Optimal and Efficient Crossover Designs for Test-Control Study When Subject Effects are Random

BY

WEI ZHENG
B.S., Zhejiang University, 2005
M.S., University of Illinois at Chicago, 2008

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Mathematics
in the Graduate College of the
University of Illinois at Chicago, 2011

Chicago, Illinois

Defense Committee:

Samad Hedayat, Chair and Advisor
Dibyen Majumdar
Jing Wang
Jie Yang
Min Yang, University of Missouri

To my parents and my wife Shujuan

ACKNOWLEDGMENTS

I would like to take this opportunity to express my appreciations to all who have warmly supported my study and research in University of Illinois at Chicago (UIC) in various ways.

First, I want to express my deepest gratitude to my advisor, Professor Samad Hedayat. He was being extremely supportive and helpful for my research and career by giving me a lot of suggestions and supporting me through research assistantship. I would not be able to be what I am now without him. But yet, I understand there is still a long way to go. Second, I would like to give my sincere thanks to Professors Dibyen Majumdar, Jing Wang, and Jie Yang in my department, and Professor Min Yang from University of Missouri for serving on my dissertation committee and for their comments on this dissertation. Third, I would like to thank Professors Emad El-Neweihi, Samad Hedayat, Dibyen Majumdar, Jing Wang, Junhui Wang, and Jie Yang for teaching me knowledge and giving me suggestions for my PhD study in the past five years. Also, my special thanks go to Professor Wei-Biao Wu from University of Chicago and Professor Min Yang. Their advices and encouragements have been crucial for my growth, and have been the power of my continuing the pursue of excellence.

Last, but by no means the least, my heartful appreciation and gratitude goes to my family, especially my parents, my sister and her husband, two of my uncles, and of course my wife for their supports, encouragements, and loves.

The research work and writing for this dissertation is supported by the National Science Foundation Grants DMS-0603761 and DMS-0904125.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

AE	A-Efficiency
BLUE	Best Linear Unbiased Estimator
CCLS	Column-Complete Latin Square
LB	Lower Bound
TBIC	Totally Balanced Incomplete Crossover Designs
TBTCI	Totally Balanced Test-Control Incomplete (Crossover Designs)
Tr	Trace
OA	Orthogonal Array
OA_I	Orthogonal Array of Type I
Ω	The class of all competing designs with the same number treatments, subjects, and periods.
Λ	A subset of Ω in which each design should allow no treatment to be preceded by itself and its control treatment is equally replicated in each period.

SUMMARY

In this dissertation, I focus on optimal crossover designs under A-criterion and MV-criterion when assuming the subject effects in the model to be random. An A-optimal design minimizes the average of the variances of the contrast estimates while A MV-optimal design minimizes the maximum of those variances. By simple arguments, the MV-efficiency of a design with completely symmetric information matrix would be at least as high as its A-efficiency. Thus, an A-optimal design with completely symmetric information matrix would also be MV-optimal design. Since symmetry is generally a desirable property, it is usually feasible to target on finding the A-optimal designs first and then check to see if its information matrix is completely symmetric, which approach is also adopted in this dissertation.

The assumption of the randomness of the subject effects could be argued in two-folds. First, most practical experiments consider subjects to be a random sample from a large population, and the main interest is to do the inference of the effects of the treatments on the whole population rather than the particular subjects in study. Second, as we know the choice of designs could be made by merely looking at the information matrix. However, mathematical derivation shows that the information matrix under random assumption would converge to the information matrix under fixed subject effect assumption when the randomness goes to infinity. Hence, not only it is more plausible in practical meaning in assuming randomness, but also the fixed subject effects model could be considered as a special case when we work on the random subject effects model.

My research in such direction is presented in five chapters.

Chapter 1 introduces the backgrounds and motivations of my research. It is written in a way such that one could easily understand it with basic concepts of statistics. By reading this chapter, one would know what do we mean by good designs, what is Crossover designs exactly and why and when do we need them, and also a review of the works done in the history.

Chapter 2 gives answers for optimal and efficient designs under A-criterion and MV-criterion. An A-optimal design minimizes the average of the variances of the estimates of parameters of interest while a MV-optimal design minimizes the maximum of those variances. I was working on the model which is not typically studied in literature in that the model assumes the subject effects to be random. As argued later in the chapter, it is not only more reasonable in practice but also covers the traditional model as a special case. In fact, results in the chapter covers the results in the literature as a special case.

Chapter 3 further considers the evaluation of the A-efficiency of any given design if it is not optimal. By A-efficiency, I mean the ratio of the A-criterion values of the optimal design and the given design. Since the optimal design is not found while we have need to evaluate the efficiency of a candidate design in general, the A-criterion value of the so called optimal design have to be replaced by some tight bound and hence the value of A-efficiency of a design is actually the lower bound of its A-efficiency. The core context of this chapter is to find the bound that is easy to calculate and at the same time very close to the A-value of the true optimal design. As for MV-efficiency, note that the designs that I proposed in Chapter 2 has their MV-efficiency no smaller than their own A-efficiency due to some symmetric structure in them.

Chapter 4 proposes some methods of constructing the totally balanced test-control incomplete (TBTCI) crossover designs. Chapter 2 has shown that such designs are very highly efficiency and robust to the change of the amount of randomness in subject effects. This motivates the studies in Chapter 4.

Chapter 5 summarizes the results of this dissertation and compared them with existing results. Some comments are made, and future problems are proposed.

As part of my PhD study, I have wrote a paper on asymptotic of sample covariances for long memory process with Professor Wei-Biao Wu and his student Yinxiao Huang, which is not included in this dissertation.

CHAPTER 1

INTRODUCTION

There are always two basic steps involved in a statistical problem, namely collecting the data and analyzing the data. In recent years, statisticians' focus are mostly on the latter. While the collection part is being less emphasized, it is equally important as the analysis part. Once the method of analyzing the data is fixed, a good strategy of collecting the data, which is called a *design*, enables us to do more precise inference under limited resource or save resource under some requirement of the precision of inference.

The linear model still retains its popularity in application not only for its simplicity but also most mechanism in reality can be satisfactorily approximated by the linear model. While people knows well how to analysis the data by assuming the linear model, the problem of selecting the design points for such model is still unsolved in many scenarios. Particularly, the problem will become more complicated when we are only interested in part of the parameters in the model, not to mention that the optimality criterion is not unique.

This thesis will focus on the problem of Crossover designs in which each subject will repeated take treatments at different periods. In such designs, the treatment taken at the previous periods will still have residual effects at the current period. In most applications in clinical trial, we are only interested in the direct effects of the treatments. Besides the residual effects, other nuisance parameters typically includes subject effects and period effects. Though complicated as a design problem, the Crossover designs is very popular in the pharmaceutical industry in comparing treatments, and hence it is worth the effort to study the problem of Crossover designs.

1.1 Optimality Criteria

Consider a simple linear model

$$\mathbb{E}Y = X\beta, \qquad Var(Y) = I \tag{1.1}$$

It is well known (Markov-Guess Theorem) that the best linear unbiased estimator (BLUE) for β is $\hat{\beta} = (X'X)^{-1}X'Y$ when $X'X$ is full rank. Here the notation $'$ indicate the transpose of a matrix or vector. The variance of the BLUE is $Var(\hat{\beta}) = (X'X)^{-1}$. If we further assume normal distribution of Y, due to Cao-Crammer Lemma, the variance matrix of any unbiased estimator (could be non-linear) of β would be greater than $(X'X)^{-1}$ in Loewner ordering since $X'X$ is the Fisher Information Matrix. Without normality assumption, we call the inverse of variance of BLUE in a linear model as *information matrix* in the rest of this thesis. Hence the information matrix for Model (1.1) is $X'X$.

Most frequently, we are interested in just part of the parameters in the model while parameters of no interest is called nuisance parameters. Now consider

$$\mathbb{E}Y = X_d\beta + Z_d\gamma, \qquad Var(Y) = I \tag{1.2}$$

where β is the parameter of interest, and γ is the nuisance parameter. The subscript d for matrices X and Z is to emphasize the taste of experimental design. The information matrix of β in Model (1.2) is then

$$C_d = X' pr^{\perp}(Z) X, \quad where \quad pr^{\perp}(Z) = I - Z(Z'Z)^{-}Z' \tag{1.3}$$

with symbol $^{-}$ indicating the generalized inverse of a matrix. Here the operator $pr^{\perp}$ is called projection. To understand this, note that a linear combination of the parameters of interest $\ell'\beta$ is estimable if and only if the vector ℓ is in the column space of the information matrix C_d. When estimable, the BLUE of $\ell'\beta$ has the variance of

$$Var(\widehat{\ell'\beta}) \quad = \quad \ell' C_d^{-} \ell \tag{1.4}$$

In the following, the hat notation $\hat{}$ always means the BLUE of the corresponding parameter. If C_d is of full rank, then the whole vector β would be estimable since the identity matrix is in the column space of C_d, and we would have

$$Var(\hat{\beta}) \quad = \quad C_d^{-1} \tag{1.5}$$

It can be seen that a design d with *larger* matrix C_d yields smaller variance of the estimate of $\boldsymbol{\tau}$, and hence the inference would be more precise. The central problem in design of experiment is to found the design d such that the information matrix C_d is as *large* as possible. However, matrices can not be directly compared in general. To make the information matrices for different designs comparable, one has to define functions of C_d, which maps from the matrix space to one dimensional space. Such functions are called criterion of optimality in the statistical society. Following are some examples of different criteria with the illustration under the case when the information matrix is of full rank. Modified but similar functions could be defined under the same name of the criterion when C_d is not of full rank.

- ◇ A-optimality: Seeks to minimize the trace of the inverse of the information matrix. This criterion results in minimizing the average variance of the estimates of the regression coefficients.
- ◇ D-optimality: Seeks to maximize $|C_d|$. This criterion results in maximizing the differential Shannon information content of the parameter estimates.
- ◇ E-optimality: Seeks to maximize the minimum eigenvalue of the information matrix.
- ◇ T-optimality: Seeks to maximize the trace of the information matrix.
- ◇ G-optimality: Seeks to minimize the maximum variance of the predicted values.
- ◇ I-optimality: Seeks to minimize the average prediction variance over the design space.
- ◇ V-optimality: Seeks to minimize the average prediction variance over a set of m specific points

Each criterion is defined for specific purpose, and hence different criteria would result in different choices of designs. On the other hand, however, these criteria have the same target of minimizing the variance of the estimates, hence there could be some occasions where there exists a design which is optimal under all these criteria. Specifically, Keifer (1975) gave the sufficient designs to guarantee the *universal* optimality of a design:

- $\diamond$ The trace of the information matrix is maximized.
- $\diamond$ The information matrix is complete symmetric.

Almost all the literatures used these two simple conditions to prove universal optimality.

1.2 Parallel designs or Crossover designs?

Typically in pharmaceutical industry, people need to do experiments in order to compare effects of different treatments. The main concern is their effects on a certain population, however an experiment could only recruit a limited number of subjects from the whole population. Hence, certain assumptions need to be made in order to infer the effectiveness of these treatments based on the data from the experiment. It is a common practice to assume that the response could be explained by an additive model, which involve effects of the treatments, subjects, other necessary factors, and finally an error term to include anything else that we can not explain.

There are basically two ways of conducting the experiment, namely Parallel design and Crossover Design. In Parallel design, each subject takes only one treatment just once and leave. Such experiment would become problematic when there are significant differences between subjects. In other words, the treatment effect would confound with the subject effect when the latter exists and is treated as fixed effect, and hence it is impossible to estimate the treatment effect. To solve the problem, one has to get the information regarding the treatment from the within-subject comparison. Hence, one need to carry out the Crossover design, in which each subject repeatedly takes treatments in a sequence of periods. The disadvantage of Crossover design is that the washout time between successive two periods are usually not long enough so that the nuisance parameter of carryover effect (effects of a treatment from previous periods) is brought into the model, and make the estimation of the direct treatment effect less efficient.

1.2.1 Introduction to Crossover designs

We denote by $\Omega_{t,n,p}$ the set of all of designs where n subjects are used in $p \geq 2$ occasions, called periods, for the purpose of evaluating and studying $t \geq 2$ treatments, usually labeled as $\{1, 2, ..., t\}$. Note that we consider designs in which each subject takes treatment at the same sequence of period. This is for technical convenience and is adopted in almost all literature on Crossover designs unless the case of dropout is considered (Low, Lewis, and Prescott (1999), Majumdar, Dean, and Lewis (2008) etc.).

For a continuous response Y, a plausible and useful linear model can be written as

$$Y_{dku} = \mu + \alpha_k + \beta_u + \tau_{d(k,u)} + \rho_{d(k-1,u)} + \varepsilon_{ku}, \tag{1.6}$$

Here, Y_{dku} denotes the response from subject u in period k to which treatment $d(k,u) \in \{1, 2, ..., t\}$ was assigned by design $d \in \Omega_{t,n,p}$ $k = 1, ..., p$, and $u = 1, ..., n$. Furthermore, μ is the general mean, α_k is the kth period effect, β_u is the uth subject effect, $\tau_{d(k,u)}$ is the (direct) treatment effect of treatment $d(k,u)$, and $\rho_{d(k-1,u)}$ is the (first-order) carryover or residual effect of treatment $d(k-1,u)$ that subject u received in the previous period (by convention

$\rho_{d(0,u)} = 0)$. Also we assume the error terms ε_{ku}'s are independent and have mean 0 and common variance σ^2. It is not necessary to make any distributional assumptions on the error terms. For example, if we have such a Crossover design:

4 4 4 2 3 1 2 3 1

1 2 3 4 4 4 1 2 3

2 3 1 1 2 3 4 4 4

Then the observation Y_{35} from the trial on subject 5 at the 3rd period could be modeled as

$$Y_{35} = \mu + \alpha_3 + \beta_5 + \tau_2 + \gamma_4 + \varepsilon_{35}$$

and we also have $n = 9, p = 3, t = 3$, $d(3,5) = 2$, and $d(2,5) = 4$. In general, writing the $np \times 1$ response vector as $Y_d = (Y_{d11}, Y_{d21}, ...Y_{dp1}, Y_{d12}, ..., Y_{dpn})'$, we have

$$Y_d \quad = \quad 1_{np}\mu + P\boldsymbol{\alpha} + U\boldsymbol{\beta} + T_d\boldsymbol{\tau} + F_d\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \tag{1.7}$$

where $\boldsymbol{\tau} = (\tau_1, ..., \tau_t)'$ contains the direct treatment effects, which is of interest. However, $\boldsymbol{\gamma} = (\gamma_1, ..., \gamma_t)'$, $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_p)'$, $\boldsymbol{\beta} = (\beta_1, ..., \beta_n)'$ are nuisance parameters which represents the carryover, period, and subject effects. Here $P = 1_n \otimes I_p$, $U = I_p \otimes 1_p$, and T_d and F_d denote the treatment/subject and carryover/subject incidence matrices. The notation $\otimes$ represents the Kronecker product; 1_s represents the column vector of length s with all its entries as 1; I_s represents the $s \times s$ identity matrix.

There are discussions on whether the subject effects should be assumed to be random or fixed effects. A heuristic argument is that: If we are interested in drawing the information about the effectiveness of the treatment on the whole population, and the subjects in study are believed to be randomly drawn from the whole population, we need to assume subject effects to be random; On the other hand, if there are limited level of subject effects and all such levels are included in the study, it is reasonable to assume subject effects to be fixed. Obviously, assuming subject effects to be random is more applicable in most applications.

Now let us assume subject effects β_u's are independent and have zero mean and a common variance σ_β^2, also they are independent with the error terms. Writing the $np \times 1$ response vector as $Y_d = (Y_{d11}, Y_{d21}, ...Y_{dp1}, Y_{d12}, ..., Y_{dpn})'$, we have

$$E(Y_d) = 1_{np}\mu + P\boldsymbol{\pi} + T_d\boldsymbol{\tau} + F_d\boldsymbol{\rho}, \qquad Var(Y_d) = \sigma^2 V, \tag{1.8}$$

where $V = I_n \otimes (I_p + \theta J_p)$ with $\theta = \sigma_\beta^2/\sigma^2$ and $J_s = 1_s 1_s'$. Note that θ represents the relative size of the variation between subjects as compared to the variation between the error terms. The general argument in the beginning of this section indicate that the advantage of Crossover designs over the parallel designs would be more obvious when θ is larger, which will also be validated by mathematical derivation later on.

The parameter of interest $\boldsymbol{\tau}$ is not estimable since each component of it confounds with the general mean μ. However, for any vector ℓ of length t satisfying $\ell' 1_t = 0$, $\ell'\boldsymbol{\tau}$ would be estimable, and its best linear unbiased estimator (BLUE) solved from the normal equation (Guess-Markov Theorem) would have the variance of

$$Var(\widehat{\ell'\boldsymbol{\tau}}) \quad = \quad \sigma_\epsilon^2 \ell' C_d^- \ell, \tag{1.9}$$

The matrix C_d in (1.9) is called the information matrix for $\boldsymbol{\tau}$ and is of the form

$$C_d = T_d' V^{-1/2} pr^\perp (V^{-1/2}[1_{np}|P|F_d]) V^{-1/2} T_d, \tag{1.10}$$

where $pr^\perp$ is the projection operator as in (1.3) and $[\cdot|\cdot|\cdot]$ represents the juxtaposition of three matrices of the same number of rows.

1.2.2 When to use Crossover designs

There are basically two types of circumstances in which Crossover designs is preferred against Parallel designs.

1. Practical concern: Experimental subjects are scarce or expensive to recruit and they have to be used repeatedly.
2. Statistical concern: Parallel designs become very inefficient for estimating treatment effects when the differences among subjects dominate the observations.

The argument of practical concern is straightforward. Hedayat and Afsarinejad (1978) mentioned examples as small clinics, large military systems, and some experiments where the subjects need to be trained over a long period of time. In the following, I would like to illustrate the second point by comparing the information matrices for these two type of designs. To make the comparison fair enough, we need to make the same assumptions for both design. Hence, the period and subject effects should also exist for Parallel designs. Also, for any given Crossover design d, we would compare it with the Parallel design which considers each sequence of p treatments by a subject in the corresponding Crossover design being taken by p different subjects. For the Parallel design, we have the model

$$Y_{dku} = \mu + \alpha_k + \beta_{ku} + \tau_{d(k,u)} + \varepsilon_{ku}, \tag{1.11}$$

As in (1.8), we assume β_{ku}'s to be independent random variables with mean zero and variance σ_β^2. Then the information matrix for $\boldsymbol{\tau}$ is

$$C_{d,p} = T_d' pr^{\perp}(P) T_d/(1+\theta) \tag{1.12}$$

as compared to (1.10) which is the information matrix for the corresponding Crossover design. As measure of variation between subjects θ approaches zero and ∞, we have

$$\begin{aligned}
\lim_{\theta\to 0} C_d &= T_d' pr^{\perp}([1_{np}|P|F_d])T_d & (1.13)\\
\lim_{\theta\to\infty} C_d &= T_d' pr^{\perp}([1_{np}|U|P|F_d])T_d & (1.14)\\
\lim_{\theta\to 0} C_{d,p} &= T_d' pr^{\perp}(P)T_d & (1.15)\\
\lim_{\theta\to\infty} C_{d,p} &= 0 & (1.16)
\end{aligned}$$

The right hand of (1.13) is exactly the information matrix for Crossover design when there are no subject effects in the model, while the right hand of (1.14) is exactly the information matrix for Crossover design when the subject effects are treated as fixed effects in the model. By comparing the equations from (1.13) to (1.16), we have $C_d \geq C_{d,p}$ for large θ and $C_d \leq C_{d,p}$ for small θ. The inequalities are in Loewner's ordering. For example, $A \leq B$ means $B - A$ is a non-negative definite matrix.

As a general guideline, if we assume the same cost for assigning a sequence of treatments to one subject or many subjects, the Crossover design is better than parallel design when there are large variations between subjects. As to quantify the exact condition for when to use Crossover designs, we need further studies under certain assumptions of the cost.

1.3 Literature Survey

The idea of using Crossover designs to estimate the effects of treatments has a long history. The earliest example in document can be traced back to 1853 when John Bennett Lawes and Baron Justus von Liebig has disagreement on deciding which manure of two is more effective for the yield of crops (Jones and Kenward, 2003). Cochran (1939) seems to be the first to formally separate out the two sorts of treatment effects (direct and carry-over). Other famous works during this early period include, but not limited to, Simpson (1938), Yates (1938), Brandt (1938), Fieller (1940), Cochran et al. (1941), Finney (1956), Outhwaite (1955,1956), and Sampford (1957). Particularly, one significant work of mathematical state was by Williams (1949,1950) which showed how balanced designs which used the minimum number of subjects could be constructed.

The area of Crossover designs started to receive more attentions from statisticians when Hedayat and Afsarinejad (1975,1978) systematically reviewed and studied optimal Crossover designs. They proved the universal optimality of uniform balanced designs for estimating both direct and carryover treatment effects among uniform designs. Later on, Cheng and Wu (1980) proved the universal optimality of strongly balanced uniform designs and its variants in the whole class for both direct and carryover effects. They also showed that balanced uniform designs are universal optimal in a subclass in which no treatment is allowed to be immediately preceded by itself and other conditions. Kunert (1984) was the first to prove the universal optimality of balanced uniform design among the whole class of designs with the same number of subjects n, periods p and treatments t when $n = p = t \geq 3$ or $n/2 = p = t \geq 6$. Hedayat and Yang (2003) extended the universal optimality of balanced uniform designs to the case of $p = t > 2$ and $n = \lambda t$ with $\lambda \leq (t-1)/2$.

Besides the line of research on (strongly) balanced uniform designs. Other designs are also proved to be universally optimal. For example, Kunert (1983) proposed Generalized Youden designs with m_{dij} satisfying some equations; Kunert and Martin (2000) proved that type I orthogonal arrays are universal optimal among binary designs under any within subject correlation. For the special case of two periods, Hedayat and Zhao (1990) gave interesting and comprehensive answers. Also, for the special case of two treatments, there is only one contrast to estimate, and hence all criteria result in the same choices of designs. More specifically, the information matrix C_d is 2×2, and it has explicit form of $C_d = Tr(C_d)B_2$ where $B_2 = pr^{\perp}(1_2)$. The problem in this scenario reduces to finding a design which maximizes the trace of the information matrix. For details, please refer to Kunert (1991) and Kunert and Stufken (2008). The latter studied the more complex model which divide carryover effect into two categories, namely mixed and self-carryover effects. For the latter model, Kunert and Stufken (2002) considered the problem of at least three treatments.

Note that all the above work were assuming subject effect to be fixed effects. However, subjects in the study may often be thought of as representing a larger population of interest from which they are, more or less, randomly selected. More than that, the information matrix under the random subject effect model actually covers the model with fixed subject effects as a special case. For more details, please refer to Section 1.2.2. There are relatively few optimality results for this direction. Mukhopadhyay and Saha (1983) showed that some of the optimality results by Hedayat and Afsarinejad (1978), Magda (1980), and Cheng and Wu (1980) remain valid when the subject effects are assumed to be random. Jones, Kunert, and Wynn (1992) obtained additional results for a similar setup in which the carryover effects are assume to be random. However, the number of periods in these results is at least equal to the number of treatments, and some of the results are, in addition, over restricted classes of designs. For example, Laska and Meisner (1985) obtained optimal two-treatment Crossover designs, given arbitrary within-subject covariance. Carrière and Reinsel (1993) showed that strongly balanced two-period designs that are uniform on the periods are universally optimal for treatment effects in the entire class of designs. This holds true also when subject effects are fixed, as noted by Hedayat and Zhao (1990).

Recently Hedayat, Stufken, and Yang (2006) found universal optimal designs in restricted subclass where the last period has the same number of replications for each treatment, and showed their high efficiency under the criterion of $Tr(C_d)$. The restriction is not fair when we have a control treatment in the study and want to compare each of the test treatments with the control treatment. For the latter case, it is reasonable to find an A-optimal design which minimizes $\max_{i=1,\ldots,t} Var_d(\widehat{\tau_i - \tau_0})$ where τ_0 represents the control treatment and the notation $\hat{}$ represents the best linear unbiased estimator (BLUE). The efficiency measured by $Tr(C_d)$ in Hedayat, Stufken, and Yang (2006) hence could not reflect the A-efficiency. On the other hand, however, Hedayat and Yang (2005,2006) have been working on the A-efficient designs for the model with fixed subject effects. This thesis will discuss A-efficient designs under the model with random subject effects.

One major difficulty in finding the optimal design is that the design space is of discrete type. To understand this, let us consider a Crossover design as selecting n sequences with replacement from all possible $N = t^p$ sequences. If we label these sequences by index $1, 2, \ldots, N$ and use x_i to denote the number replications of the ith sequence appearing in the design. Then a design could be completely characterized by the vector $x = (x_1, x_2, \ldots, x_N)$. Any criterion can thus be written as a function of the vector x, say $\phi(x)$. Then the problem of finding the optimal Crossover design under such criterion would be the same as to optimize $\phi(x)$ under the restrictions of (i) $\sum_{i=1}^{N} x_i = n$. (ii) $x_i \geq 0$. (iii) x_i's are all integers. It is easy to see that the major difficulty come from Restriction (iii). The problem of optimizing $\phi(x)$ under (i) and (ii) only is called asymptotic designs problem. Kushner (1997) gave nice answers for the traditional model with subject effects assumed to be fixed. Later, Kunert and Martin (2000) solved the problem for an interference model, and Kunert and Stufken (2002) solved the problem when the carryover effects are divided into mixed and self-carryover effects. Again both of these two papers assume subject effects to be fixed. Such results in general does not answer what design is optimal in exact design theory, however, they at least provide an idea which could be modified to evaluate the efficiency of the designs that we propose under the random subject effects model. The details will be illustrated in Chapter 3.

