td>
<td>0.4512</td>
<td>0.4512</td>
</tr>
<tr>
<td>0.00</td>
<td>0.3333</td>
<td>0.1292</td>
<td>0.5375</td>
</tr>
<tr>
<td>0.25</td>
<td>0.0143</td>
<td>0.1637</td>
<td>0.572</td>
</tr>
<tr>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
</tr>
</tbody>
</table>

experimental region is defined as

$$[z - z_0]' H^{-2} [z - z_0] \leq 1,$$

where $z_0 = (0.3, 0.25, 0.33, 0.12)'$ and $h_1 = 0.2, h_2 = 0.2, h_3 = 0.3$, and $h_4 = 0.1$.

First construct a trend-free design in the unrestricted variables that is capable of fitting the desired model. Suppose that a second-order canonical polynomial is thought to be adequate and a linear trend may exist. Then we select a complete $3^3$ factorial design and choose generators using Theorem 5.5.1 so that the run order will be at least 1-trend-free for each main-effect and each two-factor interaction. Using the generators $a_1a_2, a_1a_3$ and $a_2a_3$ produces such an ordering. Based on this run order and using the transformation of Thompson and Myers gives the 1-trend-free design in Table 5.5.
Table 5.5: A 1-trend-free design for a second-order canonical polynomial in four components

<table>
<thead>
<tr>
<th>$z_1$</th>
<th>$z_2$</th>
<th>$z_3$</th>
<th>$z_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.309</td>
<td>0.096</td>
<td>0.419</td>
<td>0.176</td>
</tr>
<tr>
<td>0.287</td>
<td>0.237</td>
<td>0.300</td>
<td>0.176</td>
</tr>
<tr>
<td>0.265</td>
<td>0.378</td>
<td>0.181</td>
<td>0.176</td>
</tr>
<tr>
<td>0.240</td>
<td>0.191</td>
<td>0.449</td>
<td>0.120</td>
</tr>
<tr>
<td>0.218</td>
<td>0.332</td>
<td>0.330</td>
<td>0.120</td>
</tr>
<tr>
<td>0.441</td>
<td>0.228</td>
<td>0.211</td>
<td>0.120</td>
</tr>
<tr>
<td>0.172</td>
<td>0.285</td>
<td>0.479</td>
<td>0.064</td>
</tr>
<tr>
<td>0.395</td>
<td>0.182</td>
<td>0.360</td>
<td>0.064</td>
</tr>
<tr>
<td>0.372</td>
<td>0.323</td>
<td>0.241</td>
<td>0.064</td>
</tr>
<tr>
<td>0.382</td>
<td>0.168</td>
<td>0.330</td>
<td>0.120</td>
</tr>
<tr>
<td>0.360</td>
<td>0.309</td>
<td>0.211</td>
<td>0.120</td>
</tr>
<tr>
<td>0.159</td>
<td>0.272</td>
<td>0.449</td>
<td>0.120</td>
</tr>
<tr>
<td>0.313</td>
<td>0.263</td>
<td>0.360</td>
<td>0.064</td>
</tr>
<tr>
<td>0.291</td>
<td>0.404</td>
<td>0.241</td>
<td>0.064</td>
</tr>
<tr>
<td>0.335</td>
<td>0.122</td>
<td>0.479</td>
<td>0.064</td>
</tr>
<tr>
<td>0.205</td>
<td>0.319</td>
<td>0.300</td>
<td>0.176</td>
</tr>
<tr>
<td>0.428</td>
<td>0.215</td>
<td>0.181</td>
<td>0.176</td>
</tr>
<tr>
<td>0.227</td>
<td>0.177</td>
<td>0.419</td>
<td>0.176</td>
</tr>
<tr>
<td>0.454</td>
<td>0.241</td>
<td>0.241</td>
<td>0.064</td>
</tr>
<tr>
<td>0.254</td>
<td>0.204</td>
<td>0.478</td>
<td>0.064</td>
</tr>
<tr>
<td>0.231</td>
<td>0.345</td>
<td>0.360</td>
<td>0.064</td>
</tr>
<tr>
<td>0.346</td>
<td>0.296</td>
<td>0.182</td>
<td>0.176</td>
</tr>
<tr>
<td>0.146</td>
<td>0.259</td>
<td>0.419</td>
<td>0.176</td>
</tr>
<tr>
<td>0.369</td>
<td>0.155</td>
<td>0.300</td>
<td>0.176</td>
</tr>
<tr>
<td>0.278</td>
<td>0.391</td>
<td>0.211</td>
<td>0.120</td>
</tr>
<tr>
<td>0.322</td>
<td>0.109</td>
<td>0.449</td>
<td>0.120</td>
</tr>
<tr>
<td>0.300</td>
<td>0.250</td>
<td>0.330</td>
<td>0.120</td>
</tr>
</tbody>
</table>
5.8 Construction of Trend-Free Designs Mixture Experiments in Constrained Regions

Suppose that the region of interest, $S$ is a constrained subregion of the simplex defined by

$$L_j \leq z_j \leq U_j \quad \text{for all } j = 1, \ldots, p$$

and the response will be modeled by a $d$-th order canonical polynomial in the presence of a $b$-th order polynomial trend.

Recall that the procedure of Section 4.4 used a method given by Piepel (1990,1991). The following algorithm constructs trend-free run orders for a mixture experiment that are based on a trend-free design for a (fractional) factorial experiment. Section 5.5 provides a discussion on how to choose the (fractional) factorial design.

