Optimal Designs for Paired Comparison Experiments

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)

von M.Phil. Eric Nyarko

geb. am 28.02.1986 in Kumasi, Ghana

genehmigt durch die Fakultät für Mathematik
der Otto-von-Guericke-Universität Magdeburg

Gutachter: Prof. Dr. Rainer Schwabe
Prof. Dr. Heinz Holling

eingereicht am: 18.06.2019

Verteidigung am: 20.09.2019

Acknowledgements

First and foremost, I would like to acknowledge my indebtedness and render my warmest thanks to my supervisor, Prof. Dr. Rainer Schwabe. This work would have been a mirage without his support and encouragement. His friendly and astute advice have been invaluable throughout all stages of the work.

The thesis has also benefited from extended discussions with Dr. Heiko Großmann in the early-stage of my research. I take this opportunity to thank him.

I am very much grateful to the members of the Institute for Mathematical Stochastics, for providing a conducive atmosphere and support for me. This include Kerstin Altenkirch, Prof. Dr. Norbert Gaffke, as well as colleagues, like Dr. Alexander Meier, Dr. Kerstin Reckrühm and Frank Röttger, for their good heart and kindness. My sincere thanks also go to my colleagues, Osama Idais from Palestine and Parisa Parsamaram from Iran, for their encouragement during the past years.

My profound gratitude goes to Dr. Isaac K. Baidoo and Prof. Kwabena Doku-Amponsah of the Department of Statistics and Actuarial Science, University of Ghana, for initiating and supervising my master's research work in the area of discrete choice experiments.

Special thanks are due to my wife, Mrs. Mavis Nyarko, as well as my children, Nana Kwame Amankwah-Nyarko and Ohemah Tawiah Amankwah-Nyarko, for their continuous support and understanding during their stay with me in Magdeburg, Germany. My thanks are extended to my father, Mr. Isaac Amankwah as well as my mother, Mrs. Augustina Amankwah, and all my four brothers, for their constant encouragement.

I would like to also thank the German Academic Exchange Service (DAAD) for its financial support.

Summary

The aim of this thesis is to derive optimal designs for linear paired comparison models with second- or third-order interactions in an analysis of variance setup where the attributes are qualitative with the same number of levels each.

After the first introductory chapter on the problem and the literature some basic concepts are presented in the Chapter 2 about paired comparison experiments in the linear model setup, and particular emphasis is laid on the special case of the part-worth model in which the influence of the attributes is additive and consists only of main effects without interactions. The fundamentals and general descriptions of optimal designs as well as some commonly used optimality criteria are presented in Chapter 3. In Chapters 4 and 5 results are presented on optimal designs for the part-worth model and a model with first-order interactions, respectively, which have been known from the literature and where components of a single attribute are used as building blocks for the holistic approach. A powerful tool for characterizing the optimal designs in these models is given by the concept of invariance.

These concepts are extended to linear paired comparison models with second-order interactions in Chapters 6 and 7 for binary attributes and for attributes with a general common number of levels, respectively. A general statement on the maximal number of types of pairs can be formulated for optimal designs, where orbits are specified by the number of attributes in which the two alternatives differ. While for part-worth (main effects) models optimal designs consist of those alternatives which differ in all attributes and for first-order interactions they consist of those pairs of alternatives which differ in about half of the attributes, respectively, there seems to be no clear general rule in models with second-order interactions. For models with small profile strengths analytic results can be obtained for optimal designs while for larger profile strengths optimal designs have to be determined numerically. Moreover, for binary attributes optimal designs require two types of pairs in which either all attributes have distinct levels or approximately half of the attributes are distinct and the other half of the attributes coincide. For larger number of levels mostly one type of pairs is sufficient. In some exceptional cases two types of pairs are needed, and only for the full interaction case all types are required.

In Chapters 8 and 9 these results are extended to paired comparison models with third-order interactions. For binary attributes two types of pairs have to be considered for which the numbers of distinct attributes are symmetric with respect to about half of the profile strength. For larger number of levels again only one type of pair is sufficient in nearly all cases.

The thesis is concluded with a brief discussion and an outlook on future research.

Zusammenfassung

Das Ziel dieser Arbeit ist die Herleitung optimaler experimenteller Designs für Paarvergleichsmodelle unter Zugrundelegung von Linearen Modellen der Varianzanalyse mit Wechselwirkungen zweiter bzw. dritter Ordnung. Dabei setzen sich die Alternativen aus mehreren, die Entscheidungen beeinflussenden Attributen zusammen, die jeweils auf eine feste Anzahl von Ausprägungen (Stufen) eingestellt werden können.