CHAPTER 2

OPTIMAL AND EFFICIENT DESIGNS FOR A-CRITERION WHEN SUBJECT EFFECTS ARE RANDOM

2.1 The Model and Notations

In this chapter, we will study optimal and efficient designs for A-Criterion under the model of

$$Y_{dku} = \mu + \alpha_k + \beta_u + \tau_{d(k,u)} + \rho_{d(k-1,u)} + \varepsilon_{ku}, \tag{2.1}$$

such that the subject effects are assumed to be random. More specifically, β_u's are independent with zero mean and a constant variance σ_β^2. If σ^2 is the variance of the error term, $\theta = \sigma_\beta^2/\sigma^2$ measures the relative size of the variation between subjects as compared to the variance of the error term. When θ is larger, the advantage of Crossover designs against parallel designs will be more obvious. We are interested in the case when there is a control treatment and we want to compare two or more test treatments with the control. Without loss of generality, we use τ_0 to denote the effect of control treatment, and $\tau_1, \tau_2, \ldots, \tau_t$ to denote the t test treatments. Hence, the number of total treatments will be $t+1$ instead of t as we were using in previous chapter. We will still use n and p to denote the number of subjects and periods, and hence we have $1 \leq u \leq n$, $1 \leq k \leq p$, and $0 \leq d(u,k) \leq t$ as in Model (2.1). Writing the $np \times 1$ response vector as $Y_d = (Y_{d11}, Y_{d21}, ...Y_{dp1}, Y_{d12}, ..., Y_{dpn})'$, we have

$$E(Y_d) = 1_{np}\mu + P\boldsymbol{\alpha} + T_d\boldsymbol{\tau} + F_d\boldsymbol{\gamma}, \quad var(Y_d) = \sigma^2 V, \tag{2.2}$$

where $V = I_n \otimes (I_p + \theta J_p)$. Here $\boldsymbol{\beta} = (\beta_0, ..., \beta_t)'$ and $\boldsymbol{\gamma} = (\gamma_0, ..., \gamma_t)'$. All other notations are consistent with that of Section 1.2.1. The information matrix for $\boldsymbol{\tau} = (\tau_0, \ldots, \tau_t)$ would be

$$C_d = T_d' V^{-1/2} pr^{\perp}(V^{-1/2}[1_{np}|P|F_d]) V^{-1/2} T_d, \tag{2.3}$$

where $pr^{\perp}$ is a projection operator such that $pr^{\perp}A = I - A(A'A)^{-}A'$ for any matrix A.

Throughout the thesis, for each design d, we adopt the notation $n_{diu}, \tilde{n}_{diu}, l_{dik}, m_{dij}, r_{di}, \tilde{r}_{di}$, to denote the number of times that treatment i is assigned to subject u, the number of times this happens in the first $p-1$ periods associated with subject u, the number of times treatment i is assigned to period k, the number of times treatment i is immediately preceded by treatment j, the total replication of treatment i in the n experimental subjects, and the total replication of treatment i limited to the first $p-1$ periods in the design. Also, we would like to define the subclass of designs

$$\Lambda_{t+1,n,p} = \{d \in \Omega_{t+1,n,p} | l_{d0k} = r_{d0}/p, k = 1, ..., p \text{ and } m_{dii} = 0, i = 0, 1, ..., t\}.$$

Therefore, for any design in Λ, the control treatment appears equally often in all periods and no treatment is allowed to be preceded by itself. Additionally, we have the following convention: For any two square matrices (e.g. A, B) of the same size, the inequality $A \leq B$ represents the Loewner's ordering of the matrices, namely $B - A$ is a non-negative definite matrix; $Tr(A)$ represents the trace of the matrix A; $[x]$ represents the greatest integer that is not greater than x; $B_s = pr^{\perp}1_s = I_s - J_s/s$; $\{S_i, i = 1, 2, ..., t!\}$ is the set of all $t \times t$ permutation matrices and

$$\tilde{S}_i = \begin{pmatrix} 1 & 0_{1\times t} \\ 0_{t\times 1} & S_i \end{pmatrix}, \quad i = 1, ..., t!.$$

2.2 A- and MV- Criteria

Section 1.1 has introduced the concept of A-criterion when the information matrix if of full rank, that is, all component of parameter of interest is estimable. For a Crossover design with Model 1.6, the parameter vector of interest $\boldsymbol{\tau}$ can not be estimable for each component since each of them confound with the general mean. However, any contrast of the τ_i's would be estimable. By contrast, I mean $\ell'\boldsymbol{\tau}$ with $\ell' 1 = 0$. For the circumstance when we have a control treatment and want to compare two or more test treatments with the control, the contrasts $\tau_i - \tau_0, 1 \leq i \leq t$ is essentially parameters of interest instead of $\boldsymbol{\tau}$ itself. Hence, we can define the A-optimality and MV-optimality and the corresponding efficiency as follows:

Definition 1. (i) *In a class of competing designs, a design is said to be A-optimal if it minimizes* $\sum_{i=1}^{t} Var_d(\hat{\tau}_i - \hat{\tau}_0)$*, where* $\hat{\boldsymbol{\tau}} = (\hat{\tau}_0, \hat{\tau}_1, ..., \hat{\tau}_t)'$ *is the generalized least square estimate of* $\boldsymbol{\tau}$*.*

(ii) *A design is said to be MV-optimal if it minimizes* $\max_{i=1,...,t} Var_d(\hat{\tau}_i - \hat{\tau}_0)$*.*

Definition 2. *For a criterion $\phi(d)$ with preference of smaller values, we define the efficiency of a design to be $\min_d\{\phi(d)\}/\phi(d)$ where the minimum is taken over the completing class of designs.*

Since the row and column sums of C_d are both 0, we could express C_d as

$$C_d = \begin{pmatrix} 1_t' M_d 1_t & -1_t' M_d \\ -M_d 1_t & M_d \end{pmatrix}. \tag{2.4}$$

The $t \times t$ submatrix M_d is closely related to the A- and MV- optimality. let $H = (-1_t|I_t)$, then $var(H\hat{\boldsymbol{\tau}}) = \sigma^2 H C_d^- H' = \sigma^2 M_d^{-1}$, since one could choose $diag(0, M_d^{-1})$ as a generalized inverse of C_d in view of (2.4). Hence, we have the following lemmas to deal with A- and MV- Criteria.

Lemma 1. *A design d^* is A-optimal if it minimizes $Tr(M_d^{-1})$.*

Lemma 2. *If the matrix M_d is completely symmetric for a design d, then MV-efficiency of it is no less than the A-efficiency of itself. Here A- and MV- efficiencies are defined by Definition 2.*

Proof. Let ϕ_A and ϕ_{MV} to denote the A- and MV- criteria, and let d^A and d^{MV} to denote the optimal designs under A- and MV- Criteria. By simple calculation, we have $\phi_{MV}(d) \geq \phi_A(d)/t$ with the equality holds for designs with completely symmetric M_d. Then we have

$$\begin{aligned} \frac{\phi_{MV}(d^{MV})}{\phi_{MV}(d)} &\geq \frac{\phi_A(d^{MV})}{\phi_A(d)} \\ &\geq \frac{\phi_A(d^A)}{\phi_A(d)} \end{aligned}$$

$\diamondsuit$

Corollary 1. *If a design d is A-optimal and the matrix M_d is completely symmetric, then d is also MV-optimal.*

2.3 Optimal and Efficient Designs

Lemma 2 and Corollary 1 show a nice shortcut for deriving MV-efficient or even optimal designs once knowledge about A-efficient or optimal designs is given. Since M_d contains all the information needed to evaluate the A- and MV- optimality of a design and M_d in turn is a submatrix of C_d by ignoring the first row and the first column, designs with the same information matrix C_d should be equivalent in the A- and MV- sense. On the other hand, the matrix C_d is a function of the unknown variable θ, hence the determination of optimal designs depends on the value of θ. As pointed out by Hedayat, Stufken, and Yang (2006), two extreme cases are worth mentioning. The case of $\theta = 0$ corresponds to the situation of no subject effect. It is easily seen that

$$\lim_{\theta \to 0} C_d = T_d' pr^{\perp}([1_{np}|P|F_d])T_d,$$

which would indeed be precisely the information matrix for $\boldsymbol{\tau}$ if we were to ignore the subject effect. Under this special case, Theorem 1 below gives the optimal designs in Ω. The other extreme case corresponds to $\theta = \infty$, and we have

$$\lim_{\theta \to \infty} C_d = T_d' pr^{\perp}([1_{np}|P|U|F_d])T_d,$$

where $U = I_n \otimes 1_p$. This limit is precisely the information matrix that we would have obtained had we treated the subject effects as fixed. Under this special case, Hedayat and Yang (2005) gave the optimal designs in the subclass Λ, and they conjectured that optimal designs in Λ is still highly efficient in Ω. In this thesis, I derive optimal designs in Λ for any value of θ, which covers their result as a special case. We also give the explicit way of evaluating the efficiency of any design for any value of θ.

2.3.1 Special Case: No Subject Effects

Theorem 1. *If $\theta = 0$, then for any n, t, p, a design d is simultaneously A- and MV- optimal in $\Omega_{n,t,p}$ if it satisfies*

1. $m_{dii} = r_{di}\tilde{r}_{di}/pn, i = 0, 1, ..., t.$
2. $m_{dij}, m_{d0i}, m_{di0}$ *are constants across all* $1 \leq i \neq j \leq t$.
3. $\ell_{dik_1} = \ell_{dik_2}, i = 0, 1, ..., t, k_1, k_2 = 1, 2, ..., p.$
4. $r_{di} = r_{dj}, i, j = 1, 2, ..., t.$
5. $r_{d0} = \arg\min_{h \in \{1,2,...,np\}}(t/(np - h) + 1/h).$

Condition 5 is equivalent to $r_{d0} = np/(1 + \sqrt{t})$ when the latter is an integer.

We would like to give an intuitive explanation for the conditions imposed in Theorem 1. Conditions 2-4 impose some structure of symmetry among test treatments and periods. Condition 5 directly determines the replication of each treatment in conjunction with Condition 4. Finally, Condition 1 requires the exact relationship between two type of variables, which is too strong in general. When $n = 18, t = 4, p = 3$, we need $r_{d0} = np/(1 + \sqrt{t}) = 18$. Under this situation, the design below constructed by Yang and Stufken (2008) is optimal in $\Omega_{18,4,3}$ when $\theta = 0$:

$$d_1 : \begin{array}{cccccccccccccccccc} 1&1&1&2&2&2&3&3&3&4&4&4&0&0&0&0&0&0 \\ 1&2&3&2&3&4&3&4&0&4&0&0&0&0&1&0&1&2 \\ 4&0&0&0&1&2&2&1&2&3&4&0&3&3&0&4&0&1 \end{array}$$

2.3.2 General Case: Random Subject Effects

For general θ, it is hard to compare all designs in Ω. However, we succeeded in finding optimal designs in the subclass Λ. In order to introduce the result, it is necessary to give the definition of a class of designs proposed by Hedayat and Yang (2005).

Definition 3. *A design d is said to be a TBTCI design if:*

1. $m_{dii} = 0$ *for all* $0 \le i \le t$.
2. m_{d0i}, m_{di0} *and* m_{dij} *are constants across all* $1 \le i \ne j \le t$
3. $\ell_{dik_1} = \ell_{dik_2}, i = 0, 1, ..., t, k_1, k_2 = 1, 2, ..., p$.
4. $r_{di} = r_{dj}, i, j = 1, 2, ..., t$.
5. $n_{diu} = 0$ *or* 1 *for* $i = 1, 2, ..., t, u = 1, 2, ..., n$.
6. $|n_{d0u} - n_{d0v}| \le 1$ *and* $|\tilde{n}_{d0u} - \tilde{n}_{d0v}| \le 1$ *for all* $1 \le u, v \le n$.
7. $\sum_{u=1}^n n_{d0u} n_{diu}, \sum_{u=1}^n n_{diu} n_{dju}, \sum_{u=1}^n \tilde{n}_{d0u} \tilde{n}_{diu}, \sum_{u=1}^n \tilde{n}_{diu} \tilde{n}_{dju}, \sum_{u=1}^n n_{d0u} \tilde{n}_{diu}, \sum_{u=1}^n \tilde{n}_{d0u} n_{diu},$ *and* $\sum_{u=1}^n n_{diu} \tilde{n}_{dju}$, *are constants across all* $1 \le i \ne j \le t$.

Theorem 2. *(i)Suppose $p \geq 3$ and $t \geq \max(p-1, 3)$, then for any value of θ a design d is simultaneously A-optimal and MV-optimal in $\Lambda_{t+1,n,p}$ if d is a TBTCI design and*

$$r_{d0} = \operatorname*{argmin}_{h \in \{1,2,\ldots,np-1\}} f(h, \theta), \tag{2.5}$$

where

$$f(r_{d0}, \theta) = t(t-1)^2(\alpha_1 - \beta_1^2/\gamma_1)^{-1} + t(\alpha_2 - \beta_2^2/\gamma_2)^{-1},$$

$$\begin{aligned}
\alpha_1 =& t(1-\lambda_p)(np - r_{d0}) - \eta(np - r_{d0})^2 - r_{d0} + \lambda_p\chi_1 + \eta r_{d0}^2, \\
\beta_1 =& \lambda_p t(n(p-1) - \tilde{r}_{d0}) + \eta(n(p-1) - \tilde{r}_{d0})(np - r_{d0}) - \lambda_p\chi_2 - \eta r_{d0}\tilde{r}_{d0}, \\
\gamma_1 =& (t + 1 - 2/p - \lambda_p t)(n(p-1) - \tilde{r}_{d0}) - n(p-1)^2/p \\
& - \eta(n(p-1) - \tilde{r}_{d0})^2 + \eta\tilde{r}_{d0}^2 + \lambda_p\chi_3, \\
\alpha_2 =& r_{d0} - \lambda_p\chi_1 - \eta r_{d0}^2, \\
\beta_2 =& \lambda_p\chi_2 + \eta r_{d0}\tilde{r}_{d0}, \\
\gamma_2 =& \tilde{r}_{d0} - (np^2 - np)^{-1}\tilde{r}_{d0}^2 - \lambda_p\chi_3 - \eta\tilde{r}_{d0}^2,
\end{aligned}$$

$$
\begin{aligned}
\lambda_p =& \theta(1+\theta p)^{-1}, \text{ and } \eta = \lambda_p(\theta p n)^{-1}, \\
\chi_1 =& r_{d0} + (2r_{d0} - n)\lfloor r_{d0}/n \rfloor - n\lfloor r_{d0}/n \rfloor^2, \\
\chi_2 =& \{\tilde{r}_{d0} + (r_{d0} + \tilde{r}_{d0} - n)\lfloor \tilde{r}_{d0}/n \rfloor - n\lfloor \tilde{r}_{d0}/n \rfloor^2\} \boldsymbol{1}_{\{r_{d0}/n - \lfloor \tilde{r}_{d0}/n \rfloor < 1\}} \\
& + \{r_{d0} + \tilde{r}_{d0} - n + (r_{d0} + \tilde{r}_{d0} - 2n)\lfloor \tilde{r}_{d0} \rfloor - n\lfloor \tilde{r}_{d0}/n \rfloor^2\} \boldsymbol{1}_{\{r_{d0}/n - \lfloor \tilde{r}_{d0}/n \rfloor \geq 1\}}, \\
\chi_3 =& \tilde{r}_{d0} + (2\tilde{r}_{d0} - n)\lfloor \tilde{r}_{d0}/n \rfloor - n\lfloor \tilde{r}_{d0}/n \rfloor^2 \\
\tilde{r}_{d0} =& (p-1)r_{d0}/p.
\end{aligned}
$$

(ii) Suppose $p = 3$ and $t = 2$, then the conclusion in (i) is still valid if we change the class of competing designs from $\Lambda_{t+1,n,p}$ to $\{d \in \Lambda_{t+1,n,p} | r_{d^*0}/n \geq 0.6306\}$.

Theorem 2 indicates that θ influences the determination of optimal designs in Λ by deciding the value of r_{d0}. Based on this r_{d0}, we could try to find a TBTCI design. Also note that Equation (2.5) is analogues to Condition 5 of Theorem 1. Now let us compare the 7 conditions in Definition 3 with Conditions 1-4 in Theorem 1. Conditions 2-4 in Definition 3 are identical to Conditions 2-4 in Theorem 1. Conditions 5-7 in Definition 3 are about symmetry among subjects, and their association with treatments. Theorem 1 does not impose these conditions since $\theta = 0$ corresponds to the model with no subject effects. Observe that Condition 1 in Definition 3 directly contradicts with Condition 1 in Theorem 1, which indicates that TBTCI designs can not be optimal when $\theta = 0$. However, TBTCI designs satisfying Equation (2.5) are robust and highly efficient for all values of θ, which will be illustrated in the next section.

Moreover, TBTCI designs exist more commonly than the designs in Theorem 1. Hedayat and Zheng (2010) gave different methods of constructing TBTCI designs.

2.4 Main Proofs

Lemma 1 gives an explicit relationship between the optimality criteria and the information matrix C_d. In order to find optimal designs, we first need to find a function $\ell_0(\theta)$ such that

$$Tr(M_d(\theta)^{-1}) \geq \ell_0(\theta) \tag{2.6}$$

for any (d,θ). If at the same time we have $Tr(M_{d^*}(\theta^*)^{-1}) = \ell_0(\theta^*)$, then the design d^* would be optimal when $\theta = \theta^*$. To establish (2.6), we can start with maximizing C_d in the Loewner's sense.

Lemma 3. *For any design d, we have*

$$C_d \leq T_d' V^{-1/2} pr^{\perp}(1_{np}|V^{-1/2}F_d)V^{-1/2}T_d. \tag{2.7}$$

The equality in (2.7) holds for any design d in which $l_{dik} = r_{di}/p, i = 0,...,t$.

Proof. It is sufficient to prove $T_d'V^{-1/2}pr^{\perp}(1_{np}|V^{-1/2}F_d)V^{-1/2}P = 0$ when $l_{dik} = r_{di}/p, i = 0,...,t$. Let

$$A = 1_n \otimes \begin{pmatrix} 0 & 0 \\ 0 & B_{p-1} \end{pmatrix}$$

Then, we have $pr^{\perp}(1_{np}|V^{-1/2}F_d)V^{-1/2}(P-A) = 0$ and $1_{np}'V^{-1/2}A = 0$. Plus, $l_{dik} = r_{di}/p, i = 0,...,t$ implies $T_d'V^{-1}A = 0$ and $T_d'V^{-1}A = 0$. The lemma is established. $\diamondsuit$

Lemma 4. *(i) For any design $d \in \Omega_{t+1,n,p}$, we have*

$$Tr(M_d^{-1}) \geq \frac{t(t-1)^2}{x_0} + \frac{t}{y_0}, \tag{2.8}$$

where

$$x_0 = \alpha_1 - \frac{\beta_1^2}{\gamma_1},$$

and

$$\begin{aligned} y_0 &= r_{d0} - \frac{\theta}{1+\theta p}\sum_{u=1}^{n} n_{d0u}^2 - \frac{r_{d0}^2}{(1+\theta p)pn} \\ &\quad - \left\{ (n(p-1) - \tilde{r}_{d0}) \left(m_{d00} - \frac{\theta}{1+\theta p}\sum_{u=1}^{n} n_{d0u}\tilde{n}_{d0u} - \frac{1}{(1+\theta p)pn} r_{d0}\tilde{r}_{d0} \right)^2 \right. \\ &\quad \left. + \tilde{r}_{d0} \left(\frac{r_{d0}}{p} - l_{d01} - m_{d00} + \frac{\theta}{1+\theta p}\sum_{u=1}^{n} n_{d0u}\tilde{n}_{d0u} + \frac{1}{(1+\theta p)pn} r_{d0}\tilde{r}_{d0} \right)^2 \right\} \\ &\quad \times \left\{ n(p-1) \left(\tilde{r}_{d0} - \frac{\theta}{1+\theta p}\sum_{u=1}^{n} \tilde{n}_{d0u}^2 - \frac{\tilde{r}_{d0}^2}{(1+\theta p)pn} \right) - \frac{\tilde{r}_{d0}^2}{p} \right\}^{-1}. \end{aligned}$$

with

$$\alpha_1 = t(1-\frac{\theta}{1+\theta p})(np-r_{d0}) - \frac{t}{(1+\theta p)pn}\sum_{i=1}^{t} r_{di}^2 - r_{d0} + \frac{\theta}{1+\theta p}\sum_{u=1}^{n} n_{d0u}^2 + \frac{r_{d0}^2}{(1+\theta p)pn},$$

$$\beta_1 = \frac{\theta t}{1+\theta p}\sum_{i=1}^{t}\sum_{u=1}^{n} n_{diu}\tilde{n}_{diu} + \frac{t}{(1+\theta p)pn}\sum_{i=1}^{t} r_{di}\tilde{r}_{di} - \frac{\theta}{1+\theta p}\sum_{u=1}^{n} n_{d0u}\tilde{n}_{d0u} - \frac{r_{d0}\tilde{r}_{d0}}{(1+\theta p)pn},$$

$$- t\sum_{i=1}^{t} m_{dii} - \frac{r_{d0}}{p} + l_{d01} + m_{d00}$$

$$\gamma_1 = (t+1-\frac{2}{p}-\frac{\theta t}{1+\theta p})(n(p-1)-\tilde{r}_{d0}) - \frac{n}{p}(p-1)^2 - \frac{t}{(1+\theta p)pn}\sum_{i=1}^{t}\tilde{r}_{di}^2,$$

$$+ \frac{\tilde{r}_{d0}^2}{(1+\theta p)pn} + \frac{\theta}{1+\theta p}\sum_{u=1}^{n}\tilde{n}_{d0u}^2,$$

Further, the equality in (2.8) holds when the following three conditions hold:

1. n_{diu} *is either 0 or 1,* $1 \le i \le t, 0 \le u \le n$
2. $l_{dik} = r_{di}/p, i = 0, ..., t$
3. $T_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} T_d, T_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} F_d$, *and* $F_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} F_d$ *are invariant under any permutation of test treatments.*

(ii) For any design $d \in \Omega_{t+1,n,p}$ *in which* $\frac{r_{d0}}{p} - l_{d01} = m_{dii} = 0, 0 \le i \le t$*, we have (2.8) with* α_1 *and* γ_1 *therein keep unchanged, however* β_1 *and* y_0 *have the following simpler forms*

$$\beta_1 = \frac{\theta t}{1+\theta p}\sum_{i=1}^{t}\sum_{u=1}^{n} n_{diu}\tilde{n}_{diu} + \frac{t}{(1+\theta p)pn}\sum_{i=1}^{t} r_{di}\tilde{r}_{di} - \frac{\theta}{1+\theta p}\sum_{u=1}^{n} n_{d0u}\tilde{n}_{d0u} - \frac{r_{d0}\tilde{r}_{d0}}{(1+\theta p)pn},$$

$$
\begin{aligned}
y_0 =& \alpha_2 - \frac{\beta_2^2}{\gamma_2} \\
\alpha_2 =& r_{d0} - \frac{\theta}{1+\theta p}\sum_{u=1}^{n} n_{d0u}^2 - \frac{r_{d0}^2}{(1+\theta p)pn}, \\
\beta_2 =& \frac{\theta}{1+\theta p}\sum_{u=1}^{n} n_{d0u}\tilde{n}_{d0u} + \frac{r_{d0}\tilde{r}_{d0}}{(1+\theta p)pn}, \\
\gamma_2 =& \tilde{r}_{d0} - \frac{\tilde{r}_{d0}^2}{(p-1)pn} - \frac{\theta}{1+\theta p}\sum_{u=1}^{n} \tilde{n}_{d0u}^2 - \frac{\tilde{r}_{d0}^2}{(1+\theta p)pn}.
\end{aligned}
$$

The equality in (2.8) still holds under the same conditions.

The proof of Lemma 4 is similar to Lemma 4 of Hedayat and Yang (2005), and we postpone it to the Supplemental Materials. The merit of Lemma 4 is that we settle down with a very special subclass of designs such that the A-criterion of designs therein are easy to be calculated directly and at the same time the optimal design is guaranteed to be included in this subclass. For the special case of $\theta = 0$, we are ready to prove Theorem 1 in Section **??**.

Proof of Theorem 1. By Lemma 4 (i), we have $Tr(M_d^{-1}) \geq t(t-1)/(np - r_{d0}) + tnp/(r_{d0}(np - r_{d0})) = t/r_{d0} + t^2/(np - r_{d0})$ for any design with the equality holds when the design satisfies Conditions 1-4. Hence the A-optimality is established. Furthermore these conditions are sufficient for M_d to be completely symmetric. $\diamondsuit$

For general value of θ, we first need the following 4 preliminary lemmas:

Lemma 5. *For any $d \in \Lambda_{t+1,n,p}$, we have $r_{d0} \leq p[\frac{n}{2}], \tilde{r}_{d0} \leq (p-1)[\frac{n}{2}]$ and $l_{d0k} \leq [\frac{n}{2}]$.*

Proof. If $l_{d0k} > [\frac{n}{2}]$, it will conflict with the condition that $m_{d00} = 0$. The other two inequalities follow immediately by noting that $l_{d0k} = l_{d0k'}$ for any $1 \leq k \neq k' \leq p$ $\diamondsuit$

Lemma 6. *For any design $d \in \Lambda_{t+1,n,p}$, we have*

$$\sum_{u=1}^{n} n_{d0u}\tilde{n}_{d0u} \leq \frac{1}{4}np(p-1). \tag{2.9}$$

Consequently, we have $\sum_{u=1}^{n} n_{d0u}\tilde{n}_{d0u} \leq t(n(p-1)-\tilde{r}_{d0})$ as long as $2t \geq p$.