First, form a $b$-trend-free $OA(N, p - 1, d', t)$ or a $b$-trend-free $d'^{p-1}$ fractional factorial design, where $d' \geq d$ and the resolution $t + 1 = 2d + 1$. Assign $p - 1$ of the mixture components to the $p - 1$ columns using the criteria advanced in Section 5.5. For example, the most important components may be assigned to the first few columns and the least important to the last columns. For each of the $p - 1$ components, associate the $d'$ levels with proportions, the largest proportion not exceeding the component’s upper bound and the smallest not lower than the component’s lower bound.

Create a $p$-th column whose entries are taken to be $z_p = 1 - \sum_{j=1}^{p-1} z_j$ for each row in the array. If $z_p < L_p$ or $z_p > U_p$, set $z_p$ to be the closer of $L_p$ or $U_p$, breaching the constraint that the row sum be one. When adjustments are required, start with $z_{p-1}$ and try to adjust it so that the row sum is one, keeping $z_{p-1}$ within its constraint.
If this cannot be accomplished using $z_{p-1}$ alone, work backwards toward $z_1$, making similar adjustments to each column until the row sum is one, keeping all of the $z_j$'s within their respective constraints. This adjustment should be repeated for each row in the array, as needed.

We exploit the fact that if $OA(N, p - 1, d', 2d)$ is $b$-trend-free for a $d$-th order polynomial in the unrestricted variables, then the resulting design for the mixture experiment should be $b$-trend-free for a $d$-th order canonical polynomial because the patterns in high and low levels should be nearly the same as for a trend-free run order of the corresponding factorial experiment. The qualifier should be is used because the procedure's success depends on the structure of the region $S$. In particular, the efficiency of the design depends on the adjusting that is done to the columns associated with $z_1, \ldots, z_{p-1}$.

Piepel was concerned with the ($D$-, $G$-, $V$-) optimality properties of these designs. He found that as more columns required adjustment, there was degradation in the performance of designs with respect to these optimality criteria. The same is true for orthogonality to deterministic trends. As columns other than the $p$-th are changed to meet the row sum constraint, the orthogonality between the trend columns and the columns corresponding to the canonical polynomial parameters is destroyed. He notes that two worst cases exist: when the first $(p - 1)$ mixture components are set at their upper bounds and the $p$-th component is below its lower bound and when the first $(p - 1)$ mixture components are set at their lower bounds and the $p$-th component is above its upper bound. The degree to which this is problematic depends on how
far above or below the \( p \)-th component is from its closes bound. Let

\[
D_p = \max_i \{0, 1 - \sum_{j=1}^{p-1} z_j - U_p, L_p - (1 - \sum_{j=1}^{p-1} z_j)\}.
\]

As \( D_p \) increases, so will the number of columns that will have to be altered to meet the constraint that the rows must sum to one and thus damaging any optimality properties of the design, including being trend-free. See Section 4.4 for additional comments.

Piepel recommends that these designs be used only for \( p \geq 7 \) mixture components because with fewer mixture components, the possibility of degradation of optimality criteria is increased because there is a higher likelihood that extreme vertices of \( S \) will be replicated. However, in the present usage, replication of vertices is of less concern since we are interested in their performance with respect to trends. Too much replication may lead to too few distinct points, creating a singular model matrix \( Z'(I - P_U)Z \). From this author’s experience with using this method, the designs seem to perform well, even with as few as four mixture components.

To illustrate the algorithm’s workings, consider the following example taken from Cornell (1990, p.256). It was of interest to determine the effects of several gasoline components on octane ratings by fitting a first-order canonical polynomial. However, for purposes of illustration, suppose that a linear trend was suspected. These components had the following constraints placed upon them and have been renumbered in order of range, \( R_j \):

\[
\begin{align*}
0 \leq z_1 &\leq 0.08 & 0 \leq z_2 &\leq 0.12 \\
0 \leq z_3 &\leq 0.12 & 0 \leq z_4 &\leq 0.21 \\
0 \leq z_5 &\leq 0.62 & 0 \leq z_6 &\leq 0.62
\end{align*}
\]
We generate a two-level fractional factorial design for $p-1$ factors in sixteen runs according to the fraction $I = ABCF = ADEF = BCDE$ which is 1-trend-free. It is impossible to find an eight run design in which all six main effects are 1-trend-free, so the quarter fraction was used. This design is of resolution IV so we can estimate all main effects free of two-factor interactions and can estimate linear combinations of two-factor interactions. Once a fractional factorial design has been selected, a suitable run order needs to be found. The run order in Table 5.6 was induced by the generators $abe, acd, abef$, and $abcdef$. Application of part 1 of Theorem 5.5.1 shows that factor $A$ is 3-trend-free, factors $B$ and $C$ are 2-trend-free, and factors $D$, $E$, and $F$ are 1-trend-free. From part 2, we have that the interaction terms $AF + BC$, $AD + EF$, and $AE + DF$ are 1-trend-free; $BD + CE$ is 2-trend-free.

Assign factors $A - F$ to $z_1 - z_6$ and for each factor associate $-1$ with $L_j$ and 1 with $U_j$. Determine $z_7$ by subtracting the row sum from one. For many mixtures this was adequate and no further adjustment was required. However, rows in Table 5.7 with *'s need adjustment.

To understand the adjustment portion of the algorithm, focus on the rows having *'s where adjustment has taken place. For all other rows, $z_7$ is obtained by subtracting the sum of the first six mixture component proportions from one. In row one, the first six mixture component proportions are at their low levels, zero. We try to set $z_7 = 1$, but this violates the constraints on $z_7$. Instead set $z_7 = U_7 = 0.74$ and try to adjust $z_6$; we do so, setting $z_6 = 0.26$, completing our adjustments for this row. Row nine poses a different problem. The first six components are at their high level, breaching the constraint that the row should sum to one when $z_6$ is set to 0.62. To
resolve this, set \( z_7 \) and \( z_8 \) to their low levels (zero) and adjust \( z_5 \) so that the row sum is one. The other rows marked by '*'s are treated similarly.