Nach einem einleitenden Kapitel in die Problemstellung und die Literatur werden im zweiten Kapitel die grundlegenden Konzepte für Paarvergleiche im Linearen Modell eingeführt. Dabei wird der klassische Spezialfall des Teilwertmodells, in dem nur Haupteffekte der Attribute und keine Wechselwirkungen auftreten, gesondert behandelt. Darauf folgen im drtitten Kapitel grundlegende Erläuterungen zu optimalen Designs sowie zu üblicherweise verwendeten Optimalitätskriterien.

Die Interaktionsmodelle zweiter und dritter Ordnung, welche das Teilwertmodell mit Komponenten eines einzelnen Attributs als Bausteine für die Resultate benutzen, sowie die Interaktionsmodelle erster Ordnung, werden in Kapitel 4 und 5 beschrieben. Von besonderer Relevanz ist dabei das Konzept der Invarianz.

In den Kapiteln 6 und 7 werden diese Konzepte auf lineare Paarvergleichsmodelle mit Interaktionen zweiter Ordnung und binären Attributen, sowie Attributen von identischer Stufenanzahl, erweitert. Hierbei kann ein allgemeingültiges Resultat über die maximale Anzahl benötigter Typen von Paaren für optimale Designs formuliert werden, wobei die verschiedenen Typen von Paaren durch die Anzahl von unterschiedlichen Attributen der Alternativen spezifiziert werden. Dabei bestehen optimale Designs für Teilwertmodelle aus den Alternativen, in denen sich alle Attribute unterscheiden, während Interaktionsmodelle erster Ordnung aus Alternativen bestehen, die sich in ungefähr der Hälfte der Attribute unterscheiden. Im Fall der Interaktionsmodelle zweiter Ordnung scheint es keine derart allgemeingültige Regel zu geben. Für Modelle von kleinen Profilstärken können analytische Lösungen für optimale Designs gefunden werden, während größere Profilstärken numerische Methoden erfordern. Zudem bedürfen optimale Designs im Fall von binären Attributen zweier Typen von Paaren, in denen entweder alle Attribute unterschiedliche Stufen haben oder ungefähr jeweils die Hälfte der Attribute identisch und unterschiedlich sind. Für eine größere Anzahl von Stufen genügt in der Regel ein Typ von Paaren. In Ausnahmefällen bedarf es zweier Typen von Paaren, und nur für den Fall vollständiger Interaktion werden alle Typen von Paaren benötigt.

In den Kapiteln 8 und 9 werden die zuvor beschriebenen Resultate auf Paarvergleichsmodelle mit Interaktionen der dritten Ordnung erweitert. Für binäre Attribute müssen wieder zwei Typen von Paaren verwendet werden, für die die Anzahl der verschiedenen Attribute symmetrisch zur Hälfte der Profilstärke ist. Für größere Stufenanzahlen genügt es in der Regel erneut, nur einen Typ von Paaren zu betrachten.

Die Arbeit schließt mit einer kurzen Diskussion und einem Ausblick in zukünftige Forschungsfragen.

Contents

List of Tables

1 Introduction

Paired comparisons are closely related to experiments with choice sets of size two (sets of two competing alternatives or options), which are presented to respondents in a form of a questionnaire for them to indicate the most preferred alternative by trading off one alternative against the other. Modes of presentation are usually a paper and pencil task that is either self-administered or presented through an interviever, a full-blown multimedia event or administered by mail or through the internet. The method of paired comparison experiments was early introduced in psychophysics by Fechner (1860) to measure the perceived heaviness of vessels. Moreover, over the last few years paired comparison experiments have also received considerable attention in many other fields of applications like statistics, health economics, transport, multidimensional scaling, sports competition, marketing, and many others (Davidson and Farquhar, 1976) for learning consumer preferences towards new products or services. For example, taste testing experiments are often designed as paired comparison experiments (Scheffé, 1952).

The roots of paired comparisons can also be traced back to the work by Thurstone (1927) and Bradley and Terry (1952), who developed models under the various assumption that the probability for one of the only two possible outcomes (preference for one or the other alternative in so-called choice sets of size two) of each comparison follow a cumulative distribution function of a normal or logistic distribution, respectively.

As was already pointed out by Train (2003) the easiest and most widely used choice model is the multinomial logit model (MNL model), which was originally derived by Luce (1959). An appealing feature of the MNL model is that after a suitable parameterization it reduces to the aforementioned Bradley-Terry model for the important particular case of paired comparisons (e.g., see Großmann, Holling, and Schwabe, 2002). Also, the MNL choice probabilities which has been derived by McFadden (1974) take a closed form, which means that the traditional maximum-likelihood procedures can be applied to obtain model parameters without using simulation methods, and is easily interpretable (Bunch and Batsell, 1989; Train, 2003).

While the MNL model embody the property of independence from irrelevant alternatives (IIA), which means that the ratio of the choice probability for any two alternatives in a choice set is not affected by addition or deletion of other alternatives in the set. The consequence of this property is that the MNL model is consistent with utility maximization (Marschak, 1960).

