Bayesian Optimal Single-Stratum and Multistratum Designs with High Parameter Estimation Efficiency

by

Ying Miao

A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfilment of the requirements of the degree of

MASTER OF SCIENCE

Department of Statistics
University of Manitoba
Winnipeg

Abstract

Most optimality criteria considered in the literature are model-based criteria that rely on having an assumed model to select an optimal design for the experiment. Having a specified model prior to experimentation might not be feasible in reality. Bayesian optimality criteria has been in the literature for decades to relieve the dependence on an assumed model. In this research, we develop new Bayesian optimality criteria with high parameter estimation efficiency for multistratum designs. Examples with comparisons and sensitivity analyses are provided for selecting optimal designs in completely randomized experiments and multistratum experiments such as split-plot designs using the new criteria.

Acknowledgment

With my greatest appreciation, I would like to first and foremost thank my research supervisor, Dr. Po Yang, for her patient guidance, instructions and advice; without her, completing this thesis would never be possible. I would also thank the rest of my thesis committee members, Dr. Sauman Mandal and Dr. Zhiyang Zhou, for their valuable critiques and feedback, and Dr. Yi Xiong, for chairing my defence.

In addition, I would like to express my special thanks to the department staff members, Ms. Margaret Smith, Ms. Rosa Di Noto and Ms. Sunny Joo, for their continuous support.

Last but not least, I must express my gratitude to my parents and family members for their profound care and emotional support throughout this academic journey. They have taught me to be a persevering person. Thanks to my dearest friends who have been motivating me. Thanks to my cat Troye for his trust and company.

Dedication

To *everyone* who is reading this thesis.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Design of Experiments

Scientific experiments seek to collect data and quantify the relationship between the proposed response variable(s) and explanatory variable(s). Designing such experiments for any kinds of scientific research involves identifying the processes and objects required to conduct an experiment, including but not limited to the design type, run size, variables of interest, treatment groups, etc. Completely randomized design (CRD) is one of the simplest design types to use in experiments, requiring homogeneity among experimental units.

Multistratum designs have gained much attention recently, for their practical usage in experiments when the homogeneity requirements for CRD cannot be satisfied in many industrial and field experiments. An example of multistratum design is a split-plot design (Jones and Nachtsheim, 2009), which involves two stages of randomization. Suppose an industrial experiment involves the use of a furnace in different temperature settings ($100°C$, $300°C$) and testing sheets made with different materials (silicon, carbon fiber). It is known that the furnace temperature will be harder to change in an increment of $200°C$ than putting a different sample into the furnace. Thus, instead of waiting for the furnace temperature to cool down or heat up to the desired temperature every time the experimenter put a different testing sample in, the experimenter can first randomize the temperature setting for the furnace, then within each temperature setting, the experimenter can randomly assign the position of the testing materials in the furnace. Such design is called a split-plot design. In this experiment, the furnace temperature is called the *hard-to-change factor* or *whole-plot factor*, and the testing materials is called the *easy-to-change factor* or *subplot factor*. Split-plot design is considered a two-stratum design.

Following the ideas of split-plotting, Arnouts and Goos (2012) developed the principles for the construction of a staggered-level design, which use three strata when there are two classes of hard-to-change factors, Class-I and Class-II hard-to-change factors, and the different classes of the hard-to-change reset at different time points. Consider the previous example when the experimenter want to test material durability under the heat and an extra factor collision, with furnace temperature as the Class-I hard-to-change factor, collision as the Class-II hard-to-change factor, and testing materials as the subplot factor. The researchers can let collision level (low, high) reset in the middle of the same furnace temperature. The design in shown in Table 1.1, the temperature is reset two times, and the collision level is reset three times during the experiment.

There are three steps of randomization in this experiment compared to the split-plot designs: first randomization happens when different temperature was randomized to the furnace, the second randomization process happens when the experimenters randomize different collision levels at different time points, and the third randomization happens when the position of test materials in the furnace is randomized. A staggered-level design is considered an three-stratum design.

Table 1.1: Staggered-level Design

Temperature ($^{\circ}C$)	Collision	Material
100	Low	Carbon fiber
100	Low	Silicon
100	High	Silicon
100	High	Carbon fiber
300	High	Carbon fiber
300	High	Silicon
300	Low	Silicon
300	Low	Carbon fiber

1.2 Optimal Designs

To maximize the quality of statistical inference on the variables examined in the aforementioned types of designs, decisions have to be made by the experimenters to appropriately design the experiments. This process involves careful selections of levels of the input treatments and optimality criteria can help in making the selection (Kiefer, 1959). The use of optimality criteria is associated with the inferential goals of experimenters. In parameter estimation, small variability within the estimators is preferred, therefore D-optimality criterion was introduced by Wald (1943). It seeks to select designs that has the lowest overall variance of parameter estimates. A-optimality criterion was introduced by Chernoff (1953) with a similar target in parameter estimation. A-optimality criterion seeks to select a design that has the lowest average variance among the parameter estimates.

The I-optimality criterion, originally called Q-optimality, was first discussed by Fedorov (1972). It targets to select design that has good response prediction capabilities, therefore it looks for designs that minimizes the integrated prediction variance.

All of the aforementioned criteria were first discussed in the sense of CRDs. Due to the special design structure of multistratum designs and their multiple sources of error, the criteria had to be adjusted for multistratum designs. Jones and Goos (2012) studied and compared the D- and I-optimal split-plot designs in their work. Trinca and Gilmour (2015) extended I-optimality criterion to multistratum designs.

1.3 Bayesian Framework

The D-, A-, and I-optimality criteria mentioned in Section 1.2 all depend on an presumed model prior to experimentation, which could be infeasible in real life. The Bayesian framework was introduced to relieve the dependence.

For the Bayesian D-optimality criterion, there are primarily two approaches. The first approach to the Bayesian D-criterion was developed by Bernardo (1979). Given prior distributions of the responses and the model parameters, the Bayesian D-criterion tries to find a design that maximizes the expected Shannon information gain. For the second approach, DuMouchel and Jones (1994) categorized the model terms into two classes: the primary terms that are assumed to be active in the model, and the potential terms that may or may not be active. Assigning prior distributions to the primary terms and potential terms, they derived a posterior distribution for the model parameters, and introduced the Bayesian D-criterion based on the posterior distribution.

The Bayesian S_P-optimality criterion was developed by Borth (1975), adopting the ideas of total entropy measurement from Box and Hill (1967). It seeks to find optimal design that maximizes the total entropy measuring uncertainties of model parameters. The total entropy approach brings out challenges in calculation as the size of the model space increases, as it requires researchers to consider all models in the model space.

Although Bayesian approaches have been in discussion for decades, they are primarily developed for CRDs. Ng and Chick (2004) developed closed forms of the Bayesian D-optimality criterion under Bernardo's approach and S_P-optimal criterion for CRDs under some assumptions. There have been some development of Bayesian approaches for Multistratum designs. Lin (2018) developed the Bayesian D-optimality criterion under DuMouchel and Jones' approach for multistratum designs.

Since all the mentioned criteria such as D-, I-, A- and Bayesian optimality criteria are looking to construct designs that involves maximization or minimization of a criterion value, we need an computer algorithm to iteratively compare designs and find the optimal design which maximizes or minimizes the designated criterion value. Two major computer algorithm that are utilized in optimal design constructions are the point-exchange algorithm and the coordinate-exchange algorithm (Cook and Nachtsheim, 1980; Meyer and Nachtsheim, 1995). Considering the special design structure of multistratum designs, Trinca and Gilmour (2015) proposed a stratum-by-stratum point-exchange algorithm for multistratum designs.

In this thesis, by adopting the ideas of expected Shannon information gain from Bernardo (1979) and total entropy measurement from Borth (1975), we extend the Bayesian framework of both concepts to multistratum designs and use the modified stratum-by-stratum point-exchange algorithm to generate examples for the newly developed criteria.

1.4 Thesis Organization

This thesis is organized as follows. In Chapter 2, we introduce the definitions and model structures for single-stratum and multistratum designs, some of the currently used optimality criteria, and the algorithms that have been used in constructing optimal designs. In Chapter 3, we introduce two new Bayesian optimality criteria, the modified algorithm and the choice of parameters we use in the research. In Chapter 4, we list out multiple examples of the constructed optimal designs and discuss their performances by evaluating efficiency values. Finally, we set our concluding remarks and discussion for future development in Chapter 5.

Chapter 2

Designs, Optimality Criteria and Algorithms

This chapter introduces the definitions of single-stratum and multistratum designs. Different optimality criteria and computer algorithms to find optimal designs are introduced. In Section 2.1, CRD or single-stratum design is introduced, and an example is given to explain the design and its model structure. In Section 2.2, multistratum designs, such as split-plot design and staggered-level design, are introduced. In Section 2.3, we review several optimality criteria developed in the existing literature. Different exchange algorithms used to find optimal designs are introduced in Section 2.4.

2.1 Completely Randomized Designs

CRD is considered a single-stratum design as it requires homogeneity among experimental units.

Example 2.1.1. An 8-run CRD with two factors X_1 and X_2 is given in Table 2.1. Each factor has two levels. It is a replicated design in which all the runs are repeated once.

Table 2.1: An 8-run D-optimal CRD with two factors

Run	X_1	X_2
1	-1	-1
2	1	1
3	-1	1
4	1	-1
5	1	1
6	-1	-1
7	1	-1
8	-1	1

For an n-run m-factor CRD, a linear model studying the relationships between responses and effects in the design is defined as:

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \tag{2.1}$$

where $\boldsymbol{Y}$ is the $n \times 1$ vector of responses, $\boldsymbol{X}$ is the $n \times (1+k)$ model matrix, $\boldsymbol{\beta}$ denotes the $(1+k) \times 1$ vector of the model parameters, $\boldsymbol{\epsilon} \backsim N(\boldsymbol{0}_n, \sigma^2 \boldsymbol{I}_n)$ denotes the $n \times 1$ vector of random error.

The variance-covariance matrix of $\boldsymbol{Y}$ is

$$Var(\boldsymbol{Y}) = Var(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}) = Var(\boldsymbol{\epsilon}) = \sigma^2 \boldsymbol{I}_n,$$

and the least-square estimator of parameter vector $\boldsymbol{\beta}$ is given by

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}.$$

The model matrix is dependent on the number of main effects and interaction effects the experimenters choose to study.

Example 2.1.1 (Cont.) Consider the design in Table 2.1, the associated model matrix $\boldsymbol{X}$ for the linear model including the intercept, all main effects and two-factor interactions is listed in Table 2.2.

Table 2.2: Model matrix for Example 2.1.1

Run	Intercept	X_1	X_2	X_1X_2
1	1	-1	-1	1
2	1	1	1	1
3	1	-1	1	-1
4	1	1	-1	-1
5	1	1	1	1
6	1	-1	-1	1
7	1	1	-1	-1
8	1	-1	1	-1

2.2 Multistratum Designs

When the requirement of homogeneity cannot be satisfied in CRDs, multistratum designs come into use. Multistratum designs are commonly seen in industrial and field experiments. A multistratum design is defined as a design with multiple sources of error due to different structures of experimental units, therefore, its design and model structure is different from those of CRDs. In this thesis, we consider split-plot designs and staggered-level designs as examples of multistratum designs, where split-plot design is an example of two-stratum design, and staggered-level design is an example of three-stratum design.

2.2.1 Split-plot Designs

In split-plot designs, factors have two levels of difficulty in changing the factor settings. The hard-to-change factors are called whole-plot factors, and the easy-to-change factors are called subplot factors. There are two stages of randomization in a split-plot design, the whole-plot factors are randomized over the whole plots first, then the subplot factors are randomized over the subplots within each whole plot.

Example 2.2.1. A 12-run D-optimal split-plot design with four factors is given in Table 2.3. W_1 and W_2 are whole-plot factors and X_1 and X_2 are subplot factors. Each factor has two levels of settings. The first stage of randomization is to randomly assign levels W_1 and W_2 to the four whole plots, and the second stage of randomization happens when we randomly assign different levels of X_1, and X_2 to the subplots within each whole plot.

Table 2.3: A 12-run D-optimal split-plot design with four factors

Run		W_1	W_2	X_1	X_2
1	Whole plot 1	1	-1	1	1
2		1	-1	-1	1
3		1	-1	1	-1
4	Whole plot 2	-1	-1	1	-1
5		-1	-1	-1	1
6		-1	-1	-1	-1
7	Whole plot 3	-1	1	-1	1
8		-1	1	1	1
9		-1	1	1	-1
10	Whole plot 4	1	1	-1	1
11		1	1	-1	-1
12		1	1	1	-1

2.2.2 Staggered-level Designs

Note that the factor settings of the whole-plot factors in Example 2.2.1 must change together, Arnouts and Goos (2012) proposed a new type of multistratum design called staggered-level designs, with different classes of hard-to-change factors, and the hard-to-change factors reset settings at independent time points. In Example 2.2.2, we give an illustration of such design

Example 2.2.2. A 16-run D-optimal staggered-level design with four factors W, S, X_1, and X_2, where W and S are the hard-to-change factors and X_1 and X_2 are easy-to-change factors is given in Table 2.4. In particular, W is a Class-I hard-to-change factor and S is a Class-II hard-to-change factor. The levels of W and S can reset at different time points. Each factor has two levels of settings. The design is considered a multistratum design with three strata that display a staggered pattern for hard-to-change factors.

2.2.3 Model Structure

Consider a g-stratum multistratum design where there are b_l experimental units for the l^{th} stratum. A generalized model studying the relationships between the response and effects is defined as:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \sum_{l=1}^{g-1} \boldsymbol{U}_l \boldsymbol{\gamma}_l + \boldsymbol{\epsilon}, \tag{2.2}$$

Table 2.4: A D-optimal 16-run staggered-level design with four factors

Run	W	S	X_1	X_2
1	-1	1	1	-1
2	-1	1	-1	-1
3	-1	-1	-1	-1
4	-1	-1	1	-1
5	1	-1	1	1
6	1	-1	-1	1
7	1	1	-1	1
8	1	1	1	1
9	-1	1	-1	-1
10	-1	1	1	-1
11	-1	-1	-1	-1
12	-1	-1	1	-1
13	1	-1	1	1
14	1	-1	-1	1
15	1	1	-1	1
16	1	1	1	1

where $\boldsymbol{Y}$ is the $n \times 1$ vector of responses, $\boldsymbol{X}$ is the $n \times (1+k)$ model matrix, $\boldsymbol{\beta}$ is the $(1+k) \times 1$ vector of the the model matrix, $\boldsymbol{\gamma}_l$ is $b_l \times 1$ vector of random effects for the l^{th} stratum, where $\boldsymbol{\gamma}_l \backsim N(\boldsymbol{0}_{b_l}, \sigma_l^2 \boldsymbol{I}_{b_l})$, $\boldsymbol{\epsilon} \backsim N(\boldsymbol{0}_n, \sigma^2 \boldsymbol{I}_n)$ is the $n \times 1$ vector of random error, $\boldsymbol{U}_l = (u_{ij})$ is an $n \times b_l$ indicator matrix for the l-th stratum where u_{ij} is equal to 1 if the i-th run is in the j-th plot of the l-th stratum. When the design is a single-stratum design ($g = 1$), the model reduces to $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$.

Example 2.2.2 (Cont.) Consider the staggered-level design in Example 2.2.2. The model of the design can be written as $\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{U}_1\boldsymbol{\gamma}_1 + \boldsymbol{U}_2\boldsymbol{\gamma}_2 + \boldsymbol{\epsilon}$. Table 2.5 gives the indicator matrix $\boldsymbol{U}_1$ for the first stratum with the Class-I factor W, and Table 2.6 gives the indicator matrix $\boldsymbol{U}_2$ for the second stratum with the Class-II factor S.

Table 2.5: Indicator matrix $\boldsymbol{U}_1$ for the first stratum of the design in Table 2.4

Run	Plot 1	Plot 2	Plot 3	Plot 4
1	1	0	0	0
2	1	0	0	0
3	1	0	0	0
4	1	0	0	0
5	0	1	0	0
6	0	1	0	0
7	0	1	0	0
8	0	1	0	0
9	0	0	1	0
10	0	0	1	0
11	0	0	1	0
12	0	0	1	0
13	0	0	0	1
14	0	0	0	1
15	0	0	0	1
16	0	0	0	1

Table 2.6: Indicator matrix $\boldsymbol{U}_2$ for the second stratum of the design in Table 2.4

Run	Plot 1	Plot 2	Plot 3	Plot 4	Plot 5
1	1	0	0	0	0
2	1	0	0	0	0
3	0	1	0	0	0
4	0	1	0	0	0
5	0	1	0	0	0
6	0	1	0	0	0
7	0	0	1	0	0
8	0	0	1	0	0
9	0	0	1	0	0
10	0	0	1	0	0
11	0	0	0	1	0
12	0	0	0	1	0
13	0	0	0	1	0
14	0	0	0	1	0
15	0	0	0	0	1
16	0	0	0	0	1

The model matrix $\boldsymbol{X}$ is dependent on the type of the model the experimenters specify. There are three types of models that will be discussed for single-stratum and multistratum designs in this thesis:

Model 1. The first-order model that considers only the main effects. The design matrix is also the model matrix in this case.

Model 2. The second-order model that considers the main effects and two-factor interactions. Table 2.2 is an example of model matrix for such model.

Model 3. The second-order quadratic model that considers the main effects, quadratic effects, and two-factor interactions. This type of model is used when factors have three different levels of settings.

Example 2.2.3. Consider a 12-run D-optimal split-plot design with three factors, where W_1 is whole-plot factor and X_1 and X_2 are subplot factors. Each factor has three levels of settings. The design matrix is shown in Table 2.7. When the assumed model is Model 3, the model matrix of the design is listed in Table 2.8. The first column of trailing ones in Table 2.8 are for the model intercept.

Table 2.7: Design matrix for the D-optimal split-plot design in Example 2.2.3

Run	W_1	X_1	X_2
1	1	-1	0
2	1	-1	-1
3	1	1	1
4	-1	0	-1
5	-1	0	1
6	-1	1	1
7	0	0	1
8	0	-1	-1
9	0	1	0
10	1	1	-1
11	1	1	0
12	1	0	-1

Table 2.8: Model matrix for the design in Table 2.7

Run	Intercept	W_1	X_1	X_2	W_1^2	X_1^2	X_2^2	W_1X_1	W_1X_2	X_1X_2
1	1	1	-1	0	1	1	0	-1	0	0
2	1	1	-1	-1	1	1	1	-1	-1	1
3	1	1	1	1	1	1	1	1	1	1
4	1	-1	0	-1	1	0	1	0	1	0
5	1	-1	0	1	1	0	1	0	-1	0
6	1	-1	1	1	1	1	1	-1	-1	1
7	1	0	0	1	0	0	1	0	0	0
8	1	0	-1	-1	0	1	1	0	0	1
9	1	0	1	0	0	1	0	0	0	0
10	1	1	1	-1	1	1	1	1	-1	-1
11	1	1	1	0	1	1	0	1	0	0
12	1	1	0	-1	1	0	1	0	-1	0

2.3 Optimality Criteria

Optimality criteria are statistical criteria that are developed to help construct optimal designs that display the best value in one or some statistical properties.

2.3.1 Optimality Criteria for CRD

As CRD is one of the simplest design, there are many optimality criteria tackling different needs of the experimenters. We present the following traditional and Bayesian optimality criteria for CRDs.

D-Optimality

One of the most popular optimality criterion is D-optimality criterion. A D-optimal design seeks to find a design that has the best parameter estimation efficiency. The D-optimality criterion is defined to construct designs that minimize the determinant of the covariance matrix of the generalized least square estimates $\hat{\boldsymbol{\beta}}$, which is equivalent to maximizing $|\boldsymbol{X}'\boldsymbol{X}|$ (Kiefer and Wolfowitz, 1959). Let ξ be any design and $\boldsymbol{X}(\xi)$ be the corresponding model matrix, and ξ_1 be the D-optimal design and $\boldsymbol{X}(\xi_1)$ be the corresponding model matrix. The D-efficiency of ξ can be measured by

$$D_{eff} = \left[\frac{|\boldsymbol{X}(\xi)'\boldsymbol{X}(\xi)|}{|\boldsymbol{X}(\xi_1)'\boldsymbol{X}(\xi_1)|}\right]^{\frac{1}{1+k}}. \tag{2.3}$$

In this thesis, we use (2.3) to calculate the D-efficiency of the optimal CRD.

A-Optimality

An A-optimal design seeks designs that minimize average variance of parameter estimates $\text{tr}[(\boldsymbol{X}'\boldsymbol{X})^{-1}]$ (Wong and Masaro, 1984). To evaluate the parameter estimation quality of a design ξ, let ξ_2 be the A-optimal design, the A-efficiency of ξ can be measured by

$$A_{eff} = \left[\frac{\text{tr}\{(\boldsymbol{X}(\xi_2)'\boldsymbol{X}(\xi_2))^{-1}\}}{\text{tr}\{(\boldsymbol{X}(\xi)'\boldsymbol{X}(\xi))^{-1}\}}\right]. \tag{2.4}$$

In this thesis, we use (2.4) to calculate the A-efficiency of the optimal CRD.

I-Optimality

An I-optimal design displays good prediction capabilities. The prediction variance of a design is $f'(\mathbf{x})(\boldsymbol{X}'\boldsymbol{X})^{-1}f(\mathbf{x})$, where $f(\mathbf{x})$ is the model expansion for the design point $\mathbf{x}$ (Trinca and Gilmour, 2015). The I-optimality criterion is defined to construct designs that minimize the average predication variance over an experimental region. The average prediction variance of a design can be calculated as (Hardin and Sloane, 2009):

$$\text{Average Prediction Variance} = \frac{\int_\chi f'(\mathbf{x})(\boldsymbol{X}'\boldsymbol{X})^{-1}f(\mathbf{x})d\mathbf{x}}{\int_\chi d\mathbf{x}}$$

over the experimental region χ. Since the prediction variance is a scalar, it can be expressed as:

$$f'(\mathbf{x})(\boldsymbol{X}'\boldsymbol{X})^{-1}f(\mathbf{x}) = \text{tr}[(\boldsymbol{X}'\boldsymbol{X})^{-1}f(\mathbf{x})f'(\mathbf{x})].$$

It follows that

$$\int_{\chi} f'(\mathbf{x})(\boldsymbol{X}'\boldsymbol{X})^{-1}f(\mathbf{x})d\mathbf{x} = \int_{\chi} \text{tr}[(\boldsymbol{X}'\boldsymbol{X})^{-1}f(\mathbf{x})f'(\mathbf{x})]d\mathbf{x}$$

$$= \text{tr}\left[\int_{\chi}(\boldsymbol{X}'\boldsymbol{X})^{-1}f(\mathbf{x})f'(\mathbf{x})d\mathbf{x}\right]$$

$$= \text{tr}\left[(\boldsymbol{X}'\boldsymbol{X})^{-1}\int_{\chi} f(\mathbf{x})f'(\mathbf{x})d\mathbf{x}\right].$$

If the experimental region is $[-1,+1]^m$, then the volume of the experimental region is $\int_{\chi} d\mathbf{x} = 2^m$. Therefore, the average prediction variance can be rewritten as:

$$\text{Average Prediction Variance} = \frac{\int_{\chi} f'(\mathbf{x})(\boldsymbol{X}'\boldsymbol{X})^{-1}f(\mathbf{x})d\mathbf{x}}{\int_{\chi} d\mathbf{x}}$$

$$= 2^{-m}\text{tr}\left[(\boldsymbol{X}'\boldsymbol{X})^{-1}\int_{\chi} f(\mathbf{x})f'(\mathbf{x})d\mathbf{x}\right]$$

$$= \text{tr}[(\boldsymbol{X}'\boldsymbol{X})^{-1}\mathbf{B}], \qquad (2.5)$$

where

$$\mathbf{B} = 2^{-m}\int_{\chi} f(\mathbf{x})f'(\mathbf{x})d\mathbf{x},$$

When the model is a full quadratic model, the matrix $\mathbf{B}$ has a specific structure, which is

$$\mathbf{B} = \begin{bmatrix} 1 & \mathbf{0}'_m & \mathbf{0}'_{m^*} & \frac{1}{3}\mathbf{1}'_m \\ \mathbf{0}_m & \frac{1}{3}\boldsymbol{I}_m & \mathbf{0}_{m\times m^*} & \mathbf{0}_{m\times m} \\ \mathbf{0}_{m^*} & \mathbf{0}_{m^*\times m} & \frac{1}{9}\boldsymbol{I}_{m^*} & \mathbf{0}_{m^*\times m} \\ \frac{1}{3}\mathbf{1}_m & \mathbf{0}_{m\times m} & \mathbf{0}_{m\times m^*} & \frac{1}{45}(4\boldsymbol{I}_m + 5\boldsymbol{J}_m) \end{bmatrix},$$

where $m^* = m(m-1)/2$ is the number of two-factor interactions, and $\boldsymbol{J}_m$ is the $m \times m$ matrix of ones. Let ξ_3 be an I-optimal design. The I-efficiency of ξ can be measured by

$$I_{eff} = \left[\frac{\text{tr}\{(\boldsymbol{X}(\xi_3)'\boldsymbol{X}(\xi_3))^{-1}\mathbf{B}\}}{\text{tr}\{(\boldsymbol{X}(\xi)'\boldsymbol{X}(\xi))^{-1}\mathbf{B}\}}\right]. \tag{2.6}$$

Bayesian D-Optimality

The D-, I-, and A-optimality criteria require an assumed model before experimentation. This is often considered the largest drawback of the traditional criteria, as having a specified model might not be feasible prior to experimentation, therefore the constructed design using these criteria might be susceptible to model misspecification. Bayesian optimal designs are introduced to relieve the reliance on having an specified model prior to experimentation.