The design given in Table 5.7 is for all practical purposes 1-trend-free. For the first-order canonical polynomial, \(| Z'_1(I - P_L)Z_1 | = 0.000148\), and the efficiency for \( \tau \) is \( \text{eff}(\tau) = 0.996 \), and the parameter efficiencies [see Definition (2.4.2)] are

\[
\begin{align*}
\text{eff}(r_1) &= 0.999 \\
\text{eff}(r_2) &= 0.998 \\
\text{eff}(r_3) &= 0.999 \\
\text{eff}(r_4) &= 0.996 \\
\text{eff}(r_5) &= 0.997 \\
\text{eff}(r_6) &= 0.984 \\
\text{eff}(r_7) &= 0.983
\end{align*}
\]

Notice that as expected, the efficiencies tail off somewhat as we approach the last column, due to the amount of adjustment that was needed to make the row sum to one.

In general, as more adjustment is done to the columns corresponding to \( z_1, \ldots, z_{p-1} \), the orthogonality between the columns for trend and the columns for the mixture model may be destroyed. This generalization is entirely situation dependent.

Bayne and Ma (1987) attempt to optimize the solvent composition for use in high performance thin-layer chromatography. They examine a solvent consisting of five mixture components which are constrained to lie in the subregion of the simplex
Bayne and Ma assume that a second-order canonical polynomial will be adequate to model the response. In order to fit the second-order polynomial, we use an $OA(27, 4, 3, 3)$. The generators used to determine a run order were $a_1a_3, a_2a_3a_4^2,$ and $a_1^2a_2a_4$. Using Theorem 5.5.1, we can conclude that $A_1, A_2, A_3, A_4, A_1A_2, A_1A_3, A_1A_4, A_2A_3,$ and $A_3A_4$ are all 1-trend-free. The design is given in Table 5.8. Notice that this is indeed a worst case scenario; many points have been replicated and $z_5$ is zero in
Table 5.7: 1-trend-free design for a first-order canonical polynomial in seven mixture components

<table>
<thead>
<tr>
<th>$z_1$</th>
<th>$z_2$</th>
<th>$z_3$</th>
<th>$z_4$</th>
<th>$z_5$</th>
<th>$z_6$</th>
<th>$z_7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.08</td>
<td>0.12</td>
<td>0</td>
<td>0.62</td>
<td>0</td>
<td>0.26</td>
<td>0.74*</td>
</tr>
<tr>
<td>0.12</td>
<td>0.12</td>
<td>0.21</td>
<td>0</td>
<td>0.55</td>
<td>0</td>
<td>0.74</td>
</tr>
<tr>
<td>0</td>
<td>0.12</td>
<td>0.12</td>
<td>0.21</td>
<td>0.62</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>0</td>
<td>0.12</td>
<td>0.12</td>
<td>0.21</td>
<td>0.62</td>
<td>0.09</td>
<td>0.05</td>
</tr>
<tr>
<td>0.08</td>
<td>0.12</td>
<td>0</td>
<td>0.62</td>
<td>0.26</td>
<td>0</td>
<td>0.05</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.62</td>
<td>0.62</td>
<td>0.26</td>
<td>0</td>
<td>0.05</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.62</td>
<td>0.62</td>
<td>0</td>
<td>0.17</td>
<td>0.05</td>
</tr>
<tr>
<td>0.08</td>
<td>0.12</td>
<td>0.21</td>
<td>0</td>
<td>0.62</td>
<td>0</td>
<td>0.17</td>
</tr>
<tr>
<td>0.08</td>
<td>0.12</td>
<td>0.21</td>
<td>0</td>
<td>0.62</td>
<td>0.06</td>
<td>0.05</td>
</tr>
<tr>
<td>0.08</td>
<td>0.12</td>
<td>0</td>
<td>0.62</td>
<td>0.02</td>
<td>0.74</td>
<td></td>
</tr>
</tbody>
</table>

all but eight of the twenty-seven mixtures. When translated into mixtures using the procedure outlined above, we do not fare as well. The efficiency is only 92.5%. Once again though, many of the first-order parameters are, for all practical purposes, 1-trend-free and selected second-order parameters have high efficiencies:

\[

eff(\tau_1) = 0.999 \quad \text{eff}(\tau_{12}) = 0.986 \quad \text{eff}(\tau_{24}) = 0.999 \\
\text{eff}(\tau_2) = 0.999 \quad \text{eff}(\tau_{13}) = 0.739 \quad \text{eff}(\tau_{25}) = 0.944 \\
\text{eff}(\tau_3) = 0.998 \quad \text{eff}(\tau_{14}) = 0.991 \quad \text{eff}(\tau_{34}) = 0.741 \\
\text{eff}(\tau_4) = 0.985 \quad \text{eff}(\tau_{15}) = 0.951 \quad \text{eff}(\tau_{35}) = 0.959 \\
\text{eff}(\tau_5) = 0.971 \quad \text{eff}(\tau_{23}) = 0.997 \quad \text{eff}(\tau_{45}) = 0.712.
\]
Table 5.8: A 1-trend-free design for the Bayne and Ma (1987) example