Proof. To maximize $\sum_{u=1}^{n} n_{d0u}\tilde{n}_{d0u}$, it is necessary for r_{d0} to attain its maximum, which is $p[\frac{n}{2}]$ by Lemma 5. (i) When n is an even number, then $r_{d0} = np/2$ and the distribution of control treatment in the design fall into one type: $n/2$ of the subjects take the control treatment at all even periods; the remaining half of the subjects take the control treatment at all odd periods. Then the inequality trivially holds. (ii) When n is an odd number, the maximum of $\sum_{u=1}^{n} n_{d0u}\tilde{n}_{d0u}$ will be attained when one of the subjects does not take the control treatment while the remaining $n-1$ subjects take the control treatment in the same way as in (i). To see this, suppose Subject 1 has the smallest value of $\tilde{n}_{d0u}$. If $\tilde{n}_{d01} > 0$, without loss of generality let us suppose Subject 1 takes the control treatment at the second period. There always exist a subject (say 2) who takes test treatments in the second period as well as in the neighboring periods, i.e. periods 1 and 3, since n is odd. Then we can exchange the treatments between these two subjects at the second period so that m_{d00} is still 0. By this exchange, the decrement of $n_{d01}\tilde{n}_{d01}$ is at most $2\tilde{n}_{d01}$, while the corresponding increment of $n_{d02}\tilde{n}_{d02}$ is at least $2\tilde{n}_{d02}+1$. Since $\tilde{n}_{d02} \geq \tilde{n}_{d01}$, $\sum_{u=1}^{n} n_{d0u}\tilde{n}_{d0u}$ is increased. Hence we have, by the argument in (i), that

$$\sum_{u=1}^{n} n_{d0u}\tilde{n}_{d0u} \leq \frac{1}{4}(n-1)p(p-1) \leq \frac{1}{4}np(p-1).$$

$\diamondsuit$

Lemma 7. *For any design $d \in \Omega_{t+1,n,p}$, we have*

$\sum_{i=1}^{t} \sum_{u=1}^{n} n_{diu}\tilde{n}_{diu} \geq n(p-1) - \tilde{r}_{d0}$.

Proof. $\sum_{i=1}^{t} \sum_{u=1}^{n} n_{diu}\tilde{n}_{diu} \geq \sum_{i=1}^{t} \sum_{u=1}^{n} \tilde{n}_{diu} = \sum_{i=1}^{t} \tilde{r}_{di} = n(p-1) - \tilde{r}_{d0}$ $\diamondsuit$

Lemma 8. *For any design $d \in \Lambda_{t+1,n,p}$, we have $\beta_1 + \gamma_1 \geq 0$ for any value of θ.*

Proof. If we consider $\beta_1 + \gamma_1$ as a function of θ, then by the simple derivative equations

$$\tfrac{d}{d\theta}\left(\tfrac{\theta}{1+\theta p}\right) = \tfrac{1}{(1+\theta p)^2}, \quad \tfrac{d}{d\theta}\left(\tfrac{1}{1+\theta p}\right) = -\tfrac{p}{(1+\theta p)^2} \tag{2.10}$$

we have

$$\frac{d(\beta_1 + \gamma_1)}{d\theta} = \frac{K_d}{(1+\theta p)^2} \tag{2.11}$$

where K_d is a constant which depends on the design d only. That means $\beta_1 + \gamma_1$ is either nondecreasing or nondecreasing with respective to θ when the design d is fixed. Hence, it is enough to prove the inequality for $\theta = 0$ and $\theta = \infty$. For $\theta = 0$, we have, in view of Lemma 5, that

$$\begin{aligned}
\beta_1 + \gamma_1 &= \left(t + 1 - \frac{2}{p}\right)(n(p-1) - \tilde{r}_{d0}) - \frac{n}{p}(p-1)^2 + \frac{t}{pn}\sum_{i=1}^{t} \tilde{r}_{di} l_{dip} - \frac{\tilde{r}_{d0} l_{d0p}}{pn} \\
&\geq \left(t - \frac{1}{p}\right) n(p-1) - \left(t + 1 - \frac{2}{p} + \frac{l_{d0p}}{pn}\right)\tilde{r}_{d0} \\
&\geq \frac{tp-1}{p} n(p-1) - \left(t + 1 - \frac{2}{p} + \frac{1}{2p}\right)\frac{n(p-1)}{2} \\
&= \frac{n(p-1)}{2p}\left(p(t-1) - \frac{1}{2}\right) \geq 0
\end{aligned}$$

For $\theta = \infty$, we have, in view of Lemmas 5 and 7, that

$$\begin{aligned}
\beta_1 + \gamma_1 &\geq \left(t + 1 - \frac{2}{p}\right)(n(p-1) - \tilde{r}_{d0}) - \frac{n}{p}(p-1)^2 - \frac{1}{p}\tilde{r}_{d0} \\
&\geq \left(t - \frac{1}{p}\right) n(p-1) - \left(t + 1 - \frac{1}{p}\right)\tilde{r}_{d0} \\
&\geq \frac{n(p-1)}{2p}(p(t-1) - 1) \geq 0
\end{aligned}$$

$\diamondsuit$

In the following, we shall search for designs which minimizes the right hand side of (2.8), where the components $\sum_{u=1}^{n} n_{d0u}^2, \sum_{u=1}^{n} \tilde{n}_{d0u}^2$, $\sum_{u=1}^{n} n_{d0u}\tilde{n}_{d0u}$, $\sum_{i=1}^{t}\sum_{u=1}^{n} n_{diu}\tilde{n}_{diu}$, $\sum_{i=1}^{t} r_{di}^2$, $\sum_{i=1}^{t} \tilde{r}_{di}^2$, and $\sum_{i=1}^{t} r_{di}\tilde{r}_{di}$ are related to each other. The latter four components are investigated by Lemmas 9 and 10 while the remaining three are investigated by Lemma 11. Since

$\sum_{i=1}^t \sum_{u=1}^n n_{diu}\tilde{n}_{diu}$, $\sum_{i=1}^t r_{di}^2, \sum_{i=1}^t \tilde{r}_{di}^2$ and $\sum_{i=1}^t r_{di}\tilde{r}_{di}$ can not be minimized simultaneously, we need the following decomposition:

$$\begin{aligned} \sum_{i=1}^t r_{di}\tilde{r}_{di} = & \sum_{i=1}^t \tilde{r}_{di}^2 + \sum_{i=1}^t \tilde{r}_{di} l_{dip}, \\ \sum_{i=1}^t r_{di}^2 = & \sum_{i=1}^t \tilde{r}_{di}^2 + 2\sum_{i=1}^t \tilde{r}_{di} l_{dip} + \sum_{i=1}^t l_{dip}^2. \end{aligned} \tag{2.12}$$

By (2.12), we transform $\sum_{i=1}^t r_{di}^2$, $\sum_{i=1}^t \tilde{r}_{di}^2$ and $\sum_{i=1}^t r_{di}\tilde{r}_{di}$ into $\sum_{i=1}^t \tilde{r}_{di} l_{dip}, \sum_{i=1}^t \tilde{r}_{di}^2$, and $\sum_{i=1}^t l_{dip}^2$. The advantage is that the latter two terms are independent of each other, and thus we can first find the minimum of $\sum_{i=1}^t \tilde{r}_{di} l_{dip}$ for fixed values of the two independent terms as shown in Lemma 9.

Lemma 9. *For any design $d \in \Omega_{t+1,n,p}$, when a and b are the values of $\sum_{i=1}^t \tilde{r}_{di}^2$ and $\sum_{i=1}^t l_{dip}^2$ respectively, we have*

$$\sum_{i=1}^t \tilde{r}_{di} l_{dip} \geq \frac{1}{t}(n(p-1) - \tilde{r}_{d0})(n - l_{d0p}) - \sqrt{(a - \frac{(n(p-1) - \tilde{r}_{d0})^2}{t})(b - \frac{(n - l_{d0p})^2}{t})}, \tag{2.13}$$

with the equality holds when $\tilde{r}_{di} = \tilde{r}_{dj}$ and $l_{dip} = l_{djp}$ for any $1 \leq i \neq j \leq t$.

Proof. For notational simplicity we will let $x_i = \tilde{r}_{di}$ and $y_i = l_{dip}$. The proof reduces to minimizing $\sum_{i=1}^t x_i y_i$ under the restrictions of $\sum_{i=1}^t x_i = n(p-1) - \tilde{r}_{d0}$, $\sum_{i=1}^t x_i^2 = a$, $\sum_{i=1}^t y_i = n - l_{d0p}$ and $\sum_{i=1}^t y_i^2 = b$. Let $\bar{x} = \sum_{i=1}^t x_i/t = (n(p-1) - \tilde{r}_{d0})/t$ and $\bar{y} = \sum_{i=1}^t y_i/t = (n - l_{d0p})/t$, then $\sum_{i=1}^{t} x_i y_i = t\bar{x}\bar{y} + \sum_{i=1}^{t} u_i v_i = (n(p-1) - \tilde{r}_{d0})(n - l_{d0p})/t + \sum_{i=1}^{t} u_i v_i$, where $u_i = x_i - \bar{x}$ and $v_i = y_i - \bar{y}$. However,

$$\begin{aligned}\sum_{i=1}^{t} u_i v_i &\geq -\left(\sum_{i=1}^{t} u_i^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{t} v_i^2\right)^{\frac{1}{2}} \\ &= -\sqrt{(a - \frac{(n(p-1) - \tilde{r}_{d0})^2}{t})(b - \frac{(n - l_{d0p})^2}{t})}\end{aligned}$$

with the equality holds if and only if $x_i = -c_0 y_i, i = 1, ..., t$, with

$$c_0 = \sqrt{\frac{ta - (n(p-1) - \tilde{r}_{d0})^2}{tb - (n - l_{d0p})^2}}.$$

Hence the lemma is established. $\diamondsuit$

Lemma 10. *If $p \leq t+1$ and $\theta \geq 0$, we have, for any design $d \in \Lambda_{t+1,n,p}$, that*

$$\frac{t(t-1)^2}{x_0} + \frac{t}{y_0} \geq \frac{t(t-1)^2}{\tilde{x}_0} + \frac{t}{y_0} \tag{2.14}$$

where

$$\tilde{x}_0 = \tilde{\alpha}_1 - \frac{\tilde{\beta}_1^2}{\tilde{\gamma}_1} \tag{2.15}$$

with

$$\begin{aligned}
\tilde{\alpha}_1 =& t(1-\frac{\theta}{1+\theta p})(np-r_{d0}) - \frac{(np-r_{d0})^2}{(1+\theta p)pn} - r_{d0} + \frac{\theta}{1+\theta p}\sum_{u=1}^{n} n_{d0u}^2 + \frac{1}{(1+\theta p)pn} r_{d0}^2 \\
\tilde{\beta}_1 =& \frac{\theta t}{1+\theta p}(n(p-1)-\tilde{r}_{d0}) + \frac{(n(p-1)-\tilde{r}_{d0})(np-r_{d0})}{(1+\theta p)pn} - \frac{\theta}{1+\theta p}\sum_{u=1}^{n} n_{d0u}\tilde{n}_{d0u} - \frac{r_{d0}\tilde{r}_{d0}}{(1+\theta p)pn} \\
\tilde{\gamma}_1 =& (t+1-\frac{2}{p}-\frac{\theta t}{1+\theta p})(n(p-1)-\tilde{r}_{d0}) - \frac{n}{p}(p-1)^2 - \frac{(n(p-1)-\tilde{r}_{d0})^2}{(1+\theta p)pn} \\
&+ \frac{\tilde{r}_{d0}^2}{(1+\theta p)pn} + \frac{\theta}{1+\theta p}\sum_{u=1}^{n} \tilde{n}_{d0u}^2.
\end{aligned}$$

The equality in (2.14) holds when all n_{diu}'s are binary and $\tilde{r}_{di} = \tilde{r}_{dj}$, $l_{dip} = l_{djp}$ for any $1 \le i \ne j \le t$.

Proof. Let a, b and c to be the values of $\sum_{i=1}^{t} \tilde{r}_{di}^2$, $\sum_{i=1}^{t} l_{dip}^2$ and $\sum_{i=1}^{t} \tilde{r}_{di} l_{dip}$ respectively. In view of the decomposition (2.12), and Lemma 8, we have

$$\frac{\partial x_0}{\partial c} = -\frac{2}{\gamma_1}(\beta_1+\gamma_1) \le 0.$$

Thus, we have $x_0 \le x_0'$, where x_0' is obtained from x_0 by replacing $\sum_{i=1}^{t} \tilde{r}_{di} l_{dip}$ by its lower bound in the right side of (2.13). Similarly, denote by β_1' the new term obtained from β_1 by the same replacement. Obviously, we have $\beta_1 \ge \beta_1'$. In the following, the notation $u \propto v$ stands for $u = f(\theta, t, n, p, b, l_{d0p}, \gamma_1)v$ with the function $f > 0$ for any $d \in \Omega_{t+1,n,p}$ and $\theta \geq 0$, and the explicit form of f may vary from line to line. By direct calculation, we have

$$\begin{aligned}
\frac{\partial x_0'}{\partial b} &\propto \left(1 + \frac{\beta_1'}{\gamma_1}\right)\sqrt{\frac{ta - (n(p-1) - \tilde{r}_{d0})^2}{tb - (n - l_{d0p})^2}} - 1 \\
&\propto (\beta_1' + \gamma_1)\sqrt{ta - (n(p-1) - \tilde{r}_{d0})^2} - \gamma_1\sqrt{tb - (n - l_{d0p})^2} \\
&= \Psi_1\sqrt{ta - (n(p-1) - \tilde{r}_{d0})^2} - \Psi_2\sqrt{tb - (n - l_{d0p})^2}, \qquad (2.16)
\end{aligned}$$

where

$$\begin{aligned}
\Psi_1 =& (t + 1 - \frac{2}{p} - \frac{\theta t}{1 + \theta p})(n(p-1) - \tilde{r}_{d0}) + \frac{\theta t}{1 + \theta p}\sum_{i=1}^{t}\sum_{u=1}^{n} n_{diu}\tilde{n}_{diu} \\
&- \frac{n}{p}(p-1)^2 + \frac{(n(p-1) - \tilde{r}_{d0})(n - l_{d0p})}{(1 + \theta p)pn} - \frac{l_{d0p}\tilde{r}_{d0}}{(1 + \theta p)pn} \\
&- \frac{\theta}{1 + \theta p}\sum_{u=1}^{n} n_{d0u}\tilde{n}_{d0u} + \frac{\theta}{1 + \theta p}\sum_{u=1}^{n}\tilde{n}_{d0u}^2 \\
\Psi_2 =& (t + 1 - \frac{2}{p} - \frac{\theta t}{1 + \theta p})(n(p-1) - \tilde{r}_{d0}) - \frac{n}{p}(p-1)^2 - \frac{(n(p-1) - \tilde{r}_{d0})^2}{(1 + \theta p)pn} \\
&+ \frac{\tilde{r}_{d0}^2}{(1 + \theta p)pn} + \frac{\theta}{1 + \theta p}\sum_{u=1}^{n}\tilde{n}_{d0u}^2.
\end{aligned}$$

Let $r = \Psi_1/\Psi_2$, then (2.16) implies

$$\frac{\partial x_0'}{\partial b} = 0 \text{ iff } \sqrt{tb - (n - l_{d0p})^2} = r\sqrt{ta - (n(p-1) - \tilde{r}_{d0})^2} \qquad (2.17)$$

$$\frac{\partial x_0'}{\partial b} > 0 \text{ iff } \sqrt{tb - (n - l_{d0p})^2} < r\sqrt{ta - (n(p-1) - \tilde{r}_{d0})^2} \qquad (2.18)$$

$$\frac{\partial x_0'}{\partial b} < 0 \text{ iff } \sqrt{tb - (n - l_{d0p})^2} > r\sqrt{ta - (n(p-1) - \tilde{r}_{d0})^2} \qquad (2.19)$$

Note that r is a positive constant with respect to the values of a and b. (2.17)-(2.19) imply that the optimal vector $(a, b) \in \mathbb{R}^2$ maximizing x_0' has to satisfy the right side of (2.17). By replacing b as a function of a using the equality in (2.17), we denote the updated expression of β_1' and x_0' by β_1'' and x_0'' respectively. Note that the latter two could also be obtained directly by plugging the following equations into β_1 and x_0 respectively:

$$\sum_{i=1}^{t} r_{di}\tilde{r}_{di} = (1-r)\left(a - \frac{(n(p-1)-\tilde{r}_{d0})^2}{t}\right) + \frac{(np - r_{d0})(n(p-1)-\tilde{r}_{d0})}{t}$$
$$\sum_{i=1}^{t} r_{di}^2 = (1-r)^2\left(a - \frac{(n(p-1)-\tilde{r}_{d0})^2}{t}\right) + \frac{(np - r_{d0})^2}{t}$$

Notice that β_1'', γ_1 and x_0'' are functions of a without b involved. By direct calculation, we have

$$\frac{\partial x_0''}{\partial a} \propto -\left(1 - r + \frac{\beta_1''}{\gamma_1}\right)^2 \leq 0. \tag{2.20}$$

Hence, x_0'' is maximized when a equals to its minimum, $(n(p-1)-\tilde{r}_{d0})^2/t$. Then, by (2.13) and (2.17), the point of $(a, b, c) \in \mathbb{R}^3$ given by the following equations will attain the maximum of x_0:

$$\begin{aligned} a =& (n(p-1)-\tilde{r}_{d0})^2/t \\ b =& (n - l_{d0p})^2/t \\ c =& (n(p-1)-\tilde{r}_{d0})(n - l_{d0p})/t. \end{aligned} \tag{2.21}$$

Note that the preceding equations will be satisfied if

$$\tilde{r}_{di} = \tilde{r}_{dj} \text{ and } l_{dip} = l_{djp} \text{ for any } 1 \le i \ne j \le t. \tag{2.22}$$

By plugging the equations of (2.21) into β_1 and x_0, we denote the new terms derived by β_1''' and x_0''' respectively. Obviously, $x_0 \le x_0'''$ with equality holds when (2.22) holds. By Lemmas 6 and 7 in the supplemental material, we have

$$\begin{aligned}\beta_1''' =& \frac{\theta}{1+\theta p}\left(t\sum_{i=1}^{t}\sum_{u=1}^{n} n_{diu}\tilde{n}_{diu} - \sum_{u=1}^{n} n_{d0u}\tilde{n}_{d0u}\right) + \frac{(n(p-1)-\tilde{r}_{d0})(np-r_{d0}) - r_{d0}\tilde{r}_{d0}}{(1+\theta p)pn} \\ \ge& 0.\end{aligned}$$

That indicates that x_0''' increases when $\sum_{i=1}^{t}\sum_{u=1}^{n} n_{diu}\tilde{n}_{diu}$ decreases. Notice that $\tilde{x}_0$ is obtained from x_0''' by replacing $\sum_{i=1}^{t}\sum_{u=1}^{n} n_{diu}\tilde{n}_{diu}$ by its minimum, $n(p-1)-\tilde{r}_{d0}$, in view of Lemma 7. The conclusion is obtained. $\diamondsuit$

Remark 1. *In maximizing the x_0 in (20), the conditions $\tilde{r}_{di} = \tilde{r}_{dj}$, $l_{dip} = l_{djp}$ minimize $\sum_{i=1}^{t} r_{di}^2$ and $\sum_{i=1}^{t} \tilde{r}_{di}^2$, however not $\sum_{i=1}^{t} r_{di}\tilde{r}_{di}$. Suppose $(n(p-1)-\tilde{r}_{d0})/(t-1)$ is an integer, let $r_{d1} = 0$, $l_{d1p} = n - l_{d0p}$, and $\tilde{r}_{di} = (n(p-1)-\tilde{r}_{d0})/(t-1)$ for $i = 2, ..., t$, then we have*

$$\sum_{i=1}^{t} r_{di}\tilde{r}_{di} = (n(p-1)-\tilde{r}_{d0})^2/(t-1) \tag{2.23}$$

$$< (np - r_{d0})(n(p-1)-\tilde{r}_{d0})/t, \tag{2.24}$$

for $d \in \Lambda$ *whenever* $p < t$.

To work on the right hand of (2.14) we now only need to investigate how it is influenced by $\sum_{u=0}^{n} n_{d0u}^2$, $\sum_{u=0}^{n} n_{d0u}\tilde{n}_{d0u}$, and $\sum_{u=0}^{n} \tilde{n}_{d0u}^2$.

Lemma 11. *(i) When* $p \geq 3$ *and* $t \geq \max(p-1, 3)$, *for any design* $d \in \Lambda_{t+1,n,p}$, *we have*

$$\frac{t(t-1)^2}{\tilde{x}_0} + \frac{t}{y_0} \geq \frac{t(t-1)^2}{x_0^*} + \frac{t}{y_0^*}, \tag{2.25}$$

where x_0^* *and* y_0^* *are derived from* $\tilde{x}_0$ *and* y_0 *respectively by replacing* $\sum_{u=0}^{n} n_{d0u}^2$, $\sum_{u=0}^{n} n_{d0u}\tilde{n}_{d0u}$, *and* $\sum_{u=0}^{n} \tilde{n}_{d0u}^2$ *therein by their minimum with given* r_{d0}. *Automatically, the equality in (2.25) holds for a design* d^* *which minimizes those three terms.*

(ii) When $p = 3$ *and* $t = 2$, *the above conclusion is still valid if* $r_{d0}/n \geq 0.6306$.

Proof. By Lemma 6, we have

$$\sum_{u=1}^{n} n_{d0u}\tilde{n}_{d0u} \leq t(n(p-1) - \tilde{r}_{d0})$$

which implies $\tilde{\beta}_1 \geq 0$ for any $\theta \geq 0$. Denote by ξ_1, ξ_2 and ξ_3 the values of $\sum_{u=1}^{n} n_{d0u}^2$, $\sum_{u=1}^{n} n_{d0u}\tilde{n}_{d0u}$ and $\sum_{u=1}^{n} \tilde{n}_{d0u}^2$. Among the restrictions by the nature of the design, we have

$$\xi_1 \geq \max(\tfrac{r_{d0}^2}{n}, r_{d0}), \quad t(n(p-1) - \tilde{r}_{d0}) \geq \xi_2 \geq \tilde{r}_{d0}, \quad \xi_3 \geq \tilde{r}_{d0}. \tag{2.26}$$

We can write $\tilde{\alpha}_1$ as a function $\tilde{\alpha}_1(\xi_1, \xi_2, \xi_3, \theta)$ with the value r_{d0} fixed. The same argument could be applied to $\tilde{\beta}_1, \tilde{\gamma}_1, \alpha_2, \beta_2, \gamma_2, \tilde{x}_0, y_0$, and we define

$$H(\xi_1, \xi_2, \xi_3, \theta) = \frac{t(t-1)^2}{\tilde{x}_0(\xi_1, \xi_2, \xi_3, \theta)} + \frac{t}{y_0(\xi_1, \xi_2, \xi_3, \theta)}. \tag{2.27}$$

For notational convenience, we also define

$$\Theta_1(\xi_1, \xi_2, \xi_3, \theta) = \tfrac{\tilde{\beta}_1(\xi_1, \xi_2, \xi_3, \theta)}{\tilde{\gamma}_1(\xi_1, \xi_2, \xi_3, \theta)}, \quad \Theta_2(\xi_1, \xi_2, \xi_3, \theta) = \tfrac{\beta_2(\xi_1, \xi_2, \xi_3, \theta)}{\gamma_2(\xi_1, \xi_2, \xi_3, \theta)}. \tag{2.28}$$

In the following, we will omit the variables $\xi_1, \xi_2, \xi_3, \theta$ for the functions defined when there is no ambiguity. We have

$$\begin{aligned}
\frac{\partial H}{\partial \xi_1} &= \frac{\theta t}{1+\theta p}\left(\frac{1}{y_0^2} - \frac{(t-1)^2}{\tilde{x}_0^2}\right) \\
\frac{\partial H}{\partial \xi_2} &= \frac{2\theta t}{1+\theta p}\left(\frac{\Theta_2}{y_0^2} - \frac{(t-1)^2 \Theta_1}{\tilde{x}_0^2}\right) \\
\frac{\partial H}{\partial \xi_3} &= \frac{\theta t}{1+\theta p}\left(\frac{\Theta_2^2}{y_0^2} - \frac{(t-1)^2 \Theta_1^2}{\tilde{x}_0^2}\right).
\end{aligned}$$

Note that the derivative $\frac{\partial H}{\partial \xi_i}$ increases with ξ_j for any $1 \leq i, j \leq 3$. Now to establish

$$\frac{\partial H}{\partial \xi_i} \geq 0, i = 1, 2, 3, \tag{2.29}$$

for any value of $\xi_i, i = 1, 2, 3$ and θ, it is enough to show

$$\frac{\partial H(\frac{r_{d0}^2}{n}, \tilde{r}_{d0}, \tilde{r}_{d0}, \theta)}{\partial \xi_i} \geq 0, i = 1, 2, 3, \tag{2.30}$$

or

$$\frac{\partial H(r_{d0}, \tilde{r}_{d0}, \tilde{r}_{d0}, \theta)}{\partial \xi_i} \geq 0, i = 1, 2, 3, \tag{2.31}$$

for any value of θ. Propositions 1-5 in the Supplemental Materials finish the proof. $\diamondsuit$

Proof of Theorem 2. Combining Lemmas 1, 4 (ii), 10, and 11, it is enough to prove that a totally balanced test-control incomplete Crossover design satisfies:

- $\diamond$ n_{diu} is either 0 or 1, $1 \leq i \leq t, 0 \leq u \leq n$
- $\diamond$ $l_{dik} = r_{di}/p, i = 0, ..., t$
- $\diamond$ $T_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} T_d, T_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} F_d$, and $F_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} F_d$ are invariant under any permutation of test treatments.
- $\diamond$ $\tilde{r}_{di} = \tilde{r}_{dj}$, $l_{dip} = l_{djp}$ for any $1 \leq i \neq j \leq t$.
- $\diamond$ $\sum_{u=0}^n n_{d0u}^2$, $\sum_{u=0}^n n_{d0u}\tilde{n}_{d0u}$, and $\sum_{u=0}^n \tilde{n}_{d0u}^2$ are minimized with respect to fixed r_{d0}.