There are two definitions of the Bayesian D-optimality criterion. Consider the linear regression model (2.1). Let $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma^2)$ be the unknown parameters with prior density function $P(\boldsymbol{\theta}|\boldsymbol{Y})$. Following Lindley (1956), Bernardo (1979) suggested to select designs that maximizes the expected gain in Shannon information, which is

$$BD = \int\int P(\boldsymbol{Y})P(\boldsymbol{\theta}|\boldsymbol{Y})\log\left[\frac{P(\boldsymbol{\theta}|\boldsymbol{Y})}{P(\boldsymbol{\theta})}\right]d\boldsymbol{\theta}d\boldsymbol{Y}. \tag{2.7}$$

By Bayes' theorem, $P(\boldsymbol{Y}, \boldsymbol{\theta}) = P(\boldsymbol{Y})P(\boldsymbol{\theta}|\boldsymbol{Y}) = P(\boldsymbol{\theta})P(\boldsymbol{Y}|\boldsymbol{\theta})$. Therefore, by Fubini's theorem, assuming the integrals exist and are independent of the design, then the order of integration does not matter. Therefore, (2.7) can be written as

$$\begin{aligned} BD &= \int\int P(\boldsymbol{Y})P(\boldsymbol{\theta}|\boldsymbol{Y})\log[P(\boldsymbol{\theta}|\boldsymbol{Y})]d\boldsymbol{\theta}d\boldsymbol{Y} - \int\int P(\boldsymbol{Y})P(\boldsymbol{\theta}|\boldsymbol{Y})\log[P(\boldsymbol{\theta})]d\boldsymbol{\theta}d\boldsymbol{Y} \\ &= \int\int P(\boldsymbol{Y})P(\boldsymbol{\theta}|\boldsymbol{Y})\log[P(\boldsymbol{\theta}|\boldsymbol{Y})]d\boldsymbol{\theta}d\boldsymbol{Y} - \int P(\boldsymbol{\theta})\log[P(\boldsymbol{\theta})]\int P(\boldsymbol{Y}|\boldsymbol{\theta})d\boldsymbol{Y}d\boldsymbol{\theta} \\ &= \int\int P(\boldsymbol{Y})P(\boldsymbol{\theta}|\boldsymbol{Y})\log[P(\boldsymbol{\theta}|\boldsymbol{Y})]d\boldsymbol{\theta}d\boldsymbol{Y} - \int \log[P(\boldsymbol{\theta})]P(\boldsymbol{\theta})d\boldsymbol{\theta}. \qquad (2.8) \end{aligned}$$

Ng and Chick (2004) provided a closed form of (2.8). Let $\sigma^2\boldsymbol{K}$ be the variance-covariance matrix of $\boldsymbol{\beta}$. Under some prior assumptions for $\boldsymbol{\beta}$ and σ^2, Ng and Chick (2004) showed that (2.8) can be simplified as

$$BD = \frac{1}{2}\log|\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{K}^{-1}| - \frac{1}{2}\log|\boldsymbol{K}^{-1}|.$$

The Bayesian D-optimality criterion selects designs that maximize

$$\frac{1}{2}\log|\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{K}^{-1}| - \frac{1}{2}\log|\boldsymbol{K}^{-1}|. \qquad (2.9)$$

Bayesian D-Optimality: DuMouchel and Jones' Approach

The second definition of the Bayesian D-optimality criterion was proposed by DuMouchel and Jones (1994), which is the modified version of the Bayesian D-criterion considered by Bernardo (1979). They split the model matrix $\boldsymbol{X}$ into $(\boldsymbol{X}_1, \boldsymbol{X}_2)$, and the parameter $\boldsymbol{\beta}$ into $(\boldsymbol{\beta}_1', \boldsymbol{\beta}_2')'$, where $\boldsymbol{X}_1$ is the $n \times p_1$

matrix of primary terms and $\boldsymbol{\beta}_1$ is the $p_1 \times 1$ vector of the corresponding parameters, and $\boldsymbol{X}_2$ is the $n \times p_2$ matrix of potential terms with $\boldsymbol{\beta}_2$ as the $p_2 \times 1$ vector of the corresponding parameters.

The primary terms are considered as active terms, therefore the coefficients of parameters are given an prior mean of $\mathbf{0}_{p_1}$ and prior variance $\psi^2 \boldsymbol{I}_{p_1}$ tending to infinity, that is, $\boldsymbol{\beta}_1 \backsim \mathrm{N}(\mathbf{0}_{p_1}, \psi^2 \boldsymbol{I}_{p_1})$, where $\psi^2 \to \infty$. The potential terms may or may not be active, thus, they are assigned a prior distribution of mean $\mathbf{0}_{p_2}$ and a finite variance $\tau^2 \boldsymbol{I}_{p_2}$, that is, $\boldsymbol{\beta}_2 \backsim \mathrm{N}(\mathbf{0}_{p_2}, \tau^2 \boldsymbol{I}_{p_2})$, where τ is a predetermined finite value by the experimenter. The joint prior distribution of the model parameters for active and potential terms is $\boldsymbol{\beta} \backsim \mathrm{N}(\mathbf{0}_{p_1+p_2}, \boldsymbol{R})$, where

$$\boldsymbol{R} = \begin{bmatrix} \psi^2 \boldsymbol{I}_{p_1 \times p_1} & \mathbf{0}_{p_1 \times p_2} \\ \mathbf{0}_{p_2 \times p_1} & \tau^2 \boldsymbol{I}_{p_2 \times p_2} \end{bmatrix}.$$

Therefore $\boldsymbol{R}^{-1} = \frac{1}{\tau^2} \boldsymbol{Q}$ as $\psi^2 \to \infty$, where

$$\boldsymbol{Q} = \begin{bmatrix} \mathbf{0}_{p_1 \times p_1} & \mathbf{0}_{p_1 \times p_2} \\ \mathbf{0}_{p_2 \times p_1} & \boldsymbol{I}_{p_2 \times p_2} \end{bmatrix}.$$

If we assume that $\boldsymbol{Y}|\boldsymbol{\beta} \backsim \mathrm{N}(\boldsymbol{X}\boldsymbol{\beta}, \sigma^2 \boldsymbol{I})$ the posterior distribution of the model parameters $\boldsymbol{\beta}$ is $\boldsymbol{\beta}|\boldsymbol{Y} \backsim \mathrm{N}\left[(\boldsymbol{X}'\boldsymbol{X} + \frac{\boldsymbol{Q}}{\tau^2})^{-1} \boldsymbol{X}'\boldsymbol{Y}, \sigma^2 (\boldsymbol{X}'\boldsymbol{X} + \frac{\boldsymbol{Q}}{\tau^2})^{-1}\right]$. The Bayesian D-optimality seeks a design that minimizes $|\boldsymbol{X}'\boldsymbol{X} + \frac{\boldsymbol{Q}}{\tau^2}|^{-1}$ or maximizes $|\boldsymbol{X}'\boldsymbol{X} + \frac{\boldsymbol{Q}}{\tau^2}|$.

Bayesian S_p-Optimality: A Total Entropy Approach

According to Box and Hill (1967), the expected change in entropy as a measure of information between input and output is defined as:

total entropy = entropy at input - entropy at output.

Extended from the definition of the entropy, Borth (1975) defined the total entropy criterion S_P measuring the the uncertainty about the model parameters corresponding to each model in the model space, which is

$$\begin{aligned} S_P &= -\sum_{i=1}^{s} P(M_i) \int P(\boldsymbol{\theta}_i|M_i)\log\, P(\boldsymbol{\theta}_i|M_i) d\boldsymbol{\theta}_i \\ &+ \Big[\int \sum_{i=1}^{s} P(M_i|\boldsymbol{Y}) \int P(\boldsymbol{\theta}_i|\boldsymbol{Y}, M_i)\log\, P(\boldsymbol{\theta}_i|\boldsymbol{Y}, M_i) d\boldsymbol{\theta}_i \\ &\times \sum_{i=1}^{s} P(M_i) P(\boldsymbol{Y}|M_i) d\boldsymbol{Y} \Big], \end{aligned} \tag{2.10}$$

where s is the total number of models in the model space, M_i, $i = 1, 2, \cdots, s$, is a model in the model space, $P(M_i)$ is the prior model probability, $P(\boldsymbol{\theta}_i|M_i)$ is the prior density function for the parameters $\boldsymbol{\theta}_i = \{\beta_{i0}, \beta_{i1}, \cdots, \beta_{ik}\}$, $P(M_i|\boldsymbol{Y})$ is the posterior model probability, $P(\boldsymbol{\theta}_i|\boldsymbol{Y}, M_i)$ is the posterior density function for the parameters, and $P(\boldsymbol{Y}|Mi)$ is the probability density function of $\boldsymbol{Y}$ under model M_i.

Ng and Chick (2004) showed that (2.10) can be simplified as

$$S_P = \frac{1}{2}\sum_{i=1}^{s} P(M_i)\log|\boldsymbol{X_i}'\boldsymbol{X}_i + \boldsymbol{K}_i^{-1}| + \boldsymbol{E},$$

where $\boldsymbol{E}$ is a matrix independent of the design. The entropy criterion S_P selects designs that maximizes

$$\sum_{i=1}^{s} P(M_i)\log|\boldsymbol{X_i}'\boldsymbol{X}_i + \boldsymbol{K}_i^{-1}|. \tag{2.11}$$

2.3.2 Optimality Criteria for Multistratum Designs

Due to the different error structure of multistratum designs, the optimality criteria in Section 2.3.1 have to be modified.

D-Optimality

Let $\boldsymbol{\Sigma}$ be the variance-covariance matrix of $\boldsymbol{Y}$. Then the generalized least squares estimates of $\boldsymbol{\beta}$ for a multistratum design is given by $(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}$ and the variance-covariance matrix of $\hat{\boldsymbol{\beta}}$ is given by $(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}$ (Trinca and Gilmour, 2001). The D-optimality finds the optimal multistratum designs by minimizing the variance of $\hat{\boldsymbol{\beta}}$, which is equivalent of maximizing $|\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X}|$ (Jones and Goos, 2012). Let ξ be any multistratum design and ξ_4 be the

D-optimal multistratum design, the D-efficiency of ξ can be measured by

$$D_{eff} = \left[\frac{|\boldsymbol{X}(\xi)'\boldsymbol{\Sigma}^{-1}\boldsymbol{X}(\xi)|}{|\boldsymbol{X}(\xi_4)'\boldsymbol{\Sigma}^{-1}\boldsymbol{X}(\xi_4)|} \right]^{\frac{1}{1+k}}. \tag{2.12}$$

A-Optimality

Since A-optimal designs seeks to minimize the average variance of parameter estimates (Jones et al., 2021), an A-optimal multistratum design is defined to minimize the average variance of parameter estimates defined by $\mathrm{tr}[(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}]$. Let ξ_5 be the A-optimal multistratum design of the same experimental setting. The A-efficiency of ξ can be measured by

$$A_{eff} = \left[\frac{\mathrm{tr}\{(\boldsymbol{X}(\xi_5)'\boldsymbol{\Sigma}^{-1}\boldsymbol{X}(\xi_5))^{-1}\}}{\mathrm{tr}\{(\boldsymbol{X}(\xi)'\boldsymbol{\Sigma}^{-1}\boldsymbol{X}(\xi))^{-1}\}} \right]. \tag{2.13}$$

I-Optimality

The average prediction variance of a multistratum design can be calculated as (Trinca and Gilmour, 2015):

$$\text{Average Prediction Variance} = \frac{\int_{\chi} f'(\mathbf{x})(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1} f(\mathbf{x}) d\mathbf{x}}{\int_{\chi} d\mathbf{x}},$$

which is found to be $\mathrm{tr}\{(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\mathbf{B}\}$. An I-optimal multistratum design is defined to minimize $\mathrm{tr}\{(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\mathbf{B}\}$.

Let ξ_6 be the I-optimal multistratum design, the I-efficiency of ξ can be measured by

$$I_{eff} = \left[\frac{\text{tr}\{(\boldsymbol{X}(\xi_6)'\boldsymbol{\Sigma}^{-1}\boldsymbol{X}(\xi_6))^{-1}\mathbf{B}\}}{\text{tr}\{(\boldsymbol{X}(\xi)'\boldsymbol{\Sigma}^{-1}\boldsymbol{X}(\xi))^{-1}\mathbf{B}\}} \right]. \tag{2.14}$$

Bayesian D-Optimality: DuMouchel and Jones' Approach

The Bayesian D-optimality criterion for multistratum design was developed by Lin (2018). Assume the model (2.2). According to Lin (2018), $\boldsymbol{X}$ is split as $(\boldsymbol{X}_1, \boldsymbol{X}_2)$ and the parameters $\boldsymbol{\beta}$ is split as $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', \boldsymbol{\beta}_2')'$, where $\boldsymbol{X}_1$ is the $n \times p_1$ matrix of primary terms and $\boldsymbol{\beta}_1$ is the $p_1 \times 1$ vector of the corresponding parameters, and $\boldsymbol{X}_2$ is the $n \times p_2$ matrix of potential terms and $\boldsymbol{\beta}_2$ is the $p_2 \times 1$ vector of the corresponding parameters. Assume that $\boldsymbol{\beta}_1 \backsim \text{N}(\mathbf{0}_{p_1}, \psi^2 \boldsymbol{I}_{p_1})$, where $\psi^2 \to \infty$ and $\boldsymbol{\beta}_2 \backsim \text{N}(\mathbf{0}_{p_2}, \tau^2 \boldsymbol{I}_{p_2})$, where τ is a predetermined value.

Lin (2018) obtained that

$$\boldsymbol{\beta}|\boldsymbol{Y} \backsim \text{N}\left[(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X} + \frac{\boldsymbol{Q}}{\tau^2})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}, (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X} + \frac{\boldsymbol{Q}}{\tau^2})^{-1}\right]$$

as the posterior distribution, where $\boldsymbol{Q} = \begin{bmatrix} \mathbf{0}_{p_1 \times p_1} & \mathbf{0}_{p_1 \times p_2} \\ \mathbf{0}_{p_2 \times p_1} & \boldsymbol{I}_{p_2 \times p_2} \end{bmatrix}$. Therefore, the Bayesian D-optimality criterion selects the multistratum design that maximizes $|\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X} + \frac{\boldsymbol{Q}}{\tau^2}|^{\frac{1}{p_1+p_2}}$.

2.4 Algorithms

The optimality criteria in Section 2.3 and Section 2.4 require a computer algorithm to generate designs. Two major exchange algorithms are coordinate exchange algorithm and point exchange algorithm (Meyer and Nachtsheim, 1995; Fedorov, 1972).

For coordinate exchange algorithm, a single design coordinate is exchanged with another possible setting of the corresponding variable in each step. For point exchange algorithm, a row of the design point is exchanged with an nonidentical row from the candidate set in each step. Due to the special structure of multistratum designs, modified versions of coordinate exchange and point exchange algorithm have been proposed (Jones and Goos, 2012). In this thesis we use a modified version of point exchange algorithm called stratum-by-stratum point exchange to generate the designs (Trinca and Gilmour, 2001, 2015).

Chapter 3

Proposed New Bayesian Criteria

We will introduce new Bayesian optimality criteria which can be used to select optimal multistratum designs with high parameter estimation efficiency. In Section 3.1, we introduce the necessary prior information and the newly developed criteria. In Section 3.2, we describe the algorithm used to generate optimal multistratum designs. The choice of parameters is discussed in Section 3.3.

3.1 Proposed Optimality Criteria

3.1.1 Notations and Prior Assumptions

Under the model (2.2) and following Lin and Yang (2022), we define $\sigma_w^2 = \sum_{l=1}^{g-1} \sigma_l^2 + \sigma^2$ and let $\boldsymbol{\rho} = (\rho_1,\ \rho_2, \cdots,\ \rho_g)$, where $\rho_l = \frac{\sigma_l^2}{\sigma_w^2}$ for $l = 1, \cdots g-1$,

$\rho_g = \frac{\sigma^2}{\sigma_w^2}$. Thus, $\rho_l \in (0,1)$ and $\sum_{l=1}^{g} \rho_l = 1$. The variance-covariance matrix of $\boldsymbol{Y}$ is

$$
\begin{aligned}
Var(\boldsymbol{Y}) &= Var(\boldsymbol{X\beta} + \sum_{l=1}^{g-1} \boldsymbol{U}_l \boldsymbol{\gamma}_l + \boldsymbol{\epsilon}) \\
&= Var(\sum_{l=1}^{g-1} \boldsymbol{U}_l \boldsymbol{\gamma}_l) + Var(\boldsymbol{\epsilon}) \\
&= \sum_{l=1}^{g-1} Var(\boldsymbol{U}_l \boldsymbol{\gamma}_l) + \sigma^2 \boldsymbol{I}_n \\
&= \sum_{l=1}^{g-1} \frac{\sigma_l^2}{\sigma_w^2} \sigma_w^2 \boldsymbol{U}_l \boldsymbol{U}_l' + \frac{\sigma^2}{\sigma_w^2} \sigma_w^2 \boldsymbol{I}_n \\
&= \sum_{l=1}^{g-1} \sigma_w^2 \rho_l \boldsymbol{U}_l \boldsymbol{U}_l' + \rho_g \sigma_w^2 \boldsymbol{I}_n \\
&= \sigma_w^2 (\sum_{l=1}^{g-1} \rho_l \boldsymbol{U}_l \boldsymbol{U}_l' + \rho_g \boldsymbol{I}_n) \\
&= \sigma_w^2 \boldsymbol{V}, \qquad (3.1)
\end{aligned}
$$

where $\boldsymbol{V} = \sum_{l=1}^{g-1} \rho_l \boldsymbol{U}_l \boldsymbol{U}_l' + \rho_g \boldsymbol{I}_n$.

We assign non-informative prior distribution to β_0, and assume that $P(\sigma_w^2) \propto \frac{1}{\sigma_w^2}$. The model parameters $\boldsymbol{\beta}$ are assigned the prior distribution $\boldsymbol{\beta} \backsim \mathrm{N}(\mathbf{0}_{1+k}, \sigma_w^2 \boldsymbol{K}^{-1})$, where $\boldsymbol{K} = \frac{1}{\tau^2} \begin{bmatrix} 0 & \mathbf{0}_{1\times k} \\ \mathbf{0}_{k\times 1} & \boldsymbol{I}_{k\times k} \end{bmatrix}$.

3.1.2 Generalized Bayesian D-Optimality Criterion

We assume the prior information introduced in Section 3.1.1 and follow the Bayesian approach described by Bernardo (1979) to develop our optimality criterion.

Theorem 3.1.1 Under the prior assumptions of $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma_w^2, \boldsymbol{\rho})$ in Section 3.1.1, for any multistratum design with the model form (2.2), (2.8) can be simplified as

$$\int_{\boldsymbol{\rho}} P(\boldsymbol{\rho}) \left[\frac{\log|\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X} + \boldsymbol{K}|}{2} - \frac{\log|\boldsymbol{K}|}{2} \right] d\boldsymbol{\rho}.$$

where $P(\boldsymbol{\rho})$ is the prior probability density function of $\boldsymbol{\rho}$.

The proof of Theorem 3.1.1 is in A.1 in Appendix A. The generalized Bayesian D (GBD)-optimality criterion is defined to choose designs that maximizes

$$\int_{\boldsymbol{\rho}} P(\boldsymbol{\rho}) \left[\frac{\log|\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X} + \boldsymbol{K}|}{2} - \frac{\log|\boldsymbol{K}|}{2} \right] d\boldsymbol{\rho}. \tag{3.2}$$

Note that we have assumed that $\boldsymbol{\rho}$ follows an unknown distribution $P(\boldsymbol{\rho})$. In this thesis, we assume predetermined values for $\boldsymbol{\rho}$. Therefore, (3.2) can be simplified to

$$\frac{\log|\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X} + \boldsymbol{K}|}{2} - \frac{\log|\boldsymbol{K}|}{2}, \tag{3.3}$$

and we select designs that maximize (3.3). The sensitivity study on $\boldsymbol{\rho}$ will be conducted in this thesis.

3.1.3 Generalized Bayesian S_P-Optimality Criterion

Considering all the models in the model space, we adopt the idea in Borth (1975) to calculate the total entropy measuring the uncertainty about model parameters and develop the generalized Bayesian S_P (GS_P)-optimality criterion for multistratum designs. Assume that factors appearing in the model are independent of one another, and the number of active factors involved in the model M_i is f_i, Box and Meyer (1993) suggested the prior probability of M_i,

$$P(M_i) = \pi^{f_i}(1-\pi)^{m-f_i}, \tag{3.4}$$

where π is the probability of a factor in the model.

Theorem 3.1.2 Under the assumptions for $\boldsymbol{\theta} = (\boldsymbol{\beta}_i, \sigma_w^2, \boldsymbol{\rho})$ in Section 3.1.1. For any multistratum design with the model form (2.2), (2.10) can be simplified as

$$\frac{1}{2}\left[\sum_{i=1}^{s} P(M_i)\int_{\boldsymbol{\rho}} P(\boldsymbol{\rho})\log|\boldsymbol{X}_i'\boldsymbol{V}^{-1}\boldsymbol{X}_i + \boldsymbol{K}_i|d\boldsymbol{\rho}\right] + \boldsymbol{E},$$

where $\boldsymbol{E}$ is some matrix independent of the design.

The proof of Theorem 3.1.2 is in A.2 in Appendix A. The generalized Bayesian S_P (GS_P)-optimality criterion is defined to select designs that maximize

$$\frac{1}{2}\left[\sum_{i=1}^{s} P(M_i) \int_{\boldsymbol{\rho}} P(\boldsymbol{\rho})\log|\boldsymbol{X}_i'\boldsymbol{V}^{-\mathbf{1}}\boldsymbol{X}_i + \boldsymbol{K}_i| d\boldsymbol{\rho}\right] + \boldsymbol{E}, \tag{3.5}$$

.

Similar to the GBD-optimality criterion, in this thesis, we assume values of $\boldsymbol{\rho}$ for design selection. Therefore, (3.5) can be simplified to

$$\frac{1}{2}\left[\sum_{i=1}^{s} P(M_i)\log|\boldsymbol{X}_i'\boldsymbol{V}^{-\mathbf{1}}\boldsymbol{X}_i + \boldsymbol{K}_i|\right] + \boldsymbol{E}, \tag{3.6}$$

and we select designs that maximize (3.6).

We consider the entirety of model space in this thesis. For example, for a second-order model with m main effects, m quadratic effects, and $\binom{m}{2}$ two-factor interaction effects, there are $s = 2^{m+m+\binom{m}{2}}$ models considered for the selection of GS_P-optimal design. Detailed choices of parameters including $\boldsymbol{\rho}$ and π are discussed in Section 3.3.

3.2 Algorithm

We will adapt and modify the stratum by stratum point exchange algorithm in Trinca and Gilmour (2015). For the n-run, m-factor, g-stratum experiment, let $\mathbf{C}_i$ be the candidate set for the l-th stratum with b_l units. For a specified number of random starts r, the stratum-by-stratum point exchange algorithm is as follows:

1. Start with the highest stratum (lowest number) $l = 1$ in $l \in \{1, \ldots, g\}$ with the most hard-to-change factors in.

   1.1. Start from the first random start $j = 1$ for $j \in \{1, 2, \ldots, r\}$. Generate a random starting design ξ_0 with b_l runs by randomly choosing rows from the candidate set $\boldsymbol{C_l}$ considering the factors in this stratum only. Treat this design as the optimal design ξ_{opt} in this step.

   1.2. Evaluate the design ξ_0 with an optimality criterion GS_P or GBD. For GS_p, we will need to modify the criterion function to sum over all the models the model space. We will store this initial criterion value as Crit_0. The optimal criterion value of the optimal design is stored as $\text{Crit}_{\text{opt}} = \text{Crit}_0$.

   1.3. Start point exchange by replacing each run in ξ_0 with a non-identical row from $\boldsymbol{C_l}$ while holding all other rows constant. Store the new design as ξ_0^*.

   1.4. Evaluate the design ξ_0^* with the optimality criterion, the criterion value will be stored as Crit_0^*.

   1.5. Compare Crit_{opt} and Crit_0^*. If $\text{Crit}_0^* > \text{Crit}_{\text{opt}}$, store ξ_0^* as ξ_{opt}, and store Crit_0^* as Crit_{opt}. If $\text{Crit}_0^* \leq \text{Crit}_0$, go to the next point exchange.

   1.6. Repeat step 1.5 for each point exchange, perform the point exchange algorithm until each of the b_l runs have been exchanged with all non-identical rows from the candidate set $\boldsymbol{C}_l$. Store the final optimal design

ξ_{opt} as ξ_1, and the optimality criterion value Crit_{opt} as Crit_1 for $j = 1$.

1.7. Set $j = j + 1$, Perform the point exchange algorithm from steps 1.1 to 1.6 until $j = r$. Obtain a list of locally optimal designs $\xi_1, \xi_2, \ldots, \xi_r$ with corresponding optimality values $\text{Crit}_1, \text{Crit}_2, \ldots, \text{Crit}_r$. Compare $\text{Crit}_1, \text{Crit}_2, \ldots, \text{Crit}_r$ and set the design with greatest optimality value as the global optimal design.

2. The optimal design found in Step 1 has b_l runs or b_l units. The design is augmented by multiplying each unit by $\frac{n}{b_l}$. For example, to construct an optimal split-plot design, we obtain optimal design in Step 1 with b_l runs, which is considered as b_l units or whole plots. By multiplying $\frac{n}{b_l}$, the run size matches with n. Set $l = l + 1$ and construct an optimal design and treat the augmented design from stratum $l - 1$ as blocks. Following the point exchange algorithm from steps 1.1 to 1.5, use the candidate set $\boldsymbol{C}_l$ involving the factors in this stratum and previous strata to find a global optimal design.

3. If $l > 2$, rearrange the units created within $(l-2)$th stratum by interchanging the units in the $(l-2)$th stratum, and re-evaluate the optimal designs to see if re-arranging helps maximizing the optimality value. If any improvement is found, store the design with interchanged units as the global optimal design.

4. Evaluate the current stratum number l and when $l > g$, we have reached the

end of algorithm and found optimal design considering strata $l = 1, \ldots, g$, stop the algorithm and output the final optimal design searched.

3.3 Choice of Parameters

3.3.1 Choice of τ

Following Lin (2018), the tuning parameter τ is chosen to be assessed at $\tau = 1, 10, 14$ for CRD, split-plot designs, and staggered-level designs, respectively.