<table>
<thead>
<tr>
<th></th>
<th>$z_1$</th>
<th>$z_2$</th>
<th>$z_3$</th>
<th>$z_4$</th>
<th>$z_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>0.035</td>
<td>0.000</td>
<td>0.450</td>
<td>0.000</td>
<td>0.515</td>
<td></td>
</tr>
<tr>
<td>0.070</td>
<td>0.000</td>
<td>0.900</td>
<td>0.000</td>
<td>0.030</td>
<td></td>
</tr>
<tr>
<td>0.000</td>
<td>0.300</td>
<td>0.450</td>
<td>0.520</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.035</td>
<td>0.300</td>
<td>0.665</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.070</td>
<td>0.300</td>
<td>0.000</td>
<td>0.630</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.000</td>
<td>0.600</td>
<td>0.400</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.035</td>
<td>0.600</td>
<td>0.000</td>
<td>0.365</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.070</td>
<td>0.600</td>
<td>0.330</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.070</td>
<td>0.300</td>
<td>0.000</td>
<td>0.630</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.000</td>
<td>0.300</td>
<td>0.450</td>
<td>0.250</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.035</td>
<td>0.300</td>
<td>0.665</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.070</td>
<td>0.600</td>
<td>0.330</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.000</td>
<td>0.600</td>
<td>0.400</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.035</td>
<td>0.600</td>
<td>0.000</td>
<td>0.365</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.070</td>
<td>0.000</td>
<td>0.900</td>
<td>0.030</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.930</td>
<td>0.070</td>
<td></td>
</tr>
<tr>
<td>0.035</td>
<td>0.000</td>
<td>0.450</td>
<td>0.515</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.035</td>
<td>0.600</td>
<td>0.000</td>
<td>0.365</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.070</td>
<td>0.600</td>
<td>0.330</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.000</td>
<td>0.600</td>
<td>0.400</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.035</td>
<td>0.450</td>
<td>0.515</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.070</td>
<td>0.000</td>
<td>0.900</td>
<td>0.030</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.930</td>
<td>0.070</td>
<td></td>
</tr>
<tr>
<td>0.035</td>
<td>0.300</td>
<td>0.665</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>0.000</td>
<td>0.300</td>
<td>0.000</td>
<td>0.000</td>
<td>0.700</td>
<td></td>
</tr>
<tr>
<td>0.000</td>
<td>0.300</td>
<td>0.450</td>
<td>0.000</td>
<td>0.250</td>
<td></td>
</tr>
</tbody>
</table>
5.9 Algorithmic Methods

This section examines two algorithmic approaches for finding trend-free orderings for mixture experiments.

We propose a method that presupposes nothing about the region of interest (i.e. additional constraints on the components) in the simplex. It is intended to be used when the mixtures have already been selected according to some specifications. In this case, there is no opportunity to select the mixtures as in Sections 5.7 and 5.8 since the matrix $Z$ is fixed.

Instead, the proposed methodology tries to find an optimal ordering of the mixtures at hand according to some meaningful criterion. In light of Definition 5.1.1, any criterion used should make $U^T Z$ small. However, the choice of how to make $U^T Z$ small is not entirely obvious. As in Chapter 4, the objective function considered is $|Z'(I - P_U)Z|$. This criterion has the interpretations that the mixture ordering is $b$-trend-free if and only if:

$$|Z'(I - P_U)Z| = |Z'Z|.$$

Thus, we maximize $|Z'(I - P_U)Z|$.

Note that an “optimal” ordering always exists, an ordering that produces the maximum value of the criterion.

The next two sections give optimization algorithms for the $|Z'(I - P_U)Z|$ criterion, one random, the other deterministic.
5.9.1 Random Sampling of Run Orders to Maximize $|Z'(I - P_U)Z|$ 

The first method of optimization involves selecting a random sample of size $M$ from the distribution of mixture orders and using the ordering producing the largest value of $|Z'(I - P_U)Z|$. 

The algorithm used is like that of Joiner and Campbell (1976) who were concerned with constructing run orders of factorial experiments in the presence of deterministic trends. However, as they noted, the algorithm is more general and can apply to any type of covariate, including blocks as in Section 4.6.2. 

Orders of the mixtures are randomly sampled from the set of all such orders. The user determines sample size. For each ordering in the sample, $|Z'(I - P_U)Z|$ is computed and the ordering having the maximum value is selected as the "best" design. Depending on the sample size chosen, this approach can require less computation than the interchange algorithm discussed in the next section. Joiner and Campbell also made provisions for more complicated cost considerations that are not considered here. 

Because we maximize $|Z'(I - P_U)Z|$, the model must be specified at the outset. Also, we must have at least $N = s + b$ mixtures, where $b$ is the degree of the expected trend. These mixtures must be selected such that $\gamma$ and $\tau$ are estimable. Of course if one is interested in some other linear parametric function of $\tau$, say $\Delta \tau$, the criterion has an obvious modification and the mixtures must be chosen such that this function is estimable. 

Generate $J$ random orderings of the numbers $1, 2, \ldots, N$. For each sample put the mixtures in the places indicated by the ordering. Find $|Z'(I - P_U)Z|$ for each of the $J$ random orderings. Choose the ordering with the maximum value of
| $Z'(I - P_U)Z$ |. Hopefully, this ordering will have an acceptable value and efficiency; this is not guaranteed since this is only a random sample. As with any random sample, the larger the sample, the more precise inference we can make about the maximum of the entire population of orderings.

If desired the algorithm can be repeated, hopefully increasing the chances of finding the best design. However, as with all other design algorithms, this is not guaranteed to find the optimal ordering. The probability that a random sample captures an optimal ordering depends on the number of orderings which have the optimal value of $| Z'(I - P_U)Z |$.