By comparing these conditions to the 7 conditions in Definition 3, we conclude the theorem. $\diamondsuit$

CHAPTER 3

EVALUATING THE EFFICIENCY OF THE DESIGNS

For any θ, let us denote by $d(\theta)$ the corresponding optimal design, then we can define the A-efficiency of a design d at this θ to be $AE(d,\theta) = Tr(M_{d(\theta)}^{-1})/Tr(M_d^{-1})$. We would be interested in deriving $AE(d,\cdot)$ for any design d, which should be resorted to deriving $Tr(M_{d(\cdot)}^{-1})$. However, $d(\cdot)$ is generally not known except when $\theta = 0$. In this chapter, we instead derive a lower bound curve $\ell(\cdot) \le Tr(M_{d(\cdot)}^{-1})$, then $LB(d,\cdot) = \ell(\cdot)/Tr(M_d^{-1})$ serves as the lower bound of $AE(d,\cdot)$. In the following, we will simply call $LB(d,\cdot)$ the efficiency of design d. Section 2.3 gave optimal designs in Ω when $\theta = 0$, and optimal designs in Λ for general θ. We would like to evaluate the efficiencies of these designs. For each design, we have

$$
\begin{aligned}
C_d &= T_d'V^{-1/2}pr^{\perp}(V^{-1/2}[1_{np}|P|F_d])V^{-1/2}T_d \\
&\le T_d'V^{-1/2}pr^{\perp}(V^{-1/2}[1_{np}|F_d])V^{-1/2}T_d \\
&\le T_d'V^{-1/2}pr^{\perp}(V^{-1/2}[1_{np}|F_dB_{t+1}])V^{-1/2}T_d \\
&\le \sum_{i=1}^{t!} \tilde{S}_i'T_d'V^{-1/2}pr^{\perp}(V^{-1/2}[1_{np}|F_dB_{t+1}])V^{-1/2}T_d\tilde{S}_i/t! \\
&\le \overline{C}_d^{(1)} - \overline{C}_d^{(2)}\left(\overline{C}_d^{(3)}\right)^{-}\overline{C}_d^{(2)} \qquad (3.1)
\end{aligned}
$$

where $\overline{C}_d^{(i)} = \sum_{i=1}^{t!} \tilde{S}_i' C_d^{(i)} \tilde{S}_i / t!, i = 1, 2, 3$, and

$$\begin{aligned}
C_d^{(1)} &= T_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} T_d \\
C_d^{(2)} &= T_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} F_d B_{t+1} \\
C_d^{(3)} &= B_{t+1} F_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} F_d B_{t+1}
\end{aligned}$$

As in (2.4), we have the representation

$$\overline{C}_d^{(i)} = \begin{pmatrix} 1_t'\overline{M}_d^{(i)} 1_t & -1_t'\overline{M}_d^{(i)} \\ -\overline{M}_d^{(i)} 1_t & \overline{M}_d^{(i)} \end{pmatrix}, \quad C_d^{(i)} = \begin{pmatrix} 1_t' M_d^{(i)} 1_t & -1_t' M_d^{(i)} \\ -M_d^{(i)} 1_t & M_d^{(i)} \end{pmatrix}, \quad i = 1, 2, 3. \tag{3.2}$$

Further, we have the relationship

$$\overline{M}_d^{(i)} = \left(\frac{Tr(M_d^{(i)})}{t-1} - \frac{1_t' M_d^{(i)} 1_t}{t(t-1)} \right) B_t + \frac{1_t' M_d^{(i)} 1_t}{t^2} J_t. \tag{3.3}$$

For any $t \times t$ matrix M, define the functions $\phi(M) = 1_t' M 1_t$ and $\varphi(M) = Tr(M) - 1_t' M 1_t / t$. By (3.1), (3.2), and (3.3), we have

$$\begin{aligned}
Tr(M_d^{-1}) &\geq \frac{t}{\phi(M_d^{(1)}) - \phi(M_d^{(2)})^2/\phi(M_d^{(3)})} + \frac{(t-1)^2}{\varphi(M_d^{(1)}) - \varphi(M_d^{(2)})^2/\varphi(M_d^{(3)})} \\
&= 1/q_d
\end{aligned} \tag{3.4}$$

where $q_d = \min_{(x,y,z)\in\mathbb{R}^3} Q_d(x,y,z)$, and

$$\begin{aligned} Q_d(x,y,z) &= (\phi(M_d^{(1)}) + 2\phi(M_d^{(2)})z + \phi(M_d^{(3)})z^2)x^2/t \\ &\quad +(\varphi(M_d^{(1)}) + 2\varphi(M_d^{(2)})y + \varphi(M_d^{(3)})y^2)(1+x)^2/(t-1)^2 \end{aligned} \tag{3.5}$$

By (A.3), we have $C_d^{(2)} = C_{d+}^{(2)} - C_{d-}^{(2)}$, where

$$\begin{aligned} C_{d+}^{(2)} &= T_d' \left(I_{np} - \lambda_p I_n \otimes J_p\right) F_d B_{t+1}, \\ C_{d-}^{(2)} &= \eta T_d' J_{np} F_d B_{t+1}, \end{aligned}$$

and λ_p and η are as defined in Theorem 2 in Section 2.3.2. Then $M_{d+}^{(2)}$ and $M_{d-}^{(2)}$ would be the corresponding submatrices of $C_{d+}^{(2)}$ and $C_{d-}^{(2)}$, and hence $M_{d+}^{(2)} = M_{d+}^{(2)} - M_{d+}^{(2)}$. In the same manner, we have the same decomposition $M_{d+}^{(i)} = M_{d+}^{(i)} - M_{d+}^{(i)}$ for $i = 1, 3$. We can now define $Q_{d+}(x,y,z)$ (*resp.* $Q_{d-}(x,y,z)$) to be the function when we replace $M_d^{(i)}, i = 1,2,3$ in (3.5) by $M_{d+}^{(i)}$ (*resp.* $M_{d-}^{(i)}$). Hence, we have

$$Q_d(x,y,z) = Q_{d+}(x,y,z) - Q_{d-}(x,y,z) \tag{3.6}$$

since $Q_d(x,y,z)$ is a linear function of $M_d^{(i)}, i = 1,2,3$. By direct calculation, we have

$$\begin{aligned}
Q_{d-}(x,y,z)/\eta &= ((\tilde{r}_{d0} - n(p-1)/(t+1))z + r_{d0})^2 x^2/t \\
&\quad + \sum_{i=1}^{t} ((\tilde{r}_{di} - n(p-1)/(t+1))y + r_{di})^2 (1+x)^2/(t-1)^2 \\
&\quad - ((\tilde{r}_{d0} - n(p-1)/(t+1))y + r_{d0})^2 (1+x)^2/(t(t-1)^2) \\
&\geq ((\tilde{r}_{d0} - n(p-1)/(t+1))z + r_{d0})^2 x^2/t \\
&\quad + np(np - 2r_{d0} - 2(\tilde{r}_{d0} - n(p-1)/(t+1))y)(1+x)^2/(t(t-1)^2)
\end{aligned} \tag{3.7}$$

Let T_d^u (*resp.* F_d^u) be the portion of T_d (*resp.* F_d) corresponding to the uth subject, $C_{d+}^{(1u)} = T_d^{u\prime}(I_p - \lambda_p J_p) T_d^u$, $C_{d+}^{(2u)} = T_d^{u\prime}(I_p - \lambda_p J_p) F_d^u B_{t+1}$, and $C_{d+}^{(3u)} = B_{t+1} F_d^{u\prime}(I_p - \lambda_p J_p) F_d^u B_{t+1}$, $M_{d+}^{(iu)}$ be the $t \times t$ submatrix of $C_{d+}^{(iu)}, i = 1,2,3$ by ignoring the first row and the first column of the latter. If we further denote by $Q_{d+}^u(x,y,z)$ the analogues of $Q_{d+}(x,y,z)$ with its components $M_{d+}^{(i)}, i = 1,2,3$ replaced by $M_{d+}^{(iu)}$, then we have

$$Q_{d+}(x,y,z) = \sum_{u=1}^{n} Q_{d+}^u(x,y,z) \tag{3.8}$$

Let $H_d^u(x,y,z) = Q_{d+}^u(x,y,z) - (p - 2n_{d0u} - 2(\tilde{n}_{d0u} - (p-1)/(t+1))y)(1+x)^2/(t(t-1)^2(1+\theta p))$, then by (3.6), (3.7) and (3.8) we have

$$Q_d(x,y,z) = \sum_{u=1}^{n} H_d^u(x,y,z) - \eta((\tilde{r}_{d0} - n(p-1)/(t+1))z + r_{d0})^2 x^2/t \tag{3.9}$$

Even though the notation $H_d^u(\cdot)$ has both the superscript u and the subscript d, this function actually depends only on the sequence, based on which the uth subject in design d is taking the treatments. Hence the design influences the summation in (3.9) through choosing n sequences from all $(t+1)^p$ possible sequences with replacement.

Definition 4. *We define two sequences $i_1 i_2 ... i_p$ and $j_1 j_2 ... j_p$ of the same length p to be TC-equivalent if $\pi(i_k) = j_k, k = 1, 2, ..., p$ for some permutation π with $\pi(0) = 0$.*

For example, the four sequences $0234, 0194, 0267$, and 0854 are TC-equivalent since all of them start with the control treatment 0, which is then followed by three distinct test treatments. Observe that two TC-equivalent sequences should result in the same function $H(\cdot)$. By classifying all the $(t+1)^p$ sequences into say J classes according to TC-equivalence and denoting by $H_j(\cdot)$ the function for the jth class, $j = 1, 2, ..., J$, we can write (3.9) in the form of

$$Q_d(x, y, z) = n \sum_{j=1}^{J} w_{dj} H_j(x, y, z) - \eta((\tilde{r}_{d0} - n(p-1)/(t+1))z + r_{d0})^2 x^2/t, \tag{3.10}$$

where w_{dj} is the proportion of the number of sequences from class j in design d, and hence $\sum_{j=1}^{J} w_{dj} = 1$. Now we have

$$\begin{aligned}
\max_d q_d &= \max_d \min_{x,y,z} Q_d(x,y,z) \\
&\leq \min_{x,y,z} \max_d Q_d(x,y,z) \\
&= \min_{x,y,z} \max_{h_1,h_2} \max_{\substack{d\,:\,r_{d0}=h_1 \\ \tilde{r}_{d0}=h_2}} \left\{ n \sum_{j=1}^{J} w_{dj} H_j(x,y,z) - \frac{\eta}{t} \left(\left(\tilde{r}_{d0} - \frac{n(p-1)}{t+1} \right) z + r_{d0} \right)^2 x^2 \right\} \\
&= \min_{x,y,z} \max_{h_1,h_2} \left\{ \max_{\substack{d\,:\,r_{d0}=h_1 \\ \tilde{r}_{d0}=h_2}} n \sum_{j=1}^{J} w_{dj} H_j(x,y,z) - \frac{\eta}{t} \left(\left(h_2 - \frac{n(p-1)}{t+1} \right) z + h_1 \right)^2 x^2 \right\},
\end{aligned} \tag{3.11}$$

where

$$\max_{h_1,h_2} := \max_{max(0,h_1-n) \leq h_2 \leq h_1 \leq np}$$

By (3.4) and (3.11), we have

Theorem 3. *For any n, t, p, and θ, the reciprocal of the right hand side of (3.11) is a lower bound of $Tr(M_d^{-1})$ for any $d \in \Omega_{n,t+1,p}$*

To implement Theorem 3, there are three levels of maximization/minimization: (I) Maximization over the designs in $\{d \in \Omega | r_{d0} = h_1, \tilde{r}_{d0} = h_2\}$ for given (h_1, h_2) and (x, y, z). (II) Maximization over (h_1, h_2) for given (x, y, z). (III) Minimization over (x, y, z).

For level (I) maximization, the problem could be reduced to the linear programming problem:

Maximize $\sum_{j=1}^{J} w_j H_j$

Subject to $n \sum w_j a_j = h_1; n \sum w_j b_j = h_2; \sum w_j = 1; w_j \geq 0, j = 1, 2, ..., J$

Here a_j (*resp.* b_j) is the number of 0's in the sequence (*resp.* the first $p-1$ symbols of the sequence) from class j. For example, when $p = 3, t \geq 3$, we have $J = 13$ and

	0	0	0	1	0	0	1	1	1	1	1	1	1
	0	0	1	0	1	1	0	0	2	1	1	2	2
	0	1	0	0	2	1	2	1	0	0	2	1	3
a_j	3	2	2	2	1	1	1	1	1	1	0	0	0
b_j	2	2	1	1	1	1	1	1	0	0	0	0	0

For level (II) maximization, we could reduce the amount of search for (h_1, h_2) by the following procedure.

1. Find a rough lower bound $\ell_1(r_{d0}, \tilde{r}_{d0}) \leq Tr(M_d^{-1}), d \in \Omega$.
2. Calculate $Tr(M_{d_0}^{-1})$ for a particular design d_0 (e.g. TBTCI design).
3. If $\ell_1(h_1, h_2) \leq Tr(M_{d_0}^{-1})$, continue with the linear programming in level (I) maximization; Otherwise, jump to next (h_1, h_2).

Lemma 12. *A lower bound of* $Tr(M_d^{-1})$ *is* $\ell_1(r_{d0}, \tilde{r}_{d0}) = t(t-1)^2/x_1 + t/y_1$*, where*

$$\begin{aligned}
x_1 &= t(1-\lambda_p)(np - r_{d0}) - r_{d0} - \eta np(np - 2r_{d0}) + \lambda_p \chi_1, \\
y_1 &= r_{d0} - \eta r_{d0}^2 - \lambda_p \chi_1 - \chi_4/n, \\
\chi_4 &= \tilde{r}_{d0} + (2\tilde{r}_{d0} - p + 1)\lfloor \tilde{r}_{d0}/(p-1) \rfloor - (p-1)\lfloor \tilde{r}_{d0}/(p-1) \rfloor^2 + (r_{d0} - \tilde{r})^2 - r_{d0}^2/p
\end{aligned}$$

and χ_1 *is defined as in Theorem 2 in Section 2.3.2.*

Proof. By ignoring F_d in C_d, we have

$$\begin{aligned}
C_d &\leq T_d' V^{-1/2} pr^{\perp}(1_{np}|V^{-1/2}P) V^{-1/2} T_d \\
&= T_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} T_d - T_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} P P_0^- P V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} T_d,
\end{aligned}$$

where $P_0 = PV^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} P = nB_p$, by choosing $P_0^- = I_p/n$, we have

$$V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} P P_0^- P V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} = n^{-1} J_n \otimes B_p.$$

By the $\tilde{S}_i$ argument as in Lemma 4 in Section 2.4, we can obtain

$$Tr(M_d^{-1}) \geq t(t-1)^2/\tilde{x}_1 + t/\tilde{y}_1, \tag{3.12}$$

where

$$\begin{aligned}
\tilde{x}_1 &= t(1-\lambda_p)(np - r_{d0}) - r_{d0} - \eta np(np - 2r_{d0}) + \lambda_p \sum_{u=1}^{n} n_{d0u}^2 \\
\tilde{y}_1 &= r_{d0} - \eta r_{d0}^2 - \lambda_p \sum_{u=1}^{n} n_{d0u}^2 - n^{-1} \sum_{k=1}^{p} (l_{d0k} - r_{d0}/p)^2.
\end{aligned}$$

In deriving (3.12), we used the fact $\sum_{i=1}^{t}(l_{dik} - r_{di}/p) \geq (l_{d0k} - r_{d0}/p)^2/t, k = 1, 2, ..., p$. The lemma is now concluded by noting $\sum_{u=1}^{n} n_{d0u}^2 \geq \chi_1$ and $\sum_{k=1}^{p}(l_{d0k} - r_{d0}/p)^2 \geq \chi_4$. $\diamondsuit$

For level (III), we could use Newton-type algorithm.

When $n = 36, p = 3, t = 4$, a TBTCI design d_2 with $r_{d0} = 36$ would be constructed through the method in Hedayat and Zheng (2010), also $d_3 = (1, 1) \otimes d_1$ is the optimal design when $\theta = 0$. The corresponding lower bound $\ell(\cdot)$ as specified by Theorem 3 could be calculated by the methods mentioned above; Note that the screening procedure for level (II) maximization saved the calculating time by 77.89%. The total time needed to calculate $\ell(\theta)$ at one value of θ is 386.50 seconds (CPU: Intel 2 Duo 1.80GHz; Software: R), while the direct search for optimal designs would require 1.11×10^{26} years.

$$d_2: \begin{array}{cccccccccccccccccccccccccccccccccccc}
0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&2&2&2&3&3&3&4&4&4&1&1&1&2&2&2&3&3&3&4&4&4\\
1&1&1&2&2&2&3&3&3&4&4&4&0&0&0&0&0&0&0&0&0&0&0&0&2&3&4&1&3&4&1&2&4&1&2&3\\
2&3&4&1&3&4&1&2&4&1&2&3&2&3&4&1&3&4&1&2&4&1&2&3&0&0&0&0&0&0&0&0&0&0&0&0
\end{array}$$

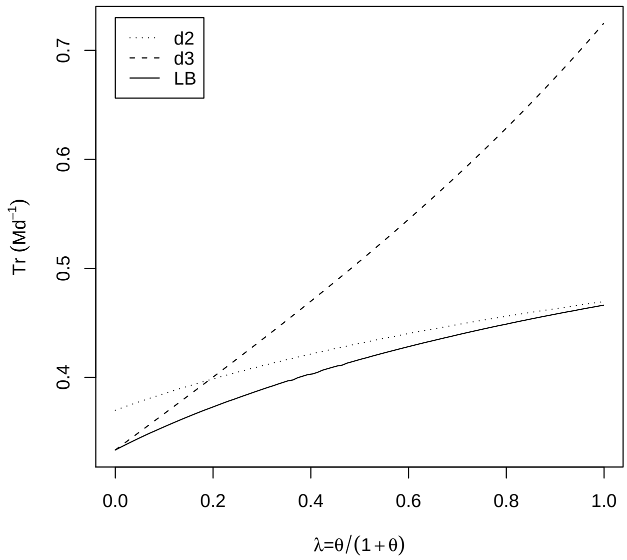

Figure 1. LB represents $\ell(\cdot)$, $\lambda = \theta/(1+\theta)$ transforms the range of θ from $[0, \infty]$ to $[0, 1]$.

$$d_3: \begin{array}{c} 1\,1\,1\,2\,2\,2\,3\,3\,3\,4\,4\,4\,0\,0\,0\,0\,0\,0\,1\,1\,1\,2\,2\,2\,3\,3\,3\,4\,4\,4\,0\,0\,0\,0\,0\,0 \\ 1\,2\,3\,2\,3\,4\,3\,4\,0\,4\,0\,0\,0\,0\,1\,0\,1\,2\,1\,2\,3\,2\,3\,4\,3\,4\,0\,4\,0\,0\,0\,0\,1\,0\,1\,2 \\ 4\,0\,0\,0\,1\,2\,2\,1\,2\,3\,4\,0\,3\,3\,0\,4\,0\,1\,4\,0\,0\,0\,1\,2\,2\,1\,2\,3\,4\,0\,3\,3\,0\,4\,0\,1 \end{array}$$

Figure Figure 1 shows the performance of d_2 and d_3 with respective to $\ell(\cdot)$. There should not be any surprise in seeing that d_3 is more efficient than d_2 when θ is small (d_3 is optimal when $\theta = 0$, and also C_d is continuous in θ). However, d_3 becomes very inefficient when θ becomes large. Instead, the TBTCI design d_3, even though not optimal, is always highly efficient for each θ. Also observe that the lower bound here is very tight since it is so close to the A-criterion of an existing design when $\theta = 0$ or ∞. Hence it is proper to use this lower bound to

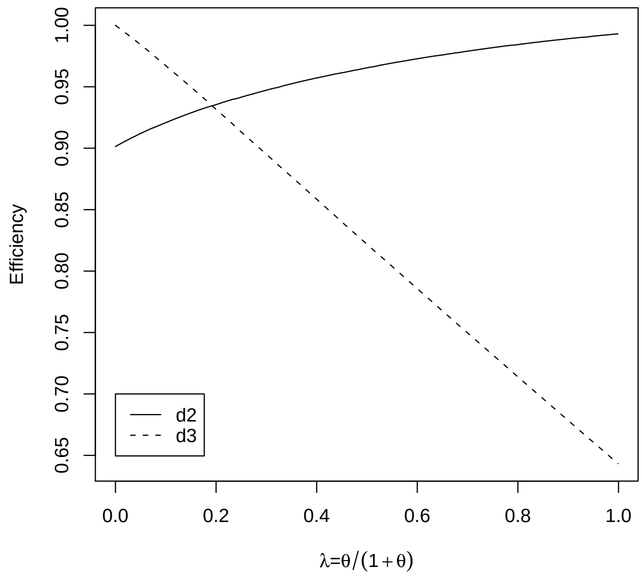

Figure 2. The TBTCI design d_2 is quite robust with the efficiency ≥ 0.9.

calculate the efficiency of designs. Figure Figure 2 shows that their efficiencies calculated by $\ell(\theta)/Tr(M_{d_i}^{-1}), i = 2, 3, \theta \in [0, \infty]$.

By Lemma 4 in Section 2.4, the A-criterion of a TBTCI design d has the form $Tr(M_d^{-1}) = n^{-1}g_{p,t}(r_{d0}/n, \theta)$ for some function g which depends on the values of p and t. Hence the A-criterion of a TBTCI design is proportional to n^{-1} as long as the ratio r_{d0}/n and other parameters are fixed. Actually, the lower bound $\ell(\cdot)$ has a similar property. To see this, let $\ell^*(\cdot)$ be the reciprocal of the right hand of (3.11) when we allow h_1 and h_2 to be real numbers instead of integers under the maximization. Then $\ell(\cdot)$ should be very close to $\ell^*(\cdot)$ especially when n is large, and we have $\ell^*(\theta) \leq \ell(\theta)$ for any θ in general. Also, we have $\ell^*(\theta) = n^{-1}\tilde{g}_{p,t}(\theta)$, hence $\ell^*(\theta)/Tr(M_d^{-1})$ only depends on the ratio r_{d0}/n. Figure Figure 3 gives the more conservative

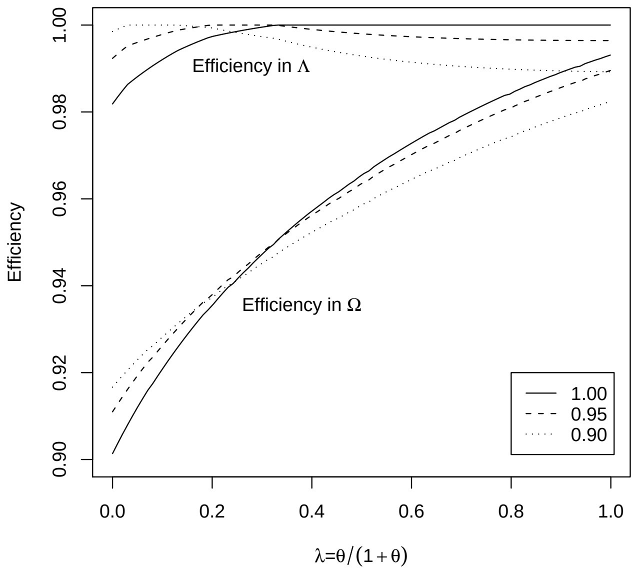

Figure 3. Efficiency in $\Lambda : min_h f(h, \theta)/Tr(M_d^{-1})$.

efficiencies of TBTCI designs with the ratio being $1, 0.95, 0.9$ (if they exist) in both Ω and Λ when $p = 3$ and $t = 4$. This figure gives the guidance regarding what kind of TBTCI design should be chosen. Similar work could be carried out for other configurations of p and t. We postpone the discussion of possible improvements on current results to Chapter 5.

To evaluate the MV-efficiencies of TBTCI designs and the type of designs proposed in Theorem 1, it is useful to note that the matrix M_d is completely symmetric for them. Hence by Corollary 1 its MV-efficiency is at least as large as its A-efficiency regardless of the competing class of designs.