3.3.2 Choice of π

The assignment of the value for π in (3.4) is up to experimenter's choice. Bingham and Chipman (2007) proposed a way of incorporating prior expectation of active effects into the calculation of π. Let π be the probability that a specific factor main effect is active. The probability, $P_{AB,q}$, that a specific interaction is active, given q factor main effect parents are active, is

$$p_{AB,q} = \begin{cases} c_1\pi \text{ if } q = 0 \\ c_2\pi \text{ if } q = 1 \\ c_3\pi \text{ if } q = 2, \end{cases} \tag{3.7}$$

where c_1, c_2 and c_3 are tuning parameters for the probabilities. Conditional on f active main effects, we define the expected number of active effects (main effects, quadratic effects and two-factor interactions) as

$$E(\text{active effects}|f \text{ active main effects})$$
$$= f + f + \binom{f}{2} c_3\pi + f(m-f)c_2\pi + \binom{m-f}{2} c_1\pi. \tag{3.8}$$

Since f follows the Binomial distribution with m trials and probability of success π, we have that $E(f) = \pi m$ and $E(f^2) = \pi m(1 - \pi + \pi m)$. Taking the expectation of (3.8) with respect to f yields

$$E(\text{active effects}) = 2\pi m + \pi \binom{m}{2} \left[c_1 + 2\pi(c_2 - c_1) + \pi^2(c_1 - 2c_2 + c_3)\right]. \tag{3.9}$$

In this thesis, we adopt the prior specification for two-factor interaction from Bingham and Chipman (2007), which chooses $c_1 = 0.01, c_2 = 0.5$ and $c_3 = 1$, that is,

$$p_{AB,i} = \begin{cases} 0.01\pi & \text{if } q = 0 \\ 0.5\pi & \text{if } q = 1 \\ \pi & \text{if } q = 2. \end{cases} \tag{3.10}$$

Then we can simply (3.9) as

$$E(\text{active effects}) = 2\pi m + \pi m(m-1)\{0.005 + 0.49\pi + 0.005\pi^2\}. \tag{3.11}$$

Given m, we could solve for π using (3.11). For example, For example, let $m = 2$ be the number of factors, then the number of the two-factor interactions considered in the model is $\binom{2}{2} = 1$. If $E(\text{active effects}) = 4$, then

$$\begin{aligned} E(\text{active effects}) &= 2\pi m + \pi m(m-1)\{0.005 + 0.49\pi + 0.005\pi^2\} \\ 4 &= 4\pi + 2\pi\{0.005 + 0.49\pi + 0.005\pi^2\} \\ 4 &= 0.01\pi^3 + 0.98\pi^2 + 2.01\pi. \end{aligned}$$

By Newton's Method, $\pi \approx 0.828$. Similarly, we can modify (3.8) according the effects considered in the model and obtain the corresponding π. For models with main effects only,

$$E(\text{active effects}) = \pi m. \tag{3.12}$$

For models with main effects and quadratic terms,

$$E(\text{active effects}) = 2\pi m. \tag{3.13}$$

For models with main effects and two-factor interaction effects,

$$E(\text{active effects}) = \pi m + \pi m(m-1)\{0.005 + 0.49\pi + 0.005\pi^2\}. \tag{3.14}$$

3.3.3 Choice of $\boldsymbol{\rho}$

The developed GBD- and GS_P-optimality criteria provides a generalized Bayesian solution by integrating over $\boldsymbol{\rho}$. We would like to study how the different choices of values for ρ_l affect the performance of the constructed design. We would pick out the following scenarios for split-plot designs and staggered-level designs.

Choice of ρ_1 and ρ_2 for Split-Plot Designs

Since the fully expanded matrix form for $\boldsymbol{V}$ for split-plot design as a two-stratum design ($g = 2$) is $\boldsymbol{V} = \sum_{l=1}^{g-1} \rho_l \boldsymbol{U}_l \boldsymbol{U}_l' + \rho_g \boldsymbol{I}_n = \rho_1 \boldsymbol{U}_1 \boldsymbol{U}_1' + \rho_2 \boldsymbol{I}_n$, we would like to assume the following sets of values for $\boldsymbol{\rho} = \{\rho_1, \rho_2\}$.

Table 3.1: Choice of values for ρ_1 and ρ_2 for split-plot designs

Scenario	ρ_1	ρ_2
$\rho_1 < \rho_2$	0.1	0.9
	0.2	0.8
	0.3	0.7
	0.4	0.6
$\rho_1 = \rho_2$	0.5	0.5
$\rho_1 > \rho_2$	0.6	0.4
	0.7	0.3
	0.8	0.2
	0.9	0.1

Choice of ρ_1, ρ_2, and ρ_3 for Staggered-Level Designs

Similarly, the fully expanded matrix form for $\boldsymbol{V}$ for staggered-level design as a three-stratum design ($g = 3$) is $\boldsymbol{V} = \sum_{l=1}^{g-1} \rho_l \boldsymbol{U}_l \boldsymbol{U}_l' + \rho_g \boldsymbol{I}_n = \rho_1 \boldsymbol{U}_1 \boldsymbol{U}_1' + \rho_2 \boldsymbol{U}_2 \boldsymbol{U}_2' + \rho_3 \boldsymbol{I}_n$. We would like to assume the following sets of values for $\boldsymbol{\rho} = \{\rho_1, \rho_2, \rho_3\}$. For staggered-level designs, we assume that the two classes of hard-to-change factors have higher variability than the easy-to-change factors.

Table 3.2: Choice of values for ρ_1, ρ_2, and ρ_3 for staggered-level designs

	ρ_1	ρ_2	ρ_3
$\rho_1 > \rho_2 > \rho_3$	0.6	0.3	0.1
$\rho_1 > \rho_2 = \rho_3$	0.6	0.2	0.2
$\rho_1 > \rho_3 > \rho_2$	0.6	0.1	0.3
$\rho_2 > \rho_1 > \rho_3$	0.3	0.6	0.1
$\rho_2 > \rho_1 = \rho_3$	0.2	0.6	0.2
$\rho_2 > \rho_3 > \rho_1$	0.1	0.6	0.3

Chapter 4

Examples of Optimal Single-Stratum and Multistratum Designs

In this chapter, we consider optimal single-stratum and multi-stratum designs based on the criteria developed in Chapter 3. We will compare the performance of the designs constructed by the criteria. Sensitivity analysis of the parameters will be discussed. In Section 4.1, CRDs are considered. In Section 4.2, optimal split-plot designs are considered. Staggered-level designs are considered in Section 4.3.

4.1 CRD

Though Bayesian D-optimality criterion and S_P criterion for CRDs are in the literature for decades, detailed study on the optimal designs does not exist, as the computation burden for S_P is heavy for a large model space. In this section, we will use (2.9) and (2.11) to generate Bayesian D- and S_P-optimal CRDs. Since CRDs are single-stratum ($g = 1$) designs, we do not need to assume values for $\boldsymbol{\rho}$.

We consider four examples for CRDs. For both Example 4.1.1 and Example 4.1.2 as 2-level designs, we consider the models that contain main effects only. For the 3-level designs in Example 4.1.3 and Example 4.1.4, we consider the second-order model that contains the main effects and quadratic effects.

Example 4.1.1. We will use BD and S_P criteria to construct 5-run 2-level optimal saturated design with 4 factors. Table 4.1 lists out the BD-optimal design.

Table 4.1: A 5-run 2-level BD-optimal design with 4-factors

Run	X_1	X_2	X_3	X_4
1	-1	-1	1	1
2	1	1	-1	1
3	1	-1	-1	-1
4	1	1	1	-1
5	-1	1	-1	-1

For the S_P designs, we take on the π values in Table 4.2 by plugging in $E(\text{active effects}) = 1, 2, 3$ and solve for π from (3.12). Table 4.3 lists out the S_P-optimal designs with $\pi = 0.25, 0.50$ and 0.75.

Table 4.2: Choice of π for the S_P-optimality criterion in Example 4.1.1

E(active effects)	π
1	0.25
2	0.50
3	0.75

Table 4.3: 5-run 2-level S_P-optimal designs with 4 factors

(a) $\pi = 0.25$

Run	X_1	X_2	X_3	X_4
1	-1	-1	-1	1
2	-1	1	-1	-1
3	1	1	-1	1
4	-1	1	1	1
5	1	-1	1	-1

(b) π=0.50

Run	X_1	X_2	X_3	X_4
1	1	1	1	1
2	1	1	-1	-1
3	-1	1	-1	1
4	1	-1	-1	1
5	-1	-1	1	-1

(c) π=0.75

Run	X_1	X_2	X_3	X_4
1	1	1	1	-1
2	1	-1	-1	-1
3	1	-1	1	1
4	-1	-1	1	-1
5	-1	1	-1	1

We compare the D- and A-efficiencies of the designs in Table 4.1 and Table 4.3. Table 4.4 shows the result. All of the BD- and S_P-optimal designs have high parameter estimation efficiency, a value of 1 in D-efficiency and A-efficiency demonstrates that the designs are both D- and A-optimal.

Table 4.4: D- and A-efficiencies of the designs in Table 4.1 and Table 4.3

Design	D_{eff}	A_{eff}
BD	1	1
$S_P, \pi = 0.25$	1	1
$S_P, \pi = 0.50$	1	1
$S_P, \pi = 0.75$	1	1

Example 4.1.2. We construct 7-run 2-level saturated Bayesian D- and S_P-optimal designs with 6 factors. Table 4.5 list out the BD-optimal design.

Table 4.5: A 7-run 2-level BD-optimal design with 6 factors

Run	X_1	X_2	X_3	X_4	X_5	X_6
1	-1	1	1	1	-1	1
2	-1	1	1	1	1	-1
3	1	-1	-1	1	1	1
4	1	-1	1	-1	1	1
5	1	-1	1	1	-1	-1
6	1	1	-1	-1	-1	-1
7	-1	-1	-1	-1	-1	-1

By plugging in $E(\text{active effects}) = 1, 2, 3, 4, 5$ and solve for π from (3.12), we obtain the π values in Table 4.6. Table 4.7 lists out the S_P-optimal designs with $\pi = 0.1667, 0.3333, 0.5000, 0.6777$ and 0.8333.

Table 4.6: Choice of π for the S_P-optimality criterion in Example 4.1.2

E(active effects)	π
1	0.1667
2	0.3333
3	0.5000
4	0.6667
5	0.8333

Table 4.7: 7-run 2-level S_P-optimal designs with 6 factors

(a) $\pi = 0.1667$

Run	X_1	X_2	X_3	X_4	X_5	X_6
1	1	1	-1	1	-1	-1
2	1	-1	1	1	-1	1
3	1	-1	-1	-1	1	1
4	-1	-1	-1	-1	-1	-1
5	-1	1	1	-1	-1	1
6	1	1	1	-1	1	-1
7	-1	1	-1	1	1	1

(b) $\pi = 0.3333$

Run	X_1	X_2	X_3	X_4	X_5	X_6
1	1	1	-1	1	1	-1
2	-1	-1	-1	1	-1	1
3	1	-1	1	-1	1	1
4	1	-1	1	1	-1	-1
5	-1	-1	-1	-1	1	-1
6	-1	1	1	1	1	1
7	-1	1	1	-1	-1	-1

(c) $\pi = 0.5000$

Run	X_1	X_2	X_3	X_4	X_5	X_6
1	-1	-1	-1	1	-1	1
2	-1	1	1	1	1	1
3	-1	1	-1	-1	1	-1
4	1	1	-1	-1	-1	1
5	1	-1	1	-1	1	1
6	-1	-1	1	-1	-1	-1
7	1	1	1	1	-1	-1

(d) $\pi = 0.6667$

Run	X_1	X_2	X_3	X_4	X_5	X_6
1	-1	-1	-1	1	1	1
2	-1	1	1	1	-1	-1
3	-1	1	-1	-1	1	-1
4	1	1	-1	1	-1	1
5	1	-1	-1	-1	-1	-1
6	1	-1	1	1	1	-1
7	1	1	1	-1	1	1

(e) $\pi = 0.8333$

Run	X_1	X_2	X_3	X_4	X_5	X_6
1	-1	1	1	-1	1	1
2	1	1	-1	1	1	1
3	-1	-1	-1	1	1	-1
4	-1	-1	1	1	-1	1
5	1	1	1	1	-1	-1
6	1	-1	1	-1	1	-1
7	-1	1	-1	-1	-1	-1

Table 4.8 shows the D- and A-efficiencies of the BD- and S_P-optimal designs, the S_P-optimal designs are high in D-efficiencies, but relatively lower in A-efficiencies, while BD-optimal designs present strong D-efficiency and A-efficiency. In fact, the BD-optimal design is 100% D- and A-efficient. For S_P-optimal designs, the values of D- and A- efficiencies are invariant with different π values which indicates the S_P-optimal designs is capable of parameter

estimation regardless of the π value chosen.

Table 4.8: D- and A-efficiencies of the designs in Table 4.5 and Table 4.7

Design	D_{eff}	A_{eff}
GBD	1	1
$GS_P, \pi = 0.1667$	0.9669076	0.7301587
$GS_P, \pi = 0.3333$	0.9669076	0.7301587
$GS_P, \pi = 0.5000$	0.9669076	0.7301587
$GS_P, \pi = 0.6777$	0.9669076	0.7301587
$GS_P, \pi = 0.8333$	0.9669076	0.7301587

To further study the BD- and S_P-optimal designs, we present the absolute correlation heatmaps for the A-, D-, and BD-optimal designs, and the S_P-optimal design with $\pi = 0.8333$ in Figure 4.1. The A- and D-optimal designs are listed in Table 4.9. We can see that in terms of pairwise correlations. The BD-optimal design shows more correlations between variables than the S_P-optimal design for the correlations between X_1 and X_2, X_3 and X_4, X_5 and X_6, respectively. The S_p-optimal design shows only moderate correlations between the variables. The BD-optimal design itself has similar performance with the D-optimal design, but stronger correlation between variables compared to the A-optimal design.

Table 4.9: 7-run 2-level A- and D-optimal designs with 6 factors

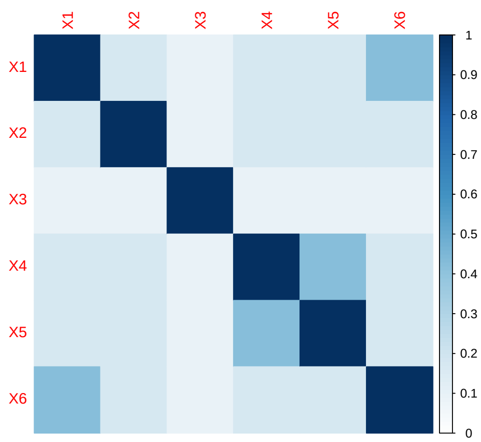
(a) A-optimal

Run	X_1	X_2	X_3	X_4	X_5	X_6
1	-1	-1	1	-1	-1	-1
2	1	1	-1	1	-1	-1
3	-1	-1	-1	1	1	1
4	1	1	1	1	1	-1
5	-1	1	1	-1	-1	1
6	-1	1	-1	-1	1	-1
7	1	-1	-1	-1	1	1

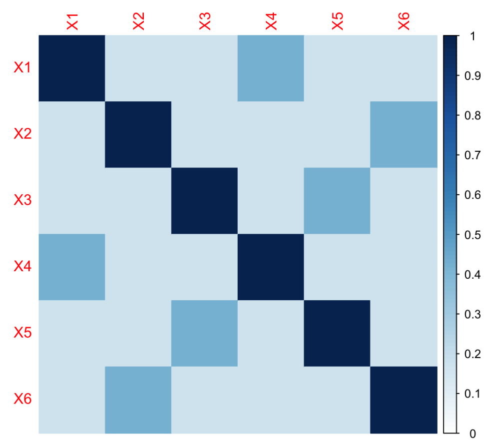
(b) D-optimal

Run	X_1	X_2	X_3	X_4	X_5	X_6
1	1	-1	-1	-1	1	-1
2	-1	1	-1	1	1	-1
3	-1	-1	1	1	1	1
4	-1	-1	-1	1	-1	1
5	-1	1	-1	-1	1	1
6	-1	-1	1	-1	-1	-1
7	1	1	1	1	-1	1

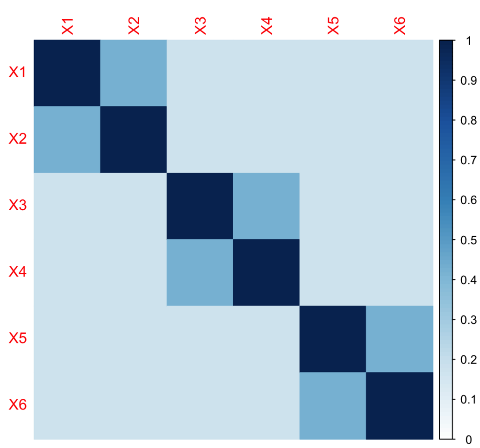
Figure 4.1: Absolute correlation maps of the A-, D-, BD- and S_P-optimal CRDs in Table 4.9, Table 4.5, and Table 4.7e



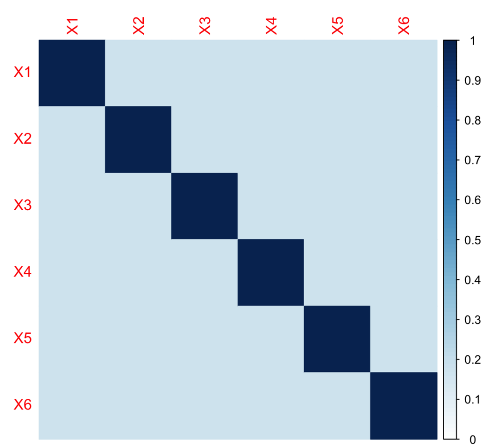
(a) A



(b) D



(c) BD



(d) $S_P, \pi = 0.8333$

Now we study the CRDs that have three different levels of settings. For such designs, we study the model that has both main effects and quadratic effects. It is noticeable that as the number of model parameter increases, the model space that is related to the calculation of S_P-optimality criterion increases exponentially. For example, for a design with four factors, when the model includes only main effects, the model space contains 2^4 models, however, when the model includes main and quadratic effects, which is the case in Example 4.1.3, the model space includes 2^8 models.

Example 4.1.3. We construct 9-run 3-level saturated designs with four factors, where the model contains four main effects and four quadratic effects. Table 4.10 list out the BD-optimal design.

Table 4.10: A 9-run 3-level BD-optimal design with 4 factors

Run	X_1	X_2	X_3	X_4
1	-1	1	1	-1
2	-1	0	-1	1
3	0	1	-1	0
4	1	-1	-1	-1
5	1	0	1	0
6	0	0	0	-1
7	0	-1	1	1
8	-1	-1	0	0
9	1	1	0	1

By plugging in $E(\text{active effects}) = 1, 2, 3, 4, 5, 6, 7$ and solve for π from (3.13), we obtain the π values in Table 4.11 which willl be used for the construction of S_P-optimal designs. Table 4.12 lists out the corresponding S_P-optimal designs.

Table 4.11: Choice of π for S_P-optimal designs in Example 4.1.3

E(active effects)	π
1	0.125
2	0.250
3	0.375
4	0.500
5	0.625
6	0.750
7	0.875

We compare the D- and I-efficiencies of the designs in Table 4.10 and Table 4.12, and the D- and I-efficiency values are listed in Table 4.13. We can see that when π has the lowest value of 0.125, the S_P-optimal design does not perform well in terms of both D- and I-efficiencies, which are 0.6874 and 0.4616, respectively. However, the BD and S_P designs at $\pi = 0.250, 0.375, 0.500, 0.625,$ and 0.750 performed well in terms D- and I-efficiencies, having efficiencies values of ones.

Table 4.13: D- and I-efficiencies of the designs in Table 4.10 and Table 4.12

Design	D_{eff}	I_{eff}
GBD	1	1
$GS_P, \pi = 0.125$	0.6874	0.4616
$GS_P, \pi = 0.250$	1	1
$GS_P, \pi = 0.375$	1	1
$GS_P, \pi = 0.500$	1	1
$GS_P, \pi = 0.625$	1	1
$GS_P, \pi = 0.750$	1	1
$GS_P, \pi = 0.875$	1	1

Table 4.12: 9-run 3-level S_P-optimal designs with 4 factors

(a) $\pi = 0.125$

Run	X_1	X_2	X_3	X_4
1	0	0	-1	-1
2	0	-1	-1	0
3	-1	0	0	0
4	1	-1	0	-1
5	0	0	0	0
6	0	1	1	1
7	1	1	1	0
8	1	0	-1	1
9	-1	1	0	1

(b) $\pi = 0.250$

Run	X_1	X_2	X_3	X_4
1	0	-1	1	1
2	-1	0	1	-1
3	0	0	-1	0
4	1	0	0	1
5	1	-1	-1	-1
6	1	1	1	0
7	-1	-1	0	0
8	-1	1	-1	1
9	0	1	0	-1

(c) $\pi = 0.375$

Run	X_1	X_2	X_3	X_4
1	0	-1	1	0
2	1	0	0	0
3	0	1	0	1
4	-1	1	-1	0
5	1	1	1	-1
6	-1	-1	0	-1
7	0	0	-1	-1
8	1	-1	-1	1
9	-1	0	1	1

(d) $\pi = 0.500$

Run	X_1	X_2	X_3	X_4
1	0	1	-1	-1
2	1	-1	1	-1
3	-1	0	0	-1
4	-1	1	1	0
5	-1	-1	-1	1
6	0	-1	0	0
7	1	0	-1	0
8	0	0	1	1
9	1	1	0	1

(e) $\pi = 0.625$

Run	X_1	X_2	X_3	X_4
1	0	1	0	-1
2	-1	-1	0	1
3	1	1	1	1
4	1	-1	-1	-1
5	-1	0	1	-1
6	0	-1	1	0
7	-1	1	-1	0
8	0	0	-1	1
9	1	0	0	0

(f) $\pi = 0.750$

Run	X_1	X_2	X_3	X_4
1	1	0	1	-1
2	1	-1	-1	1
3	0	1	1	1
4	-1	1	-1	-1
5	-1	0	0	1
6	-1	-1	1	0
7	1	1	0	0
8	0	-1	0	-1
9	0	0	-1	0

(g) $\pi = 0.875$

Run	X_1	X_2	X_3	X_4
1	0	-1	1	-1
2	1	0	0	-1
3	-1	-1	0	1
4	1	1	1	1
5	1	-1	-1	0
6	0	1	0	0
7	-1	0	1	0
8	0	0	-1	1
9	-1	1	-1	-1

We further study the BD- and S_P-optimal designs by comaring them with other optimal designs. Since three-level designs are often used for the response prediction, we compare our optimal designs with the D- and I-optimal designs, which are listed in Table 4.14. We then plot the absolute correlation heatmaps for the D-, I-, BD-optimal designs, and the S_P-optimal design with $\pi = 0.875$. It can be seen that none of the generated optimal designs shows signs of pairwise correlations between any two variables, which indicates that all the designs are orthogonal.

Table 4.14: 9-run 3-level D- and I- optimal designs with 4 factors

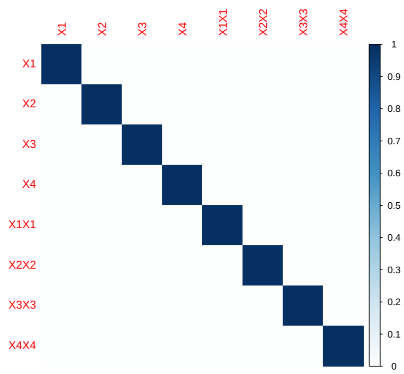
(a) D-optimal

Run	X_1	X_2	X_3	X_4
1	1	-1	-1	0
2	-1	1	1	0
3	1	0	1	-1
4	0	1	-1	-1
5	-1	-1	0	-1
6	0	-1	1	1
7	1	1	0	1
8	-1	0	-1	1
9	0	0	0	0

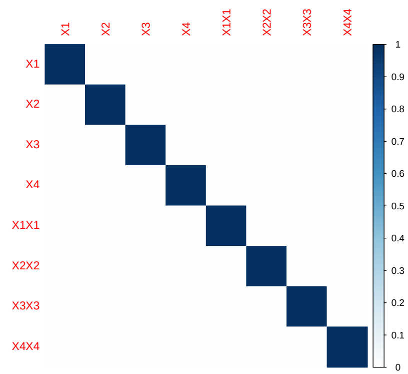
(b) I-optimal

Run	X_1	X_2	X_3	X_4
1	-1	-1	-1	1
2	-1	0	0	0
3	0	-1	0	-1
4	0	1	-1	0
5	1	1	0	1
6	-1	1	1	-1
7	0	0	1	1
8	1	0	-1	-1
0	1	-1	1	0

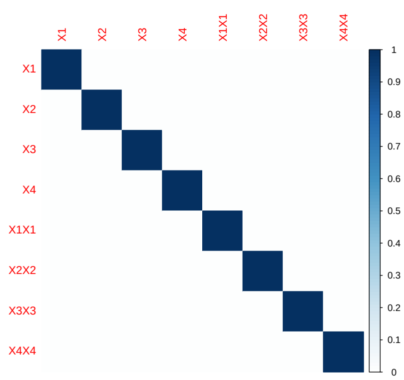
Figure 4.2: Absolute correlation maps for D-, I-, BD- and S_P-optimal CRDs in Table 4.14, Table 4.10, and Table 4.12g



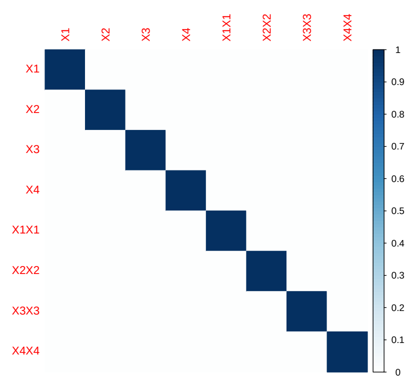
(a) D



(b) I



(c) BD



(d) $S_P, \pi = 0.875$

Example 4.1.4. We construct 11-run 3-level saturated BD- and S_P-optimal designs with five factors, where the model being considered contains five main effects and five quadratic effects. Table 4.15 lists out the BD-optimal design.

Table 4.15: A 11-run 3-level BD-optimal design with 5 factors

Run	X_1	X_2	X_3	X_4	X_5
1	1	-1	1	-1	1
2	-1	0	-1	-1	-1
3	1	1	1	1	-1
4	-1	-1	1	0	0
5	0	0	1	1	1
6	1	-1	-1	1	0
7	0	1	0	-1	0
8	-1	1	-1	1	1
9	0	-1	-1	0	-1
10	-1	-1	0	1	-1
11	1	0	0	0	1

By plugging in $E(\text{active effects}) = 1, 2, 3, 4, 5, 6, 7, 8, 9$ and solve for π from (3.13), we obtain that $\pi = 0.1, 0.2, 0.3, 0.3, 0.5, 0.6, 0.7, 0.8$, and 0.9, which are listed in Table 4.16. Table 4.17 lists out the corresponding S_P-optimal designs.