Consider an example of Draper and Lawrence (1965). They found the minimum bias design for $p = 3$. There appears to be no other method of finding a trend-free arrangement other than to use some algorithmic method. Here we use the sampling algorithm outlined above with $J = 40$. This gave the result in Table 5.9.

Table 5.9: A nearly 2-trend-free arrangement for the minimum integrated bias design of Draper and Lawrence (1965)

<table>
<thead>
<tr>
<th>$z_1$</th>
<th>$z_2$</th>
<th>$z_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.256</td>
<td>0.622</td>
<td>0.122</td>
</tr>
<tr>
<td>0.591</td>
<td>0.075</td>
<td>0.334</td>
</tr>
<tr>
<td>0.081</td>
<td>0.081</td>
<td>0.837</td>
</tr>
<tr>
<td>0.075</td>
<td>0.591</td>
<td>0.334</td>
</tr>
<tr>
<td>0.482</td>
<td>0.482</td>
<td>0.036</td>
</tr>
<tr>
<td>0.622</td>
<td>0.256</td>
<td>0.122</td>
</tr>
<tr>
<td>0.410</td>
<td>0.045</td>
<td>0.545</td>
</tr>
<tr>
<td>0.241</td>
<td>0.678</td>
<td>0.081</td>
</tr>
<tr>
<td>0.045</td>
<td>0.410</td>
<td>0.545</td>
</tr>
<tr>
<td>0.678</td>
<td>0.241</td>
<td>0.081</td>
</tr>
<tr>
<td>0.184</td>
<td>0.184</td>
<td>0.631</td>
</tr>
</tbody>
</table>
The efficiency for this design is 0.954, the determinant value is $2.185 \times 10^{-6}$. Efficiency is measured as in Section 2.4. Considering that the efficiency uses an unattainable bound as its yardstick, 0.954 does not seem to be too bad. The individual parameters efficiencies are all over 95% except $\tau_{13}$ and $\tau_{23}$ with 89% and 87%, respectively. Again, these efficiencies all seem to be quite high. In practice, it is best to generate several samples to see if any improvement can be made. Since this algorithm is not computationally intensive, such tries should be relatively inexpensive.

The interchange algorithm will be applied to this design in the next section to see if any improvement will result and show how the algorithms can work in tandem.

5.9.2 An Interchange Algorithm to Maximize $|Z'(I - P_U)Z|$

In contrast to the random sampling approach of Section 5.9.1, a wholly deterministic approach is now considered.

Cook and Nachtsheim (1989) proposed an algorithm for blocking response surface designs. The method given here is a modification of Cook and Nachtsheim's interchange algorithm, tailored to our present problem of finding optimum run orders. Effectively, this increases the scope of the algorithm because interchanges need to be considered for each distinct pair of mixtures, rather than only among mixtures in different blocks.

As with the random sampling algorithm, the user must select a design for which $\gamma$ and $\tau$ are estimable under the specified model. Also, a convergence criterion, say $\zeta$, must be specified.

The algorithm may be summarized as follows. For each mixture, consider interchanges with all other mixtures, i.e. consider switching mixture 1 with all
other mixtures, then mixture 2 with all other mixtures, etc. For each interchange, 
\[ |Z'(I - P_U)Z| \] is calculated. If an improvement results from the interchange, immediately incorporate it and update \[ |Z'(I - P_U)Z| \] and \( Z \). When these interchanges have been exhausted, compare \[ |Z'(I - P_U)Z| \] with its starting value, i.e. is
\[
\frac{|Z'(I - P_U)Z|}{|Z^{(0)}(I - P_U)Z^{(0)}|} < \zeta,
\]
where \( Z^{(0)} \) is the model matrix at the previous stage of the algorithm. If this condition is untrue, then the next stage of interchanges begins, starting with the new mixture 1; if true, the algorithm stops. The criterion \( \zeta \) should be some fraction much smaller than one. The smaller \( \zeta \) is, the longer the algorithm will take to converge, finding a better design than for a larger value of \( \zeta \). Of course there is a point of diminishing return.

Because there is no probabilistic element in the interchange algorithm, the changes to be made are predetermined and advantageous interchanges requiring more than the interchange of a group of variables may not be found. As a result, it may be unable to find optimal orderings that may be found by the sampling algorithm. For instance, when beginning with the order presented in Draper and Lawrence (1965a), the interchange algorithm produced an improved ordering but an ordering found by the random sampling method was superior. This flaw in the interchange algorithm points to a need to introduce a stochastic element.

There are several ways that a random element may be included in the interchange algorithm. Foremost, and simplest, involves using the interchange and random sampling algorithms in tandem. This induces a random element by choosing the starting design to be the best ordering found by the random sampling algorithm. Table 5.10 gives an ordering generated by the interchange algorithm starting with the best or-
dering found by the random sampling algorithm, superior to any ordering found by individual use of either technique. The determinant was marginally improved to $2.24 \times 10^{-6}$. Another way in which randomness may be introduced into the interchange algorithm is to consider an individual interchange to be a Bernoulli trial; the interchange is incorporated with a probability set by the user. This would allow the algorithm to possibly break out of its deterministic pattern, finding advantageous changes that would otherwise be missed, hopefully finding a better ordering.