CHAPTER 4

CONSTRUCTION OF TBTCI DESIGNS

4.1 Introduction

This chapter aims to introduce techniques for constructing two-way arrays of a special type, which is very useful in the area of experimental design. In this chapter, a two-way array , say d, would be called a *design*. Here is an example:

$$
\begin{array}{ccccccccccccc}
 & 1 & 2 & 3 & 3 & 2 & 0 & 1 & 3 & 0 & 2 & 1 & 0 \\
d_1: & 2 & 3 & 1 & 2 & 0 & 3 & 3 & 0 & 1 & 1 & 0 & 2 \\
 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 3
\end{array}
$$

In d_1, each column represents an ordered sequence of treatments for one subject to take. Along the experiment, responses of interest like blood pressure will be measured at each of these $3 \times 12 = 36$ runs so that statistical inference could be carried out to estimate the effects of the treatments on the responses. To do this, we need to understand what sources contribute to the variation of the response measurements besides uncontrollable random errors. Here, the response at the circled run could depend on the effect of treatment 3, the physical condition of the 2nd subject, the time (period) of this run, and even the effect of treatment 2 from the previous period when the washout time is not very long. Below is a reasonable model which formulates the ideas.

$$
Y_{dku} = \mu + \alpha_k + \beta_u + \tau_{d(k,u)} + \gamma_{d(k-1,u)} + \epsilon_{ku}, \;\; k = 1, 2, ..., p, \; u = 1, 2, ..., n. \tag{4.1}
$$

First, the design $d : (k, u) \mapsto i$ decides treatment i to be applied to subject u at period k, then Y_{dku} and ϵ_{ku} are the corresponding response measurement and uncontrollable random error. α_k is the effect of period k; β_u is the effect of subject u; $\tau_{d(k,u)}$ is the direct effect of treatment $d(k, u)$; and $\gamma_{d(k-1,u)}$ is the carryover effect of treatment $d(k-1, u)$ from the previous period(by convention $\gamma_{d(0,u)} = 0$). Typical works in finding designs under Model (4.1) include Cheng and Wu (1980), Hedayat and Afarinejad (1975, 1978), Hedayat and Yang (2003, 2004), Kunert (1984), Kunert and Martin (2000), Kushner (1997, 1998), and Stufken (1991, 1996) among others.

Here, we assume ϵ_{ku} and β_u to be random with $E(\epsilon_{ku}) = E(\beta_u) = 0$, $Var(\epsilon_{ku}) = \sigma_\epsilon^2 < \infty$, and $Var(\beta_u) = \sigma_\beta^2 < \infty$. Any two of these random components are mutually independent. Other factors in the right-hand side of the model are assumed to be non-random. By writing $\mathbf{Y}_d = (Y_{d11}, Y_{d21}, ..., Y_{dpn})'$ with the index arranged in colexicographical order, we can express Model (1.1) in matrix notation as

$$\begin{aligned} E(\mathbf{Y}_d) = & \quad 1_{np}\mu + P\boldsymbol{\alpha} + T_d\boldsymbol{\tau} + F_d\boldsymbol{\gamma}, \\ var(\mathbf{Y}_d) = & \quad \sigma^2(I_n \otimes (I_p + \theta J_\rho)), \end{aligned} \tag{4.2}$$

where $\theta = \sigma_\beta^2/\sigma_\epsilon^2 \geq 0$, $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_p)'$, $\boldsymbol{\tau} = (\tau_0, ..., \tau_t)'$, $\boldsymbol{\gamma} = (\gamma_0, ..., \gamma_t)'$, $P = 1_n \otimes I_p$ with $\otimes$ to be the Kronecker product, and T_d and F_d denote the treatment and carryover incidence matrices. Let $C_d = T_d'V^{-1/2}pr^\perp(V^{-1/2}[1_{np}|P|F_d])V^{-1/2}T_d$, where $pr^\perp(A) = I - A(A'A)^-A'$ and $V = I_n \otimes (I_p + \theta J_p)$. Then C_d would serve as the information matrix for $\boldsymbol{\tau}$ in the sense that $Var(B\hat{\boldsymbol{\tau}}) = \sigma_\epsilon^2 BC_d^{-1}B'$ for any matrix B of $t+1$ columns with $B1 = 0$, where $\hat{\boldsymbol{\tau}}$ is the generalized least square estimate of $\boldsymbol{\tau}$. Thus, C_d carries all the information about the design d necessary to evaluate the accuracy of estimating the effects of treatments.

In the context of comparing test treatments, $\{1, 2, ..., t\}$, with a control treatment, $\{0\}$, the most frequently used optimality criterion is A-optimality which minimizes $\sum_{i=1}^{t} Var_d(\hat{\tau}_i - \hat{\tau}_0)$. Let $M_d = QC_dQ'$, where $Q = [0_{t\times 1}|I_t]$. Then, a design which minimizes $Tr(M_d^{-1})$ is A-optimal. Another optimality criterion is MV-optimality which minimizes $\max_{i=1,\ldots,t} Var_d(\hat{\tau}_i - \hat{\tau}_0)$. It is well known that an A-optimal design is also an MV-optimal design if M_d is completely symmetric.

Note that $\lim_{\theta\to 0} C_d = T_d' pr^{\perp}([1_{np}|P|F_d])T_d$ is the information matrix for $\boldsymbol{\tau}$ when $\beta_u = 0$ almost surely (everywhere) and $\lim_{\theta\to\infty} C_d = T_d' pr^{\perp}([1_{np}|P|U|F_d])T_d$, where $U = I_n \otimes 1_p$, is the information matrix for $\boldsymbol{\tau}$ when β_u is non-random. Hence, specifying β_u to be random enables us to cover a wide range of models and θ will play a very important role in identifying optimal designs. See Hedayat, Stufken, and Yang (2006) for detailed arguments.

For further discussion, we define $n_{diu} = \sum_{k=1}^{p} I_{[d(k,u)=i]}$, $\tilde{n}_{diu} = \sum_{k=1}^{p-1} I_{[d(k,u)=i]}$, $l_{dik} = \sum_{u=1}^{n} I_{[d(k,u)=i]}$, $m_{dij} = \sum_{u=1}^{n}\sum_{k=1}^{p-1} I_{[d(k,u)=i,d(k+1,u)=j]}$, $r_{di} = \sum_{u=1}^{n}\sum_{k=1}^{p} I_{[d(k,u)=i]}$, $\tilde{r}_{di} = \sum_{u=1}^{n}\sum_{k=1}^{p-1} I_{[d(k,u)=i]}$, and $\Gamma_i = \{u : d(p,u) = i\}$. Let $\Omega_{t+1,n,p}$ be the collection of all designs with n subjects, p periods, $t+1$ treatments. We also define $\Omega^1_{t+1,n,p} = \{d \in \Omega_{t+1,n,p} : l_{d0k} = r_{d0}/p, k = 1, 2, ...p\}$ and $\Lambda_{t+1,n,p} = \{d \in \Omega^1_{t+1,n,p} : m_{dii} = 0, i = 0, 1, ..., t\}$. Ideally, we want to find optimal designs in $\Omega_{t+1,n,p}$ for comparing test treatments to the control. Unfortunately, no such work has been carried out yet. Hedayat and Yang (2006), and Yang and Park (2007) derived properties of A-optimal designs within $\Omega^1_{t+1,n,p}$ when $\theta = \infty$ and p, t satisfies (i)

$p = 3$ and $3 \leq t \leq 20$ or (ii) $p \geq 4$, $(p-3)(p-2)+2 \leq t \leq (p-2)(p-1)+1$, $n \geq p(p-1)/2$. For more general values of t, p, θ, Chapter 2 established both A- and MV- optimality of certain type of designs within $\Lambda_{t+1,n,p}$. Concrete examples shows that optimal designs in $\Lambda_{t+1,n,p}$ will usually be highly A-efficient, sometimes even optimal, in $\Omega_{t+1,n,p}$. The following design first proposed by Hedayat and Yang (2005) plays the central role in constructing optimal designs.

Definition 5. *A design $d \in \Lambda_{t+1,n,p}$ for comparing test treatments, $\{1, 2, ..., t\}$, with the control treatment, $\{0\}$, is called a totally balanced test-control incomplete Crossover (TBTCI) design if it satisfies*

1. $|n_{d0u} - n_{d0v}| \leq 1$ *and* $|\tilde{n}_{d0u} - \tilde{n}_{d0v}| \leq 1$ *for all* $1 \leq u, v \leq n$.
2. $n_{diu} = 0$ *or* 1 *for all* $1 \leq i \leq t$ *and* $1 \leq u \leq n$.
3. l_{dik} *is a constant across all* $1 \leq i \leq t$ *and* $1 \leq k \leq p$
4. m_{d0i}, m_{di0} *and* m_{dij} *are constants across all* $1 \leq i \neq j \leq t$ *for all* $0 \leq i \leq t$.
5. $\sum_{u=1}^{n} \tilde{n}_{d0u}\tilde{n}_{diu}, \sum_{u\in\Gamma_0} \tilde{n}_{diu}, \sum_{u\in\Gamma_i} \tilde{n}_{d0u}, \sum_{u=1}^{n} \tilde{n}_{diu}\tilde{n}_{dju}, \sum_{u\in\Gamma_i} \tilde{n}_{dju}$ *are constants across all* $1 \leq i \neq j \leq t$.

Even though the conditions for a design to be a TBTCI seems complex, the existence of such designs is not uncommon and note that d_1 is a TBTCI design. For each design $d \in \Lambda_{t+1,n,p}$, we define the function

$$l(t, n, p, \theta, r_{d0}) = t(t-1)^2(\alpha_1 - \beta_1^2/\gamma_1)^{-1} + t(\alpha_2 - \beta_2^2/\gamma_2)^{-1} \tag{4.3}$$

where

$$
\begin{aligned}
\alpha_1 =& t(1-\lambda_p)(np-r_{d0}) - \eta(np-r_{d0})^2 - r_{d0} + \lambda_p S_1 + \eta r_{d0}^2 \\
\beta_1 =& \lambda_p t(n(p-1)-\tilde{r}_{d0}) + \eta(n(p-1)-\tilde{r}_{d0})(np-r_{d0}) - \lambda_p S_2 - \eta r_{d0}\tilde{r}_{d0} \\
\gamma_1 =& (t+1-2/p-\lambda_p t)(n(p-1)-\tilde{r}_{d0}) - n(p-1)^2/p - \eta(n(p-1)-\tilde{r}_{d0})^2 + \eta\tilde{r}_{d0}^2 + \lambda_p S_3. \\
\alpha_2 =& r_{d0} - \lambda_p S_1 - \eta r_{d0}^2, \\
\beta_2 =& \lambda_p S_2 + \eta r_{d0}\tilde{r}_{d0}, \\
\gamma_2 =& \tilde{r}_{d0} - (np^2-np)^{-1}\tilde{r}_{d0}^2 - \lambda_p S_3 - \eta\tilde{r}_{d0}^2.
\end{aligned}
$$

with

$$
\begin{aligned}
\lambda_p =& \theta(1+\theta p)^{-1}, \text{ and } \eta = \lambda_p(\theta p n)^{-1} \\
S_1 =& r_{d0} + 2(r_{d0}-n)\lfloor r_{d0}/n \rfloor - n\lfloor r_{d0}/n \rfloor^2 \\
S_2 =& \{\tilde{r}_{d0} + (r_{d0}+\tilde{r}_{d0}-n)\lfloor \tilde{r}_{d0}/n \rfloor - n\lfloor \tilde{r}_{d0}/n \rfloor^2\}\mathbf{1}_{\{r_{d0}/n - \lfloor \tilde{r}_{d0}/n \rfloor < 1\}} \\
&+ \{r_{d0} + \tilde{r}_{d0} - n + (r_{d0}+\tilde{r}_{d0}-2n)\lfloor \tilde{r}_{d0} \rfloor - n\lfloor \tilde{r}_{d0}/n \rfloor^2\}\mathbf{1}_{\{r_{d0}/n - \lfloor \tilde{r}_{d0}/n \rfloor \geq 1\}} \\
S_3 =& \tilde{r}_{d0} + 2(\tilde{r}_{d0}-n)\lfloor \tilde{r}_{d0}/n \rfloor - n\lfloor \tilde{r}_{d0}/n \rfloor^2
\end{aligned}
$$

Since $\tilde{r}_{d0} = (p-1)r_{d0}/p$ for any $d \in \Lambda_{t+1,n,p}$, the function l depends on d only through r_{d0}, the total replication of the control treatment. Now we are ready to rewrite a result from Chapter 2 by the following theorem for the purpose of construction issue. The corollary follows easily since M_d is completely symmetric whenever d is a TBTCI design.

Theorem 4. *When $t \geq 3$ and $t+1 \geq p \geq 3$, for any given θ, $Tr(M_d^{-1}) \geq l(t,n,p,\theta,r_{d0})$ for any design $d \in \Lambda_{t+1,n,p}$. The equality holds whenever d is a TBTCI design. When $p = 3$ and $t = 2$, the conclusion still holds but only within a subclass of $\Lambda_{t+1,n,p}$ in which $r_{d0}/n \geq 0.6306$.*

Corollary 2. *When $t \geq 3$ and $t+1 \geq p \geq 3$, for any given θ, a design d^* is simultaneously A- and MV-optimal among designs in $\Lambda_{t+1,n,p}$ if it is a TBTCI and $l(t,n,p,\theta,r_{d^*0}) = \min_{d\in\Lambda_{t+1,n,p}} l(t,n,p,\theta,r_{d0})$. When $p = 3$ and $t = 2$, the conclusion still holds but only within a subclass of $\Lambda_{t+1,n,p}$ in which $r_{d0}/n \geq 0.6306$.*

Remark 2. *Hedayat and Zhao (1990) discovered optimal designs for the special case of $p = 2$.*

With respect to the criterion of A-optimality, we define the efficiency of a design $d^* \in \Lambda_{t+1,n,p}$ to be $\min_{d\in\Lambda_{t+1,n,p}} Tr(M_d^{-1})/Tr(M_{d^*}^{-1})$. Automatically, the efficiency of an optimal design is 1. Let $\mathcal{I}_{t,n,p} = \{r_{d0} : d \in \Lambda_{t+1,n,p}\}$, then $\mathcal{I}_{t,n,p} = \{mp : 1 \leq m < n/2, m \text{ is an integer}\}$. Also let $\mathcal{S}_{t,n,p,\theta} = \{r \in \mathcal{I}_{t,n,p} : l(t,n,p,\theta,r) = \min_{d\in\Lambda_{t+1,n,p}} l(t,n,p,\theta,r_{d0})\}$. For given t,n,p,θ, if we can find a TBTCI design d with $r_{d0} \in \mathcal{S}_{t,n,p,\theta}$, this design is A- and MV-optimal by Corollary 2 and the efficiency is 1. If there is no such design, a TBTCI design d^* with r_{d^*0} close to elements in $\mathcal{S}_{t,n,p,\theta}$ would still be plausible. In this case, we could use the lower bound of the efficiency $l(t,n,p,\theta,r)/Tr(M_d^{-1}), r \in \mathcal{S}_{t,n,p,\theta}$ instead to evaluate its efficiency.

Denote by $|\mathcal{S}_{t,n,p,\theta}|$ the number of elements in $\mathcal{S}_{t,n,p,\theta}$. In this chapter, all values of t,n,p,θ involved in calculation yields $|\mathcal{S}_{t,n,p,\theta}| = 1$. In general, we can define $g(t,n,p,\theta) = \max \mathcal{S}_{t,n,p,\theta}$. However, the function g does not have a closed form due to the discrete nature of r_{d0}. However, we can visualize the function g by fixing values of t,p,θ and draw the equation of $r_{d^*0} = h_{t,p,\theta}(n) \equiv g(t,n,p,\theta)$. We like to emphasize that in practice small values of p such as $3, 4$ and 5

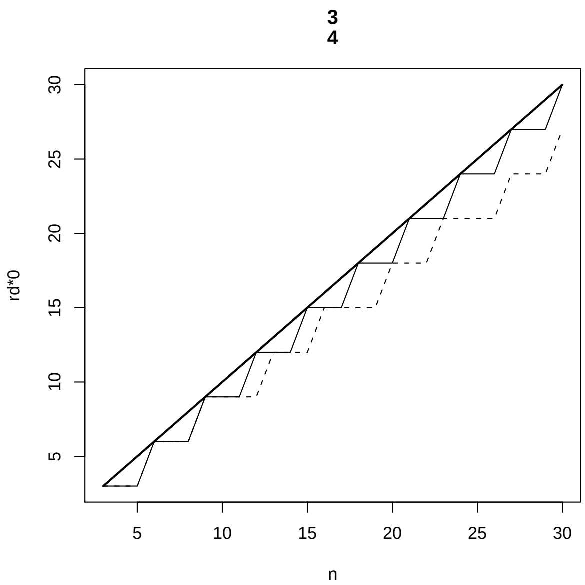

Figure 4. The two curves: $r_{d^*0} = h_{4,3,\infty}(n)$ (solid) and $r_{d^*0} = h_{4,3,0}(n)$ (dashed)

are used very frequently. In Figure 1, the bold straight line represents the reference line by the equation of $r_{d^*0} = n$, the solid curve represents r_{d^*0} as a function of n when we fix $p = 3, t = 4$ and $\theta = \infty$, and the dashed curve corresponds to the case of $p = 3, t = 4$ and $\theta = 0$. For general values of $\theta > 0$, the corresponding curves would be between the solid and dashed curves. In constructing a design, t, n, p are known, while θ is unknown. Thus, we don't know whether a design is optimal or how efficient it is. Instead, we could figure out its efficiency for each value of θ. From the figure, a TBTCI design d with r_{d0}/n equals to or slightly smaller than 1, depending on values of t, p, θ, would behave reasonably well.

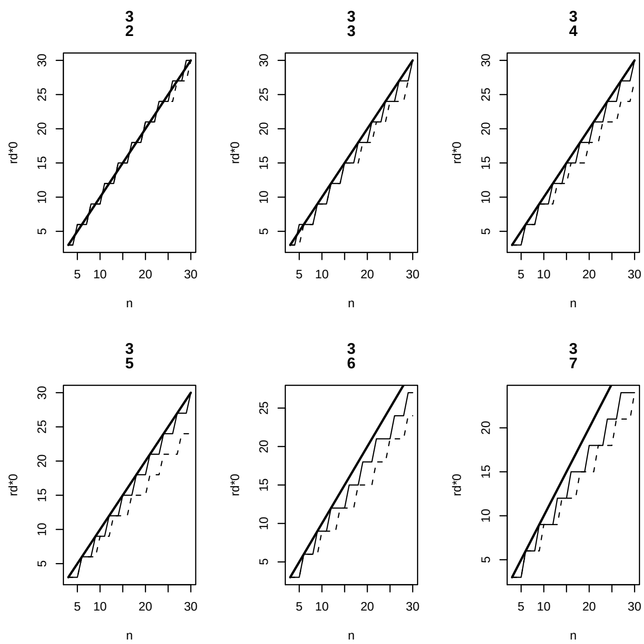

Figure 5. $r_{d^*0} = h_{t,3,\infty}(n)$ (solid) and $r_{d^*0} = h_{t,3,0}(n)$ (dashed), $t = 2, 3, ..., 7$

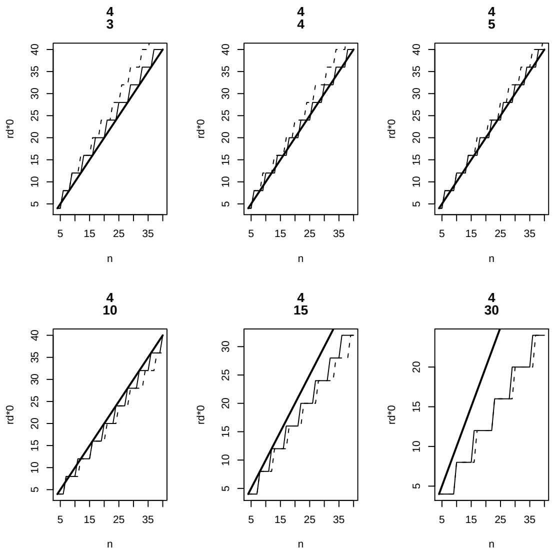

Figure 6. $r_{d^*0} = h_{t,4,\infty}(n)$ (solid) and $r_{d^*0} = h_{t,4,0}(n)$ (dashed), $t = 3, 4, 5, 10, 15, 30$

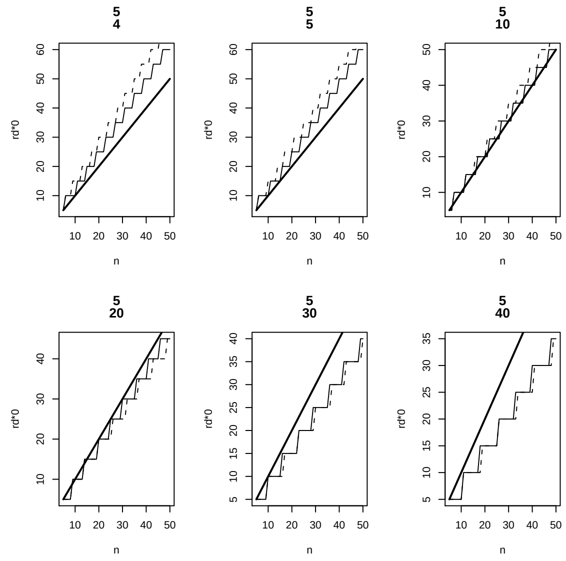

Figure 7. $r_{d^*0} = h_{t,5,\infty}(n)$ (solid) and $r_{d^*0} = h_{t,5,0}(n)$ (dashed), $t = 4, 5, 10, 20, 30, 40$

Similarly, we present Figure 2 for $p = 3$ and $t = 2, 3, ..., 7$. Notice that the curves become flatter while t becomes larger, it's because whenever new test treatments is introduced, they will force other treatments including the control to reduce their replications since the total number of the runs remain unchanged. This trend remains true for any other values of p. However, from Figures 3-6, we can see one remarkable difference for the cases of $p \geq 4$. That is, the curves are somehow above the reference line of $r_{d^*0} = n$ in the beginning (for small t) and go below the line eventually when t is large enough.

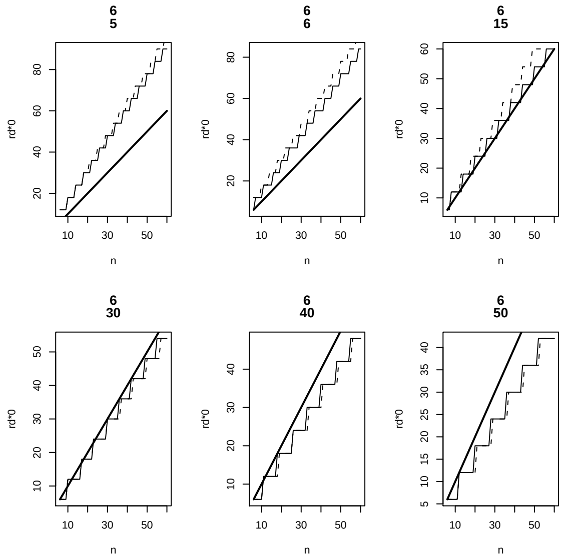

Figure 8. $r_{d^*0} = h_{t,6,\infty}(n)$ (solid) and $r_{d^*0} = h_{t,6,0}(n)$ (dashed), $t = 5, 6, 15, 30, 40, 50$

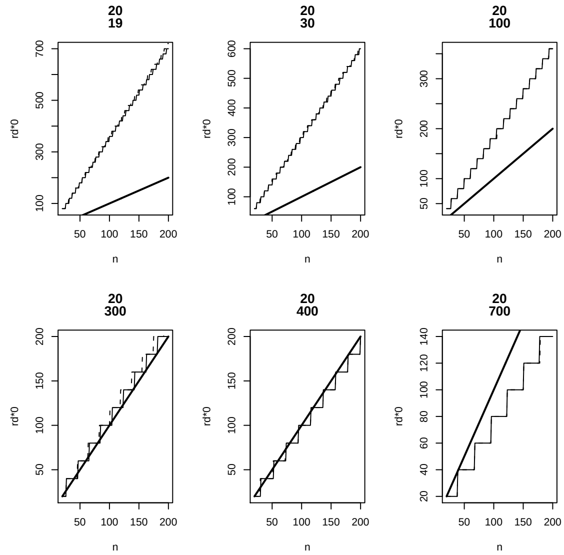

Figure 9. $r_{d^*0} = h_{t,20,\infty}(n)$ (solid) and $r_{d^*0} = h_{t,20,0}(n)$ (dashed), $t = 19, 30, 100, 200, 400, 700$

4.2 Construction Tools

In this section, we will use terms of column, row, and symbol in normal meaning to replace the terms of subject, period, and treatment in Section 4.1. We shall introduce a new class of designs which would be very useful for the construction of TBTCI designs.

Definition 6. *A $p \times n$ array with symbols from $\{1, 2, ..., t\}$ is called a totally balanced incomplete Crossover (TBIC) design denoted by $TBIC(t, n, p)$ if it satisfies:*

1. *$n_{diu} = 0$ or 1 for all $1 \leq i \leq t$ and $1 \leq u \leq n$.*

2. *l_{dik} is a constant cross all $1 \leq i \leq t$ and $1 \leq k \leq p$*

3. m_{dij} *is a constant across all* $1 \le i \ne j \le t$ *for all* $1 \le i \le t$.

4. $\sum_{u=1}^{n} \tilde{n}_{diu}\tilde{n}_{dju}, \sum_{u \in \Gamma_i} \tilde{n}_{dju}$ *are constants across all* $1 \le i \ne j \le t$.