Table 4.16: Choice of π for S_P-optimal designs in Example 4.1.4

E(active effects)	π
1	0.1
2	0.2
3	0.3
4	0.4
5	0.5
6	0.6
7	0.7
8	0.8
9	0.9

Table 4.17: 11-run 3-level S_P-optimal designs with 5 factors

(a) $\pi = 0.1$

Run	X_1	X_2	X_3	X_4	X_5
1	0	0	1	0	0
2	0	-1	0	0	1
3	1	0	1	0	-1
4	-1	0	0	0	0
5	0	0	-1	-1	-1
6	-1	1	-1	0	1
7	0	0	0	1	1
8	0	1	1	1	0
9	-1	-1	0	-1	0
10	1	1	0	-1	-1
11	1	-1	-1	1	0

(b) $\pi = 0.2$

Run	X_1	X_2	X_3	X_4	X_5
1	1	0	1	0	0
2	0	0	0	-1	1
3	0	-1	1	0	1
4	0	1	1	-1	0
5	-1	0	1	1	1
6	-1	-1	0	0	0
7	0	0	0	0	-1
8	-1	1	-1	0	-1
9	1	-1	-1	-1	-1
10	0	0	-1	1	0
11	1	1	0	1	1

(c) $\pi = 0.3$

Run	X_1	X_2	X_3	X_4	X_5
1	0	1	1	-1	0
2	0	-1	1	0	1
3	0	0	0	0	0
4	1	0	1	1	0
5	0	1	-1	1	-1
6	1	1	0	0	1
7	-1	0	-1	-1	1
8	1	-1	0	-1	-1
9	-1	0	1	0	-1
10	-1	-1	0	1	0
11	1	-1	-1	0	0

(d) $\pi = 0.4$

Run	X_1	X_2	X_3	X_4	X_5
1	-1	1	-1	-1	-1
2	1	1	-1	0	0
3	0	1	1	1	1
4	0	-1	1	-1	0
5	1	0	0	-1	1
6	1	-1	0	1	-1
7	1	0	1	0	-1
8	0	1	0	0	-1
9	-1	-1	-1	0	1
10	0	0	-1	1	-1
11	-1	0	0	1	0

(e) $\pi = 0.5$

Run	X_1	X_2	X_3	X_4	X_5
1	-1	-1	0	0	0
2	1	-1	-1	1	0
3	0	-1	1	-1	-1
4	1	0	1	-1	0
5	-1	0	-1	0	-1
6	0	0	0	1	1
7	-1	1	1	1	-1
8	1	1	0	-1	-1
9	-1	1	-1	-1	1
10	1	1	1	0	1
11	0	1	-1	0	0

(f) $\pi = 0.6$

Run	X_1	X_2	X_3	X_4	X_5
1	1	1	-1	0	1
2	0	0	-1	-1	0
3	1	0	0	-1	-1
4	0	0	1	1	1
5	-1	1	1	-1	-1
6	1	-1	-1	1	-1
7	1	-1	1	0	0
8	-1	1	0	1	0
9	-1	-1	-1	-1	1
10	-1	0	-1	0	-1
11	0	-1	0	0	-1

(g) $\pi = 0.7$

Run	X_1	X_2	X_3	X_4	X_5
1	1	-1	0	-1	-1
2	1	1	1	-1	1
3	-1	0	1	-1	-1
4	-1	1	-1	-1	0
5	0	-1	1	1	0
6	1	1	-1	1	-1
7	1	0	0	0	0
8	0	1	0	0	-1
9	0	0	-1	-1	1
10	-1	-1	-1	0	1
11	-1	1	0	1	1

(h) $\pi = 0.8$

Run	X_1	X_2	X_3	X_4	X_5
1	1	1	-1	-1	1
2	1	-1	-1	1	0
3	-1	1	0	1	1
4	-1	-1	-1	-1	-1
5	-1	1	1	-1	0
6	0	1	-1	0	-1
7	-1	0	-1	1	1
8	0	-1	1	-1	1
9	1	0	1	1	-1
10	1	0	0	-1	0
11	1	-1	1	0	1

(i) $\pi = 0.9$

Run	X_1	X_2	X_3	X_4	X_5
1	1	-1	-1	-1	-1
2	1	1	1	1	0
3	-1	-1	0	1	-1
4	0	1	1	-1	-1
5	1	1	0	-1	1
6	1	0	1	0	-1
7	-1	1	-1	1	-1
8	-1	0	-1	-1	0
9	-1	1	-1	0	1
10	-1	-1	1	-1	1
11	0	-1	-1	1	1

The D- and I-efficiency values of the BD- and S_P-optimal designs are listed in Table 4.18. The best overall performance we can see from the optimal designs is with S_P-optimal design at $\pi = 0.7$, which are 99.16% D-efficient and 88.94% I-efficient. As we can see, the I and D efficiencies of S_P-optimal designs drops at the minimum and maximum values of π at $0.1, 0.2, 0.8, 0.9$, but is stabilized when π is chosen at 0.3 to 0.7.

Table 4.18: D- and I-efficiencies of the designs in Table 4.15 and Table 4.17

Design	D_{eff}	I_{eff}
BD	0.9916	0.7892
$S_P, \pi = 0.1$	0.7063	0.3014
$S_P, \pi = 0.2$	0.8956	0.7630
$S_P, \pi = 0.3$	0.9828	0.9231
$S_P, \pi = 0.4$	0.9916	0.7606
$S_P, \pi = 0.5$	0.9916	0.8471
$S_P, \pi = 0.6$	0.9828	0.8375
$S_P, \pi = 0.7$	0.9916	0.8894
$S_P, \pi = 0.8$	0.9023	0.4698
$S_P, \pi = 0.9$	0.8223	0.2201

The absolute correlation heatmaps of the D-, I-, and BD-optimal designs, and the S_P-optimal design with $\pi = 0.7$ are shown in Figure 4.3. The D- and I-optimal designs are listed in Table 4.19. For the main effects, D-, I, and S_P-optimal designs have similar performance. We only see weak pairwise correlations between any two main effects for the D-optimal design. For the BD-optimal design, we only see weak correlations between two main effects, except that no pairwise correlation is found for X_1 and X_2, X_1 and X_4, X_2 and

X_3, X_4 and X_5, respectively. For the S_P-optimal design, we observe similar behaviour, as we see weak correlations between two main effects, except that no pairwise correlation is found for X_1 and X_2, X_1 and X_4, X_2 and X_5, X_4 and X_5, respectively. For the quadratic effects, the BD- and D-optimal designs have similar performance, showing only weak pairwise correlations between the quadratic effects. For the S_P-optimal design, we see moderate correlations between ${X_3}^2$ (labelled as X_3X_3) and X_4^2 (labelled as X_4X_4). We do not observe any moderate or severe correlations among quadratic effects and main effects. In general, BD- and S_P-optimal designs show less correlations between effects than D- and I-optimal designs.

Table 4.19: 11-run 3-level D- and I- optimal designs with 4 factors

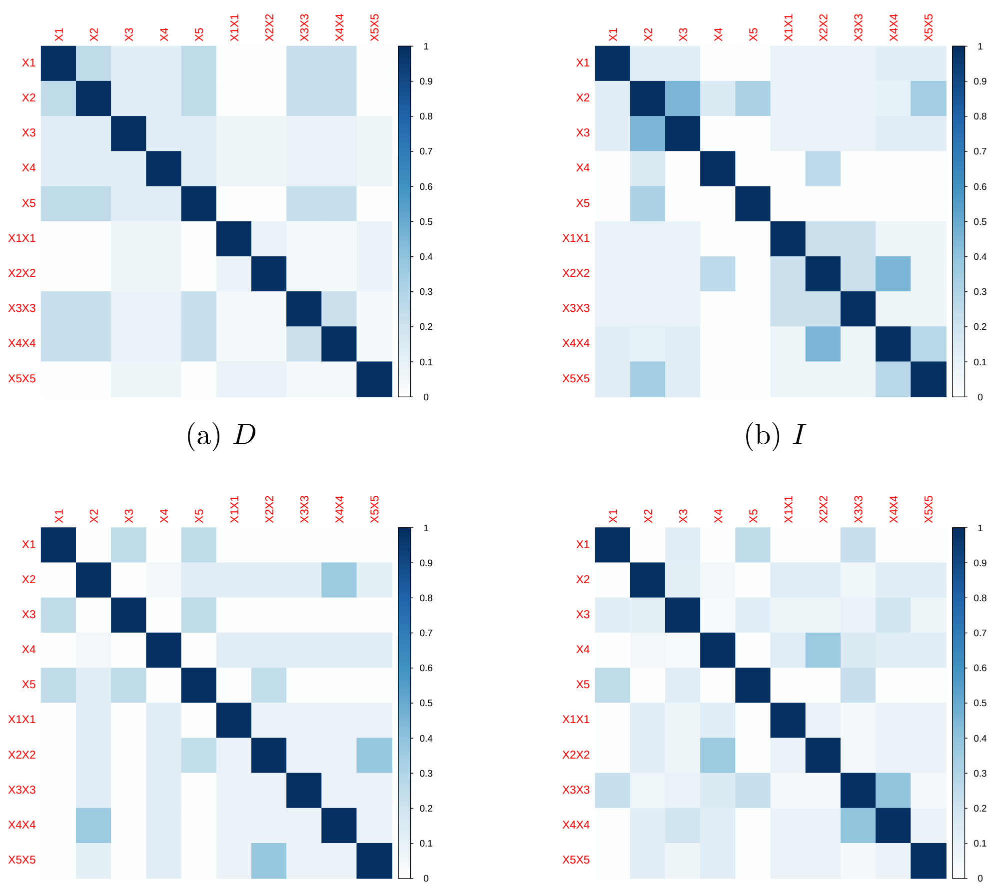
(a) D-optimal

Run	X_1	X_2	X_3	X_4	X_5
1	0	1	0	1	-1
2	1	0	0	0	-1
3	-1	0	1	1	1
4	-1	-1	-1	1	-1
5	1	-1	0	-1	1
6	1	1	-1	0	1
7	-1	1	1	-1	-1
8	-1	1	0	0	0
9	1	-1	1	1	0
10	0	0	-1	-1	0
11	0	-1	1	0	1

(b) I-optimal

Run	X_1	X_2	X_3	X_4	X_5
1	1	1	-1	0	0
2	-1	0	1	0	0
3	0	0	-1	0	-1
4	1	0	0	0	1
5	-1	1	0	-1	-1
6	0	-1	0	0	0
7	1	1	1	1	-1
8	-1	-1	-1	1	1
9	0	0	0	1	0
10	0	1	1	-1	1
11	1	-1	-1	-1	0

Figure 4.3: Absolute correlation maps for D-, I-, BD- and S_P-optimal CRDs in Table 4.19, Table 4.15 and Table 4.17g



(a) D

(b) I

(c) BD

(d) $S_P, \pi = 0.7$

4.2 Split-Plot Designs

For split-plot designs, the GBD- and GSP-optimal designs will be constructed using (3.3) and (3.6). For two-level designs, we consider the model containing the main effects and two-factor interactions. For three-level designs, we consider the model containing the main effects, quadratic effects, and two-factor interactions. For the choice of $\boldsymbol{\rho}$, we consider Table 3.1 for reference.

Example 4.2.1. We construct 18-run 2-level split-plot GBD- and GS_P-optimal designs with three factors where each factor has two levels. The model contains three main effects and $\binom{3}{2} = 3$ two-factor interaction effects. For GS_P-optimal designs, π can be calculated by plugging in $E(\text{active effects}) = 1, 2, 3, 4, 5$ into (3.14) and solve for π, we obtain the π values in Table 4.20. When we fix π at 0.8844 and when $\rho_1 = 0.5$ and $\rho_2 = 0.5$, Table 4.21b lists out the GS_P-optimal design, and Table 4.21a shows the generated GBD-optimal design.

Table 4.20: Choice of π for GS_P-optimal designs in Example 4.2.1

E(active effects)	π
1	0.2628
2	0.4567
3	0.6176
4	0.7581
5	0.8844

The D- and A-efficiencies for the designs in Table 4.21 are shown in Table 4.22. Both GBD- and GS_P-optimal designs are 100% A-efficient and D-efficient.

Table 4.21: 18-run 2-level GBD- and GS_P-optimal split-plot designs with three factors when $\rho_1 = 0.5$ and $\rho_2 = 0.5$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	-1	-1	1
2	-1	-1	-1
3	-1	1	1
4	-1	1	-1
5	-1	-1	-1
6	-1	-1	1
7	-1	-1	-1
8	-1	1	1
9	-1	1	-1
10	1	1	1
11	1	-1	-1
12	1	-1	1
13	1	1	1
14	1	-1	1
15	1	1	-1
16	1	-1	-1
17	1	-1	1
18	1	1	-1

(b) GS_P-optimal design with $\pi = 0.8844$

Run	W_1	X_1	X_2
1	-1	-1	-1
2	-1	-1	1
3	-1	1	-1
4	1	-1	1
5	1	1	1
6	1	1	-1
7	1	1	1
8	1	-1	1
9	1	-1	-1
10	-1	1	1
11	-1	1	-1
12	-1	-1	-1
13	-1	1	1
14	-1	1	-1
15	-1	-1	1
16	1	-1	1
17	1	-1	-1
18	1	1	-1

Table 4.22: D- and A-efficiencies of the GBD and GS_P-optimal split-plot designs in Table 4.21

Efficiency	GBD	GSP
D_{eff}	1.0000	1.0000
A_{eff}	1.0000	1.0000

We would like to study the effect of changing π for GS_P-optimal designs. When we fix $\rho_1 = 0.5$ and $\rho_2 = 0.5$, we evaluate the GS_P-optimal designs for the π values in Table 4.20. Table B.1 in Appendix A list out the split-plot GS_P-optimal designs for $\pi = 0.2628, 0.4567, 0.6176, 0.7581$, and 0.8844. Table 4.23

lists out the D- and A-efficiencies of the corresponding GS_P-optimal designs. We can see that the generated GS_P-optimal split-plot designs have consistent performance over different values of π in terms of D-efficiency and A-efficiency. In particular, all of the designs are D- and A-optimal, which demonstrates good parameter estimation efficiency under any values of π being taken into consideration.

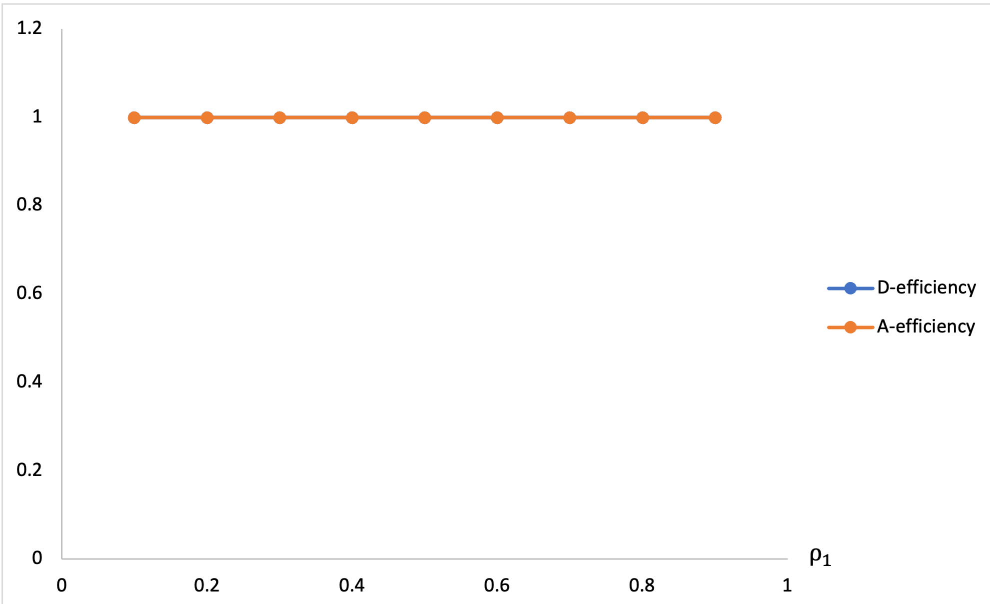
Table 4.23: D- and A-efficiencies of the the GS_P-optimal split-plot designs in Table B.1

π	D_{eff}	A_{eff}
0.2628	1.0000	1.0000
0.4567	1.0000	1.0000
0.6176	1.0000	1.0000
0.7581	1.0000	1.0000
0.8844	1.0000	1.0000

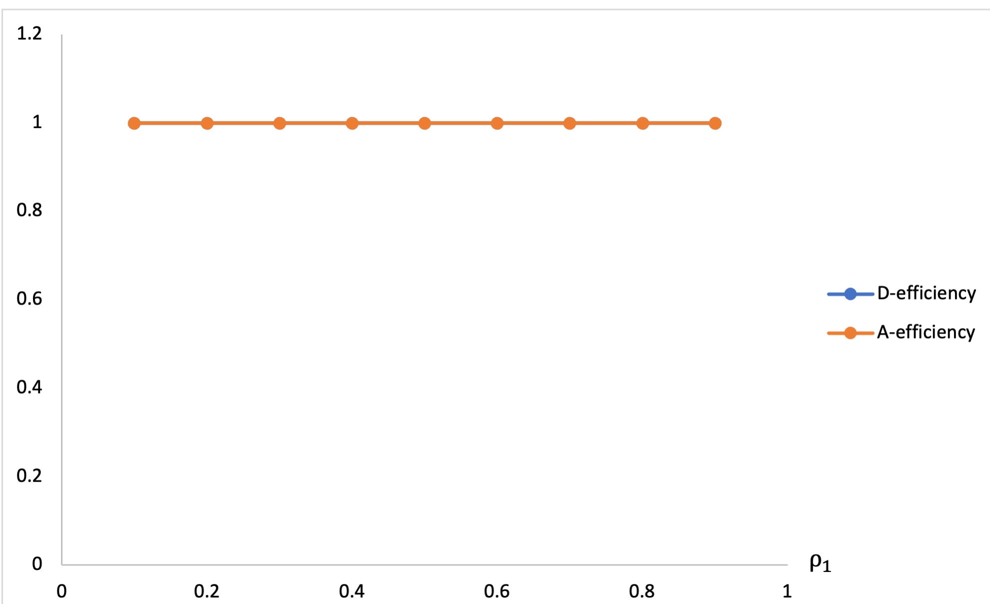
Next, we would like to study how the optimal designs perform with changing variance ratios $\boldsymbol{\rho}$ has on both GBD- and GS_P-optimal designs. When we fix $\pi = 0.8844$, we would test the efficiency changes when $\boldsymbol{\rho}$ changes according to Table 3.1. Table B.2 - Table B.10 in Appendix A list out all the generated GBD- and GS_P-optimal designs. The D- and A-efficiencies of the GBD- and GS_P-optimal designs are plotted in Figure 4.4. We notice that all optimal designs are both 100% D- and A-efficient.

Next, we examine the absolute correlation heatmaps for the GBD- and GS_P-optimal split-plot designs in Table 4.24. Figure 4.5 shows the heatmaps, where the D- and A-optimal designs are listed in Table 4.25. From Figure 4.5, all of the D-, A-, GBD-, and GS_P-optimal designs show only weak pairwise

Figure 4.4: D- and A-efficiency of the GBD- and GS_P-optimal split-plots designs in Table B.2 - Table B.10 when $\rho_1 = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$



(a) D- and A-efficiencies of the GBD-optimal split-plot designs in Table B.2a-Table B.10a



(b) D- and A-efficiencies of the GS_P-optimal split-plot designs in Table B.2b - Table B.10b when $\pi = 0.8844$

correlations between effects. The A- and GS_P-optimal designs show the same performance that there are weak correlations between main effects, and between the interaction effects, but there is no correlation between main effects and interaction effects. For the GBD-optimal design, we see a scattered correlation pattern compared to other optimal designs. The GBD-optimal design performs similar to D-optimal design in terms of pairwise correlations between the effects, but performs better than the GS_P-optimal design, as smaller correlations between effects are observed in GBD-optimal design.

Table 4.24: 18-run 2-level GBD- and GS_P-optimal split-plot designs with 3 factors when $\rho_1 = 0.5$ and $\rho_2 = 0.5$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	-1	-1	1
2	-1	-1	-1
3	-1	1	1
4	-1	1	-1
5	-1	-1	-1
6	-1	-1	1
7	-1	-1	-1
8	-1	1	1
9	-1	1	-1
10	1	1	1
11	1	-1	-1
12	1	-1	1
13	1	1	1
14	1	-1	1
15	1	1	-1
16	1	-1	-1
17	1	-1	1
18	1	1	-1

(b) GS_P-optimal design with $\pi = 0.8844$

Run	W_1	X_1	X_2
1	-1	-1	-1
2	-1	-1	1
3	-1	1	-1
4	1	-1	1
5	1	1	1
6	1	1	-1
7	1	1	1
8	1	-1	1
9	1	-1	-1
10	-1	1	1
11	-1	1	-1
12	-1	-1	-1
13	-1	1	1
14	-1	1	-1
15	-1	-1	1
16	1	-1	1
17	1	-1	-1
18	1	1	-1

Table 4.25: 18-run 2-level D- and A-optimal split-plot designs with 3 factors when $\rho_1 = 0.5$ and $\rho_2 = 0.5$

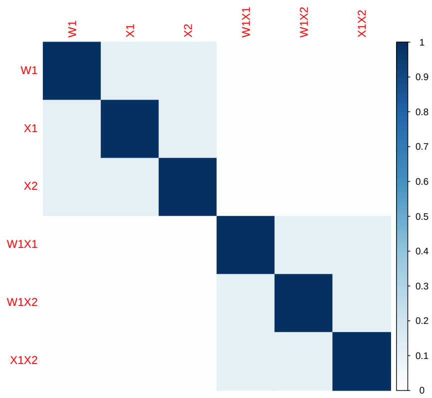
(a) D-optimal design

Run	W_1	X_1	X_2
1	1	-1	-1
2	1	1	-1
3	1	1	1
4	1	1	-1
5	1	-1	-1
6	1	-1	1
7	1	1	-1
8	1	-1	1
9	1	1	1
10	-1	1	1
11	-1	-1	-1
12	-1	-1	1
13	-1	1	1
14	-1	1	-1
15	-1	-1	-1
16	-1	1	-1
17	-1	-1	-1
18	-1	-1	1

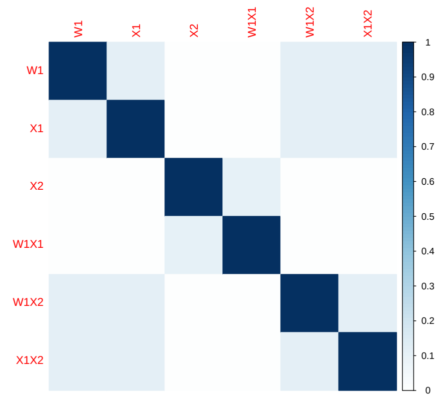
(b) A-optimal design

Run	W_1	X_1	X_2
1	-1	-1	1
2	-1	-1	-1
3	-1	1	-1
4	-1	-1	1
5	-1	1	-1
6	-1	1	1
7	-1	1	1
8	-1	1	-1
9	-1	-1	-1
10	1	-1	1
11	1	1	1
12	1	1	-1
13	1	-1	1
14	1	-1	-1
15	1	1	1
16	1	1	-1
17	1	-1	-1
18	1	-1	1

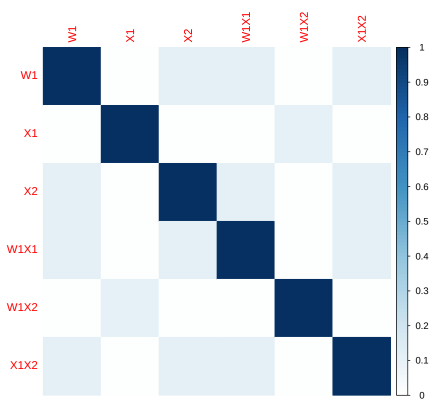
Figure 4.5: Absolute correlation maps for A-, D-, GBD- and GS_P-optimal split-plot designs in Table 4.24 and Table 4.25



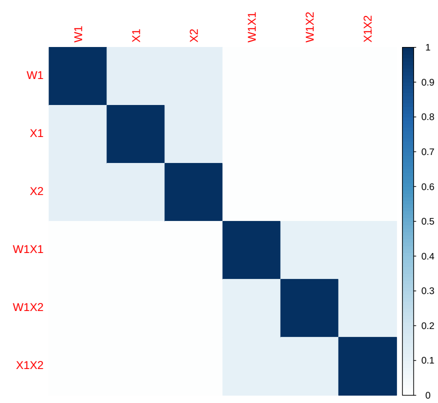
(a) A



(b) D



(c) GBD



(d) $GS_P, \pi = 0.8844$

Next, we would like to study the optimal design generated for more complex experimental settings, when each factor has three levels of settings, and study their performance. For GS_P-optimal designs that are considering the entire model space, we consider the model space that contains all main effects, two-factor interaction effects, and quadratic effects.

Example 4.2.2. We construct 18-run 3-level optimal split-plot designs with 3 factors. Since there are three factors, the model includes the three main effects, three quadratic effects, and $\binom{3}{2} = 3$ two-factor interactions. We construct GS_P-optimal split-plot design first. We choose the π values by plugging in $E(\text{active effects}) = 1, 2, 3, 4, 5, 6, 7, 8$ into (3.11), The obtained π values are shown in Table 4.26. Table B.11 in Appendix A lists the corresponding GS_P-optimal split-plot designs with $\rho_1 = 0.5, \rho_2 = 0.5$.

Table 4.26: Choice of π for GS_P-optimal designs in Example 4.2.2

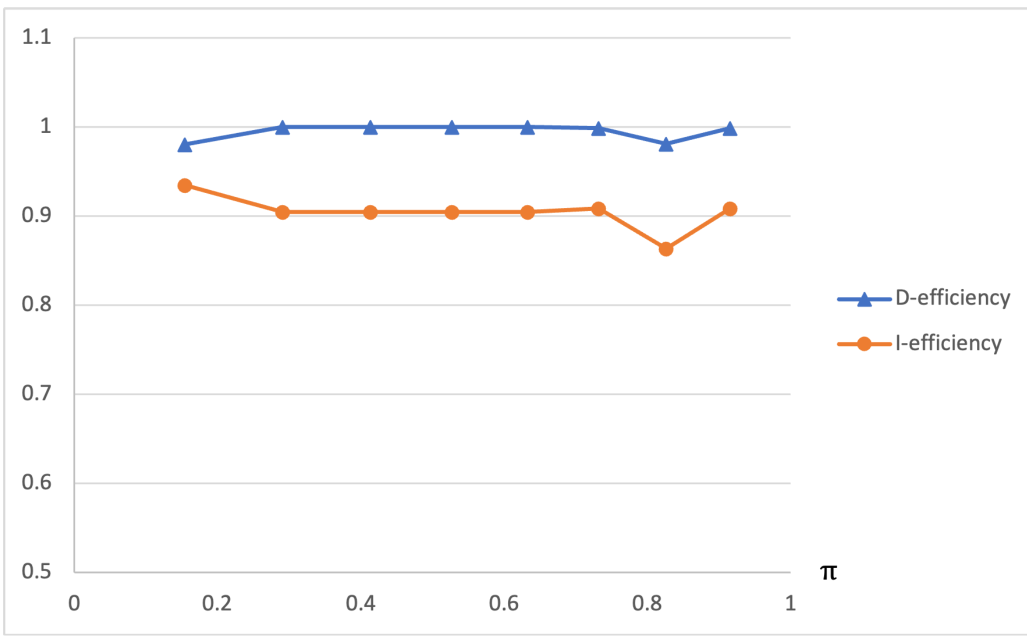
E(active effects)	π
1	0.1542
2	0.2904
3	0.4137
4	0.5271
5	0.6327
6	0.7319
7	0.8257
8	0.9148

The D- and I-efficiencies are shown in Table 4.27 with line plot demonstration in Figure 4.6. The GS_P-optimal split-plot designs show high D-efficiency values and has good parameter estimation performance. The I-efficiency value is highest ($I_{eff} = 0.9349$) when $\pi = 0.1542$, and lowest ($I_{eff} = 0.8635$) when $\pi = 0.8257$. Overall, all of the GS_P-optimal designs show good prediction capabilities.