Table 5.10: A nearly 2-trend-free arrangement for the minimum integrated bias design of Draper and Lawrence (1965)

<table>
<thead>
<tr>
<th>$z_1$</th>
<th>$z_2$</th>
<th>$z_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.256</td>
<td>0.622</td>
<td>0.122</td>
</tr>
<tr>
<td>0.591</td>
<td>0.075</td>
<td>0.334</td>
</tr>
<tr>
<td>0.081</td>
<td>0.081</td>
<td>0.837</td>
</tr>
<tr>
<td>0.622</td>
<td>0.256</td>
<td>0.122</td>
</tr>
<tr>
<td>0.482</td>
<td>0.482</td>
<td>0.036</td>
</tr>
<tr>
<td>0.075</td>
<td>0.591</td>
<td>0.334</td>
</tr>
<tr>
<td>0.410</td>
<td>0.045</td>
<td>0.545</td>
</tr>
<tr>
<td>0.241</td>
<td>0.678</td>
<td>0.081</td>
</tr>
<tr>
<td>0.045</td>
<td>0.410</td>
<td>0.545</td>
</tr>
<tr>
<td>0.678</td>
<td>0.241</td>
<td>0.081</td>
</tr>
<tr>
<td>0.184</td>
<td>0.184</td>
<td>0.631</td>
</tr>
</tbody>
</table>

The interchange algorithm can also be used for "design repair;" given a non-singular starting design or ordering, subject the ordering to the interchange algorithm to see if an improvement can be effected. Recall that the Bayne and Ma (1987) example of Section 5.8 was somewhat of a failure with respect to the construction method formulated in that section. Since the design was fixed, it was subjected to the interchange algorithm to see if a more advantageous order for the mixtures could
be found. The interchange algorithm was able to find an ordering that is for all intensive purposes, 1-trend-free.

Finally, with a modification of the algorithm, trend-free or nearly-trend-free block designs may be found.

Consider the model

\[ y = U_1 \gamma_1 + U_2 \gamma_2 + Z \tau + e, \]

where \( \gamma_1 \) is a vector of block effects and \( \gamma_2 \) is a vector of orthogonal polynomials corresponding to a common deterministic trend within blocks. Then, Bradley and Yeh (1980) define a trend-free block design as a block design for which

\[ SS(\tau, \mu, \gamma_1) = SS(\tau, \mu, \gamma_1, \gamma_2). \]

They demonstrate that a necessary and sufficient condition for a block design to be a trend-free block design is for \( U_2'Z = 0 \).

Thus, by modifying the interchange algorithm into a two step process, using the interchange algorithm of Section 4.6.1 with the algorithm of this section we can find trend-free or nearly-trend-free block designs. These two algorithms work in tandem. First, use the interchange algorithm for blocks to find the best allocation of runs to blocks. Then, apply the interchange algorithm for trends to the individual blocks to find the best mixture order within each block.

5.10 Conclusions

Goel (1980) initiated the study of trend-free designs for mixture experiments by providing a method of constructing orderings for the mixtures for which the estimator of \( \tau \) is uncorrelated with the estimator of \( \gamma \).
Building on Goel's work, this chapter presented a variety of methods for finding trend-free designs. The methods described in Sections 5.7 and 5.8 gave ways of selecting both the mixtures and the ordering of the mixtures, creating a trend-free design. Sections 5.9.1 and 5.9.2 found nearly-optimum orderings in situations where the mixtures are already specified.

The algorithmic methods should be viewed as complementary to one another and to the other methods in this chapter. Both can be used to see if a design can be improved by finding a better ordering of the mixtures for a specified polynomial trend. Regarding flexibility, the algorithmic methods get the edge. Both the random sampling algorithm and the interchange algorithm can be easily adapted to any existing set of mixtures and for any model for the mixture components, including trend-free block designs. It should also be noted that both techniques have applications beyond the arena of mixture experiments.

The method using transformations is also quite ubiquitous, applicable to any canonical polynomial and when the region is a constrained subregion of the simplex. Trend-free designs can be easily found by borrowing from the theory of trend-free factorial designs, whereas the computer methods produce designs that are often only nearly-trend-free. However, the transformation method suffers from the same limitations discussed in Chapter 3.

Finally, using orthogonal arrays in constrained regions can be successful—depending on the characteristics of the subregion in the simplex. Of the four methods discussed in this chapter, the orthogonal array method seems to be the easiest to implement. The trick is to select the appropriate orthogonal array or fraction for the hypothesized model, sometimes not an easy or possible task. What is needed is a further develop-
ment of orthogonal arrays of strength $t+$ [Hedayat (1990)] which will allow estimation of main effects and certain interactions. Otherwise the size of the orthogonal arrays of the required strength become unwieldy.
CHAPTER 6. CONCLUSIONS

The successful application of experimental design is facilitated by the development of general classes of designs fitting a large variety of experimental scenarios. The design of mixture experiments is no exception; the work of Scheffé, Snee, Cornell and others provides researchers with many options that fit a variety of situations.

This paper attempts to add to these options in the less well-developed areas of block designs and trend-free designs for mixture experiments. In a real sense, the methods in this paper are “something old, something new, something borrowed and something blue.” A little bit of everything was used in attempting to find sensible designs for mixture experiments in the presence of blocking variables and time trends. Techniques from factorial experiments, combinatorial design theory, transformations, and computational methods were all used at different points in time and all with success.

Chapter 3 dredged the backwaters of transformations in mixture experiments to find useful results. Nothing startling was obtained, but it is hoped that the role of transformations was put into the proper perspective. The one new result was the invariance of orthogonality to deterministic trends under non-singular linear transformations, a fact used in Chapter 5. Also, some results on induced matrices were collected and elucidated, something that this author has not found in searches on the
Both blocking and trend-free run orders have been well explored for factorial experiments. In factorial experiments, algebraic techniques have yielded results because of their group theoretic nature. However, no algebraic approaches seem natural when dealing with mixture experiments, mixture mates excepting. Because of this, many of the results obtained here have attempted to utilize known results for factorial experiments where applicable. Other results have imposed a factorial structure upon mixture experiments when the region of interest is a constrained subregion of the simplex and had some success, but also some problems with the practical limitations of such designs. The applications to blocking and deterministic trends have been successful, sharing the same caveats, such as difficulty with transformation methods and the use of fractional factorials and orthogonal arrays as templates for extreme vertices designs.