Note that Conditions (1) – (4) in Definition 6 is equivalent to Conditions (1) – (5) in Definition 5 plus the condition of $r_{d0} = 0$. Also, A totally balanced design defined by Kunert and Stufken (2002) reduces to a totally balanced incomplete Crossover design if and only if $p \le t$. Note that in a TBIC design the symbol 0 is not included in labelling the treatments. For a TBIC design d, we could let $l = l_{dik}$ and $m = m_{dij}$ for any i, j, k. By Condition (3) in Definition 2.1, the existence of a TBIC design requires

$$l(p-1) = m(t-1). \tag{4.4}$$

Since $n = lt$, a TBIC design is said to be of *minimal* size if the corresponding m and l are relatively prime and satisfy (2.1).

A special class of TBIC designs, which is well known, is worth pointing out separately. A Latin square L of order p is said to be $column-complete$, and is denoted by $CCLS(p)$, if the ordered pairs $(L_{ij}, L_{i+1,j})$ are all distinct for $1 \le i \le p-1$ and $1 \le j \le p$. A TBIC design will reduce to a $column-complete$ Latin square when $t = n = p$. A $CCLS(p)$ exists whenever p is a composite number. In case p is prime, we could find two Latin squares $L^{(1)}$ and $L^{(2)}$ of order p each, such that every ordered pair of distinct elements from $\{1, 2, ..., p\}$ appears twice in the collection of $(L_{ij}^{(k)}, L_{i+1,j}^{(k)}), 1 \le i \le p-1, 1 \le j \le p, k = 1, 2$. Readers interested in the details of the argument is refered to Williams (1949), Gordon (1961) and Higham (1998). Now, it's convenient to introduce another well known type of designs.

A design is said to be a uniform balanced design if it satisfied Conditions (2) and (3) in Definition 2.1 as well as the condition that n_{diu} is a constant across all $1 \leq i \leq t$ and $1 \leq u \leq n$. By the latter condition, both p and n have to be a multiple of t. It is easy to verify that a uniform balanced design with $p = t$ is a TBIC design. Thus, a uniform balanced design in $\Omega_{t,\alpha t,t}$ could be obtained from α copies of a $CCLS(t)$ whenever t is a composite number and α is an integer, or from $\alpha/2$ copies of $L^{(1)}$, $L^{(2)}$ described above whenever t is a prime integer and α is an even number.

4.2.1 Construction of TBTCI Designs Using TBIC Designs

Note that the optimality of certain TBTCI designs established by Theorem 4 is only applicable when $3 \leq p \leq t+1$. According to Condition (3) in Definition 5 and the definition of $\Lambda_{t+1,n,p}$, $n \geq t+1$ is necessary for the existence of TBTCI designs. Actually, there exists TBTCI designs under the boundary condition of $n = t+1$. Therefore, the discussion concerning the construction of TBTCI designs in the sequel always assumes the condition

$$3 \leq p \leq t+1 \leq n. \tag{4.5}$$

In reality, the choices for the numbers t, n, p are decided by the nature of experiments themselves, and our job is to prepare TBTCI designs with *proper* values of r_{d0} as indicated by Corollary 2 for all possible configurations of t, n, p. Indeed, the value of r_{d0}/n is more convenient for discussion than r_{d0} itself based on the figures presented earlier. It is obvious that not all triples (t, n, p) admits the existence a TBTCI design. From the point of view of construction, we would construct TBTCI designs with as many as possible different values of n for every configuration of p and t. Based on Theorem 5 and Corollary 3 below, there are two ways to go for when t and p are chosen: (i) Construct a TBTCI design d with the minimal possible n, say n_0, and desirable value of r_{d0} (ii) Construct a TBTCI design d with n not a multiple of n_0 and desirable value of r_{d0}.

Theorem 5. *The juxtaposition of any finite many TBTCI designs with the common number of rows and treatments would still be a TBTCI design as long as* $|n_{d0u} - n_{d0v}| \leq 1$ *and* $|\tilde{n}_{d0u} - \tilde{n}_{d0v}| \leq 1$ *where* u *and* v *represent two different columns in the resulting design.*

Hereafter, we denote by $TBTCI_{t,p}(n, r)$ a TBTCI design with t test treatments, p rows, n columns and $r_{d0} = r$. When only the number of rows (p) and the number of test treatments (t) are specified, we will use $TBTCI_{t,p}$ to denote the design. Then, we have

Corollary 3. *The juxtaposition of* q *copies of a* $TBTCI_{t,p}(n, r_{d0})$ *is a* $TBTCI_{t,p}(qn, qr_{d0})$.

Now, we start with a simple way of constructing TBTCI designs. Given any $TBIC(t + 1, n, p)$, we directly obtains a $TBTCI_{t,p}(n, np/(t + 1))$ design by relabelling the treatment $t + 1$ by the control treatment 0. Thus, we have $r_{d0}/n = p/(t + 1) \leq 1$ for this family of designs.

When n is a multiple of t, we can construct a $TBTCI_{t,p}(n, np/t)$ based on an arbitrary $TBIC(t, n/t, p)$. Denote by $D(i), 1 \leq i \leq t$ the design obtained from the $TBIC(t, n/t, p)$ when treatment i is relabelled by the control treatment 0. The juxtaposition of $D(1)$ to $D(t)$

will give a $TBTCI(n, p, t, np/(t+1))$. According to Condition (1) of Definition 6, the existence of a $TBIC(t, n/t, p)$ mentioned earlier requires the condition of $p \leq t$. Thus, for the $TBTCI_{t,p}(n, np/t)$ constructed here, we have $r_{d0}/n = p/t \leq 1$. Hedayat and Yang (2005) used the same idea of expansion based on uniform balanced designs with $p = t$, which is a special type of TBIC designs. However, their method could only produce TBTCI designs in which the number of test treatments is identical to the number of rows. With our generalization, we can deal with any values of p and t satisfying $p \leq t$.

The TBTCI designs constructed above all have the property that $r_{d0} \leq n$. To construct a $TBTCI_{t,p}$ design d with $r_{d0} \geq n$, we first need to prepare a $TBIC(t, n, p-1)$ and a $TBIC(t, n, p-2)$. Then we add one row of 0's to the $TBIC(t, n, p-1)$, and denote the resulting design by $A(k), 1 \leq k \leq p$ if it has the 0's in the kth row of itself. When adding two rows of 0's to the $TBIC(t, n, p-2)$, we label the resulting design by $B(k_1, k_2)$ if it has the 0's in the k_1th and k_2th rows. We need to construct a collection of those designs such that: (i) The numbers $1, 2, ..., p$ appears exactly once in the parentheses of either A or B type designs, (ii) the number p has to appear in the parentheses of a B type design. Then, by Theorem 5, the resulting design obtained by juxtaposing this collection of designs will be a TBTCI design. The number of columns of the resulting design is the product of n and the number of the smaller designs of both A and B types. There are two extreme cases worth mentioning here. When all of designs juxtaposed are A type designs, then we only need a $TBIC(t, n, p-1)$ and will obtain a TBTCI design with $r_{d0} = n$. When p is even and all of designs juxtaposed are B type designs, then we obtain a TBTCI design with $r_{d0} = 2n$. The value r_{d0}/n for the designs constructed by this method is between 1 and 2. Here is an example of a $TBTCI_{3,4}(12, 24)$:

$$d_2 : \begin{array}{cccccccccccc} 0&0&0&0&0&0&2&3&1&3&1&2 \\ 2&3&1&3&1&2&0&0&0&0&0&0 \\ 0&0&0&0&0&0&1&1&2&2&3&3 \\ 1&1&2&2&3&3&0&0&0&0&0&0 \end{array}$$

The design is derived from the trivial $TBIC(3, 6, 2)$:

$$\begin{array}{cccccc} 2&3&1&3&1&2 \\ 1&1&2&2&3&3 \end{array}$$

Currently, our methods of constructing TBTCI designs rely on using TBIC designs. Thus, the issue of the number n would be extensively discussed during the construction of TBIC designs in Subsections 4.2.2–4.2.4.

4.2.2 Method 1 of Constructing TBIC Designs

A type I orthogonal array $OA_I(n, p, t, s)$ is a $p \times n$ array based on t symbols, where the columns of any $s \times n$ subarray contains all $t!/(t - s)!$ permutations of s distinct symbols the same number of times. Here s is said to be the strength of the type I orthogonal array. Suppose the t symbols are $\{1, 2, ..., t\}$, then an $OA_I(n, p, t, 2)$ is a $TBIC(t, n, p)$. Actually, a type I orthogonal array imposes more structures than a TBIC design, however it is easier to construct the former. We first illustrate our method of constructing this type of designs for $p = 3$, and the method could be easily generalized to cases of $p \geq 4$.

A *transversal* in a Latin square of order n is a collection of n positions which exhaust all of the n different symbols, rows and columns. It is well known that a Latin square with a transversal always exists as long as the order is not 2. By permuting the rows and columns of such a Latin square of order t with symbols from $\{1, ..., t\}$, we can always obtain an *idempotent* Latin square L, in which $L_{i,i} = i$. Now we label the rows of the newly obtained Latin square by $1, ..., t$ from the first row to the last row respectively. We also label the columns in the same manner. Each position in the Latin square could indicate a three-dimension column vector with entries filled with the corresponding row label, column label, and the symbol. Then the juxtaposition of these vectors corresponding to all of the off-diagonal positions will yield a type I orthogonal array. Here is an example when $t = 5$:

$$\begin{array}{ccccc} ① & 5 & 4 & 3 & 2 \\ 4 & ② & 5 & 1 & 3 \\ 5 & 3 & 2 & ④ & 1 \\ 2 & 1 & ③ & 5 & 4 \\ 3 & 4 & 1 & 2 & ⑤ \end{array} \longmapsto \begin{array}{ccccc} ① & 5 & 4 & 3 & 2 \\ 4 & ② & 5 & 1 & 3 \\ 2 & 1 & ③ & 5 & 4 \\ 5 & 3 & 2 & ④ & 1 \\ 3 & 4 & 1 & 2 & ⑤ \end{array}$$

$$\longmapsto d_3 : \begin{array}{l} 5\ 4\ 3\ 2\ 4\ 5\ 1\ 3\ 2\ 1\ 5\ 4\ 5\ 3\ 2\ 1\ 3\ 4\ 1\ 2 \text{ Symbol} \\ 2\ 3\ 4\ 5\ 1\ 3\ 4\ 5\ 1\ 2\ 4\ 5\ 1\ 2\ 3\ 5\ 1\ 2\ 3\ 4 \text{ Column} \\ 1\ 1\ 1\ 1\ 2\ 2\ 2\ 2\ 3\ 3\ 3\ 3\ 4\ 4\ 4\ 4\ 5\ 5\ 5\ 5 \text{ Row} \end{array}$$

Two Latin squares of the same order are said to be *orthogonal* if no pair of corresponding elements occurs more than once when one square is superimposed onto the other. A set of Latin squares of the same order is *mutually orthogonal* if every pair of Latin squares from the set is orthogonal. In order to construct TBTCI designs with $p \geq 4$ rows, we need $p - 2$

mutually orthogonal Latin squares with a common transversal. In practical implementation, it is more convenient to start with $p-1$ mutually orthogonal Latin squares if they exist. Then one of the Latin squares will be sacrificed to locate a common transversal of other $p-2$ Latin squares. Each position could produce a p-dimension column vector with entries filled with the corresponding row label, column label, and $p-2$ symbols. Then the juxtaposition of these vectors corresponding to all of the off-diagonal positions will yield a type I orthogonal array. Here is an example of constructing an OA_I with $p = t = 5$, via the following 4 mutually orthogonal Latin squares of order 5:

$$L_1: \begin{array}{ccccc} 1&2&3&4&5\\ 2&3&4&5&1\\ 3&4&5&1&2\\ 4&5&1&2&3\\ 5&1&2&3&4 \end{array} \quad L_2: \begin{array}{ccccc} 1&2&3&4&5\\ 3&4&5&1&2\\ 5&1&2&3&4\\ 2&3&4&5&1\\ 4&5&1&2&3 \end{array} \quad L_3: \begin{array}{ccccc} 1&2&3&4&5\\ 4&5&1&2&3\\ 2&3&4&5&1\\ 5&1&2&3&4\\ 3&4&5&1&2 \end{array} \quad L_4: \begin{array}{ccccc} 1&2&3&4&5\\ 5&1&2&3&4\\ 4&5&1&2&3\\ 3&4&5&1&2\\ 2&3&4&5&1 \end{array}$$

L_4 has all of the positions of the main diagonal filled with symbol 1. Automatically, L_1, L_2 and L_3 have the main diagonal as a common transversal and we could rename the symbols in the first three squares in independent ways to obtain the following three mutually orthogonal Latin squares of order 5 with a common transversal on the diagonal:

$$L_1': \begin{array}{ccccc} 1&4&2&5&3\\ 4&2&5&3&1\\ 2&5&3&1&4\\ 5&3&1&4&2\\ 3&1&4&2&5 \end{array} \quad L_2': \begin{array}{ccccc} 1&3&5&2&4\\ 5&2&4&1&3\\ 4&1&3&5&2\\ 3&5&2&4&1\\ 2&4&1&3&5 \end{array} \quad L_3': \begin{array}{ccccc} 1&5&4&3&2\\ 3&2&1&5&4\\ 5&4&3&2&1\\ 2&1&5&4&3\\ 4&3&2&1&5 \end{array}$$

When we go through all of the off-diagonal positions, we obtain an $OA_I(20,5,5,2)$:

$$d_4: \begin{array}{l} 4\ 2\ 5\ 3\ 4\ 5\ 3\ 1\ 2\ 5\ 1\ 4\ 5\ 3\ 1\ 2\ 3\ 1\ 4\ 2\ L'_1 \\ 3\ 5\ 2\ 4\ 5\ 4\ 1\ 3\ 4\ 1\ 5\ 2\ 3\ 5\ 2\ 1\ 2\ 4\ 1\ 3\ L'_2 \\ 5\ 4\ 3\ 2\ 3\ 1\ 5\ 4\ 5\ 4\ 2\ 1\ 2\ 1\ 5\ 3\ 4\ 3\ 2\ 1\ L'_3 \\ 2\ 3\ 4\ 5\ 1\ 3\ 4\ 5\ 1\ 2\ 4\ 5\ 1\ 2\ 3\ 5\ 1\ 2\ 3\ 4\ Column \\ 1\ 1\ 1\ 1\ 2\ 2\ 2\ 2\ 3\ 3\ 3\ 3\ 4\ 4\ 4\ 4\ 5\ 5\ 5\ 5\ Row \end{array}$$

Any $q \times 20$ subarray of the above array would consists of an $OA_I(20,q,5,2)$ for $q < 5$. Theorem 2.4 below and its related proof could be concluded from the above process of construction.

Theorem 6. *We can always construct an $OA_I(t(t-1),p,t,2)$ for all $p \leq m+1$ when there exits m mutually orthogonal Latin squares of order t.*

Remark 3. *The number of mutually orthogonal Latin squares of order t is at most $t-1$ and this upper bound could be reached whenever t is a prime power.*

All of the designs constructed in this subsection will have the relationship of $n = t(t-1)$, which means $l = t-1$, and thus $m = p-1$ by (4.4). Therefore, the designs constructed here would be of minimal size if and only if $t-1$ and $p-1$ are relatively prime. When $t-1$ and $p-1$ have common factors, the following two methods will provide some answers for the constrution of the designs of minimal size.

4.2.3 Method 2 of Constructing TBIC Designs

A design with t treatments, n columns (blocks) and p rows is said to be a balanced incomplete block (BIB) design, denoted by $BIB(t,n,p)$, if (i) $p < t$ (ii) $n_{diu} = 0$ or 1 for $1 \leq i \leq t$ and

$1 \leq u \leq n$ (iii) $\sum_{u=1}^{n} n_{diu}n_{dju}$ is a constant, denoted by ν, across all $1 \leq i \neq j \leq t$. For the treatment i, we have $r_{di}(p-1) = \nu(t-1)$, thus $r_{di} = r_{dj}$ for all $i \neq j$. So we can denote r_{di} by the same r for any i, and hence

$$r(p-1) = \nu(t-1). \tag{4.6}$$

Another equality we have is $np = rt$. One significant feature of BIB designs compared to previous designs is that, these designs are invariant to rearranging the positions of treatments within each column. However, columns of $BIB(t, n, p)$ coupled with a CCLS(p) would give a $TBIC(t, np, p)$ as follows:

Without loss of generality, we assume the treatments of a BIB design are denoted by $1, 2, ...t$. Suppose the ith column of the $BIB(t, n, p)$ contains treatments from $\{a_1, ..., a_p\} \subset \{1, 2, ...t\}$. We use these symbols to construct a $CCLS(p)$, denoted by LS_i. Then the juxtaposition of LS_1 to LS_n will give a $TBIC(t, np, p)$.

For the TBIC designs constructed above, we have $m = \nu$ and $l = r$ according to the process of construction. Then (4.4) holds either due to the fact that it is a TBIC design or based on (4.6). Meanwhile, whenever there is a $BIB(t, n, p)$ in which $\nu = np(p-1)/(t(t-1))$ and $r = np/t$ are relatively prime, the corresponding TBIC design would be of minimal size. A comprehensive list of the existence of BIB designs with different parameters could be found in the Handbook of Combinatorial Designs edited by Colbourn and Dinitz (2007).

4.2.4 Method 3 of Constructing TBIC Designs

The following theorem gives a sufficient condition for the existence of one type of TBIC designs of minimal size. The details of construction is included in the proof of the theorem.

Theorem 7. *A TBIC(t,n,p) of minimal size (i.e. m and l in (2.1) are relatively prime) exists if there exists l vectors of the form $(a_{i1}, a_{i2}, ..., a_{i,p-1}), i = 1, 2, ..., l$ with $a_{ij} \in \{1, 2, ..., t-1\}$ such that*

1. 0 *is not contained in the collection* $\sum_{j=j_1}^{j_2} a_{ij}$ *(mod t),* $1 \le j_1 \le j_2 \le p-1$ *and* $1 \le i \le l$.

2. *Each number from* $\{1, 2, ..., t-1\}$ *appears* m *times in the collection* a_{ij}, $1 \le j_1 \le j_2 \le p-1$ *and* $1 \le i \le l$.

3. *Each number from* $\{1, 2, ..., t-1\}$ *appears the same number of times in the collection* $\pm \sum_{j=j_1}^{j_2} a_{ij}$ *(mod t),* $1 \le j_1 \le j_2 \le p-1$ *and* $1 \le i \le l$.

4. *Each number from* $\{1, 2, ..., t-1\}$ *appears the same number of times in the collection* $\pm \sum_{j=j_1}^{p-1} a_{ij}$ *(mod t),* $1 \le j_1 \le p-1$ *and* $1 \le i \le l$.

Proof. Based on the ith vector $(a_{i1}, a_{i2}, ..., a_{i,p-1})$, we can construct a $p \times t$ array, $B_i = (b^i_{jk})$, in which b^i_{jk} is the entry in the jth row and the kth column of B_i, such that the first row $(b^i_{11}, b^i_{12}, ..., b^i_{1t})$ is any permutation of $\{0, 1, ..., t-1\}$ and $b^i_{jk} = b^i_{1k} + \sum_{h=1}^{j-1} a_{ih}$ (mod t) for $2 \le j \le p$. Then we can obtain a $TBIC(t, lt, p)$ by juxtaposing B_1 to B_l and map the treatments in $\{0, 1, ..., t-1\}$ to $\{1, 2, ..., t\}$ in any injective manner. Specifically, Condition (1) here implies the Condition (1) in Definition 6; Condition (2) here implies the Condition (3) in

Definition 6; Conditions (3) and (4) here imply the Condition (4) in Definition 6. The Condition (2) therein is satisfied since each row of each B_i is simply a permutation of the treatments. $\diamondsuit$

Here we illustrate the idea of construction in Theorem 7 for the case of $p = 4$ and $t = 7$. Since $2(p-1) = t-1$, we can construct a TBIC of minimal size, i.e. $m = 1$ and $l = 2$, if the sufficient conditions in the theorem exists. In fact, the vectors $(a_{11}, a_{12}, a_{13}) = (1, 2, 3)$ and $(a_{21}, a_{22}, a_{23}) = (6, 5, 4)$ satisfies all the four conditions in the theorem. Based on those two vectors, we will have the following arrays respectively:

0	1	2	3	4	5	6		0	1	2	3	4	5	6
1	2	3	4	5	6	0		6	0	1	2	3	4	5
3	4	5	6	0	1	2		4	5	6	0	1	2	3
6	0	1	2	3	4	5		1	2	3	4	5	6	0

It is easy to verify by Definition 6 that we will obtain a TBIC design if we juxtapose these two designs and replace the symbol 0 by 7.

4.2.5 Examples

Since we could make the statistical inference as precise as possible by increasing the number of runs, it is only fair when we compare $Tr(M_d^{-1})$ of designs with the same values of t, n, p. Unfortunately, there may not be any design guaranteed to be optimal since different θ will require different values of r_{d0} for a TBTCI design to be optimal. For given values of n, p, t and θ, Figures 2-6 give some idea about what value of r_{d0} should a TBTCI design possess to be an optimal design. Indeed, for a complete comparison between designs, we need to compare their efficiencies at all values of θ.

For each TBTCI design d constructed in this chapter, the conditions $C1 : r_{d0} < n$, $C2 : r_{d0} = n$, and $C3 : r_{d0} > n$ imply $n_{d0u} \leq 1$, $n_{d0u} = 1$ and $n_{d0u} \geq 1$ respectively for $1 \leq u \leq n$. Based on Theorem 5, the juxtaposition of any two TBTCI designs would still be a TBTCI design except when one of them satisfies Condition $C1$ and the other satisfies Condition $C3$. In general, the juxtaposition of finite many TBTCI designs would still be a TBTCI design as long as at least one of the conditions $C1$ and $C3$ is not satisfied by any of these designs.

Suppose now that we want to construct efficient designs with 4 periods and 3 test treatments. Based on the first picture in Figure Figure 6, we need TBTCI designs with r_{d0}/n slightly greater than 1. d_2 in Section 4.2.1 as a $TBTCI_{3,4}(12, 24)$ satisfies $r_{d0}/n = 2$, and this number is obviously too big. On the other hand, a $CCLS(4)$ exists since 4 is a composite number. By relabelling, we immediately obtain a $TBTCI_{3,4}(4, 4)$ with $r_{d0}/n = 1$ from the $CCLS(4)$, and this number is ideal for some cases of t, n, p, θ, but too small for most of the cases. With multiple copies of the latter design as well as $0, 1$ and 2 copies of the former design, we can have $TBTCI_{3,4}(224, 224)$, $TBTCI_{3,4}(224, 236)$ and $TBTCI_{3,4}(224, 248)$ respectively. Figure Figure 10 gives a comparison of these three designs:

Since the range of θ is $[0, \infty)$ which is hard to cover in a figure, we use the monotone transformation of $\lambda \equiv \lambda_1 = \theta/(1 + \theta) \in [0, 1]$. As expected from Figure 3, $TBTCI_{3,4}(224, 224)$, $TBTCI_{3,4}(224, 236)$ and $TBTCI_{3,4}(224, 248)$ wins in turn when θ is large, moderate, and small respectively. By providing a complete comparison for all possible values of θ, practitioners could decide which design to choose according to their priori knowledge of θ. If no priori knowledge is available, robust criteria such as $\max_d \min_\theta Eff(d, \theta)$ with $Eff(d, \theta) =$

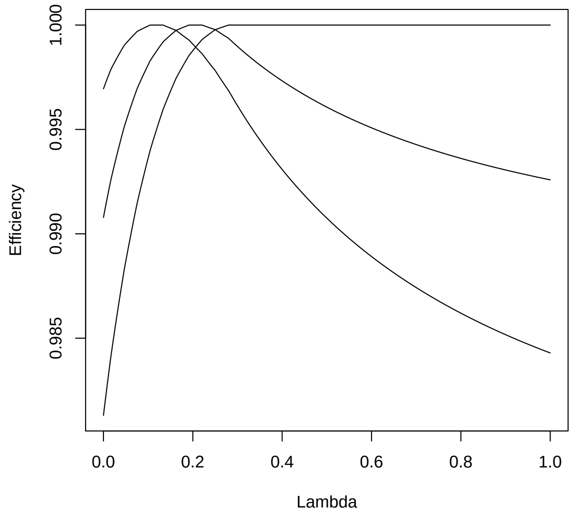

Figure 10. From left: $TBTCI_{3,4}(224, 248)$, $TBTCI_{3,4}(224, 236)$ and $TBTCI_{3,4}(224, 224)$

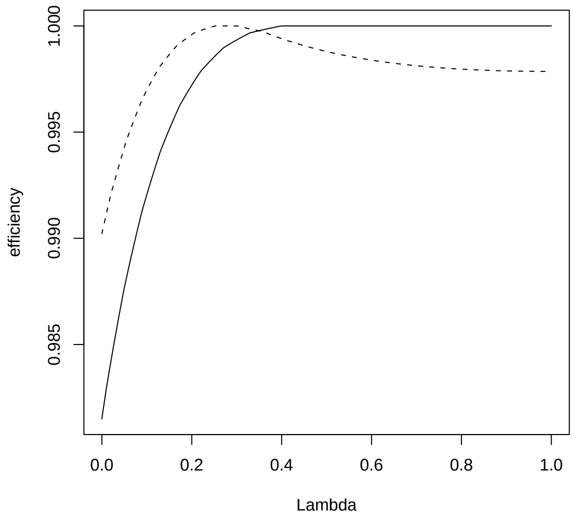

Figure 11. $TBTCI_{4,3}(180, 180)$ (solid curve) and $TBTCI_{4,3}(200, 192)$ (dashed curve)

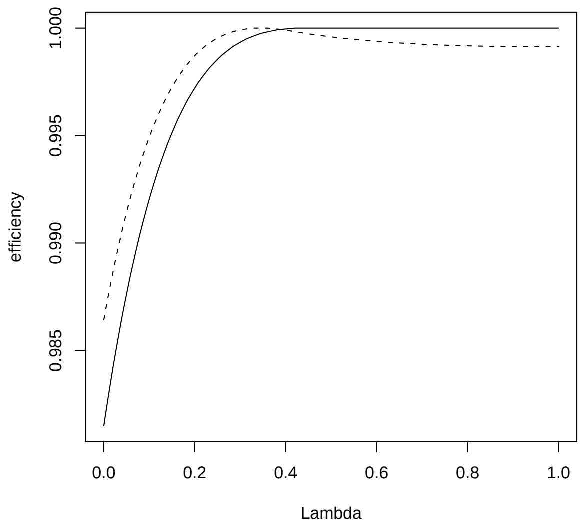

Figure 12. $TBTCI_{4,3}(360, 360)$ (solid curve) and $TBTCI_{4,3}(380, 272)$ (dashed curve)

$\min_{d' \in \Lambda_{t+1,n,p}} Tr(M_{d'}^{-1})/Tr(M_d^{-1})$ could be proposed. Based on this robust criterion, $TBTCI_{3,4}(224, 236)$ will be the best choice.