Table 4.27: D- and I-efficiencies of the GS_P-optimal split-plot designs in Table B.11

Choice of π	D_{eff}	I_{eff}
0.1542	0.9806	0.9349
0.2904	1.0000	0.9047
0.4137	1.0000	0.9047
0.5271	1.0000	0.9047
0.6327	1.0000	0.9047
0.7319	0.9987	0.9085
0.8257	0.9813	0.8635
0.9148	0.9987	0.9085

Figure 4.6: D- and I-efficiencies of the GS_P-optimal split-plot designs of Table B.11 when $\pi = 0.1542, 0.2904, 0.4137, 0.5271, 0.6327, 0.7319, 0.8257, 0.9148$ and ρ_1=$\rho_2 = 0.5$



Now we would like to examine if the performance is consistent for different choice of $\boldsymbol{\rho}$ listed in Section 3.3 for both GBD- and GS_P-optimal three-level split-plot designs. Table B.12 - Table B.20 in Appendix A list out all the GBD-optimal designs and GS_P-optimal designs with $\pi = 0.7319$. Table 4.28

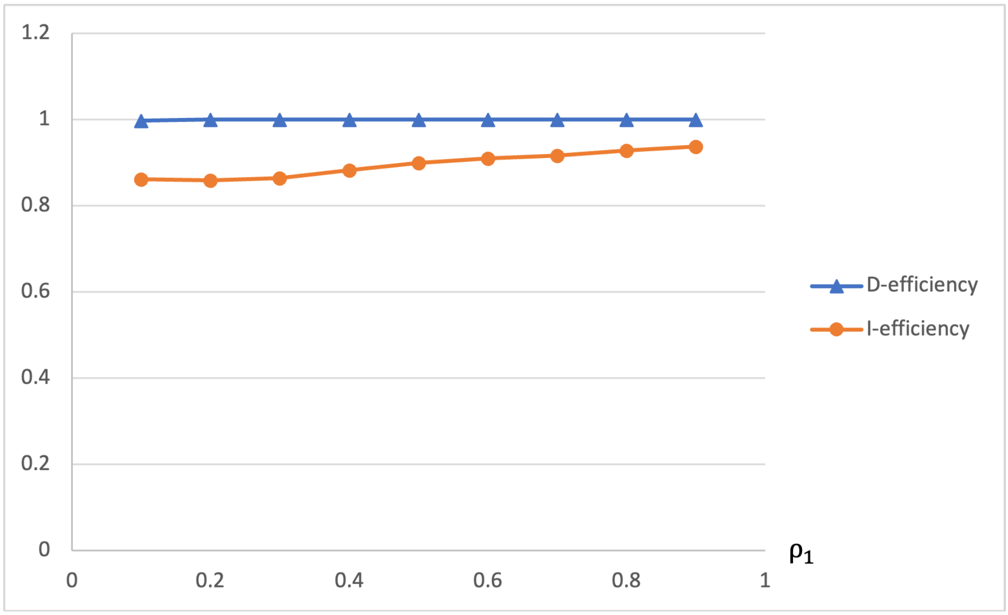
show the corresponding D- and I-efficiencies of the designs. The line plots are shown in Figure 4.7. We can see that the GBD-optimal split-plot designs have good parameter estimation efficiencies in terms of D-efficiency values, and as the variance ratio ρ_1 increases from 0.1 to 0.9, the corresponding I-efficiency increases as well. The GS_P-optimal designs are also D-optimal except for the case that $rho_1 = 0.1$, and $\rho_2 = 0.9$, where the D-efficiency is 99.08%, which is still very high. The I-efficiency decreases from $\rho_1 = 0.1$ to $\rho_1 = 0.4$, and increases afterwards. In general, GS_P-optimal designs perform better than GBD-optimal designs in terms of D- and I-efficiencies.

Table 4.28: D- and I-efficiencies for GBD- and GS_P-optimal split-plot designs in Table Table B.12 - Table B.20

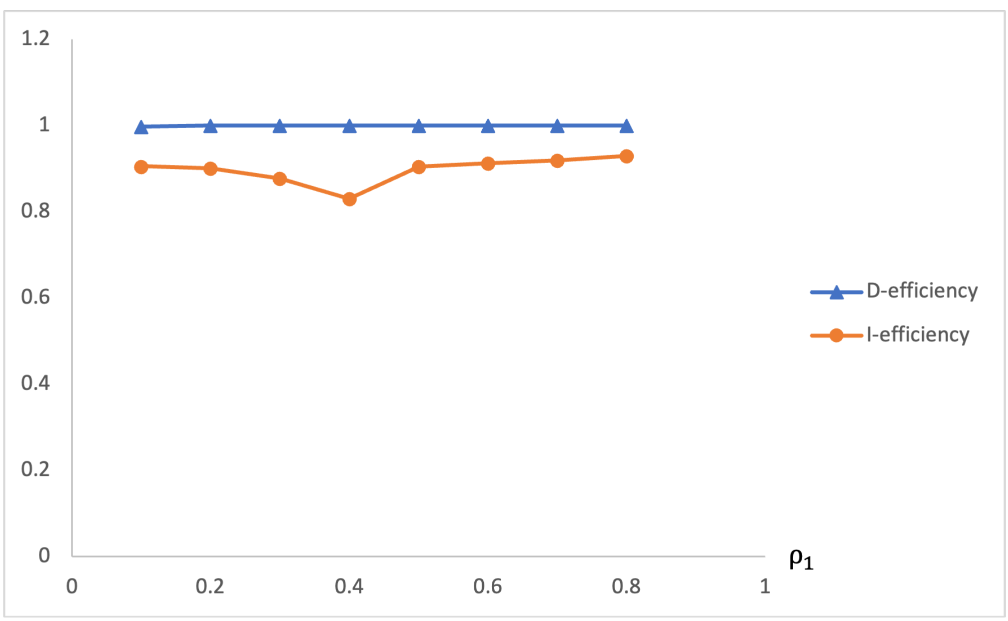
ρ_1	ρ_2	GBD,D_{eff}	GS_P,D_{eff}	GBD,I_{eff}	GS_P,I_{eff}
0.1	0.9	0.9908	0.9982	0.8614	0.9057
0.2	0.8	0.9904	1.0000	0.8592	0.9012
0.3	0.7	0.9967	1.0000	0.8646	0.8778
0.4	0.6	0.9880	1.0000	0.8828	0.8307
0.5	0.5	0.9987	1.0000	0.8999	0.9047
0.6	0.4	0.9980	1.0000	0.9104	0.9134
0.7	0.3	0.9975	1.0000	0.9167	0.9190
0.8	0.2	0.9970	1.0000	0.9285	0.9301
0.9	0.1	0.9966	1.0000	0.9376	0.9384

Finally, we examine the absolute correlation heatmaps of the GBD-optimal design and GS_P-optimal design with π at 0.7319 in Table 4.29, and compare them with the D- and I-optimal designs, which are listed in in Table 4.30. Surprisingly, the GBD- and D-optimal design show the same correlation between effects, while the GS_P-optimal design displays the similar, but weaker

Figure 4.7: D- and I-efficiencies of the GBD- and GS_P-optimal split-plot designs in Table B.12-Table B.20



(a) D- and I-efficiencies of the GBD-optimal split-plot designs in Table B.12a-Table B.20a when $\rho_1 = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$



(b) D- and I-efficiencies of the GS_P-optimal split-plot designs in Table B.12b-Table B.20b when $\rho_1 = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9$, and $\pi = 0.7319$

correlations than the GBD- and D-optimal designs. Although the I-optimal design is an orthogonal design, it shows stronger correlations between quadratic effects than other optimal designs.

Table 4.29: 18-run 3-level optimal split-plot designs with 3 factors when $\rho_1 = 0.5$ and $\rho_2 = 0.5$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	-1	-1	-1
2	-1	1	-1
3	-1	0	1
4	-1	1	1
5	-1	-1	1
6	-1	0	-1
7	1	1	1
8	1	-1	0
9	1	1	-1
10	0	0	0
11	0	1	1
12	0	-1	1
13	0	-1	0
14	0	1	-1
15	0	0	1
16	1	-1	1
17	1	1	0
18	1	-1	-1

(b) GS_P-optimal design with $\pi = 0.7319$

Run	W_1	X_1	X_2
1	0	-1	-1
2	0	0	0
3	0	1	1
4	1	1	1
5	1	-1	1
6	1	0	-1
7	1	-1	-1
8	1	1	-1
9	1	0	1
10	0	1	-1
11	0	-1	0
12	0	0	1
13	-1	-1	-1
14	-1	-1	1
15	-1	1	0
16	-1	1	-1
17	-1	1	1
18	-1	-1	0

Table 4.30: 18-run 3-level D- and I-optimal split-plot designs with 3 factors when $\rho_1 = 0.5$ and $\rho_2 = 0.5$

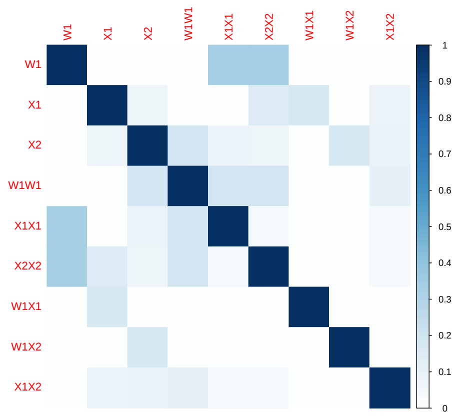
(a) D-optimal design

Run	W_1	X_1	X_2
1	-1	1	-1
2	-1	-1	-1
3	-1	0	1
4	0	0	1
5	0	1	-1
6	0	-1	0
7	1	-1	-1
8	1	1	0
9	1	-1	1
10	0	1	1
11	0	-1	1
12	0	0	0
13	1	-1	0
14	1	1	1
15	1	1	-1
16	-1	-1	1
17	-1	0	-1
18	-1	1	1

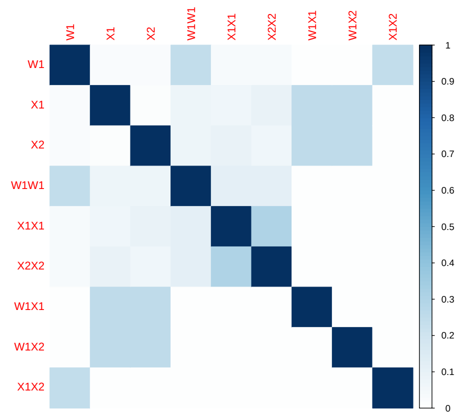
(b) I-optimal design

Run	W_1	X_1	X_2
1	0	0	0
2	0	-1	0
3	0	1	1
4	-1	-1	1
5	-1	1	0
6	-1	0	-1
7	0	1	-1
8	0	0	1
9	0	0	0
10	-1	1	1
11	-1	0	0
12	-1	-1	-1
13	0	0	0
14	0	-1	-1
15	0	1	-1
16	1	1	0
17	1	-1	1
18	1	0	-1

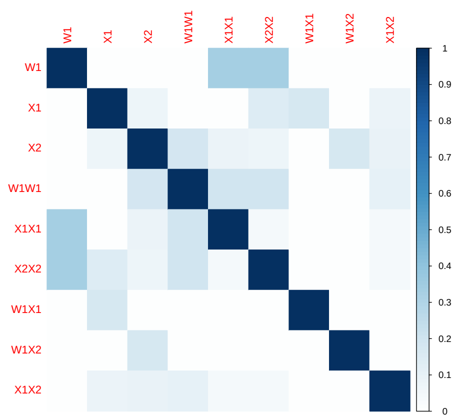
Figure 4.8: Absolute correlation maps for D-, I-, GBD- and GS_P-optimal 2-level split-plot designs in Table 4.30 and Table 4.29



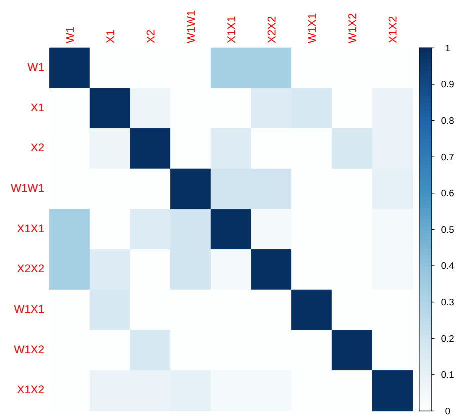
(a) D



(b) I



(c) GBD



(d) $GS_P, \pi = 0.7319$

4.3 Staggered-Level Designs

For the staggered-level designs, we consider both two-level and three-level designs similar to split-plot designs. We need to consider the values of ρ_1, ρ_2 and ρ_3 for staggered-level designs at different stages of the stratum-by-stratum algorithm. We construct the designs considering the choice of $\boldsymbol{\rho}$ values in Table 3.2. Note that in the second stage of stratum-by-stratum point-exchange algorithm, only the first two strata are being constructed, therefore we honour the value ρ_1 in Table 3.2, and use $1 - \rho_1$ as the value for ρ_2 when we are constructing the second stratum.

Example 4.3.1. We construct 20-run 2-level staggered-level designs with 3 factors and use the model consisting main effects and two-factor interaction effects. π values that we use are in Table 4.20, which are the same as the ones for split-plot designs. Table B.21 in Appendix A lists out the GS_P-optimal designs for $\pi = 0.2628, 0.4567$, $0.6176, 0.7581$, and 0.8844 when $\rho_1 = 0.6, \rho_2 = 0.3$ and $\rho_3 = 0.1$.

Table 4.31 showsthe D- and A-efficiencies of the GS_P-optimal designs. It shows that the GS_P-optimal designs are 100% D-efficient and A-efficient, which indicates good parameter estimation efficiencies of the GS_P-optimal staggered-level designs for all choices of π.

Next, we would like to compare the D- and A-efficiencies for GBD- and GS_P-optimal staggered-level designs at different values of $\boldsymbol{\rho} = \{\rho_1, \rho_2, \rho_3\}$. The

GBD-optimal designs and GS_P-optimal designs with $\pi = 0.8844$ corresponding to the different cases of $\boldsymbol{\rho}$ values in Table 3.2 are listed out in Table B.22 - Table B.27 in Appendix A. Table 4.32 lists out the D- and A-efficiencies of the GBD and GS_P-optimal designs, which indicates all of the GBD- and GS_P-optimal designs are capable of good parameter estimation in terms of having both D- and A-efficiencies close to 1, and we note that all of the six GS_P-optimal designs are also 100% D-efficient.

Table 4.31: D- and A-efficiencies of the the GS_P-optimal staggered-level designs in Table B.21

π	D_{eff}	A_{eff}
0.2628	1.0000	1.0000
0.4567	1.0000	1.0000
0.6176	1.0000	1.0000
0.7581	1.0000	1.0000
0.8844	1.0000	1.0000

Table 4.32: D- and A-efficiencies for GBD- and GS_P-optimal staggered-level designs in Table B.22 - Table B.27

Case	ρ_1	ρ_2	ρ_3	GBD, D_{eff}	GS_P, D_{eff}	GBD, A_{eff}	GS_P, A_{eff}
$\rho_1 > \rho_2 > \rho_3$	0.6	0.3	0.1	1.0000	1.0000	1.0000	1.0000
$\rho_1 > \rho_2 = \rho_3$	0.6	0.2	0.2	1.0000	1.0000	0.9919	0.9919
$\rho_1 > \rho_3 > \rho_1$	0.6	0.1	0.3	1.0000	1.0000	0.9990	0.9990
$\rho_2 > \rho_1 > \rho_3$	0.3	0.6	0.1	1.0000	1.0000	1.0000	1.0000
$\rho_2 > \rho_1 = \rho_3$	0.2	0.6	0.2	0.9851	1.0000	1.0000	0.9661
$\rho_2 > \rho_3 > \rho_1$	0.1	0.6	0.3	0.9552	1.0000	0.9419	0.9714

We present the absolute correlation maps in Figure 4.9 for the D- and A-optimal designs listed in Table 4.33, and GBD-optimal design and GS_P-optimal design with $\pi = 0.8844$ listed in Table 4.34. We see the same correlations performance in all four types of optimal designs. For main effects, there are only weak correlations between the two hard-to-change factors, W_1 and W_2. The two-factor interaction W_1W_2 is weakly correlated with W_1 and W_2, and W_1X_1 is weakly correlated with X_1 and W_2X_1, and W_2X_1 has weak correlation with X_1. There is no correlation between other effects.

Table 4.33: 20-run 2-level D- and A-optimal staggered-level designs with 3 factors when $\rho_1 = 0.6$, $\rho_2 = 0.3$, and $\rho_3 = 0.1$

(a) D-optimal design

Run	W_1	W_2	X_1
1	-1	1	1
2	-1	1	-1
3	-1	-1	-1
4	-1	-1	1
5	1	-1	-1
6	1	-1	1
7	1	1	-1
8	1	1	1
9	-1	1	-1
10	-1	1	1
11	-1	-1	1
12	-1	-1	-1
13	1	-1	1
14	1	-1	-1
15	1	1	-1
16	1	1	1
17	-1	1	-1
18	-1	1	1
19	-1	1	1
20	-1	1	-1

(b) A-optimal design

Run	W_1	W_2	X_1
1	1	1	1
2	1	1	-1
3	1	1	1
4	1	1	-1
5	-1	1	1
6	-1	1	-1
7	-1	-1	-1
8	-1	-1	1
9	1	-1	1
10	1	-1	-1
11	1	1	-1
12	1	1	1
13	-1	1	1
14	-1	1	-1
15	-1	-1	-1
16	-1	-1	1
17	1	-1	1
18	1	-1	-1
19	1	1	1
20	1	1	-1

Table 4.34: 20-run 2-level GBD- and GS_P-optimal staggered-level designs with 3 factors when $\rho_1 = 0.6$, $\rho_2 = 0.3$, and $\rho_3 = 0.1$

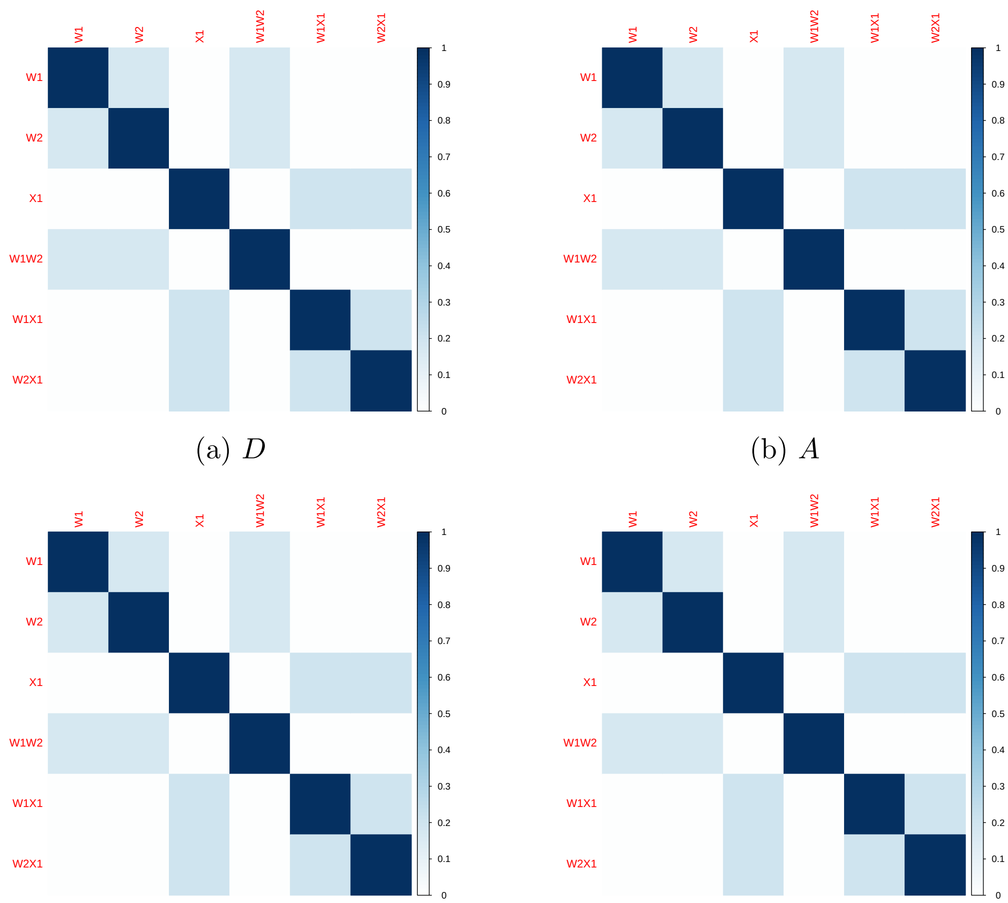
(a) GBD-optimal design

Run	W_1	W_2	X_1
1	1	-1	-1
2	1	-1	1
3	1	1	-1
4	1	1	1
5	-1	1	-1
6	-1	1	1
7	-1	-1	-1
8	-1	-1	1
9	1	-1	1
10	1	-1	-1
11	1	1	-1
12	1	1	1
13	-1	1	1
14	-1	1	-1
15	-1	-1	-1
16	-1	-1	1
17	1	-1	1
18	1	-1	-1
19	1	-1	1
20	1	-1	-1

(b) GS_P-optimal design with π=0.8844

Run	W_1	W_2	X_1
1	1	-1	-1
2	1	-1	1
3	1	1	-1
4	1	1	1
5	-1	1	1
6	-1	1	-1
7	-1	-1	-1
8	-1	-1	1
9	1	-1	1
10	1	-1	-1
11	1	1	-1
12	1	1	1
13	-1	1	-1
14	-1	1	1
15	-1	-1	-1
16	-1	-1	1
17	1	-1	-1
18	1	-1	1
19	1	-1	-1
20	1	-1	1

Figure 4.9: Absolute correlation maps for D-, A-, GBD- and GS_P-optimal staggered-level designs in Table 4.34 and Table 4.33



(a) D

(b) A

(c) GBD

(d) $GS_P, \pi = 0.7319$

Next, we consider three-level GBD- and GS_P-optimal staggered-level designs in Example 4.3.2 and examine their performance for different values of π for GS_P-optimal designs and different cases of $\boldsymbol{\rho}$ for both types of designs.

Example 4.3.2. We construct 20-run three-level GBD- and GS_P-optimal staggered-level design with three factors. We consider the second-order model consisting the three main effects, three quadratic effects, three interaction effects in this example.

The π values for the GS_P-optimal designs that we choose are in Table 4.26. We first fix $\boldsymbol{\rho}$ at $\rho_1 = 0.6, \rho_2 = 0.3$, and $\rho_3 = 0.1$ for the GS_P-optimal designs to study the performance of the designs under π = 0.1542, 0.2904, 0.4137, 0.5271, 0.6327, 0.7319, 0.8257, and 0.9148. Table B.28 in Appendix A lists out the corresponding GS_P-optimal designs. We compare the D- and A-efficiencies in Table 4.35 for the GS_P-optimal designs. We see that all of the designs maintain high performance in high D-efficiency values, while the D-efficiencies for the GS_P-optimal designs with $\pi > 0.4137$ are greater than that of the GS_P-optimal designs with $\pi = 0.1542$ and 0.2904. Also, the I-efficiencies are high when $\pi = 0.1542$ and 0.2904, but drop for the GS_P-optimal designs when $\pi > 0.2904$.

Table 4.35: D- and I-efficiencies of the GS_P-optimal designs in Table B.28

Design	D_{eff}	I_{eff}
$\pi = 0.1542$	0.9248	0.9013
$\pi = 0.2904$	0.9248	0.9013
$\pi = 0.4137$	0.9573	0.7439
$\pi = 0.5271$	0.9997	0.7980
$\pi = 0.6327$	0.9573	0.7439
$\pi = 0.7319$	0.9997	0.7980
$\pi = 0.8257$	0.9997	0.7980
$\pi = 0.9148$	0.9647	0.7303

Next, we would like to compare GBD-optimal designs and GS_P-optimal designs with $\pi = 0.8257$ and evaluate them at different choices of $\boldsymbol{\rho}$ values listed in Table 3.2. Table B.29 - Table B.34 in Appendix A list out the GBD- and GS_P- optimal designs.

Table 4.36 shows the D- and I-efficiencies of the GBD- and GS_P-optimal designs. It indicates that all the GBD- and GS_P-optimal designs maintain high D-efficiencies. The I-efficiencies are relatively high for both GBD- and GS_P-optimal designs except for when $\rho_1 > \rho_2 > \rho_3$, and $\rho_2 > \rho_3 > \rho_1$, where the I-efficiency values drop under 0.80 for both types of designs.

Table 4.36: D- and I-efficiencies for GBD- and GS_P-optimal staggered-level designs in Table B.29 - Table B.34

Case	ρ_1	ρ_2	ρ_3	GBD, D_{eff}	GS_P, D_{eff}	GBD, I_{eff}	GS_P, I_{eff}
$\rho_1 > \rho_2 > \rho_3$	0.6	0.3	0.1	0.9573	0.9997	0.7439	0.7980
$\rho_1 > \rho_2 = \rho_3$	0.6	0.2	0.2	0.9750	1.0000	0.9348	0.9968
$\rho_1 > \rho_3 > \rho_1$	0.6	0.1	0.3	1.0000	0.9997	0.9874	0.9503
$\rho_2 > \rho_1 > \rho_3$	0.3	0.6	0.1	0.9585	1.0000	0.9981	0.9940
$\rho_2 > \rho_1 = \rho_3$	0.2	0.6	0.2	0.9695	0.9722	0.8289	0.8199
$\rho_2 > \rho_3 > \rho_1$	0.1	0.6	0.3	0.9959	0.9959	0.7972	0.7972

We compare the absolute correlation heatmaps of the D- and I-optimal designs listed in Table 4.37, and GBD-optimal design and GS_P-optimal design with $\pi = 0.8257$ listed in Table 4.38 when $\rho_1 = 0.6$, $\rho_2 = 0.2$, and $\rho_3 = 0.2$. Figure 4.10 shows the absolute correlation heatmaps of the designs. The D-, GBD- and GS_P-optimal designs present similar patterns in terms of pairwise correlation between the main effects, where stronger correlations are observed between quadratic effects, which are stronger than those of the I-optimal designs.