In the arena of block designs, existing necessary and sufficient conditions for orthogonal blocking were examined in detail and extended to include a larger class of models, including all \( d \)-th order canonical polynomials. Based on these conditions two flexible structures called symbolic and integral mixture mates of strength \( t \) have been developed. Both structures generalize current notions about mixture mates, allowing for a larger variety of structures and also partially decoupling the idea that only Latin squares can be used to produce orthogonal block designs for mixture experiments.

While developing new methods to construct integral mixture mates of strength \( t \), we have also begun the development of trade-off methods for \( m \)-ary designs. Nothing on this subject has appeared in the literature; it is hoped that the intimate connections between trade-off and integral mixture mates will spur research in trade-off
theory. The results provided encompass known results on binary designs as special cases.

Non-orthogonal block designs were studied in an attempt to get around the stringent requirements of orthogonal blocking and still produce useful designs. This was done by using mixture mates in conjunction with asymmetrical orthogonal arrays, shedding further light on the utility of mixture mates in block designs.

Regarding trend-free designs for mixture experiments, this appears to be the first work since Goel (1980) tackled the issue. As in the case of blocking, factorial designs have been utilized in several different manners as a primary method to create trend-free designs for mixture experiments. Trend-free (fractional) factorial designs have been used with the transformation techniques of Chapter 3 and as templates in constrained subregions of the simplex to produce trend-free designs for mixtures.

Finally, when all else fails and the structure of the problem just does not fit into any of the other options, computational methods for finding designs minimizing the correlation between estimators of the parameters in the mixture polynomial and estimators of the covariate parameters have been suggested. This has been utilized with success when dealing with both blocking variables and deterministic trends.
BIBLIOGRAPHY


Khosrovshahi, G. B. and Ajoonian-Namini, S. A New Basis for Trades. Technical


Tocher, K. D. Design and Analysis of Block Experiments. *Journal of the Royal


Wu, C. F. J., Runchu Zhang, and Renguan Wang. Construction of Asymmetrical Orthogonal Arrays of the type $OA(s^k, s^m(s^{s_1})^{n_1} \cdots (s^{s_t})^{n_t})$. Statistica Sinica, 2:203–219, 1992.
ACKNOWLEDGEMENTS

I'd like to thank three people who had to deal with me over the last four years or so.

First, my thanks go to Dr. John Stufken who, at times, must have felt that he had another child asking "Are we done yet?" More seriously, without his guidance and insight, this dissertation would never have come about. He pushed me when needed and reigned me in when I pushed too much.

Next to my wife Sue, who must have felt as if she was single again for these last six months as I pushed to finish. Her love, moral support, patience, and ability to bring me back to reality kept me on course in a non-academic sense.

Finally, I thank my parents and family for their love and support over these last twenty-two long years of education.
APPENDIX

In this Appendix, the optimality criterion discussed in Chapter 3 are described. These criteria include alphabetic optimality criteria (D-, G-, and V-optimality), rotatability, and criteria accounting for misspecification of the systematic portion of the model (minimum integrated bias and minimum integrated mean-squared error).

To facilitate discussion of the various criteria in a general setting, the following generic model notation is adopted:

\[ y = R\theta + e \]
\[ = S\alpha + T\delta + e, \quad (A.1) \]

where \( S \) is \( N \times t \) and \( T \) is \( N \times q - t \). Suppose that \( \text{rank}(R) = q \) and \( \text{rank}[T'(I - P_S)T] = q - t \) are attainable for some designs, where \( P_S = S(S'S)^{-1}S' \). Let \( \mathcal{R} \) be the design space.

All alphabetic optimality criteria assume that the model is specified correctly (no systematic errors are present).

Let \( \xi \) be a design and \( \Xi \) be the class of competing designs for a particular model. \( \Xi \) may consist of all \( N \) run designs with runs from \( \mathcal{R} \) or a smaller class of designs. Under model A.1, let

\[ \hat{\theta}_\xi = (R'_\xi R_\xi)^{-1} R'_\xi y \]
be the ordinary least squares estimator of $\theta$. Then,

$$\text{var}(\hat{\theta}_\xi) = \sigma^2(R'_\xi R_\xi)^{-1}$$

**Definition A.1**: A design $\xi^*$ is said to be D-optimal in $\Xi$ for $\theta$ if it minimizes

$$| (R'_\xi R_\xi)^{-1} | = \prod_{i=1}^{q} l_{\xi_i}^{-1}$$  \hspace{1cm} (A.2)

over the class of designs $\Xi$. Here, $l_{\xi_i}$'s are the $q$ eigenvalues of the matrix $R'_\xi R_\xi$. If $R'_\xi R_\xi$ is singular for some design, then $| (R'_\xi R_\xi)^{-1} |$ taken as infinite.

This type of D-optimality is unsuitable if we are only interested in $\delta$, often the case when $\alpha$ corresponds to a covariate. Instead of D-optimality, consider the following criterion.

**Definition A.2**: A design $\xi^*$ is said to be $D_\delta$-optimal in $\Xi$ if it minimizes

$$| (T'_\xi(I - P_{S_t})T_\xi)^{-1} | = \prod_{i=1}^{q-t} \tilde{l}_{\xi_i}^{-1}$$  \hspace{1cm} (A.3)

over the class of designs $\Xi$, where $\tilde{l}_{\xi_i}$'s are the $q-t$ positive eigenvalues of $T'_\xi(I - P_{S_t})T_\xi$.