When $p = 3$ and $t = 4$, we can start with $TBTCI_{4,3}(20, 12)$ derived from d_3 and $TBTCI_{4,3}(36, 36)$ as given by Hedayat and Yang (2005). The juxtaposition of copies of these two designs will give rise to a $TBTCI_{4,3}(180, 180)$ and a $TBTCI_{4,3}(200, 192)$. The comparison is given by Figure Figure 11. Note that the two designs in comparison have different numbers of columns. In calculating the efficiencies, the minimization in $\min_d l(t, n, p, \theta, r_{d0})$ should be taken within $\Lambda_{5,180,3}$ and $\Lambda_{5,200,3}$. However, this difference in n is small enough to make the comparison meaningful. Similarly, using $TBTCI_{4,3}(20, 12)$ and $TBTCI_{4,3}(36, 36)$, we can construct $TBTCI_{4,3}(380, 372)$ and $TBTCI_{4,3}(360, 360)$. Figure Figure 12 depicts the comparison.

CHAPTER 5

CONCLUSION

In mixed linear model with random subject effects, the ratio (θ) of the variance of the subject effect to the variance of the error plays a crucial role in deciding which design should be used. Note that in estimating the treatment effects, $\theta = 0$ corresponds to the model without the subject effects while $\theta = \infty$ corresponds to the model with fixed subject effects. For the latter model, Hedayat and Yang (2005) found optimal designs in the subclass Λ and conjectured that Λ will not exclude too many good designs so that optimal designs in this subclass is still highly efficient or even optimal in Ω. As for how efficient these designs will be, they didn't carry out the investigation.

In this thesis, I dealt with general $\theta \geq 0$ and found optimal designs in Λ, which naturally covered the result in Hedayat and Yang (2005) as a special case when $\theta = \infty$. Moreover, we found optimal designs in Ω when $\theta = 0$. Further, I gave the algorithm to calculate the lower bound of $Tr(M_d^{-1})$, which turns out to be very close to the value of optimal designs. Hence it is proper to use this lower bound to evaluate the efficiencies of the designs proposed. Specifically, I evaluated the behavior of two types of designs. One is the optimal designs when $\theta = 0$, whose efficiency decreases dramatically as θ increases. The other is the TBTCI designs with properly chosen value of r_{d0}, which are shown to be robust across different values of θ. Since we usually don't know the value of θ in application, this robustness is critical. Also, it is gratifying that

TBTCI designs commonly exist, while the optimal designs for $\theta = 0$ is not so common to exist due to the first condition of Theorem 1 in Section 2.3.1.

In studying universal optimality, Hedayat, Stufken, and Yang (2006) proved the high efficiency of the totally balanced designs, which is essentially a special type of TBTCI designs with the property of $r_{d0} = r_{d1}$. Particularly, they proved the universal optimality of the totally balanced designs in $\Omega^2 = \{d \in \Omega | l_{dip} = n/(t+1), m_{dii} = 0, i = 0, 1, ..., t\}$, which is certainly true for A- and MV- optimality. However, as pointed out by one of the referees, optimality frameworks based on functions of M_d result in placing overriding weight on the control and hence higher control replication, which is shown in Section 2.3 and Chapter 3. Actually, the optimality of the totally balanced designs in Ω^2 established by Hedayat, Stufken, and Yang (2006) could be explained by the fact that the class Ω^2 rules out any designs with unequal replications of treatments. Also, the high efficiency of the totally balanced designs in Ω established therein based on the trace approach could not be carried over to A- and MV- criteria.

At the same time, we observe from Figure Figure 3 that the gap between optimal designs in Λ and Ω is more obvious for small values of θ. I conjecture that this is due to the restriction of $m_{dii} = 0, i = 0, 1, ..., t$ in view of Condition 1 in Theorem 1 in Section 2.3.1. When $\theta = \infty$, Hedayat and Yang (2006) and Yang and Park (2007) extended the class of competing designs from Λ to $\Omega^1 = \{d \in \Omega | l_{d0k} = r_{d0}/p, k = 1, 2, ..., p\}$ when either of the following conditions satisfies (i) $p = 3$ and $3 \leq t \leq 20$; or (ii) $p \geq 4$, $(p-3)(p-2) + 2 \leq t \leq (p-2)(p-1) + 1$ and $n \geq p(p-1)/2$. They proposed designs allowing part of the subjects in the study to have identical treatments in the last two periods, which are A-better than TBTCI designs. However, we note that Conditions (i) and (ii) here imposed on n, t, p are quite restrictive, therefore it would be essential if we could find optimal designs in Ω^1 or even Ω for wider ranges of the parameters n, t, p. Another possible direction for future research is to study the case when $p > t + 1$, for which we have not seen any work carried out yet. I believe some designs with similar structures as TBTCI designs would be highly efficient in this case.

The class of totally balanced test-control incomplete (TBTCI) Crossover designs proved to be an important class of designs for comparing two or more test treatments with a standard control treatment. The author is the first to seriously consider the construction of these designs. Our approach is to derive these designs from a simpler class of designs with no control treatment involved, which is called totally balanced Crossover (TBIC) designs. TBIC designs are equivalent to a special type of TBTCI designs. More importantly, TBIC designs could be used as building blocks to construct TBTCI designs. Thus, various methods of constructing TBIC designs were presented. In the process, it is shown that the concepts of type I orthogonal arrays, complete column Latin squares, balanced incomplete block designs and finite group were closely related to TBIC designs. Following is some future research topics to be investigated.

In this thesis, the construction of TBTCI designs relies heavily on the existence of TBIC designs. We would like to point out that there is a lot of examples in Hedayat and Yang (2005) which have nothing to do with TBIC designs. However, those examples are derived by computer search and lack generalization to other configuration of t, n and p. The work on these designs will have the advantage that n could be smaller. This is very important due to Theorem 5. References related to this topic include Mendelsohn (1968), Dey (1986) and Linder and Rodger

(1997) among others. In constructing TBIC designs, Theorem 6 gives sufficient conditions for the existence of TBIC deigns of minimal size and the corresponding methods of construction. However, these sufficient conditions are complex themselves and need to be further investigated and simplified.

The optimality of TBTCI designs are established within $\Lambda_{t+1,n,p}$, a subclass of $\Omega_{t+1,n,p}$. Though there is evidence to indicate that TBTCI designs are also highly efficient or optimal among $\Omega_{t+1,n,p}$, alternative designs need to be investigated for two reasons. (1) TBTCI designs do not exist for some configurations of t, n and p. Actually, the nonexistence is more common than the existence. (2) There are better designs for some particular values of t, n and p in terms of A-optimality or MV-optimality, and these designs are actually close to TBTCI designs in structure. Correspondingly, we have two ways to go. One approach is to go for precise mathematical discovery and an alternative approach is to apply some algorithms such as Genetic Algorithm (GA) to search for desirable designs. Suppose a practitioner is in an immediate need of a design under t, n, and p for which proper TBTCI designs does not exist or we do not know the existence for these parameters. A prudent approach to help this practitioner will be to find a TBTCI design d with number of columns $n_1 (> n)$ and generate a population of designs by using n columns of d in as many ways we want. Then we can use GA together with judiciously selected genetic operations such as mating and mutation and select a good design for the practitioner.

APPENDICES

Appendix A

PROOFS OMITTED IN SECTION 2.4

Proof of Lemma 4. Since (ii) is a direct result of (i), it is sufficient to prove (i) only. Let $N = t!$, $Q = (0|I_t)'$ and $\bar{M}_d = \sum_{i=1}^N S_i' M_d S_i / N$. By convexity of A-criterion, we have

$$Tr(M_d^{-1}) \geq Tr(\bar{M}_d^{-1}). \tag{A.1}$$

Since $S_i' M_d S_i = S_i' Q' C_d Q S_i = Q' \tilde{S}_i' C_d \tilde{S}_i Q$, we have

$$N\bar{M}_d = Q' \left(\sum_{i=1}^N \tilde{S}_i' C_d \tilde{S}_i \right) Q.$$

Then by Proposition 1 of Kunert and Martin (2000) and Lemma 3, we have

$$\begin{aligned} \sum_{i=1}^N \tilde{S}_i' C_d \tilde{S}_i \leq & \left(\sum_{i=1}^N \tilde{S}_i' T_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} T_d \tilde{S}_i \right) \\ & - \left(\sum_{i=1}^N \tilde{S}_i' T_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} F_d \tilde{S}_i \right) \\ & \times \left(\sum_{i=1}^N \tilde{S}_i' F_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} F_d \tilde{S}_i \right)^{-} \\ & \times \left(\sum_{i=1}^N \tilde{S}_i' F_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} T_d \tilde{S}_i \right), \end{aligned} \tag{A.2}$$

with the equalities in (A.1) and (A.2) both hold when

Appendix A (Continued)

(i) $l_{dik} = r_{di}/p, i = 0, 1, ..., t.$

(ii) $T_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} T_d$, $T_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} F_d$ and $F_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} F_d$ are invariant under any permutation of test treatments.

Using the equality

$$V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} = I_{np} - \frac{\theta}{1+\theta p} I_n \otimes J_{p\times p} - \frac{1}{(1+\theta p)pn} J_{np\times np}, \tag{A.3}$$

it's easy to see that the matrices inside the last three parentheses in (A.2) have the same form

$$\begin{pmatrix} a_m & f_m 1_t' \\ c_m 1_t & (b_m - e_m) I_t + e_m J_t \end{pmatrix}, m = 1, 2, 3.$$

For $\sum_{i=1}^{N} \tilde{S}_i' T_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} T_d \tilde{S}_i$,

$$a_1 = N\left(r_{d0} - \frac{\theta}{1+\theta p}\sum_{u=1}^{n} n_{d0u}^2 - \frac{1}{(1+\theta p)pn} r_{d0}^2\right)$$

$$b_1 = \frac{N}{t}\left(np - r_{d0} - \frac{\theta}{1+\theta p}\sum_{i=1}^{t}\sum_{u=1}^{n} n_{diu}^2 - \frac{1}{(1+\theta p)pn}\sum_{i=1}^{t} r_{di}^2\right)$$

$$c_1 = f_1 = -\frac{a_1}{t}$$

$$e_1 = -\frac{b_1}{(t-1)} + \frac{a_1}{t(t-1)}$$

Appendix A (Continued)

and for $\sum_{i=1}^{N} \tilde{S}_i' T_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} F_d \tilde{S}_i$,

$$\begin{aligned}
a_2 &= N\left(m_{d00} - \frac{\theta}{1+\theta p}\sum_{u=1}^{n} n_{d0u}\tilde{n}_{d0u} - \frac{r_{d0}\tilde{r}_{d0}}{(1+\theta p)pn}\right)\\
b_2 &= \frac{N}{t}\left(\sum_{i=1}^{t} m_{dii} - \frac{\theta}{1+\theta p}\sum_{i=1}^{t}\sum_{u=1}^{n} n_{diu}\tilde{n}_{diu} - \frac{1}{(1+\theta p)pn}\sum_{i=1}^{t} r_{di}\tilde{r}_{di}\right)\\
c_2 &= -\frac{a_2}{t}\\
f_2 &= \frac{N}{t}\left(\frac{r_{d0}}{p} - l_{d01} - m_{d00} + \frac{\theta}{1+\theta p}\sum_{u=1}^{n} n_{d0u}\tilde{n}_{d0u} + \frac{r_{d0}\tilde{r}_{d0}}{(1+\theta p)pn}\right)\\
e_2 &= -\frac{b_2+f_2}{t-1}
\end{aligned}$$

and for $\sum_{i=1}^{N} \tilde{S}_i' F_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} F_d \tilde{S}_i$,

$$\begin{aligned}
a_3 &= N\left(\tilde{r}_{d0} - \frac{\theta}{1+\theta p}\sum_{u=1}^{n} \tilde{n}_{d0u}^2 - \frac{\tilde{r}_{d0}^2}{(1+\theta p)pn}\right)\\
b_3 &= \frac{N}{t}\left(n(p-1) - \tilde{r}_{d0} - \frac{\theta}{1+\theta p}\sum_{i=1}^{t}\sum_{u=1}^{n} \tilde{n}_{diu}^2 - \frac{1}{(1+\theta p)pn}\sum_{i=1}^{t} \tilde{r}_{di}^2\right)\\
c_3 = f_3 &= \frac{N}{t}\left(-\frac{p-1}{p}\tilde{r}_{d0} + \frac{\theta}{1+\theta p}\sum_{u=1}^{n} \tilde{n}_{d0u}^2 + \frac{\tilde{r}_{d0}^2}{(1+\theta p)pn}\right)\\
e_3 &= \frac{N}{t(t-1)}\Bigg\{-\frac{n}{p}(p-1)^2 + \frac{2(p-1)}{p}\tilde{r}_{d0} + \frac{\theta}{1+\theta p}\left(\sum_{i=1}^{t}\sum_{u=1}^{n} \tilde{n}_{diu}^2 - \sum_{u=1}^{n} \tilde{n}_{d0u}^2\right)\\
&\qquad + \frac{1}{(1+\theta p)pn}\left(\sum_{i=1}^{t} \tilde{r}_{di}^2 - \tilde{r}_{d0}^2\right)\Bigg\}.
\end{aligned}$$

Appendix A (Continued)

Notice that if $\sum_{i=1}^{t}\sum_{u=1}^{n}\tilde{n}_{diu}^2$ is decreased by some amount, the increment for the matrix $\sum_{i=1}^{N}\tilde{S}_i'F_d'V^{-1/2}pr^{\perp}(1_{np})V^{-1/2}F_d\tilde{S}_i$ is proportional to

$$\begin{pmatrix} 0 & 0 \\ 0 & pr^{\perp}(1_t) \end{pmatrix}$$

which is a nonnegative definite matrix. Thus we would like to minimize $\sum_{i=1}^{t}\sum_{u=1}^{n}\tilde{n}_{diu}^2$ over $d \in \Omega_{t+1,n,p}$. The minimum is attained when $\tilde{n}_{diu}$ is binary for $i > 0$ and the corresponding minimum is $n(p-1) - \tilde{r}_{d0}$. Similarly, we would like to replace $\sum_{i=1}^{t}\sum_{u=1}^{n} n_{diu}^2$ therein by its own minimum over $d \in \Omega_{t+1,n,p}$. The minimum of the latter is $np - r_{d0}$ and is obtained when n_{dij} is binary for $i > 0$. Four values with relevant adjustments are shown below:

$$\begin{aligned}
\tilde{b}_1 =& \frac{N}{t}\left(\left(1-\frac{\theta}{1+\theta p}\right)(np - r_{d0}) - \frac{1}{(1+\theta p)pn}\sum_{i=1}^{t} r_{di}^2\right)\\
\tilde{e}_1 =& -\frac{\tilde{b}_1}{(t-1)} + \frac{a_1}{t(t-1)}\\
\tilde{b}_3 =& \frac{N}{t}\left(\left(1-\frac{\theta}{1+\theta p}\right)(n(p-1) - \tilde{r}_{d0}) - \frac{1}{(1+\theta p)pn}\sum_{i=1}^{t} \tilde{r}_{di}^2\right)\\
\tilde{e}_3 =& \frac{N}{t(t-1)}\left\{\frac{n}{p}(p-1)^2 + \left(\frac{\theta}{1+\theta p} - \frac{2(p-1)}{p}\right)(n(p-1) - \tilde{r}_{d0})\right.\\
&\left. - \frac{\tilde{r}_{d0}^2}{(1+\theta p)pn} - \frac{\theta}{1+\theta p}\sum_{u=1}^{n}\tilde{n}_{d0u}^2 + \frac{1}{(1+\theta p)pn}\sum_{i=1}^{t}\tilde{r}_{di}^2\right\}.
\end{aligned}$$

It can be verified that $a_3 > 0, \tilde{b}_3 - \tilde{e}_3 > 0$ and $a_3\tilde{b}_3 + (t-1)a_3\tilde{e}_3 - tc_3^2 > 0$. So the updated $\sum_{i=1}^{N}\tilde{S}_i'F_d'V^{-1/2}pr^{\perp}(1_{np})V^{-1/2}F_d\tilde{S}_i$, with b_3 and e_3 replaced by $\tilde{b}_3$ and $\tilde{e}_3$ respectively, is a

Appendix A (Continued)

positive definite matrix and its inverse has the same form as itself when $a_3, \tilde{b}_3, c_3$, and $\tilde{e}_3$ are replaced by a_4, b_4, c_4, and e_4. Here

$$\begin{aligned}
a_4 &= \frac{\tilde{b}_3 + (t-1)\tilde{e}_3}{a_3\tilde{b}_3 + (t-1)a_3\tilde{e}_3 - tc_3^2} \\
b_4 &= \frac{a_3\tilde{b}_3 + (t-2)a_3\tilde{e}_3 - (t-1)c_3^2}{(\tilde{b}_3 - \tilde{e}_3)(a_3\tilde{b}_3 + (t-1)a_3\tilde{e}_3 - tc_3^2)} \\
c_4 &= \frac{-c_3}{a_3\tilde{b}_3 + (t-1)a_3\tilde{e}_3 - tc_3^2} \\
e_4 &= \frac{c_3^2 - a_3\tilde{e}_3}{(\tilde{b}_3 - \tilde{e}_3)(a_3\tilde{b}_3 + (t-1)a_3\tilde{e}_3 - tc_3^2)}.
\end{aligned}$$

The related inverse matrix can be expressed as $D + c_4 J_{t+1}$, where

$$D = \begin{pmatrix} a_4 - c_4 & 0 \\ 0 & (b_4 - e_4)I_t + (e_4 - c_4)J_t \end{pmatrix}.$$

Note that $c_4 \geq 0$ due to $c_3 \leq 0$, we have

$$\begin{aligned}
&\left(\sum_{i=1}^{N} \tilde{S}_i' T_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} F_d \tilde{S}_i\right) \\
&\times \left(\sum_{i=1}^{N} \tilde{S}_i' F_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} F_d \tilde{S}_i\right)^{-} \left(\sum_{i=1}^{N} \tilde{S}_i' F_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} T_d \tilde{S}_i\right) \\
&\geq \left(\sum_{i=1}^{N} \tilde{S}_i' T_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} F_d \tilde{S}_i\right) D \left(\sum_{i=1}^{N} \tilde{S}_i' F_d' V^{-1/2} pr^{\perp}(1_{np}) V^{-1/2} T_d \tilde{S}_i\right). \qquad \text{(A.4)}
\end{aligned}$$

Appendix A (Continued)

The equality in (A.4) will hold when $l_{di1} = r_{di}/p, i = 0, 1, ..., t$. By (A.1),(A.2) and (A.4), we have

$$Tr(M_d^{-1}) \geq Tr(\bar{M}_d^{-1}) \geq Tr(\tilde{M}_d^{-1}) \tag{A.5}$$

with both equalities hold when the three conditions in this lemma hold. Here $\tilde{M}_d = xI + yJ$, where

$$\begin{aligned} x =& \frac{1}{N}\left(\tilde{b}_1 - \tilde{e}_1 - (b_2 - e_2)^2(b_4 - e_4)\right) \\ y =& \frac{1}{N}\left(\tilde{e}_1 - c_2^2(a_4 - c_4) - e_2(b_2 - e_2 - f_2)(b_4 - e_4) - f_2^2(e_4 - c_4)\right). \end{aligned}$$

Since $\tilde{M}_d$ has eigenvalues of x with multiplicity $t-1$ and $x+ty$ with multiplicity 1, it is enough to prove $x_0 = t(t-1)x$ and $y_0 = t(x+ty)$. By direct calculation, we have

$$\begin{aligned} b_2 - e_2 =& (tb_2 + f_2)/(t-1) \\ =& \frac{N}{t(t-1)}\Bigg\{t\sum_{i=1}^{t} m_{dii} + \frac{r_{d0}}{p} - l_{d01} - m_{d00} - \frac{\theta t}{1+\theta p}\sum_{i=1}^{t}\sum_{u=1}^{n} n_{diu}\tilde{n}_{diu} \\ & - \frac{t}{(1+\theta p)pn}\sum_{i=1}^{t} r_{di}\tilde{r}_{di} + \frac{\theta}{1+\theta p}\sum_{u=1}^{n} n_{d0u}\tilde{n}_{d0u} + \frac{r_{d0}\tilde{r}_{d0}}{(1+\theta p)pn}\Bigg\}, \end{aligned} \tag{A.6}$$

Appendix A (Continued)

$$
\begin{aligned}
x+ty =& \frac{1}{N}\left(\tilde{b}_1+(t-1)\tilde{e}_1-tc_2^2(a_4-c_4)-tf_2^2(e_4-c_4)-f_2^2(b_4-e_4)\right)\\
=& \frac{\tilde{a}_1}{Nt}-\frac{tc_2^2(\tilde{b}_3+(t-1)\tilde{e}_3+c_3)+f_2^2(a_3+tc_3)}{N[a_3\tilde{b}_3+(t-1)a_3\tilde{e}_3-tc_3^2]}\\
=& \frac{1}{t}\left(r_{d0}-\frac{\theta}{1+\theta p}\sum_{u=1}^{n}n_{d0u}^2-\frac{r_{d0}^2}{(1+\theta p)pn}\right)\\
& -\frac{1}{t}\Bigg\{(n(p-1)-\tilde{r}_{d0})\left(m_{d00}-\frac{\theta}{1+\theta p}\sum_{u=1}^{n}n_{d0u}\tilde{n}_{d0u}-\frac{1}{(1+\theta p)pn}r_{d0}\tilde{r}_{d0}\right)^2\\
& +\tilde{r}_{d0}\left(\frac{r_{d0}}{p}-l_{d01}-m_{d00}+\frac{\theta}{1+\theta p}\sum_{u=1}^{n}n_{d0u}\tilde{n}_{d0u}+\frac{1}{(1+\theta p)pn}r_{d0}\tilde{r}_{d0}\right)^2\Bigg\}\\
& \times\left\{n(p-1)\left(\tilde{r}_{d0}-\frac{\theta}{1+\theta p}\sum_{u=1}^{n}\tilde{n}_{d0u}^2-\frac{\tilde{r}_{d0}^2}{(1+\theta p)pn}\right)-\frac{\tilde{r}_{d0}^2}{p}\right\}^{-1}.
\end{aligned}
\tag{A.7}
$$

the lemma follows by noting $b_4-e_4=1/(\tilde{b}_3-\tilde{e}_3)$ and $\tilde{b}_1-\tilde{e}_1=(t^2\tilde{b}_1-a_1)/(t(t-1))$. $\diamondsuit$

The following propositions would be helpful in establishing (2.29) in Lemma 11. The notations in Lemmas 10 and 11 would be adopted.