Table 4.37: 20-run 3-level D- and I-optimal staggered-level designs with 3 factors when $\rho_1 = 0.6$, $\rho_2 = 0.2$, and $\rho_3 = 0.2$

(a) D-optimal design

Run	W_1	W_2	X_1
1	1	0	0
2	1	0	1
3	1	-1	1
4	1	-1	-1
5	-1	-1	1
6	-1	-1	-1
7	-1	1	1
8	-1	1	-1
9	1	1	-1
10	1	1	1
11	1	-1	1
12	1	-1	-1
13	-1	-1	1
14	-1	-1	-1
15	-1	1	-1
16	-1	1	1
17	0	1	0
18	0	1	1
19	0	0	-1
20	0	0	0

(b) I-optimal design

Run	W_1	W_2	X_1
1	1	1	0
2	1	1	1
3	1	-1	0
4	1	-1	1
5	-1	-1	-1
6	-1	-1	1
7	-1	0	1
8	-1	0	0
9	1	0	0
10	1	0	-1
11	1	-1	1
12	1	-1	-1
13	0	-1	-1
14	0	-1	1
15	0	0	-1
16	0	0	1
17	-1	0	0
18	-1	0	1
19	-1	1	-1
20	-1	1	1

Table 4.38: 20-run 3-level GBD- and GS_P-optimal staggered-level designs with 3 factors when $\rho_1 = 0.6$, $\rho_2 = 0.2$, and $\rho_3 = 0.2$

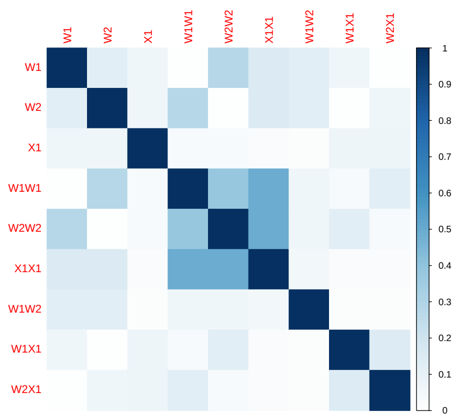
(a) GBD-optimal design

Run	W_1	W_2	X_1
1	0	-1	0
2	0	-1	-1
3	0	0	0
4	0	0	1
5	1	0	0
6	1	0	-1
7	1	1	-1
8	1	1	1
9	-1	1	-1
10	-1	1	1
11	-1	-1	1
12	-1	-1	-1
13	1	-1	1
14	1	-1	-1
15	1	1	-1
16	1	1	1
17	-1	1	1
18	-1	1	-1
19	-1	-1	-1
20	-1	-1	1

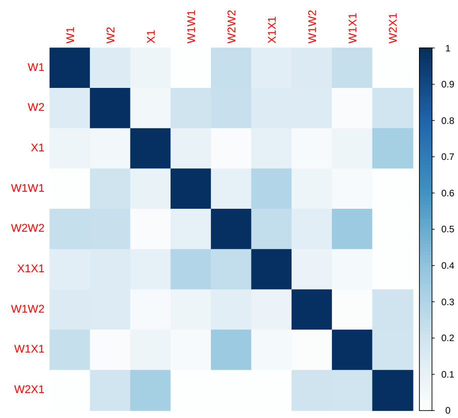
(b) GS_P-optimal design with $\pi = 0.8257$

Run	W_1	W_2	X_1
1	-1	0	-1
2	-1	0	0
3	-1	1	1
4	-1	1	-1
5	1	1	1
6	1	1	-1
7	1	-1	-1
8	1	-1	1
9	-1	-1	1
10	-1	-1	-1
11	-1	1	1
12	-1	1	-1
13	1	1	-1
14	1	1	1
15	1	-1	-1
16	1	-1	1
17	0	-1	-1
18	0	-1	0
19	0	0	0
20	0	0	1

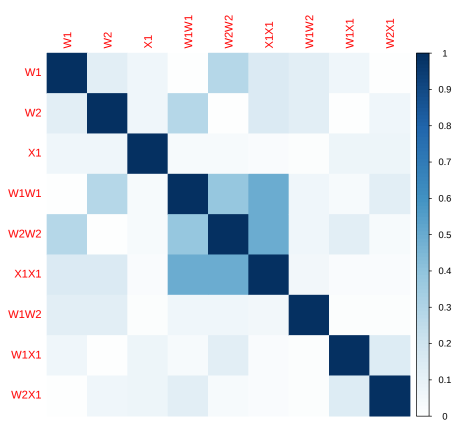
Figure 4.10: Absolute correlation maps for D-, I-, GBD- and GS_P-optimal 3-level staggered-level designs in Table 4.37 and Table 4.38



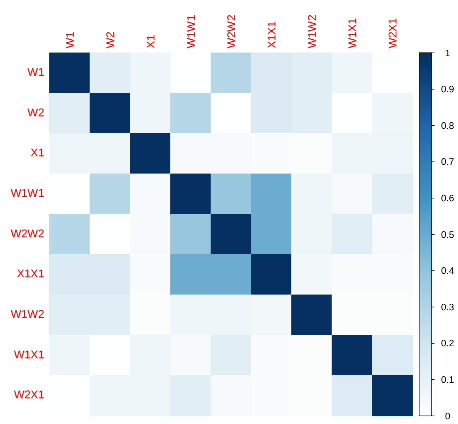
(a) D



(b) I



(c) GBD



(d) $GS_P, \pi = 0.8257$

Chapter 5

Conclusion and Future Work

Multistratum designs could be an feasible alternative when the homogeneity requirement cannot be satisfied for CRDs. To select optimal designs, we need to take into the consideration of the goal of the experimental design, whether it is aiming for better parameter estimation, or for better prediction. The traditional optimality criteria such as D-, A- and I-optimality criteria requires an assumed model. In this thesis, we consider the Bayesian approach to select optimal designs with high parameter estimation efficiencies. We first introduce the types of designs, their associated model structure, and various criteria from both traditional and Bayesian approach that are used in the thesis. Introduction of the algorithm being modified and utilized are also introduced. Next, we develop GBD- and GS_P-optimality criterion and describe the modified algorithm and our choices of values for the parameters. Finally, we provide examples for CRDs, split-plot designs and staggered-level designs with increasing numbers of strata in the design, and we provide some analysis of the examples.

Bayesian approach can be useful in generating follow-up designs to increase the precision on the analysis of results. Lin and Yang (2022) utilized the Bayesian approach for generating follow-up experiments for supersaturated multistratum designs. In this thesis, we mainly focus on generating initial designs, it is possible to extend the current work to consider follow-up designs in the future.

The current design examples are considered with three strata at maximum. Another type of three-stratum design that could be considered is split-split-plot design for experiment with three or more factors (Jones and Goos, 2009). Therefore, it could be one of the future work to apply the Bayesian optimality criteria to generate optimal split-split-plot designs. Designs with more than three strata could also be considered, however, as the number of strata increases, the model space expands with more factors included, and the structure of the variance-covariance matrix will also be complicated. To include more strata and more factors into consideration, a modification on the current computer algorithm should be implemented.

Furthermore, since GBD- and GS_P-optimality criteria are developed with possibilities to select optimal designs based on a distribution for $\boldsymbol{\rho}$, more examples can be considered where $\boldsymbol{\rho}$ are no longer selected with different values but a distribution, one possible distribution for $\boldsymbol{\rho}$ that can be considered for future work is the Dirichlet distribution, where $P(\boldsymbol{\rho}) = \frac{\Gamma(\sum_{l=1}^{g} \alpha_l)}{\prod_{l=1}^{g} \Gamma(\alpha_l)} \prod_{l=1}^{g} \rho_l^{\alpha_l - 1}$ with $(\alpha_1,\ \alpha_2, \cdots,\ \alpha_g)$ predetermined prior to the experiment.

Bibliography

Arnouts, H. and P. Goos (2012). Staggered-level designs for experiments with more than one hard-to-change factor. *Technometrics 54*(4), 355–366. (Cited on pages 2 and 12.)

Bernardo, J. M. (1979). Expected information as expected utility. *The Annals of Statistics 7*(3), 686–690. (Cited on pages 5, 6, 20, 21 and 30.)

Bingham, D. R. and H. A. Chipman (2007). Incorporating prior information in optimal design for model selection. *Technometrics 49*(2), 155–163. (Cited on pages 35 and 36.)

Borth, D. M. (1975). A total entropy criterion for the dual problem of model discrimination and parameter estimation. *Journal of the Royal Statistical Society: Series B (Methodological) 37*(1), 77–87. (Cited on pages 5, 6, 23 and 31.)

Box, G. and R. D. Meyer (1993). Finding the active factors in fractionated screening experiments. *Journal of Quality Technology 25*, 94–105. (Cited on page 31.)

Box, G. E. P. and W. J. Hill (1967). Discrimination among mechanistic models. *Technometrics 9*(1), 57–71. (Cited on pages 5 and 23.)

Chernoff, H. (1953). Locally optimal designs for estimating parameters. *The Annals of Mathematical Statistics 24*(4), 586–602. (Cited on page 4.)

Cook, R. D. and C. J. Nachtsheim (1980). A comparison of algorithms for constructing exact D-optimal designs. *Technometrics 22*(3), 315–324. (Cited on page 6.)

DuMouchel, W. and B. Jones (1994). A simple Bayesian modification of *D*-optimal designs to reduce dependence on an assumed model. *Technometrics 36*(1), 37–47. (Cited on pages 5 and 21.)

Fedorov, V. (1972). *Theory of Optimal Experiments Designs.* (Cited on pages 4 and 27.)

Hardin, R. and N. Sloane (2009). Computer-generated minimal (and larger) response-surface designs: (II) the cube. (Cited on page 18.)

Jones, B., K. Allen-Moyer, and P. Goos (2021). *A*-optimal versus *D*-optimal design of screening experiments. *Journal of Quality Technology 53*(4), 369–382. (Cited on page 25.)

Jones, B. and P. Goos (2009). D-optimal design of split-split-plot experiments. *Biometrika 96*(1), 67–82. (Cited on page 82.)

Jones, B. and P. Goos (2012). I-optimal versus D-optimal split-plot response surface designs. *Journal of Quality Technology 44*(2), 85–101. (Cited on pages 4, 24 and 27.)

Jones, B. and C. J. Nachtsheim (2009). Split-plot designs: What, why, and how. *Journal of Quality Technology 41*(4), 340–361. (Cited on page 1.)

Kiefer, J. (1959). Optimum experimental designs. *Journal of the Royal Statistical Society. Series B (Methodological) 21*(2), 272–319. (Cited on page 3.)

Kiefer, J. and J. Wolfowitz (1959). Optimum designs in regression problems. *The Annals of Mathematical Statistics 30*(2), 271–294. (Cited on page 17.)

Lin, C.-Y. (2018). Generalized bayesian D criterion for single-stratum and multistratum designs. *Quality and Reliability Engineering International 34*(8), 1566–1574. (Cited on pages 5, 26 and 35.)

Lin, C.-Y. and P. Yang (2022). Bayesian analysis and follow-up experiments for supersaturated multistratum designs. *Journal of Quality Technology 54*(5), 527–546. (Cited on pages 28 and 82.)

Lindley, D. V. (1956). On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics 27*(4), 986–1005. (Cited on page 20.)

Meyer, R. K. and C. J. Nachtsheim (1995). The coordinate-exchange algorithm for constructing exact optimal experimental designs. *Technometrics 37*(1), 60–69. (Cited on pages 6 and 27.)

Ng, S. and S. Chick (2004). Design of follow-up experiments for improving model discrimination and parameter estimation. *Naval Research Logistics 51*, 1129–1148. (Cited on pages 5, 21 and 23.)

Trinca, L. A. and S. G. Gilmour (2001). Multistratum response surface designs. *Technometrics 43*(1), 25–33. (Cited on pages 24 and 27.)

Trinca, L. A. and S. G. Gilmour (2015). Improved split-plot and multistratum designs. *Technometrics 57*(2), 145–154. (Cited on pages 4, 6, 18, 25, 27 and 32.)

Wald, A. (1943). On the Efficient Design of Statistical Investigations. *The Annals of Mathematical Statistics 14*(2), 134–140. (Cited on page 4.)

Wong, C. S. and J. C. Masaro (1984). *A*-optimal design matrices. *Discrete Mathematics 50*, 295–318. (Cited on page 17.)

Appendix A

Proofs

A.1 Proof of Theorem 3.1.1

Proof: We consider the first double integral in (2.8) first. Note that

$$P(\boldsymbol{Y})P(\boldsymbol{\theta}|\boldsymbol{Y})\log[P(\boldsymbol{\theta}|\boldsymbol{Y})] = P(\boldsymbol{Y})P(\boldsymbol{\beta}, \sigma_w^2, \boldsymbol{\rho}|\boldsymbol{Y})\log[P(\boldsymbol{\beta}, \sigma_w^2, \boldsymbol{\rho}|\boldsymbol{Y})],$$

where

$$P(\boldsymbol{\beta}, \sigma_w^2, \boldsymbol{\rho}|\boldsymbol{Y}) \propto P(\boldsymbol{\beta}|\boldsymbol{Y}, \sigma_w^2, \boldsymbol{\rho})P(\sigma_w^2|\boldsymbol{\rho})P(\boldsymbol{\rho}),$$

The first integral in (2.8) can be written as

$$\int\int\int\int P(\boldsymbol{Y})P(\boldsymbol{\beta}|\boldsymbol{Y}, \sigma_w^2, \boldsymbol{\rho})P(\sigma_w^2|\boldsymbol{\rho})P(\boldsymbol{\rho})\log P(\boldsymbol{\beta}|\boldsymbol{Y}, \sigma_w^2, \boldsymbol{\rho})d\boldsymbol{\beta}d\sigma_w^2 d\boldsymbol{\rho}d\boldsymbol{Y}$$

$$+\int\int\int\int P(\boldsymbol{Y})P(\boldsymbol{\beta}|\boldsymbol{Y}, \sigma_w^2, \boldsymbol{\rho})P(\sigma_w^2|\boldsymbol{\rho})P(\boldsymbol{\rho})\log[P(\sigma_w^2|\boldsymbol{\rho})P(\boldsymbol{\rho})]d\boldsymbol{\beta}d\sigma_w^2 d\boldsymbol{\rho}d\boldsymbol{Y}.$$

Similarly,

$$\begin{aligned} P(\boldsymbol{\theta})\log[P(\boldsymbol{\theta})] &= P(\boldsymbol{\beta}, \sigma_w^2, \boldsymbol{\rho})\log[P(\boldsymbol{\beta}, \sigma_w^2, \boldsymbol{\rho})] \\ &\propto P(\boldsymbol{\beta}|\sigma_w^2, \boldsymbol{\rho})P(\sigma_w^2|\boldsymbol{\rho})P(\boldsymbol{\rho})\log[P(\boldsymbol{\beta}|\sigma_w^2, \boldsymbol{\rho})P(\sigma_w^2|\boldsymbol{\rho})P(\boldsymbol{\rho})]. \end{aligned}$$

Therefore, the second integral of (2.8) can be written as

$$\begin{aligned} &\int\int\int P(\boldsymbol{\beta}|\sigma_w^2, \boldsymbol{\rho})P(\sigma_w^2|\boldsymbol{\rho})P(\boldsymbol{\rho})\log P(\boldsymbol{\beta}|\sigma_w^2, \boldsymbol{\rho})d\boldsymbol{\beta}d\sigma_w^2 d\boldsymbol{\rho} \\ &+\int\int\int P(\boldsymbol{\beta}|\sigma_w^2, \boldsymbol{\rho})P(\sigma_w^2|\boldsymbol{\rho})P(\boldsymbol{\rho})\log[P(\sigma_w^2|\boldsymbol{\rho})P(\boldsymbol{\rho})]d\boldsymbol{\beta}d\sigma_w^2 d\boldsymbol{\rho}. \end{aligned}$$

Then (2.8) can be expressed as

$$\begin{aligned} &\int\int\Bigg[\int\int P(\boldsymbol{Y})P(\boldsymbol{\beta}|\boldsymbol{Y}, \sigma_w^2, \boldsymbol{\rho})P(\sigma_w^2|\boldsymbol{\rho})P(\boldsymbol{\rho})\log P(\boldsymbol{\beta}|\boldsymbol{Y}, \sigma_w^2, \boldsymbol{\rho})d\boldsymbol{\beta}d\boldsymbol{Y} \\ &-\int P(\boldsymbol{\beta}|\sigma_w^2, \boldsymbol{\rho})P(\sigma_w^2|\boldsymbol{\rho})P(\boldsymbol{\rho})\log P(\boldsymbol{\beta}|\sigma_w^2, \boldsymbol{\rho})d\boldsymbol{\beta}\Bigg]d\sigma_w^2 d\boldsymbol{\rho} \\ &+\int\int P(\sigma_w^2|\boldsymbol{\rho})P(\boldsymbol{\rho})\log[P(\sigma_w^2|\boldsymbol{\rho})P(\boldsymbol{\rho})]\Bigg[\int\int P(\boldsymbol{Y})P(\boldsymbol{\beta}|\boldsymbol{Y}, \sigma_w^2, \boldsymbol{\rho})d\boldsymbol{Y}d\boldsymbol{\beta}\Bigg]d\sigma_w^2 d\boldsymbol{\rho} \\ &-\int\int P(\sigma_w^2|\boldsymbol{\rho})P(\boldsymbol{\rho})\log[P(\sigma_w^2|\boldsymbol{\rho})P(\boldsymbol{\rho})]\Bigg[\int P(\boldsymbol{\beta}|\sigma_w^2, \boldsymbol{\rho})d\boldsymbol{\beta}\Bigg]d\sigma_w^2 d\boldsymbol{\rho}. \qquad \text{(A.1)} \end{aligned}$$

Since $P(\boldsymbol{\rho})$ and $P(\sigma_w^2|\boldsymbol{\rho})$ is irrelevant of $\boldsymbol{\beta}$ and $\boldsymbol{Y}$, and $P(\boldsymbol{Y})$ is irrelevant of $\boldsymbol{\beta}$, (A.1) can be simplified as

$$\int P(\boldsymbol{\rho})\int P(\sigma_w^2|\boldsymbol{\rho})\Bigg[\int P(\boldsymbol{Y})\int P(\boldsymbol{\beta}|\boldsymbol{Y}, \sigma_w^2, \boldsymbol{\rho})\log P(\boldsymbol{\beta}|\boldsymbol{Y}, \sigma_w^2, \boldsymbol{\rho})d\boldsymbol{\beta}d\boldsymbol{Y}$$

$$- \int P(\boldsymbol{\beta}|\sigma_w^2, \boldsymbol{\rho}) \log P(\boldsymbol{\beta}|\sigma_w^2, \boldsymbol{\rho}) d\boldsymbol{\beta} \Big] d\sigma_w^2 d\boldsymbol{\rho}. \qquad \text{(A.2)}$$

Now, we will find $P(\boldsymbol{\beta}|\boldsymbol{Y}, \sigma_w^2, \boldsymbol{\rho})$. Note that

$$\begin{aligned}
& P(\boldsymbol{\beta}|\boldsymbol{Y}, \sigma_w^2, \boldsymbol{\rho}) \propto P(\boldsymbol{Y}|\boldsymbol{\beta}, \sigma_w^2, \boldsymbol{\rho}) P(\boldsymbol{\beta}|\sigma_w^2, \boldsymbol{\rho}) \\
\propto \quad & \sigma_w^{-n} |\boldsymbol{V}|^{-1/2} \times \exp\{-\frac{1}{2}\sigma_w^{-2}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta})'\boldsymbol{V}^{-1}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta})\} \sigma_w^{-k} \tau^{-k} \exp\{-\frac{1}{2}\sigma_w^{-2}(\boldsymbol{\beta}'\boldsymbol{K}\boldsymbol{\beta})\} \\
\propto \quad & \sigma_w^{-n-k} \tau^{-k} |\boldsymbol{V}|^{-1/2} \exp\Big[-\frac{1}{2}\sigma_w^{-2}\{(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta})'\boldsymbol{V}^{-1}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}) + \boldsymbol{\beta}'\boldsymbol{K}\boldsymbol{\beta}\} \Big].
\end{aligned}$$

It can be shown that,

$$\begin{aligned}
& (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta})'\boldsymbol{V}^{-1}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}) + \boldsymbol{\beta}'\boldsymbol{K}\boldsymbol{\beta} \\
= \quad & (\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})'\boldsymbol{V}^{-1}(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) + (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})'\boldsymbol{X}\boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) \\
& + (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})'\boldsymbol{K}(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}) + \hat{\boldsymbol{\beta}}'\boldsymbol{K}\hat{\boldsymbol{\beta}} \\
= \quad & S + (\boldsymbol{\beta} - \hat{\boldsymbol{\beta}})'(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X} + \boldsymbol{K})(\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}),
\end{aligned}$$

where

$$\begin{aligned}
S \quad = \quad & (\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})'\boldsymbol{V}^{-1}(\boldsymbol{Y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) + \hat{\boldsymbol{\beta}}'\boldsymbol{K}\hat{\boldsymbol{\beta}} \\
= \quad & \boldsymbol{Y}\boldsymbol{V}^{-1}\boldsymbol{Y} - \boldsymbol{Y}\boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X} + \boldsymbol{K})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{Y},
\end{aligned}$$

and

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}\boldsymbol{V}^{-1}\boldsymbol{X} + \boldsymbol{K})^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{Y}.$$

Therefore,

$$P(\boldsymbol{\beta}|\boldsymbol{Y}, \sigma_w^2, \boldsymbol{\rho}) \propto P(\boldsymbol{Y}|\boldsymbol{\beta}, \sigma_w^2, \boldsymbol{\rho}) P(\boldsymbol{\beta}|\sigma_w^2, \boldsymbol{\rho})$$

$$\propto \quad \sigma_w^{-n-k}\tau^{-k}|\boldsymbol{V}|^{-1/2}\exp\left[-\frac{1}{2}\sigma_w^{-2}\{S+(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})'(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}+\boldsymbol{K})(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})\}\right]$$

$$\propto \quad \sigma_w^{-n-k}\tau^{-k}|\boldsymbol{V}|^{-1/2}\exp(-\frac{1}{2}\sigma_w^{-2}S)\exp\left[-\frac{1}{2}\sigma_w^{-2}\{(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})'(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}+\boldsymbol{K})(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})\}\right],$$

which indicates that $\boldsymbol{\beta}|\boldsymbol{Y},\sigma_w^2,\boldsymbol{\rho} \backsim N(\hat{\boldsymbol{\beta}},\sigma_w^2(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}+\boldsymbol{K})^{-1})$. Therefore,

$$\int P(\boldsymbol{\beta}|\boldsymbol{Y},\sigma_w^2,\boldsymbol{\rho})\log P(\boldsymbol{\beta}|\boldsymbol{Y},\sigma_w^2,\boldsymbol{\rho})d\boldsymbol{\beta}$$

$$= \int P(\boldsymbol{\beta}|\boldsymbol{Y},\sigma_w^2,\boldsymbol{\rho})\log\{(2\pi)^{-\frac{(k+1)}{2}}|\sigma_w^2(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}+\boldsymbol{K})^{-1})|^{-\frac{1}{2}}$$

$$\times\exp[-\frac{1}{2}\sigma_w^{-2}\{(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})'(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}+\boldsymbol{K})(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})]\}d\boldsymbol{\beta}$$

$$= \int P(\boldsymbol{\beta}|\boldsymbol{Y},\sigma_w^2,\boldsymbol{\rho})\{\frac{-(k+1)\log(2\pi)}{2}+\frac{\log|\sigma_w^{-2}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}+\boldsymbol{K})|}{2}$$

$$+[-\frac{1}{2}\sigma_w^{-2}\{(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})'(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}+\boldsymbol{K})(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})]\}d\boldsymbol{\beta}$$

$$= \frac{-(k+1)\log(2\pi)}{2}+\frac{\log|\sigma_w^{-2}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}+\boldsymbol{K})|}{2}$$

$$-\frac{1}{2}\sigma_w^{-2}\int P(\boldsymbol{\beta}|\boldsymbol{Y},\sigma_w^2,\boldsymbol{\rho})\{(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})'(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}+\boldsymbol{K})(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})\}d\boldsymbol{\beta}$$

$$= \frac{-(k+1)\log(2\pi)}{2}+\frac{\log|\sigma_w^{-2}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}+\boldsymbol{K})|}{2}$$

$$-\frac{1}{2}\sigma_w^{-2}E[(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})'(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}+\boldsymbol{K})(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}})]$$

$$= \frac{-(k+1)\log(2\pi)}{2}+\frac{\log|\sigma_w^{-2}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}+\boldsymbol{K})|}{2}$$

$$-\frac{1}{2}\sigma_w^{-2}tr[(\boldsymbol{X'V}^{-1}\boldsymbol{X}+\boldsymbol{K})\sigma_w^2(\boldsymbol{X'V}^{-1}\boldsymbol{X}+\boldsymbol{K})^{-1})]$$

$$= \frac{-(k+1)\log(2\pi)}{2}+\frac{\log|\sigma_w^{-2}(\boldsymbol{X'V}^{-1}\boldsymbol{X}+\boldsymbol{K})|}{2}-\frac{k+1}{2}$$

$$= \frac{-(k+1)\log(2\pi)}{2}-(k+1)\log\sigma_w+\frac{\log|\boldsymbol{X'V}^{-1}\boldsymbol{X}+\boldsymbol{K}|}{2}-\frac{k+1}{2}.$$

Since $\boldsymbol{\beta}|\sigma_w^2,\boldsymbol{\rho} \backsim N(\boldsymbol{0},\sigma_w^2\boldsymbol{K}^{-1})$, we follow similar method and obtain that

$$\int P(\boldsymbol{\beta}|\sigma_w^2,\boldsymbol{\rho})\log P(\boldsymbol{\beta}|\sigma_w^2,\boldsymbol{\rho})d\boldsymbol{\beta}$$

$$= \int P(\boldsymbol{\beta}|\sigma_w^2,\boldsymbol{\rho})\log\{(2\pi)^{-\frac{(k+1)}{2}}|\sigma_w^2\boldsymbol{K}^{-1}|^{-\frac{1}{2}}\exp[-\frac{1}{2}\sigma_w^{-2}(\boldsymbol{\beta'K\beta})]\}d\boldsymbol{\beta}$$

$$= \frac{-(k+1)\log(2\pi)}{2}-(k+1)\log\sigma_w+\frac{\log|\boldsymbol{K}|}{2}-\int P(\boldsymbol{\beta}|\sigma_w^2,\boldsymbol{\rho})\frac{1}{2}\sigma_w^{-2}(\boldsymbol{\beta'K\beta})\}d\boldsymbol{\beta}$$

$$= \frac{-(k+1)\log(2\pi)}{2}-(k+1)\log\sigma_w+\frac{\log|\boldsymbol{K}|}{2}-\frac{1}{2}\sigma_w^{-2}E[\boldsymbol{\beta'K\beta}]$$

$$= \frac{-(k+1)\log(2\pi)}{2}-(k+1)\log\sigma_w+\frac{\log|\boldsymbol{K}|}{2}-\frac{1}{2}\sigma_w^{-2}tr[\sigma_w^2\boldsymbol{K}^{-1}\boldsymbol{K}]$$

$$= \frac{-(k+1)\log(2\pi)}{2}-(k+1)\log\sigma_w+\frac{\log|\boldsymbol{K}|}{2}-\frac{k+1}{2}.$$