Alternatively, $D_\delta$-optimality may be posed as a maximization problem, where

$$\prod_{i=1}^{q-t} \tilde{l}_{\xi_i}$$

is to be maximized over $\Xi$.

Box and Draper (1971) and Hedayat (1981) motivate the statistical interpretation of D- and $D_\delta$-optimality.

The next two criteria do not generalize to models having covariates which are of secondary interest. Suppose that $S$ is omitted from the model. Then, $\hat{y}(t)$ denotes the estimated mean response at the point $t \in \mathcal{R}$. 
Another alphabetic optimality criterion that is useful in mixture experiments is *V-optimality*.

The quantity

\[ V_\xi(t) = \frac{N}{\sigma^2} \text{var}(\hat{\gamma}_\xi(t)) \]  

(A.4)

is called the *variance function*.

**Definition A.3**: A design \( \xi^* \) is a minimum integrated variance design or a V-optimal or IV-optimal design if it minimizes

\[ V_\xi = \int_\mathcal{R} V_\xi(t)W(t)dt \]  

(A.5)

over the class of designs \( \Xi \), where \( W(t) \) is some weight function such that \( \int_\mathcal{R} W(t)dt = 1 \).

**Definition A.4**: A design \( \xi^* \) is said to be G-optimal if it minimizes the maximum variance of an estimated mean response over the class of designs \( \Xi \), i.e.

\[ \max_{t \in \mathcal{R}} \text{var}(\hat{\gamma}_{\xi^*}(t)) = \min_{\xi \in \Xi} \max_{t \in \mathcal{R}} \text{var}(\hat{\gamma}_\xi(t)). \]  

(A.6)

Silvey (1980) and Atkinson and Donev (1992), among others, provide more details about the alphabetic optimality criteria and algorithms to generate optimal designs.

For the following discussion of bias and mean-squared error, we continue to ignore any possible covariates to clarify discussion. Suppose that the fitted model is

\[ y = T_1 \delta_1 + e \]  

(A.7)

but the true model is

\[ y = T_2 \delta_2 + e, \]  

(A.8)
where $C(T_1) \subset C(T_2)$. The ordinary least squares estimator of $\delta_1$ under model (A.7) is $(T_1' T_1)^{-1} T_1' y$. The model matrix $T_1$ is specified by the design $\xi \in \Xi$.

Now suppose that we wish to make inference about an estimable function of $\delta$, say $\Lambda \delta$. Then define bias as

$$B_\xi = E(\Lambda \hat{\delta}) - \Lambda \delta_1$$

$$= \Lambda (T_1' T_1)^{-1} T_1' T_2 \delta_2 - \Lambda \delta_1$$

where expectation is taken with respect to the true model. This notion of bias is somewhat different than standard notions [Box and Draper (1987), Khuri and Cornell (1988)] because it does not require that the assumed model (A.7) be nested within the true model (A.8); the constraint on the column spaces is a more general requirement than nesting. This will be helpful in Chapter 3.

Consider the problem of trying to estimate the mean response at a point $t \in \mathcal{R}$, where $t_1'$ and $t_2'$ will be the corresponding rows in model matrices $T_1$ and $T_2$, respectively. In this case, the estimable function is the mean response corresponding to $\Lambda = t_1'$, so the bias is

$$B_\xi(t) = t_1' (T_1' T_1)^{-1} T_1' T_2 \delta_2 - t_1' \delta_1.$$

**Definition A.5**: A design, $\xi^*$, is a minimum integrated bias design if it minimizes

$$B_\xi = \int_{\mathcal{R}} \frac{1}{\sigma^2} B_\xi(t)^2 W(t) dt$$

(A.9)

over the class of designs $\Xi$, where $W$ is a weight function as in Definition A.3.

**Definition A.6**: A design is a minimum integrated mean-squared error design if it minimizes

$$M_\xi = V_\xi + B_\xi.$$

(A.10)
Box and Draper (1959, 1987) and Khuri and Cornell (1988) discuss the role of minimum integrated bias and minimum integrated mean-squared error designs in response surface methodology.

**Definition A.7**: A design is said to be rotatable about a reference point $t_0$ in $\mathcal{R}$ if

$$\text{var}(\hat{y}(t)) = \sigma^2 v(\rho[t, t_0]),$$

(A.11)

i.e. the variance of the estimated mean response at a point, $t$, depends only on the Euclidean distance from $t$ to some reference point in $\mathcal{R}$, $t_0$. Here, $v(*)$ is a non-negative function and $\rho[a, b]$ is the Euclidean distance between $a$ and $b$.

Sa and Edwards (1993) motivate the use of rotatability from a multiple comparisons viewpoint in response surface methodology when the goal is to use steepest ascent for process improvement. Box and Draper (1987) give other motivations for the use of rotatable designs in the context of response surface methods.

Cornell (1991) reviews the role of optimality criteria in designing mixture experiments. Atkinson and Donev (1992) devote several chapters to discussing alphabetic optimality criteria in the context of mixture experiments.

The optimality criteria in the previous section all provide guidelines for constructing designs with desirable properties. It is unwise to rely exclusively on one criterion. Instead, a good design will possess properties making it attractive with respect to several criteria. Vining, Cornell, and Myers (1993) recommend a graphical procedure called a *variance dispersion graph* to compare designs generated from the various criteria with respect to the behavior of the estimated variance of the estimated mean response.