Proposition 1. *When $\xi_2=\xi_3=\tilde{r}_{d0}$ and (2.29) holds for t_0, (2.29) will still hold for any $t>t_0$ while other parameters keep unchanged.*

Proof. First of all, (2.29) is equivalent to the following two conditions:

$$\tilde{x}_0-(t-1)y_0\geq 0 \tag{A.8}$$

$$\tilde{x}_0\Theta_2-(t-1)y_0\Theta_1\geq 0. \tag{A.9}$$

Appendix A (Continued)

By direct calculation, we have

$$\begin{aligned}
\frac{\partial \tilde{x}_0}{\partial t} &= \left(1-\frac{\theta}{1+\theta p}\right)(np-r_{d0}) - 2\frac{\theta}{1+\theta p}(n(p-1)-\tilde{r}_{d0})\frac{\tilde{\beta}_1}{\tilde{\gamma}_1} \\
&\quad + \left(1-\frac{\theta}{1+\theta p}\right)(n(p-1)-\tilde{r}_{d0})\left(\frac{\tilde{\beta}_1}{\tilde{\gamma}_1}\right)^2 \\
&= \left(1-\frac{\theta}{1+\theta p}\right)(n(p-1)-\tilde{r}_{d0})\left(\left(\frac{\tilde{\beta}_1}{\tilde{\gamma}_1}-\frac{\theta}{1+\theta p-\theta}\right)^2-\left(\frac{\theta}{1+\theta p-\theta}\right)^2+\frac{p}{p-1}\right) \\
&\geq 0
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial \Theta_1}{\partial t} &= \frac{n(p-1)-\tilde{r}_{d0}}{\tilde{\gamma}_1^2}\left(\frac{\theta}{1+\theta p}\tilde{\gamma}_1-\left(1-\frac{\theta}{1+\theta p}\right)\tilde{\beta}_1\right) \\
&= \frac{((1+\theta p)(\theta+p)-\theta)(n(p-1)-\tilde{r}_{d0})(2\tilde{r}_{d0}-n(p-1))}{(1+\theta p)^2 p\tilde{\gamma}_1^2} \\
&\leq 0.
\end{aligned}$$

By observing that y_0 and Θ_2 are invariant with respect to t, the proposition is proved. $\diamondsuit$

Proposition 2. *When $3 \leq p \leq t+1$ and $\tilde{r}_{d0} \geq n$, (2.30) holds for any value of θ.*

Appendix A (Continued)

Proof. First of all, set $\xi_1 = r_{d0}^2/n$, $\xi_2 = \tilde{r}_{d0}$ and $\xi_3 = \tilde{r}_{d0}$. In the following, the notation $u \propto v$ stands for $u = f(\theta, t, n, p, r_{d0}, \tilde{r}_{d0})v$ with $f > 0$ for any $d \in \Omega_{t+1,n,p}$ and $\theta \geq 0$. Note that the form of f may vary from line to line. Now, we have

$$
\begin{aligned}
\frac{\partial \tilde{\alpha}_1}{\partial \theta} &\propto p - t - \frac{r_{d0}}{n} \leq 0 \\
\frac{\partial \tilde{\beta}_1}{\partial \theta} &\propto (n(p-1) - \tilde{r}_{d0})(t + \frac{r_{d0}}{n} - p) + \tilde{r}_{d0}(\frac{r_{d0}}{n} - 1) \geq 0 \\
\frac{\partial \tilde{\gamma}_1}{\partial \theta} &\propto n(p-1)(p-1-t) + \tilde{r}_{d0}(t - 2p + 3) \leq 0 \\
\frac{\partial \alpha_2}{\partial \theta} &= 0 \\
\frac{\partial \beta_2}{\partial \theta} &\propto 1 - \frac{r_{d0}}{n} \leq 0 \\
\frac{\partial \gamma_2}{\partial \theta} &\propto \frac{\tilde{r}_{d0}}{n} - 1 \geq 0.
\end{aligned}
$$

Thus,

$$\tfrac{\partial \tilde{x}_0}{\partial \theta} \leq 0 \quad \tfrac{\partial y_0}{\partial \theta} \geq 0 \quad \tfrac{\partial \Theta_1}{\partial \theta} \geq 0 \quad \tfrac{\partial \Theta_2}{\partial \theta} \leq 0. \tag{A.10}$$

Consequently,

$$\frac{\partial^2 H(\frac{r_{d0}^2}{n}, \tilde{r}_{d0}, \tilde{r}_{d0}, \theta)}{\partial \xi_i \partial \theta} \leq 0, i = 1, 2, 3, \tag{A.11}$$

for any value of θ. So it is enough to show

$$\frac{\partial H(\frac{r_{d0}^2}{n}, \tilde{r}_{d0}, \tilde{r}_{d0}, \infty)}{\partial \xi_i} \geq 0, i = 1, 2, 3. \tag{A.12}$$

Which is a direct result from Propositions A.3,A.4 and A.5 in Hedayat and Yang (2005). $\diamond$

Appendix A (Continued)

Proposition 3. *When $p = t + 1 \geq 3$, $\tilde{r}_{d0} \leq n$ and $r_{d0} \geq n$, (2.30) holds for any value of θ.*

Proof. First of all, set $\xi_1 = r_{d0}^2/n$, $\xi_2 = \tilde{r}_{d0}$ and $\xi_3 = \tilde{r}_{d0}$ and let $x = r_{d0}/n$, then $x \in [1, p/(p-1)]$ and we have

$$\tfrac{\partial \tilde{\alpha}_1}{\partial \theta} \leq 0, \quad \tfrac{\partial \tilde{\beta}_1}{\partial \theta} \geq 0, \quad \tfrac{\partial \tilde{\gamma}_1}{\partial \theta} \leq 0, \quad \tfrac{\partial \beta_2}{\partial \theta} \leq 0. \tag{A.13}$$

By setting θ to be 0 or ∞ according to (A.13), we have

$$\begin{aligned}
\tilde{\alpha}_1 \geq & t(1 - \frac{1}{p})(np - r_{d0}) - r_{d0} + \frac{r_{d0}^2}{pn} \\
= & \frac{n}{p}\left((p-1)^2(p-x) - x(p-x)\right), \\
\tilde{\beta}_1 \leq & \frac{t}{p}(n(p-1) - \tilde{r}_{d0}) - \frac{\tilde{r}_{d0}}{p} \\
= & \frac{n}{p}\left((p-1)^2 - (p-1)x\right), \\
\tilde{\gamma}_1 \geq & (p - \frac{p+1}{p})(n(p-1) - \tilde{r}_{d0}) - \frac{n}{p}(p-1)^2 + \frac{\tilde{r}_{d0}}{p} \\
= & \frac{n}{p}\left(p(p-1)(p-2) - \frac{(p^2-1)(p-2)}{p}x\right), \\
\beta_2 \geq & \frac{\tilde{r}_{d0}}{p} = \frac{n}{p}\frac{p-1}{p}x.
\end{aligned}$$

Also, it is obvious that

$$\begin{aligned}
\alpha_2 = & r_{d0} - \frac{r_{d0}^2}{pn} = \frac{n}{p}(p-x)x \\
\gamma_2 \leq & \tilde{r}_{d0} = \frac{n}{p}(p-1)x.
\end{aligned}$$

Appendix A (Continued)

In the following, we try to prove the following sufficient conditions for the lemma

$$\tilde{x}_0 - (t-1)y_0 \geq 0 \tag{A.14}$$

$$\frac{\tilde{x}_0}{\Theta_1} - (t-1)\frac{y_0}{\Theta_2} \geq 0 \tag{A.15}$$

For (A.14), when $p \geq 4$, we have

$$\begin{aligned}\frac{p}{n}(\tilde{x}_0 - (t-1)y_0) \geq & (p-1)(p-1-x)(p-x) + \frac{(p-1)(p-2)}{p^2}x - \frac{((p-1)^2-(p-1)x)^2}{p(p-1)(p-2) - \frac{(p^2-1)(p-2)}{p}x} \\ \geq & (p-1)(p-1-x)(p-x-\frac{5}{2p}) + \frac{(p-1)(p-2)}{p^2}x \geq 0\end{aligned}$$

In the above, we used the inequality $p(p-1-x)/(p^2(p-2)-(p^2-p-2)x) \leq 5/2p$ for $p \geq 4$ and $x \in [1, p/(p-1)]$. When $p = 3$, we have

$$\begin{aligned}\frac{p}{n}(\tilde{x}_0 - (t-1)y_0) \geq & 2(3-x)(2-x) + \frac{2}{9}x - \frac{(4-2x)^2}{6-8x/3} \\ \geq & 2(3-x)(2-x) + \frac{2}{9}x - \frac{3}{8}(4-2x) \\ = & 10.5 - 9x + 2x^2 - \frac{x}{36} \geq 0\end{aligned}$$

Appendix A (Continued)

For (A.15), when $p \geq 4$, we have

$$\begin{aligned}\frac{p}{n}\left(\frac{\tilde{x}_0}{\Theta_1}-(t-1)\frac{y_0}{\Theta_2}\right) \geq& \frac{2}{5}p(p-1)^2(p-x)-(p-1)(p-1-x)\\ &-p(p-\frac{8}{5})(p-x)x+\frac{(p-1)(p-2)}{p}x\\ =&: f(x)\end{aligned}$$

Since

$$f'(x) \leq -\frac{2}{5}p(p-1)^2-p(p-\frac{8}{5})(p-\frac{8}{3})-1 \leq 0,$$

applying $(p-1)^2 \geq p(p-2)$, we have

$$\begin{aligned}f(x) \geq f(\frac{p}{p-1}) \geq& \frac{2}{5}p^2(p-1)(p-2)-p^2(p-\frac{8}{5})-(p-1)^2+2p-2\\ \geq& \frac{p^2}{5}(2p^2-11p+7) \geq 0\end{aligned}$$

for $p \geq 5$. When $p = 4$, direct calculation gives $f(p/(p-1)) = 19/15 > 0$. When $p = 3$, we have

$$\frac{p}{n}\left(\frac{\tilde{x}_0}{\Theta_1}-(t-1)\frac{y_0}{\Theta_2}\right) \geq \frac{168-274x+148x^2-26x^3}{12-6x} \geq 0$$

for $x \in [1, 1.5]$. $\diamondsuit$

Appendix A (Continued)

Proposition 4. *When (i)* $4 \le p = t+1$ *and* $r_{d0} \le n$ *or (ii)* $p = 3, t = 2$ *and* $0.6306 \le r_{d0}/n \le 1$, *(2.31) holds for any value of* θ.

Proof. First of all, set $\xi_1 = r_{d0}$, $\xi_2 = \tilde{r}_{d0}$ and $\xi_3 = \tilde{r}_{d0}$ and let $x = r_{d0}/n$, then $x \in (0, 1]$ and we have

$$\tfrac{\partial \tilde{\alpha}_1}{\partial \theta} \ge 0 \quad \tfrac{\partial \tilde{\beta}_1}{\partial \theta} \le 0 \quad \tfrac{\partial \tilde{\gamma}_1}{\partial \theta} \le 0 \quad \tfrac{\partial \alpha_2}{\partial \theta} \le 0 \quad \tfrac{\partial \beta_2}{\partial \theta} \ge 0 \quad \tfrac{\partial \gamma_2}{\partial \theta} \le 0 \tag{A.16}$$

By setting θ to be 0 or ∞ according to (A.16), we have

$$\begin{aligned}
\tilde{\alpha}_1 \ge & t(np - r_{d0}) - \frac{(np - r_{d0})^2}{pn} - r_{d0} + \frac{r_{d0}^2}{pn} \\
= & \frac{n}{p} p(p-2)(p-x), \\
\tilde{\beta}_1 \le & \frac{(np - r_{d0})(n(p-1) - \tilde{r}_{d0})}{pn} - \frac{r_{d0}\tilde{r}_{d0}}{pn} \\
= & \frac{n}{p}(p-1)(p-2x), \\
\tilde{\gamma}_1 \ge & (p - \frac{p+1}{p})(n(p-1) - \tilde{r}_{d0}) - \frac{n}{p}(p-1)^2 + \frac{\tilde{r}_{d0}}{p} \\
= & \frac{n}{p}\left(p(p-1)(p-2) - \frac{(p^2-1)(p-2)}{p}x\right), \\
\alpha_2 \le & r_{d0} - \frac{r_{d0}^2}{pn} = \frac{n}{p}(p-x)x, \\
\beta_2 \ge & \frac{r_{d0}\tilde{r}_{d0}}{pn} = \frac{n}{p}\frac{p-1}{p}x^2, \\
\gamma_2 \le & \tilde{r}_{d0} - \frac{\tilde{r}_{d0}^2}{n(p-1)} = \frac{n}{p}(p-1)\left(1 - \frac{x}{p}\right)x \le \frac{n}{p}(p-1)x.
\end{aligned}$$

Appendix A (Continued)

When $p \geq 4$, in view of $p(p-2x)/(p^2(p-2)-(p^2-p-2)x) \leq 2/p$, we have

$$\begin{aligned}\frac{p}{n}(x_0-(t-1)y_0) \geq& p(p-2)(p-x)-\frac{2}{p}(p-1)(p-2x)-(p-2)(p-x)x+\frac{(p-1)(p-2)}{p^2}x^3 \\ \geq& (p-2)(p-x)^2-2(p-2x) \geq 0.\end{aligned}$$

Also,

$$\begin{aligned}\frac{p}{n}(\frac{x_0}{\Theta_1}-(t-1)\frac{y_0}{\Theta_2}) \geq& \frac{p}{2}p(p-2)(p-x)-(p-1)(p-2x)-p(p-2)(p-x)+\frac{(p-1)(p-2)}{p}x^2 \\ \geq& 2p(p-x)-(p-1)(p-2x) \geq 0.\end{aligned}$$

When $p = 3$, we have

$$\begin{aligned}\frac{p}{n}(x_0-(t-1)y_0) \geq& (3-x)^2-\frac{4(3-2x)^2}{6-8x/3}+\frac{2}{9}x^3 \\ \geq& (3-x)^2-2(3-2x) \geq 0\end{aligned}$$

for $x \in (0, 1]$ and

$$\frac{p}{n}(\frac{x_0}{\Theta_1}-(t-1)\frac{y_0}{\Theta_2}) \geq \frac{-36+78x-34x^2+\frac{4}{3}x^3}{6-4x} \geq 0$$

for $x \in [0.6306, 1]$.

$\diamondsuit$

Proposition 5. *When $p = t = 3$ and $r_{d0} \leq n$, (2.31) holds with any value of θ.*

Appendix A (Continued)

Proof. First of all, set $\xi_1 = r_{d0}$, $\xi_2 = \tilde{r}_{d0}$ and $\xi_3 = \tilde{r}_{d0}$ and let $x = r_{d0}/n$, then $x \in (0, 1]$ and we have

$$\tfrac{\partial \tilde{\alpha}_1}{\partial \theta} \leq 0, \quad \tfrac{\partial \tilde{\beta}_1}{\partial \theta} \geq 0, \quad \tfrac{\partial \tilde{\gamma}_1}{\partial \theta} \leq 0, \quad \tfrac{\partial \alpha_2}{\partial \theta} \leq 0, \quad \tfrac{\partial \beta_2}{\partial \theta} \geq 0, \quad \tfrac{\partial \gamma_2}{\partial \theta} \leq 0. \tag{A.17}$$

By setting θ to be 0 or ∞ according to (A.17), we have

$$\begin{aligned}
\tilde{\alpha}_1 \geq & t(1-\frac{1}{p})(np - r_{d0}) - r_{d0} + \frac{r_{d0}}{p} = \frac{n}{p}(18 - 8x) \\
\tilde{\beta}_1 \leq & \frac{t}{p}(n(p-1) - \tilde{r}_{d0}) - \frac{\tilde{r}_{d0}}{p} = \frac{n}{p}\left(6 - \frac{8}{3}x\right) \\
\tilde{\gamma}_1 \geq & (t + 1 - \frac{t+2}{p})(n(p-1) - \tilde{r}_{d0}) - \frac{n}{p}(p-1)^2 + \frac{\tilde{r}_{d0}}{p} = \frac{n}{p}(10 - 4x) \\
\alpha_2 \leq & r_{d0} - \frac{r_{d0}^2}{pn} = \frac{n}{p}(3x - x^2) \\
\beta_2 \geq & \frac{r_{d0}\tilde{r}_{d0}}{pn} = \frac{n}{p}\frac{2}{3}x^2 \\
\gamma_2 \leq & \tilde{r}_{d0} - \frac{\tilde{r}_{d0}^2}{n(p-1)} = \frac{n}{p}(p-1)\left(1 - \frac{x}{p}\right)x \leq \frac{n}{p}2x.
\end{aligned}$$

Applying the inequalities, we have

$$\begin{aligned}
\frac{p}{n}(\tilde{x}_0 - (t-1)y_0) \geq & 18 - 8x - \frac{(6 - 8x/3)^2}{10 - 4x} - 2(3x - x^2) + \frac{4}{9}x^3 \\
\geq & 18 - 8x - \frac{3}{5}(6 - 8x/3) - 2(3x - x^2) + \frac{4}{9}x^3 \\
= & 14.4 - 12.4x + 2x^2 + \frac{4}{9}x^3 \geq 0
\end{aligned}$$

Appendix A (Continued)

for $x \in (0,1]$. Also, we have

$$\frac{p}{n}\left(\frac{\tilde{x}_0}{\Theta_1} - (t-1)\frac{y_0}{\Theta_2}\right) \geq 6 - 6x + \frac{8}{3}x + \frac{4}{3}x^2 \geq 0$$

for $x \in (0,1]$.

$\diamondsuit$

CITED LITERATURE

Brandt, A. E. (1938), Tests of significance in reversal or switchback trials. *Research Bulletin No. 234, Iowa Agricultural Experimental Station.*

Carrière, K. C. and Reinsel, G. C. (1993), Optimal two-period repeated measurement designs with two or more treatments, *Biometrika* **80** 924–929

Cheng, C. S., and Wu, C. F. (1980), Balanced repeated measurements designs, *Annals of Statistics,* **8** 1272–1283.

Cochran, W. G. (1939), Long-term agricultural experiments (with discussion), *Journal of the Royal Statistical Society, Series B* **6** 104–148

Cochran, W. G., Autrey, K. M., and Cannon, C. Y. (1941), A double change-over design for dairy cattle feeding experiments. *Journal of Dairy Science* **24** 937–951

Colbourn, C. J., and Dinitz, J. H., editors, (2007), Handbook of Combinatorial Designs, *CRC Press, Boca Raton, FL*, second edition.

Dey, A. (1986), Theory of Block Designs, *Wiley Eastern Limited, New York.*

Fieller, E. C. (1940), The biological standardization of insulin (with discussion), *Supplement of the Journal of the Royal Statistical Society* **7** 1–64

Finney, D. J. (1956), Cross-over designs in bioassay, *Proceedings of the Royal Society, Series B* **145** 42–60

Finney, D. J. and Outhwaite, A. D. (1955), Serially balanced sequences. *Nautre*, 176–748

Finney, D. J. and Outhwaite, A. D. (1956), Serially balanced sequences in bioassay. *Proceedings of the Royal Society, Series B* **145** 493–507

Gordon, B. (1961), Sequences in groups with distinct partial products, *Pacific J. Math.*, **11** 1309–1313.

Hedayat, A. S., and Afarinejad, K. (1975), Repeated measurements designs, I, *A Survey of Statistical Design and Linear Models, J. N. Srivastava, ed., North-Holland, Amsterdam,* 229–240.

Hedayat, A. S., and Afarinejad, K. (1978), Repeated measurements designs, II, *Annals of Statistics,* **6**, 619–628. Springer, New York.

Hedayat, A. S., Stufken, J., and Yang, M. (2006), Optimal and efficient Crossover designs when subject effects are random, *Journal of The American Statistical Association*, **101**, 1031–1038.

Hedayat, A. S., and Yang, M. (2003), Universal optimality of balanced uniform Crossover designs, *Annals of Statistics,* **31**, 978–983.

Hedayat, A. S., and Yang, M. (2004), Universal optimality for selected Crossover designs, *Journal of The American Statistical Association,* **99**, 461–466.

Hedayat, A. S., and Yang, M. (2005), Optimal and efficient Crossover designs for comparing test treatments with a control treatment, *Annals of Statistics,* **33**, 915–943.

Hedayat, A. S., and Yang, M. (2006), Efficient Crossover designs for comparing test treatments with a control treatment, *Discrete Mathematics,* **306**, 3112–3124.

Hedayat, A. S., and Zhao, W. (1990), Optimal two-period repeated measurements designs, *Annals of Statistics,* **18**, 1805–1816. Corrigendum: Ibid 20 (1992), p. 619.

Hedayat, A. S. and Zheng, W. (2010a) Optimal Crossover designs when subject effects are random *Journal of the American Statistical Association*, **105**, 1581-1592.

Hedayat and Zheng, W. (2010b), Totally Balanced Test-Control Incomplete Crossover Designs and Their Statistical Applications, *Combinatorics and Graphs, Contemporary Mathematics* **531** 43–58

Higham, J. (1998), Row-complete latin squares of every composite order exist, *J. Combin. Des.*, **6** 63–77.

Jones, B. and Kenward, M. G. (2003), Design and analysis of cross-over trials. Second Edition. *Chapman & HallCRC Press, New York.*

Jones, B., Kunert, J., and Wynn H. P. (1992), Information matrices for mixed effects models with applications to the optimality of repeated measurements designs, *Journal of Statistical Planning and Inference* **33** 261–274

Kunert, J. (1983), Optimal designs and refinement of the linear model with applications to repeated measurements designs. *Annals of Statistics*, **11** 247–257

Kunert, J. (1984), Optimality of balanced uniform repeated measurements designs, *Annals of Statistics*, **12**, 1006–1017.

Kunert, J. (1991), Cross-over designs for two treatments and correlated errors. *Biometrika*, **78** 315–324

Kunert, J., and Martin, R. J. (2000), On the determination of optimal designs for an interference model, *Annals of Statistics*, **28**, 1728–1742.

Kunert, J. and Stufken, J. (2008), Optimal Crossover designs for two treatments in the presence of mixed and self carryover effects. *Journal of the American Statistical Association*, **103**, 1641–1647.

Kunert, J., and Stufken, J. (2002), Optimal Crossover designs in a model with self and mixed carryover effects, *Journal of The American Statistical Association,* **97**, 898–906.

Kushner, H. B. (1997), Optimal repeated measurements designs: The linear optimality equations, *Annals of Statistics*, **25**, 2328–2344.

Kushner, H. B. (1998), Optimal and efficient repeated-measurements designs for uncorrelated observations, *Journal of The American Statistical Association*, **93**, 1176–1187.

Laska, E. M. and Meisner, M. (1985), A variational approach to optimal two-treatment Crossover designs: Application to carryover-effect models, *Journal of the American Statistical Association*, **80**, 704–710.

Lindner, C. C. and Rodger, C. A. (1997) Design Theory, *CRC Press, New York.*

Magda, G. C. (1980), Circular balanced repeated measurement designs. *Communications in Statistics–Theory and Methods* 18, A9.

Mendelsohn, N. S. (1996) Hamiltonian decomposition of the complete directed n-graph, *In P. Erdos and J. Catona, editors, Theory of Graphs, Proc. Colloq., Tihany*, 237–241. Academic Press, New York, 1968.

Mukhopadhyay, A. C. and Saha, R. (1983). Repeated measurement designs. *Calcatta Statistical Association Bulletin*, **38** 153–168

Sampford, M. R. (1957), Methods of construction and analysis of serially balanced sequence. *Journal of the Royal Society, Series B* **19** 286–304

Simpson, T. W. (1938), Experimental methods and human nutrition (with discussion), *Journal of the Royal Statistical Society, Series B* **5** 46–69

Stufken, J. (1991), Some families of optimal and efficient repeated measurements designs, *Journal of Statistical Planning and Inference* , **27**, 75–83.

Stufken, J. (1996), Optimal Crossover designs, Handbook of Statistics 13, S. Ghosh and C. R. Rao, eds, *Design and Analysis of Experiments, North-Holland, Amsterdam.* 63–90.

Williams, E. J. (1949), Experimental designs balanced for the estimation of residual effects of treatments, *J. Sci. Res. Ser. A,* **2** 149–168.

Williams, E. J. (1950), Experimental designs balanced for pairs of residual effects. *Australian Journal of Scientific Research* **3** 351–363

Yang, M., and Park, M. (2007), Efficient Crossover designs for comparing test treatments with a control treatment when p=3, *Journal of Statistical Planning and Inference,* **137**, 2056–2067.

Yang, M., and Stufken, J. (2008), Optimal and efficient Crossover designs for comparing test treatments to a control treatment under various models, *Journal of Statistical Planning and Inference* , **138**, 278–285.

Yates, F. (1938), The gain in efficiency resulting from the use of balanced designs. *Journal of the Royal Statistical Society, Series B* **5** 70–74

VITA

A.1 Education

- ◇ **Ph.D.** in Statistics, University of Illinois, Chicago 2006 - 2011
- ◇ **B.S.** in Mathematics and Statistics, Zhejiang University, Hangzhou 2001 - 2005

A.2 Research Interests

- ◇ Time Series
- ◇ Experimental Designs

A.3 Publications

1. Optimal and Efficient Crossover Designs for Test-Control Study When Subject Effects are Random. (2010) *Journal of the American Statistical Association*, Vol. 105 pp. 1581–1592. (with A. S. Hedayat)
2. Covariances Estimation for Long-Memory Process. (2010) *Advances in Applied Probability*, Vol. 42 pp. 137–157. (with Wei-Biao Wu and Yinxiao Huang)
3. Totally balanced test-control incomplete Crossover designs and their statistical applications. (2010) *Contemporary Mathematics*, Vol. 531 pp. 43–58. (with A. S. Hedayat)

A.4 Presentations

- ◇ **Invited Speaker**, International Conference On Design of Experiments, Memphis, TN 2011

- ◇ Seminar, University of Hong Kong, HK 2010
- ◇ Seminar, DePaul University, Chicago, IL 2010
- ◇ Seminar, Illinois Institute of Technology, Chicago, IL 2010
- ◇ Conference of Design and Analysis of Experiments, Columbia, Missouri 2009
- ◇ Seminar, University of Illinois, Chicago 2009
- ◇ Seminar, University of Illinois, Chicago 2007

A.5 Honors and Awards

University of Illinois at Chicago

- ◇ Dean's Scholar Award (University-Wide) 2009
- ◇ NSF Funded RA, Evaluation of Rare Events Summer 2009
- ◇ NSF Funded RA, Pathway Analysis in Terrestrial Food Webs Spring 2008
- ◇ Department Travel Awards 2008 - 2010
- ◇ Stat Lab Consulting Awards 2007

Zhejiang University

- ◇ Excellent Undergraduate Thesis (5%) 2005
- ◇ 1st Outstanding Student Scholarship (1%) 2004
- ◇ Excellent Student Award (5%) 2003