Therefore, (A.2) can be simplified to

$$\int P(\boldsymbol{\rho})\int P(\sigma_w^2|\boldsymbol{\rho})\Big[\int P(\boldsymbol{Y})\int P(\boldsymbol{\beta}|\boldsymbol{Y},\sigma_w^2,\boldsymbol{\rho})\log P(\boldsymbol{\beta}|\boldsymbol{Y},\sigma_w^2,\boldsymbol{\rho})d\boldsymbol{\beta}d\boldsymbol{Y}$$

$$-\int P(\boldsymbol{\beta}|\sigma_w^2,\boldsymbol{\rho})\log P(\boldsymbol{\beta}|\sigma_w^2,\boldsymbol{\rho})d\boldsymbol{\beta}\Big]d\sigma_w^2d\boldsymbol{\rho}$$

$$
\begin{aligned}
&= \int P(\boldsymbol{\rho}) \int P(\sigma_w^2|\boldsymbol{\rho}) \Big[\int P(\boldsymbol{Y}) \frac{\log|\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}+\boldsymbol{K}|}{2} d\boldsymbol{Y} - \frac{\log|\boldsymbol{K}|}{2} \Big] d\sigma_w^2 d\boldsymbol{\rho} \\
&= \int P(\boldsymbol{\rho}) \int P(\sigma_w^2|\boldsymbol{\rho}) \Big[\frac{\log|\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}+\boldsymbol{K}|}{2} - \frac{\log|\boldsymbol{K}|}{2} \Big] d\sigma_w^2 d\boldsymbol{\rho} \\
&= \int_{\boldsymbol{\rho}} P(\boldsymbol{\rho}) \Big[\frac{\log|\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}+\boldsymbol{K}|}{2} - \frac{\log|\boldsymbol{K}|}{2} \Big] d\boldsymbol{\rho}. \qquad \text{(A.3)}
\end{aligned}
$$

A.2 Proof of Theorem 3.1.2

Proof: Let s be the number of models in the model space. Then, by Bayes' theorem, for any model M_i, $i = 1, \cdots, s$,

$$
P(M_i|\boldsymbol{Y}) = P(\boldsymbol{Y}|M_i)P(M_i) / \sum_{j=1}^{s} P(\boldsymbol{Y}|M_j)P(M_j).
$$

Thus, (2.10) can be simplified as,

$$
\begin{aligned}
& -\sum_{i=1}^{s} P(M_i) \int P(\boldsymbol{\theta}_i|M_i) \log P(\boldsymbol{\theta}_i|M_i) d\boldsymbol{\theta}_i \\
& + \int \sum_{i=1}^{s} P(M_i|\boldsymbol{Y}) \int P(\boldsymbol{\theta}_i|\boldsymbol{Y}, M_i) \log P(\boldsymbol{\theta}_i|\boldsymbol{Y}, M_i) d\boldsymbol{\theta}_i \sum_{j=1}^{s} P(M_j)P(\boldsymbol{Y}|M_j) d\boldsymbol{Y} \\
= & -\sum_{i=1}^{s} P(M_i) \int P(\boldsymbol{\theta}_i|M_i) \log P(\boldsymbol{\theta}_i|M_i) d\boldsymbol{\theta}_i \\
& + \sum_{i=1}^{s} P(M_i) \int P(\boldsymbol{Y}|M_i) \int P(\boldsymbol{\theta}_i|\boldsymbol{Y}, M_i) \log P(\boldsymbol{\theta}_i|\boldsymbol{Y}, M_i) d\boldsymbol{\theta}_i d\boldsymbol{Y}
\end{aligned}
$$

$$= \sum_{i=1}^{s} P(M_i)\Big[-\int P(\boldsymbol{\theta}_i|M_i)\log\, P(\boldsymbol{\theta}_i|M_i)d\boldsymbol{\theta}_i$$

$$+ \int P(\boldsymbol{Y}|M_i) \int P(\boldsymbol{\theta}_i|\boldsymbol{Y}, M_i)\log\, P(\boldsymbol{\theta}_i|\boldsymbol{Y}, M_i)d\boldsymbol{\theta}_i d\boldsymbol{Y}\Big]. \quad \text{(A.4)}$$

By (2.8) and (A.3), we can further simplify (A.4) to

$$\sum_{i=1}^{s} P(M_i) \int_{\boldsymbol{\rho}} P(\boldsymbol{\rho}) \Big[\frac{\log|\boldsymbol{X}_i'\boldsymbol{V}^{-1}\boldsymbol{X}_i + \boldsymbol{K}_i|}{2} - \frac{\log|\boldsymbol{K}_i|}{2} \Big] d\boldsymbol{\rho}$$

$$= \frac{1}{2}\Big[\sum_{i=1}^{s} P(M_i) \int_{\boldsymbol{\rho}} P(\boldsymbol{\rho})\log|\boldsymbol{X}_i'\boldsymbol{V}^{-1}\boldsymbol{X}_i + \boldsymbol{K}_i| d\boldsymbol{\rho}\Big] + \boldsymbol{E}, \quad \text{(A.5)}$$

where $\boldsymbol{E}$ is some matrix independent of the design.

Appendix B

Design Examples

Table B.1: 18-run 2-level GS_P-optimal split-plot designs with 3 factors when $\rho_1 = 0.5$ and $\rho_2 = 0.5$

(a) $\pi = 0.2628$

Run	W_1	X_1	X_2
1	-1	1	-1
2	-1	1	1
3	-1	-1	1
4	1	1	-1
5	1	-1	1
6	1	-1	-1
7	-1	1	1
8	-1	-1	1
9	-1	-1	-1
10	-1	-1	-1
11	-1	1	-1
12	-1	1	1
13	1	1	1
14	1	1	-1
15	1	-1	-1
16	1	1	-1
17	1	1	1
18	1	-1	1

(b) $\pi = 0.4567$

Run	W_1	X_1	X_2
1	1	-1	1
2	1	1	-1
3	1	1	1
4	-1	-1	1
5	-1	1	-1
6	-1	-1	-1
7	-1	1	1
8	-1	-1	1
9	-1	1	-1
10	1	1	1
11	1	-1	-1
12	1	1	-1
13	-1	-1	-1
14	-1	-1	1
15	-1	1	1
16	1	1	-1
17	1	-1	1
18	1	-1	-1

(c) $\pi = 0.6176$

Run	W_1	X_1	X_2
1	1	-1	1
2	1	1	-1
3	1	1	1
4	-1	-1	1
5	-1	1	-1
6	-1	-1	-1
7	-1	1	1
8	-1	-1	1
9	-1	1	-1
10	1	1	1
11	1	-1	-1
12	1	1	-1
13	-1	-1	-1
14	-1	-1	1
15	-1	1	1
16	1	1	-1
17	1	-1	1
18	1	-1	-1

Table B.1: 18-run 2-level GS_P-optimal split-plot designs with 3 factors when $\rho_1 = 0.5$ and $\rho_2 = 0.5$ (cont.)

(d) $\pi = 0.7581$

Run	W_1	X_1	X_2
1	1	-1	1
2	1	1	-1
3	1	1	1
4	-1	-1	1
5	-1	1	-1
6	-1	-1	-1
7	-1	1	1
8	-1	-1	1
9	-1	1	-1
10	1	1	1
11	1	-1	-1
12	1	1	-1
13	-1	-1	-1
14	-1	-1	1
15	-1	1	1
16	1	1	-1
17	1	-1	1
18	1	-1	-1

(e) $\pi = 0.8844$

Run	W_1	X_1	X_2
1	1	-1	1
2	1	1	-1
3	1	1	1
4	-1	-1	1
5	-1	1	-1
6	-1	-1	-1
7	-1	1	1
8	-1	-1	1
9	-1	1	-1
10	1	1	1
11	1	-1	-1
12	1	1	-1
13	-1	-1	-1
14	-1	-1	1
15	-1	1	1
16	1	1	-1
17	1	-1	1
18	1	-1	-1

Table B.2: 18-run 2-level GBD- and GS_P-optimal split-plot designs with 3 factors when $\rho_1 = 0.1$ and $\rho_2 = 0.9$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	-1	1	-1
2	-1	1	1
3	-1	-1	1
4	1	1	-1
5	1	-1	1
6	1	-1	-1
7	-1	1	1
8	-1	-1	1
9	-1	-1	-1
10	-1	-1	-1
11	-1	1	-1
12	-1	1	1
13	1	1	1
14	1	1	-1
15	1	-1	-1
16	1	1	-1
17	1	1	1
18	1	-1	1

(b) GS_P-optimal design with $\pi = 0.8844$

Run	W_1	X_1	X_2
1	1	-1	1
2	1	1	-1
3	1	1	1
4	-1	-1	1
5	-1	1	-1
6	-1	-1	-1
7	-1	1	1
8	-1	-1	1
9	-1	1	-1
10	1	1	1
11	1	-1	-1
12	1	1	-1
13	-1	-1	-1
14	-1	-1	1
15	-1	1	1
16	1	1	-1
17	1	-1	1
18	1	-1	-1

Table B.3: 18-run 2-level GBD- and GS_P-optimal split-plot designs with 3 factors when $\rho_1 = 0.2$ and $\rho_2 = 0.8$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	1	-1	-1
2	1	1	-1
3	1	-1	1
4	-1	1	-1
5	-1	-1	-1
6	-1	-1	1
7	-1	1	1
8	-1	-1	1
9	-1	-1	-1
10	1	1	-1
11	1	1	1
12	1	-1	-1
13	1	-1	-1
14	1	-1	1
15	1	1	1
16	-1	1	-1
17	-1	1	1
18	-1	-1	1

(b) GS_P-optimal design with $\pi = 0.8844$

Run	W_1	X_1	X_2
1	-1	-1	-1
2	-1	1	-1
3	-1	-1	1
4	-1	-1	-1
5	-1	1	1
6	-1	1	-1
7	-1	-1	1
8	-1	1	1
9	-1	-1	-1
10	1	-1	-1
11	1	1	1
12	1	-1	1
13	1	1	-1
14	1	-1	1
15	1	1	1
16	1	1	1
17	1	-1	-1
18	1	1	-1

Table B.4: 18-run 2-level GBD- and GS_P-optimal split-plot designs with 3 factors when $\rho_1 = 0.3$ and $\rho_2 = 0.7$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	-1	-1	-1
2	-1	1	-1
3	-1	-1	1
4	-1	1	-1
5	-1	1	1
6	-1	-1	-1
7	1	1	1
8	1	-1	1
9	1	-1	-1
10	1	-1	-1
11	1	-1	1
12	1	1	-1
13	-1	1	1
14	-1	-1	-1
15	-1	-1	1
16	1	1	-1
17	1	-1	1
18	1	1	1

(b) GS_P-optimal design with $\pi = 0.8844$

Run	W_1	X_1	X_2
1	1	-1	-1
2	1	1	-1
3	1	-1	1
4	-1	-1	1
5	-1	-1	-1
6	-1	1	1
7	-1	1	-1
8	-1	1	1
9	-1	-1	1
10	1	-1	-1
11	1	1	-1
12	1	1	1
13	1	1	-1
14	1	1	1
15	1	-1	1
16	-1	-1	-1
17	-1	1	-1
18	-1	-1	1

Table B.5: 18-run 2-level GBD- and GS_P-optimal split-plot designs with 3 factors when $\rho_1 = 0.4$ and $\rho_2 = 0.6$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	-1	-1	-1
2	-1	1	1
3	-1	1	-1
4	1	-1	1
5	1	-1	-1
6	1	1	1
7	1	1	1
8	1	-1	-1
9	1	1	-1
10	-1	1	-1
11	-1	-1	-1
12	-1	-1	1
13	-1	-1	1
14	-1	1	-1
15	-1	1	1
16	1	-1	1
17	1	1	1
18	1	1	-1

(b) GS_P-optimal design with $\pi = 0.8844$

Run	W_1	X_1	X_2
1	1	-1	1
2	1	-1	-1
3	1	1	1
4	1	1	-1
5	1	1	1
6	1	-1	-1
7	-1	1	1
8	-1	1	-1
9	-1	-1	-1
10	-1	-1	1
11	-1	-1	-1
12	-1	1	-1
13	-1	-1	-1
14	-1	1	1
15	-1	-1	1
16	1	-1	1
17	1	1	-1
18	1	1	1

Table B.6: 18-run 2-level GBD- and GS_P-optimal split-plot designs with 3 factors when $\rho_1 = 0.5$ and $\rho_2 = 0.5$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	-1	-1	1
2	-1	-1	-1
3	-1	1	1
4	-1	1	-1
5	-1	-1	-1
6	-1	-1	1
7	-1	-1	-1
8	-1	1	1
9	-1	1	-1
10	1	1	1
11	1	-1	-1
12	1	-1	1
13	1	1	1
14	1	-1	1
15	1	1	-1
16	1	-1	-1
17	1	-1	1
18	1	1	-1

(b) GS_P-optimal design with $\pi = 0.8844$

Run	W_1	X_1	X_2
1	-1	-1	-1
2	-1	-1	1
3	-1	1	-1
4	1	-1	1
5	1	1	1
6	1	1	-1
7	1	1	1
8	1	-1	1
9	1	-1	-1
10	-1	1	1
11	-1	1	-1
12	-1	-1	-1
13	-1	1	1
14	-1	1	-1
15	-1	-1	1
16	1	-1	1
17	1	-1	-1
18	1	1	-1

Table B.7: 18-run 2-level GBD- and GS_P-optimal split-plot designs with 3 factors when $\rho_1 = 0.6$ and $\rho_2 = 0.4$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	-1	1	1
2	-1	-1	-1
3	-1	-1	1
4	-1	-1	1
5	-1	-1	-1
6	-1	1	-1
7	1	-1	1
8	1	-1	-1
9	1	1	-1
10	1	1	1
11	1	-1	-1
12	1	-1	1
13	1	1	1
14	1	1	-1
15	1	-1	-1
16	-1	1	1
17	-1	-1	1
18	-1	1	-1

(b) GS_P-optimal design with $\pi = 0.8844$

Run	W_1	X_1	X_2
1	1	-1	1
2	1	1	1
3	1	1	-1
4	-1	1	-1
5	-1	-1	-1
6	-1	1	1
7	-1	1	-1
8	-1	-1	-1
9	-1	-1	1
10	-1	-1	1
11	-1	1	-1
12	-1	1	1
13	1	1	-1
14	1	-1	1
15	1	-1	-1
16	1	1	1
17	1	-1	1
18	1	-1	-1

Table B.8: 18-run 2-level GBD- and GS_P-optimal split-plot designs with 3 factors when $\rho_1 = 0.7$ and $\rho_2 = 0.3$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	-1	-1	-1
2	-1	-1	1
3	-1	1	1
4	1	-1	1
5	1	1	-1
6	1	-1	-1
7	1	1	1
8	1	-1	-1
9	1	-1	1
10	1	1	1
11	1	1	-1
12	1	-1	-1
13	-1	1	1
14	-1	1	-1
15	-1	-1	1
16	-1	-1	-1
17	-1	-1	1
18	-1	1	-1

(b) GS_P-optimal design with $\pi = 0.8844$

Run	W_1	X_1	X_2
1	-1	1	1
2	-1	-1	1
3	-1	-1	-1
4	1	-1	-1
5	1	-1	1
6	1	1	-1
7	-1	-1	1
8	-1	1	-1
9	-1	1	1
10	1	-1	-1
11	1	1	-1
12	1	1	1
13	1	-1	-1
14	1	-1	1
15	1	1	1
16	-1	1	1
17	-1	1	-1
18	-1	-1	-1

Table B.9: 18-run 2-level GBD- and GS_P-optimal split-plot designs with 3 factors when $\rho_1 = 0.8$ and $\rho_2 = 0.2$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	-1	1	1
2	-1	-1	1
3	-1	1	-1
4	1	1	-1
5	1	-1	1
6	1	-1	-1
7	1	1	-1
8	1	1	1
9	1	-1	1
10	-1	-1	-1
11	-1	1	1
12	-1	1	-1
13	1	-1	-1
14	1	1	-1
15	1	1	1
16	-1	1	1
17	-1	-1	-1
18	-1	-1	1

(b) GS_P-optimal design with $\pi = 0.8844$

Run	W_1	X_1	X_2
1	-1	1	1
2	-1	-1	1
3	-1	1	-1
4	1	1	1
5	1	1	-1
6	1	-1	-1
7	1	1	-1
8	1	-1	-1
9	1	-1	1
10	-1	1	1
11	-1	1	-1
12	-1	-1	-1
13	1	-1	1
14	1	-1	-1
15	1	1	1
16	-1	-1	-1
17	-1	-1	1
18	-1	1	1

Table B.10: 18-run 2-level GBD- and GS_P-optimal split-plot designs with 3 factors when $\rho_1 = 0.9$ and $\rho_2 = 0.1$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	1	-1	1
2	1	1	-1
3	1	1	1
4	-1	-1	-1
5	-1	1	-1
6	-1	1	1
7	1	-1	-1
8	1	-1	1
9	1	1	1
10	1	1	1
11	1	-1	-1
12	1	1	-1
13	-1	-1	1
14	-1	1	1
15	-1	1	-1
16	-1	1	-1
17	-1	-1	-1
18	-1	-1	1

(b) GS_P-optimal design with $\pi = 0.8844$

Run	W_1	X_1	X_2
1	-1	1	1
2	-1	-1	1
3	-1	1	-1
4	1	1	1
5	1	1	-1
6	1	-1	-1
7	1	1	-1
8	1	-1	-1
9	1	-1	1
10	-1	1	1
11	-1	1	-1
12	-1	-1	-1
13	1	-1	1
14	1	-1	-1
15	1	1	1
16	-1	-1	-1
17	-1	-1	1
18	-1	1	1

Table B.11: 18-run 3-level GS_P-optimal split-plot designs with 3 factors when $\rho_1 = \rho_2 = 0.5$

(a) $\pi = 0.1542$

Run	W_1	X_1	X_2
1	-1	1	1
2	-1	-1	0
3	-1	1	-1
4	-1	1	0
5	-1	-1	1
6	-1	0	-1
7	0	0	0
8	0	-1	-1
9	0	1	-1
10	1	0	-1
11	1	1	1
12	1	-1	1
13	1	-1	-1
14	1	0	1
15	1	1	-1
16	0	1	1
17	0	-1	1
18	0	0	0

(b) $\pi = 0.2904$

Run	W_1	X_1	X_2
1	-1	-1	1
2	-1	0	-1
3	-1	1	1
4	-1	0	1
5	-1	1	-1
6	-1	-1	-1
7	0	1	-1
8	0	0	1
9	0	-1	0
10	0	-1	1
11	0	0	0
12	0	-1	-1
13	1	-1	0
14	1	1	1
15	1	1	-1
16	1	-1	1
17	1	1	0
18	1	-1	-1

(c) $\pi = 0.4137$

Run	W_1	X_1	X_2
1	0	-1	0
2	0	1	-1
3	0	0	1
4	-1	1	1
5	-1	0	-1
6	-1	-1	0
7	-1	1	0
8	-1	-1	-1
9	-1	-1	1
10	0	0	0
11	0	-1	1
12	0	-1	-1
13	1	0	-1
14	1	-1	1
15	1	1	1
16	1	1	-1
17	1	0	1
18	1	-1	-1

Table B.11: 18-run 3-level GS_P-optimal split-plot designs with 3 factors when $\rho_1 = \rho_2 = 0.5$ (cont.)

(d) $\pi = 0.5271$

Run	W_1	X_1	X_2
1	0	-1	-1
2	0	0	0
3	0	1	-1
4	-1	1	-1
5	-1	0	1
6	-1	-1	-1
7	0	-1	0
8	0	1	1
9	0	0	-1
10	1	-1	-1
11	1	-1	1
12	1	1	0
13	-1	-1	1
14	-1	0	-1
15	-1	1	1
16	1	-1	0
17	1	1	-1
18	1	1	1

(e) $\pi = 0.6327$

Run	W_1	X_1	X_2
1	1	-1	1
2	1	1	1
3	1	0	-1
4	0	-1	1
5	0	0	0
6	0	-1	-1
7	0	1	-1
8	0	0	1
9	0	-1	0
10	-1	-1	-1
11	-1	-1	1
12	-1	1	0
13	-1	-1	0
14	-1	1	1
15	-1	1	-1
16	1	0	1
17	1	1	-1
18	1	-1	-1

(f) $\pi = 0.7319$

Run	W_1	X_1	X_2
1	0	-1	-1
2	0	0	0
3	0	1	1
4	1	1	1
5	1	-1	1
6	1	0	-1
7	1	-1	-1
8	1	1	-1
9	1	0	1
10	0	1	-1
11	0	-1	0
12	0	0	1
13	-1	-1	-1
14	-1	-1	1
15	-1	1	0
16	-1	1	-1
17	-1	1	1
18	-1	-1	0

Table B.11: 18-run 3-level GS_P-optimal split-plot designs with 3 factors when $\rho_1 = \rho_2 = 0.5$ (cont.)

(g) $\pi = 0.8257$

Run	W_1	X_1	X_2
1	0	1	1
2	0	0	0
3	0	-1	-1
4	1	-1	1
5	1	1	1
6	1	1	-1
7	1	1	0
8	1	-1	-1
9	1	-1	1
10	0	-1	0
11	0	0	1
12	0	1	-1
13	-1	1	1
14	-1	-1	1
15	-1	0	-1
16	-1	0	1
17	-1	-1	-1
18	-1	1	-1

(h) $\pi = 0.9148$

Run	W_1	X_1	X_2
1	0	1	0
2	0	-1	-1
3	0	0	1
4	1	1	1
5	1	-1	0
6	1	1	-1
7	0	1	-1
8	0	0	0
9	0	-1	1
10	-1	-1	-1
11	-1	0	1
12	-1	1	-1
13	-1	1	1
14	-1	0	-1
15	-1	-1	1
16	1	-1	-1
17	1	-1	1
18	1	1	0

Table B.12: 18-run 3-level optimal split-plot designs with 3 factors when $\rho_1 = 0.1$ and $\rho_2 = 0.9$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	1	1	-1
2	1	-1	-1
3	1	-1	1
4	0	-1	1
5	0	-1	-1
6	0	0	0
7	-1	-1	1
8	-1	1	-1
9	-1	-1	-1
10	-1	-1	0
11	-1	1	1
12	-1	0	-1
13	1	0	-1
14	1	-1	0
15	1	1	1
16	0	1	-1
17	0	1	0
18	0	0	1

(b) GS_P-optimal design with $\pi = 0.7319$

Run	W_1	X_1	X_2
1	-1	-1	-1
2	-1	1	-1
3	-1	1	1
4	1	1	1
5	1	-1	-1
6	1	-1	1
7	0	0	-1
8	0	1	0
9	0	-1	1
10	0	0	0
11	0	1	1
12	0	-1	-1
13	1	0	1
14	1	1	-1
15	1	-1	0
16	-1	1	-1
17	-1	-1	1
18	-1	0	0

Table B.13: 18-run 3-level optimal split-plot designs with 3 factors when $\rho_1 = 0.2$ and $\rho_2 = 0.8$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	1	1	1
2	1	1	-1
3	1	-1	-1
4	1	1	0
5	1	0	-1
6	1	-1	1
7	-1	-1	-1
8	-1	1	1
9	-1	1	-1
10	0	1	-1
11	0	0	1
12	0	-1	0
13	0	1	1
14	0	0	0
15	0	-1	-1
16	-1	1	0
17	-1	-1	1
18	-1	0	-1

(b) GS_P-optimal design with $\pi = 0.7319$

Run	W_1	X_1	X_2
1	1	-1	1
2	1	1	1
3	1	-1	-1
4	1	1	-1
5	1	0	1
6	1	-1	0
7	0	-1	-1
8	0	0	0
9	0	1	1
10	-1	-1	-1
11	-1	1	-1
12	-1	1	1
13	0	-1	1
14	0	0	-1
15	0	1	0
16	-1	-1	1
17	-1	1	-1
18	-1	0	0

Table B.14: 18-run 3-level optimal split-plot designs with 3 factors when $\rho_1 = 0.3$ and $\rho_2 = 0.7$, and $\pi = 0.7319$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	1	1	-1
2	1	-1	1
3	1	1	1
4	-1	1	-1
5	-1	1	1
6	-1	-1	1
7	0	1	-1
8	0	-1	1
9	0	0	0
10	0	0	-1
11	0	1	1
12	0	-1	0
13	1	-1	-1
14	1	1	0
15	1	0	1
16	-1	0	1
17	-1	1	0
18	-1	-1	-1

(b) GS_P-optimal design

Run	W_1	X_1	X_2
1	1	1	-1
2	1	-1	-1
3	1	0	1
4	0	1	-1
5	0	-1	1
6	0	0	0
7	0	-1	-1
8	0	0	1
9	0	1	0
10	1	-1	1
11	1	0	-1
12	1	1	1
13	-1	-1	0
14	-1	1	1
15	-1	1	-1
16	-1	-1	-1
17	-1	-1	1
18	-1	1	0

Table B.15: 18-run 3-level optimal split-plot designs with 3 factors when $\rho_1 = 0.4$ and $\rho_2 = 0.6$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	1	-1	1
2	1	1	0
3	1	-1	-1
4	0	0	1
5	0	-1	-1
6	0	1	0
7	-1	-1	-1
8	-1	1	-1
9	-1	0	1
10	0	1	1
11	0	1	-1
12	0	0	0
13	-1	0	-1
14	-1	-1	1
15	-1	1	1
16	1	1	-1
17	1	-1	0
18	1	1	1

(b) GS_P-optimal design with $\pi = 0.7319$

Run	W_1	X_1	X_2
1	0	0	1
2	0	1	0
3	0	-1	-1
4	1	-1	-1
5	1	1	-1
6	1	-1	1
7	1	0	-1
8	1	-1	0
9	1	1	1
10	0	0	0
11	0	-1	1
12	0	1	-1
13	-1	1	1
14	-1	-1	-1
15	-1	-1	1
16	-1	-1	0
17	-1	1	1
18	-1	1	-1

Table B.16: 18-run 3-level optimal split-plot designs with 3 factors when $\rho_1 = 0.5$ and $\rho_2 = 0.5$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	-1	-1	-1
2	-1	1	-1
3	-1	0	1
4	-1	1	1
5	-1	-1	1
6	-1	0	-1
7	1	1	1
8	1	-1	0
9	1	1	-1
10	0	0	0
11	0	1	1
12	0	-1	1
13	0	-1	0
14	0	1	-1
15	0	0	1
16	1	-1	1
17	1	1	0
18	1	-1	-1

(b) GS_P-optimal design with $\pi = 0.7319$

Run	W_1	X_1	X_2
1	0	-1	-1
2	0	0	0
3	0	1	1
4	1	1	1
5	1	-1	1
6	1	0	-1
7	1	-1	-1
8	1	1	-1
9	1	0	1
10	0	1	-1
11	0	-1	0
12	0	0	1
13	-1	-1	-1
14	-1	-1	1
15	-1	1	0
16	-1	1	-1
17	-1	1	1
18	-1	-1	0

Table B.17: 18-run 3-level optimal split-plot designs with 3 factors when $\rho_1 = 0.6$ and $\rho_2 = 0.4$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	1	-1	-1
2	1	1	-1
3	1	0	1
4	1	1	1
5	1	-1	1
6	1	0	-1
7	0	1	-1
8	0	0	0
9	0	-1	-1
10	-1	1	0
11	-1	-1	-1
12	-1	-1	1
13	0	0	-1
14	0	-1	0
15	0	1	1
16	-1	-1	0
17	-1	1	-1
18	-1	1	1

(b) GS_P-optimal design with $\pi = 0.7319$

Run	W_1	X_1	X_2
1	-1	1	1
2	-1	1	-1
3	-1	-1	0
4	0	1	-1
5	0	0	0
6	0	-1	1
7	1	0	-1
8	1	-1	1
9	1	1	1
10	0	-1	-1
11	0	1	0
12	0	0	1
13	-1	-1	-1
14	-1	1	0
15	-1	-1	1
16	1	0	1
17	1	-1	-1
18	1	1	-1

Table B.18: 18-run 3-level optimal split-plot designs with 3 factors when $\rho_1 = 0.7$ and $\rho_2 = 0.3$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	1	-1	1
2	1	0	-1
3	1	1	1
4	1	1	-1
5	1	0	1
6	1	-1	-1
7	-1	-1	-1
8	-1	1	0
9	-1	-1	1
10	-1	1	1
11	-1	-1	0
12	-1	1	-1
13	0	-1	0
14	0	1	1
15	0	0	-1
16	0	0	0
17	0	-1	-1
18	0	1	-1

(b) GS_P-optimal design with $\pi = 0.7319$

Run	W_1	X_1	X_2
1	-1	-1	1
2	-1	-1	-1
3	-1	1	0
4	-1	-1	0
5	-1	1	-1
6	-1	1	1
7	1	-1	1
8	1	1	1
9	1	0	-1
10	0	-1	1
11	0	1	-1
12	0	0	0
13	0	1	0
14	0	-1	-1
15	0	0	1
16	1	-1	-1
17	1	0	1
18	1	1	-1

Table B.19: 18-run, 3-level optimal split-plot designs with 3 factors when $\rho_1 = 0.8$ and $\rho_2 = 0.2$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	1	1	0
2	1	-1	-1
3	1	-1	1
4	-1	1	-1
5	-1	0	1
6	-1	-1	-1
7	1	1	-1
8	1	1	1
9	1	-1	0
10	0	0	1
11	0	1	-1
12	0	-1	0
13	-1	-1	1
14	-1	1	1
15	-1	0	-1
16	0	1	1
17	0	-1	1
18	0	0	0

(b) GS_P-optimal design with $\pi = 0.7319$

Run	W_1	X_1	X_2
1	-1	-1	-1
2	-1	-1	1
3	-1	1	0
4	1	-1	-1
5	1	0	1
6	1	1	-1
7	1	1	1
8	1	-1	1
9	1	0	-1
10	0	1	1
11	0	-1	-1
12	0	0	0
13	-1	-1	0
14	-1	1	1
15	-1	1	-1
16	0	0	1
17	0	1	-1
18	0	-1	0

Table B.20: 18-run 3-level optimal split-plot designs with 3 factors when $\rho_1 = 0.9$ and $\rho_2 = 0.1$

(a) GBD-optimal design

Run	W_1	X_1	X_2
1	0	1	1
2	0	1	-1
3	0	0	0
4	-1	1	0
5	-1	-1	-1
6	-1	-1	1
7	1	0	-1
8	1	1	1
9	1	-1	1
10	-1	1	-1
11	-1	1	1
12	-1	-1	0
13	1	-1	-1
14	1	0	1
15	1	1	-1
16	0	0	-1
17	0	-1	1
18	0	1	0

(b) GS_P-optimal design with $\pi = 0.7319$

Run	W_1	X_1	X_2
1	0	-1	1
2	0	1	0
3	0	0	-1
4	-1	-1	1
5	-1	1	1
6	-1	0	-1
7	0	-1	-1
8	0	1	1
9	0	0	0
10	-1	1	-1
11	-1	-1	-1
12	-1	0	1
13	1	1	-1
14	1	1	1
15	1	-1	0
16	1	-1	-1
17	1	-1	1
18	1	1	0

Table B.21: 20-run 2-level GS_P-optimal staggered-level designs with 3 factors when $\rho_1 = 0.6$, $\rho_2 = 0.3$, and $\rho_3 = 0.1$

(a) $\pi = 0.2628$

Run	W_1	W_2	X_1
1	-1	1	-1
2	-1	1	1
3	-1	-1	-1
4	-1	-1	1
5	1	-1	-1
6	1	-1	1
7	1	1	1
8	1	1	-1
9	-1	1	1
10	-1	1	-1
11	-1	-1	-1
12	-1	-1	1
13	1	-1	-1
14	1	-1	1
15	1	1	1
16	1	1	-1
17	-1	1	1
18	-1	1	-1
19	-1	1	1
20	-1	1	-1

(b) $\pi = 0.4567$

Run	W_1	W_2	X_1
1	1	1	-1
2	1	1	1
3	1	-1	1
4	1	-1	-1
5	-1	-1	-1
6	-1	-1	1
7	-1	1	-1
8	-1	1	1
9	1	1	-1
10	1	1	1
11	1	-1	-1
12	1	-1	1
13	-1	-1	1
14	-1	-1	-1
15	-1	1	-1
16	-1	1	1
17	1	1	-1
18	1	1	1
19	1	1	-1
20	1	1	1

(c) $\pi = 0.6176$

Run	W_1	W_2	X_1
1	-1	1	-1
2	-1	1	1
3	-1	-1	-1
4	-1	-1	1
5	1	-1	1
6	1	-1	-1
7	1	1	-1
8	1	1	1
9	-1	1	-1
10	-1	1	1
11	-1	-1	1
12	-1	-1	-1
13	1	-1	-1
14	1	-1	1
15	1	1	-1
16	1	1	1
17	-1	1	1
18	-1	1	-1
19	-1	1	1
20	-1	1	-1

Table B.21: 20-run 2-level GS_P-optimal staggered-level designs with 3 factors when $\rho_1 = 0.6$, $\rho_2 = 0.3$, and $\rho_3 = 0.1$ (cont.)

(d) $\pi = 0.7581$

Run	W_1	W_2	X_1
1	1	-1	1
2	1	-1	-1
3	1	1	-1
4	1	1	1
5	-1	1	-1
6	-1	1	1
7	-1	-1	-1
8	-1	-1	1
9	1	-1	1
10	1	-1	-1
11	1	1	-1
12	1	1	1
13	-1	1	1
14	-1	1	-1
15	-1	-1	1
16	-1	-1	-1
17	1	-1	1
18	1	-1	-1
19	1	-1	1
20	1	-1	-1

(e) $\pi = 0.8844$

Run	W_1	W_2	X_1
1	1	-1	-1
2	1	-1	1
3	1	1	-1
4	1	1	1
5	-1	1	1
6	-1	1	-1
7	-1	-1	-1
8	-1	-1	1
9	1	-1	1
10	1	-1	-1
11	1	1	-1
12	1	1	1
13	-1	1	-1
14	-1	1	1
15	-1	-1	-1
16	-1	-1	1
17	1	-1	-1
18	1	-1	1
19	1	-1	-1
20	1	-1	1

Table B.22: 20-run 2-level GBD- and GS_P-optimal staggered-level designs with 3 factors when $\rho_1 = 0.6$, $\rho_2 = 0.3$, and $\rho_3 = 0.1$

(a) GBD-optimal design

Run	W_1	W_2	X_1
1	1	-1	-1
2	1	-1	1
3	1	1	-1
4	1	1	1
5	-1	1	-1
6	-1	1	1
7	-1	-1	-1
8	-1	-1	1
9	1	-1	1
10	1	-1	-1
11	1	1	-1
12	1	1	1
13	-1	1	1
14	-1	1	-1
15	-1	-1	-1
16	-1	-1	1
17	1	-1	1
18	1	-1	-1
19	1	-1	1
20	1	-1	-1

(b) GS_P-optimal design with π=0.8844

Run	W_1	W_2	X_1
1	1	-1	-1
2	1	-1	1
3	1	1	-1
4	1	1	1
5	-1	1	1
6	-1	1	-1
7	-1	-1	-1
8	-1	-1	1
9	1	-1	1
10	1	-1	-1
11	1	1	-1
12	1	1	1
13	-1	1	-1
14	-1	1	1
15	-1	-1	-1
16	-1	-1	1
17	1	-1	-1
18	1	-1	1
19	1	-1	-1
20	1	-1	1

Table B.23: 20-run 2-level GBD- and GS_P-optimal staggered-level designs with 3 factors when $\rho_1 = 0.6$, $\rho_2 = 0.2$, and $\rho_3 = 0.2$

(a) GBD-optimal design

Run	W_1	W_2	X_1
1	-1	1	-1
2	-1	1	1
3	-1	-1	1
4	-1	-1	-1
5	1	-1	1
6	1	-1	-1
7	1	1	-1
8	1	1	1
9	-1	1	1
10	-1	1	-1
11	-1	-1	-1
12	-1	-1	1
13	1	-1	1
14	1	-1	-1
15	1	1	-1
16	1	1	1
17	-1	1	-1
18	-1	1	1
19	-1	1	-1
20	-1	1	1

(b) GS_P-optimal design with π=0.8844

Run	W_1	W_2	X_1
1	-1	1	1
2	-1	1	-1
3	-1	-1	1
4	-1	-1	-1
5	1	-1	1
6	1	-1	-1
7	1	1	1
8	1	1	-1
9	-1	1	1
10	-1	1	-1
11	-1	-1	-1
12	-1	-1	1
13	1	-1	-1
14	1	-1	1
15	1	1	-1
16	1	1	1
17	-1	1	1
18	-1	1	-1
19	-1	1	-1
20	-1	1	1

Table B.24: 20-run 2-level GBD- and GS_P-optimal staggered-level designs with 3 factors when $\rho_1 = 0.6$, $\rho_2 = 0.1$, and $\rho_2 = 0.3$

(a) GBD-optimal design

Run	W_1	W_2	X_1
1	1	1	-1
2	1	1	1
3	1	-1	-1
4	1	-1	1
5	-1	-1	-1
6	-1	-1	1
7	-1	1	-1
8	-1	1	1
9	1	1	1
10	1	1	-1
11	1	-1	-1
12	1	-1	1
13	-1	-1	1
14	-1	-1	-1
15	-1	1	1
16	-1	1	-1
17	1	1	-1
18	1	1	1
19	1	1	1
20	1	1	-1

(b) GS_P-optimal design with π=0.8844

Run	W_1	W_2	X_1
1	1	-1	-1
2	1	-1	1
3	1	1	1
4	1	1	-1
5	-1	1	-1
6	-1	1	1
7	-1	-1	1
8	-1	-1	-1
9	1	-1	1
10	1	-1	-1
11	1	1	-1
12	1	1	1
13	-1	1	-1
14	-1	1	1
15	-1	-1	1
16	-1	-1	-1
17	1	-1	-1
18	1	-1	1
19	1	-1	1
20	1	-1	-1

Table B.25: 20-run 2-level GBD- and GS_P-optimal staggered-level designs with 3 factors when $\rho_1 = 0.3$, $\rho_2 = 0.6$, and $\rho_3 = 0.1$

(a) GBD-optimal design

Run	W_1	W_2	X_1
1	-1	1	1
2	-1	1	-1
3	-1	-1	1
4	-1	-1	-1
5	1	-1	1
6	1	-1	-1
7	1	1	1
8	1	1	-1
9	-1	1	1
10	-1	1	-1
11	-1	-1	1
12	-1	-1	-1
13	1	-1	1
14	1	-1	-1
15	1	1	-1
16	1	1	1
17	-1	1	1
18	-1	1	-1
19	-1	1	1
20	-1	1	-1

(b) GS_P-optimal design with π=0.8844

Run	W_1	W_2	X_1
1	1	-1	1
2	1	-1	-1
3	1	1	1
4	1	1	-1
5	-1	1	1
6	-1	1	-1
7	-1	-1	1
8	-1	-1	-1
9	1	-1	-1
10	1	-1	1
11	1	1	1
12	1	1	-1
13	-1	1	-1
14	-1	1	1
15	-1	-1	-1
16	-1	-1	1
17	1	-1	1
18	1	-1	-1
19	1	-1	-1
20	1	-1	1

Table B.26: 20-run 2-level GBD- and GS_P-optimal staggered-level designs with 3 factors when $\rho_1 = 0.2$, $\rho_2 = 0.6$, and $\rho_3 = 0.2$

(a) GBD-optimal design

Run	W_1	W_2	X_1
1	-1	1	-1
2	-1	1	1
3	-1	-1	-1
4	-1	-1	1
5	1	-1	-1
6	1	-1	1
7	1	1	1
8	1	1	-1
9	-1	1	-1
10	-1	1	1
11	-1	-1	1
12	-1	-1	-1
13	-1	-1	-1
14	-1	-1	1
15	-1	-1	-1
16	-1	-1	1
17	1	-1	-1
18	1	-1	1
19	1	1	-1
20	1	1	1

(b) GS_P-optimal design with π=0.8844

Run	W_1	W_2	X_1
1	1	-1	-1
2	1	-1	1
3	1	1	1
4	1	1	-1
5	-1	1	-1
6	-1	1	1
7	-1	-1	1
8	-1	-1	-1
9	1	-1	1
10	1	-1	-1
11	1	1	-1
12	1	1	1
13	-1	1	-1
14	-1	1	1
15	-1	-1	1
16	-1	-1	-1
17	1	-1	-1
18	1	-1	1
19	1	-1	1
20	1	-1	-1

Table B.27: 20-run 2-level GBD- and GS_P-optimal staggered-level designs with 3 factors when $\rho_1 = 0.1$, $\rho_2 = 0.6$, and $\rho_3 = 0.3$

(a) GBD-optimal design

Run	W_1	W_2	X_1
1	-1	1	-1
2	-1	1	1
3	-1	-1	-1
4	-1	-1	1
5	1	-1	1
6	1	-1	-1
7	1	1	1
8	1	1	-1
9	-1	1	-1
10	-1	1	1
11	-1	-1	-1
12	-1	-1	1
13	1	-1	-1
14	1	-1	1
15	1	-1	-1
16	1	-1	1
17	-1	-1	-1
18	-1	-1	1
19	-1	1	1
20	-1	1	-1

(b) GS_P-optimal design with π=0.8844

Run	W_1	W_2	X_1
1	1	-1	1
2	1	-1	-1
3	1	1	-1
4	1	1	1
5	-1	1	1
6	-1	1	-1
7	-1	-1	1
8	-1	-1	-1
9	1	-1	1
10	1	-1	-1
11	1	1	-1
12	1	1	1
13	-1	1	1
14	-1	1	-1
15	-1	-1	1
16	-1	-1	-1
17	1	-1	1
18	1	-1	-1
19	1	-1	1
20	1	-1	-1

Table B.28: 20-run 3-level GS_P-optimal staggered-level designs with 3 factors when $\rho_1 = 0.6$, $\rho_2 = 0.3$, and $\rho_3 = 0.1$

(a) $\pi = 0.1542$

Run	W_1	W_2	X_1
1	0	1	0
2	0	1	-1
3	0	0	0
4	0	0	1
5	0	0	1
6	0	0	0
7	0	-1	-1
8	0	-1	0
9	1	-1	1
10	1	-1	-1
11	1	1	-1
12	1	1	1
13	-1	1	1
14	-1	1	-1
15	-1	-1	1
16	-1	-1	-1
17	1	-1	-1
18	1	-1	1
19	1	1	1
20	1	1	-1

(b) $\pi = 0.2904$

Run	W_1	W_2	X_1
1	1	-1	-1
2	1	-1	1
3	1	1	-1
4	1	1	1
5	-1	1	1
6	-1	1	-1
7	-1	-1	-1
8	-1	-1	1
9	1	-1	-1
10	1	-1	1
11	1	1	1
12	1	1	-1
13	0	1	1
14	0	1	0
15	0	0	0
16	0	0	-1
17	0	0	0
18	0	0	-1
19	0	-1	1
20	0	-1	0

(c) $\pi = 0.4137$

Run	W_1	W_2	X_1
1	1	1	1
2	1	1	-1
3	1	-1	1
4	1	-1	-1
5	0	-1	0
6	0	-1	-1
7	0	0	1
8	0	0	0
9	-1	0	0
10	-1	0	-1
11	-1	1	-1
12	-1	1	1
13	1	1	-1
14	1	1	1
15	1	-1	1
16	1	-1	-1
17	-1	-1	1
18	-1	-1	-1
19	-1	1	-1
20	-1	1	1

Table B.28: 20-run 3-level GS_P-optimal staggered-level designs with 3 factors when $\rho_1 = 0.6$, $\rho_2 = 0.3$, and $\rho_3 = 0.1$ (cont.)

(d) $\pi = 0.5271$

Run	W_1	W_2	X_1
1	1	0	-1
2	1	0	0
3	1	-1	-1
4	1	-1	1
5	-1	-1	1
6	-1	-1	-1
7	-1	1	-1
8	-1	1	1
9	1	1	-1
10	1	1	1
11	1	-1	1
12	1	-1	-1
13	-1	-1	1
14	-1	-1	-1
15	-1	1	-1
16	-1	1	1
17	0	1	0
18	0	1	-1
19	0	0	0
20	0	0	1

(e) $\pi = 0.6327$

Run	W_1	W_2	X_1
1	1	-1	1
2	1	-1	-1
3	1	1	-1
4	1	1	1
5	0	1	1
6	0	1	0
7	0	0	-1
8	0	0	0
9	-1	0	1
10	-1	0	0
11	-1	-1	-1
12	-1	-1	1
13	1	-1	1
14	1	-1	-1
15	1	1	-1
16	1	1	1
17	-1	1	-1
18	-1	1	1
19	-1	-1	1
20	-1	-1	-1

(f) $\pi = 0.7319$

Run	W_1	W_2	X_1
1	-1	0	0
2	-1	0	-1
3	-1	-1	1
4	-1	-1	-1
5	1	-1	1
6	1	-1	-1
7	1	1	1
8	1	1	-1
9	-1	1	-1
10	-1	1	1
11	-1	-1	-1
12	-1	-1	1
13	1	-1	-1
14	1	-1	1
15	1	1	-1
16	1	1	1
17	0	1	-1
18	0	1	0
19	0	0	1
20	0	0	0

Table B.28: 20-run 3-level GS_P-optimal staggered-level designs with 3 factors when $\rho_1 = 0.6$, $\rho_2 = 0.3$, and $\rho_3 = 0.1$ (cont.)

(g) $\pi = 0.8257$

Run	W_1	W_2	X_1
1	1	0	1
2	1	0	0
3	1	-1	1
4	1	-1	-1
5	-1	-1	1
6	-1	-1	-1
7	-1	1	1
8	-1	1	-1
9	1	1	-1
10	1	1	1
11	1	-1	1
12	1	-1	-1
13	-1	-1	-1
14	-1	-1	1
15	-1	1	-1
16	-1	1	1
17	0	1	1
18	0	1	0
19	0	0	0
20	0	0	-1

(h) $\pi = 0.9148$

Run	W_1	W_2	X_1
1	-1	1	-1
2	-1	1	1
3	-1	-1	-1
4	-1	-1	1
5	1	-1	1
6	1	-1	-1
7	1	1	1
8	1	1	-1
9	-1	1	1
10	-1	1	-1
11	-1	-1	1
12	-1	-1	-1
13	1	-1	-1
14	1	-1	1
15	1	0	0
16	1	0	1
17	0	0	0
18	0	0	-1
19	0	1	1
20	0	1	0

Table B.29: 20-run 3-level GBD- and GS_P-optimal staggered-level designs with 3 factors when $\rho_1 = 0.6$, $\rho_2 = 0.3$, and $\rho_3 = 0.1$

(a) GBD-optimal design

Run	W_1	W_2	X_1
1	-1	1	1
2	-1	1	-1
3	-1	-1	1
4	-1	-1	-1
5	0	-1	0
6	0	-1	1
7	0	0	0
8	0	0	-1
9	1	0	0
10	1	0	1
11	1	1	1
12	1	1	-1
13	-1	1	-1
14	-1	1	1
15	-1	-1	-1
16	-1	-1	1
17	1	-1	-1
18	1	-1	1
19	1	1	-1
20	1	1	1

(b) GS_P-optimal design with $\pi = 0.8257$

Run	W_1	W_2	X_1
1	1	0	1
2	1	0	0
3	1	-1	1
4	1	-1	-1
5	-1	-1	1
6	-1	-1	-1
7	-1	1	1
8	-1	1	-1
9	1	1	-1
10	1	1	1
11	1	-1	1
12	1	-1	-1
13	-1	-1	-1
14	-1	-1	1
15	-1	1	-1
16	-1	1	1
17	0	1	1
18	0	1	0
19	0	0	0
20	0	0	-1

Table B.30: 20-run 3-level GBD- and GS_P-optimal staggered-level designs with 3 factors when $\rho_1 = 0.6$, $\rho_2 = 0.2$, and $\rho_3 = 0.2$

(a) GBD-optimal design

Run	W_1	W_2	X_1
1	0	-1	0
2	0	-1	-1
3	0	0	0
4	0	0	1
5	1	0	0
6	1	0	-1
7	1	1	-1
8	1	1	1
9	-1	1	-1
10	-1	1	1
11	-1	-1	1
12	-1	-1	-1
13	1	-1	1
14	1	-1	-1
15	1	1	-1
16	1	1	1
17	-1	1	1
18	-1	1	-1
19	-1	-1	-1
20	-1	-1	1

(b) GS_P-optimal design with $\pi = 0.8257$

Run	W_1	W_2	X_1
1	-1	0	-1
2	-1	0	0
3	-1	1	1
4	-1	1	-1
5	1	1	1
6	1	1	-1
7	1	-1	-1
8	1	-1	1
9	-1	-1	1
10	-1	-1	-1
11	-1	1	1
12	-1	1	-1
13	1	1	-1
14	1	1	1
15	1	-1	-1
16	1	-1	1
17	0	-1	-1
18	0	-1	0
19	0	0	0
20	0	0	1

Table B.31: 20-run 3-level GBD- and GS_P-optimal staggered-level designs with 3 factors when $\rho_1 = 0.6$, $\rho_2 = 0.1$, and $\rho_3 = 0.3$

(a) GBD-optimal design

Run	W_1	W_2	X_1
1	1	1	-1
2	1	1	1
3	1	-1	1
4	1	-1	-1
5	-1	-1	-1
6	-1	-1	1
7	-1	0	0
8	-1	0	1
9	0	0	-1
10	0	0	0
11	0	1	0
12	0	1	1
13	1	1	-1
14	1	1	1
15	1	-1	1
16	1	-1	-1
17	-1	-1	0
18	-1	-1	1
19	-1	1	1
20	-1	1	-1

(b) GS_P-optimal design with $\pi = 0.8257$

Run	W_1	W_2	X_1
1	-1	1	1
2	-1	1	-1
3	-1	-1	1
4	-1	-1	-1
5	1	-1	-1
6	1	-1	1
7	1	1	-1
8	1	1	1
9	-1	1	0
10	-1	1	1
11	-1	-1	1
12	-1	-1	-1
13	1	-1	0
14	1	-1	-1
15	1	0	-1
16	1	0	1
17	0	0	-1
18	0	0	0
19	0	1	1
20	0	1	-1

Table B.32: 20-run 3-level GBD- and GS_P-optimal staggered-level designs with 3 factors when $\rho_1 = 0.3$, $\rho_2 = 0.6$, and $\rho_3 = 0.1$

(a) GBD-optimal design

Run	W_1	W_2	X_1
1	-1	1	-1
2	-1	1	1
3	-1	-1	1
4	-1	-1	-1
5	0	-1	0
6	0	-1	1
7	0	0	0
8	0	0	-1
9	1	0	0
10	1	0	1
11	1	1	-1
12	1	1	1
13	-1	1	1
14	-1	1	-1
15	-1	-1	1
16	-1	-1	-1
17	1	-1	-1
18	1	-1	1
19	1	1	-1
20	1	1	1

(b) GS_P-optimal design with $\pi = 0.8257$

Run	W_1	W_2	X_1
1	-1	0	1
2	-1	0	0
3	-1	1	-1
4	-1	1	1
5	1	1	-1
6	1	1	1
7	1	-1	-1
8	1	-1	1
9	-1	-1	1
10	-1	-1	-1
11	-1	1	1
12	-1	1	-1
13	1	1	-1
14	1	1	1
15	1	-1	-1
16	1	-1	1
17	0	-1	1
18	0	-1	0
19	0	0	0
20	0	0	-1

Table B.33: 20-run 3-level *GBD*- and GS_P-optimal staggered-level designs with 3 factors when $\rho_1 = 0.2$, $\rho_2 = 0.6$, and $\rho_3 = 0.2$

(a) *GBD*-optimal design

Run	W_1	W_2	X_1
1	-1	-1	1
2	-1	-1	-1
3	-1	1	-1
4	-1	1	1
5	1	1	-1
6	1	1	1
7	1	-1	-1
8	1	-1	1
9	-1	-1	1
10	-1	-1	-1
11	-1	0	-1
12	-1	0	0
13	0	0	0
14	0	0	1
15	0	1	-1
16	0	1	0
17	1	1	-1
18	1	1	1
19	1	-1	1
20	1	-1	-1

(b) GS_P-optimal design with $\pi = 0.8257$

Run	W_1	W_2	X_1
1	-1	-1	1
2	-1	-1	-1
3	-1	1	-1
4	-1	1	1
5	1	1	1
6	1	1	-1
7	1	-1	1
8	1	-1	-1
9	-1	-1	-1
10	-1	-1	1
11	-1	1	1
12	-1	1	-1
13	0	1	0
14	0	1	-1
15	0	0	1
16	0	0	0
17	1	0	-1
18	1	0	0
19	1	-1	-1
20	1	-1	1

Table B.34: 20-run 3-level GBD- and GS_P-optimal staggered-level designs with 3 factors when $\rho_1 = 0.1$, $\rho_2 = 0.6$, and $\rho_3 = 0.3$

(a) GBD-optimal design

Run	W_1	W_2	X_1
1	1	-1	-1
2	1	-1	1
3	1	1	1
4	1	1	-1
5	-1	1	-1
6	-1	1	1
7	-1	-1	1
8	-1	-1	-1
9	1	-1	1
10	1	-1	-1
11	1	0	1
12	1	0	0
13	0	0	0
14	0	0	-1
15	0	1	0
16	0	1	1
17	-1	1	-1
18	-1	1	1
19	-1	-1	1
20	-1	-1	-1

(b) GS_P-optimal design with $\pi = 0.8257$

Run	W_1	W_2	X_1
1	-1	1	-1
2	-1	1	1
3	-1	-1	1
4	-1	-1	-1
5	0	-1	0
6	0	-1	-1
7	0	0	0
8	0	0	1
9	1	0	0
10	1	0	-1
11	1	1	-1
12	1	1	1
13	-1	1	-1
14	-1	1	1
15	-1	-1	1
16	-1	-1	-1
17	1	-1	1
18	1	-1	-1
19	1	1	-1
20	1	1	1