To relax the property of IIA, and as an extension of the MNL model, more flexible choice models such as the generalized extreme value (GEV) and the mixed logit model have been developed (Train, 2003, p. 54). However, there have been few implemantations or applications of these flexible models in choice experiments (see Carson et al., 1994; Train, 2003). The MNL model continues to play a key role in choice experiments, and it performs well for other models (Burgess, Street, and Wasi, 2011; Bush, 2014). A comprehensive introduction to the general area of choice experiments can be found in the literature (Louviere, Hensher, and Swait, 2000; Train, 2003; Louviere, Street, and Burgess, 2004; Großmann and Schwabe, 2015).

Another prominent technique of measuring paired comparisons which is a variant of the choice experiments is known as conjoint analysis, which originated from conjoint measurement in psychology (Luce and Tukey, 1964), and was first applied in marketing by Green and Rao (1971). Reviews of theory, methods and applications of conjoint analysis are provided (Green and Srinivasan, 1990). The main difference of conjoint analysis from choice experiments is the response formats used in the survey and the statistical models used for analyzing data. While for conjoint analysis responses are usually assessed on a rating scale (or generally respondents assign a score representing the degree of preferences) and the general linear model is used, discrete choice experiments which is non-linear draws on the aforementioned MNL model (see Großmann, Holling, and Schwabe, 2002). Some early useful contributions to these areas can be found in (Scheffé, 1952; Quenouille and John, 1971; Green, 1974; Green and Srinivasan, 1978).

The problem of optimal design constructions for paired comparison studies has been early considered in the literature (El-Helbawy and Bradley, 1978; Offen and Littell, 1987; van Berkum, 1987b). For instance, when there are three attributes each at two-levels some optimal design results incorporating main-effects and all interactions have been derived for a general mixed factorial model under the assumption that model parameters are equal (El-Helbawy and Bradley, 1978). These results have been extended to the case when there are more than three attributes (El-Helbawy and Ahmed, 1984) and to general asymetrical factorial paired comparison experiments (El-Helbawy, Ahmed, and Alharbey, 1994). Corresponding results on optimal designs for symmetric and asymmetric factorial experiments have also been obtained by El-Helbawy, Ahmed, and Alharbey (1994). Offen and Littell (1987) derived optimal paired comparison designs for the Bradley-Terry model without ties. Moreover, optimal paired comparison designs for main-effects and two-attribute interactions under the assumption of no treatment differences have been derived (van Berkum, 1987a; van Berkum, 1987b; van Berkum, 1989), and some optimal design results when there is prior information about model parameter values have also been obtained (Huber and Zwerina, 1996; Sandor and Wedel, 2001). Some results about generating locally optimal designs for the MNL model can also be found (Huber and Zwerina, 1996).

In what follows, it is worthwhile mentioning that the linear difference model considered here can be realized as a linearization of the binary response model by Bradley and Terry (1952) under the assumption that the parameter vector $\boldsymbol{\beta} = \mathbf{0}$ (e.g., see Großmann, Holling, and Schwabe, 2002). Specifically, under this indifference assumption of equal choice probabilities, the Bradley-Terry type choice experiments as in the work of Street and Burgess (2007), amongst others can be derived by considering the linear paired comparison model. In particular, this assumption simplifies the information matrix of the binary logit model because of its non-linearity and the dependence of the model parameters on the information matrix. As a consequence, and in particular for the present work optimal designs for the binary logit model can be derived by considering the linear paired comparison model, which is the approach adopted in many scientific works (Street, Bunch, and Moore, 2001; Street and Burgess, 2004; Street, Burgess, and Louviere, 2005; Graßhoff et al., 2003; Graßhoff et al., 2004; Großmann and Schwabe, 2015).

Typical with paired comparisons, respondents usually evaluate pairs of competing options (alternatives) in a hypothetical (occasionally real) setting which are generated by an experimental design as already mentioned, and are characterized by a number of attribute levels. The preferences of respondents are analyzed with a statistical model like the binary logit model or the linear paired comparison model to provide quantitative measures or utility estimates of the relative importance of each attribute. However, in applications (such as marketing), situations may arise in which practitioners may be interested in special comparison among the attributes (interactions). For example, Bradley and El-Helbawy (1976) considered a taste preference experiment on coffee with three attributes; brew strength, roast color and coffee brand, each at two levels, and up to three-attribute interactions. The various assumptions about comparison among these three attributes involving main-effects, two-attribute and three-attribute interactions are well summarized in their Tables 3 and 4 of their paper. Similar results involving three-attribute interactions can also be found in Example 3 of El-Helbawy and Bradley (1978). Elrod, Louviere, and Davey (1992) also considered a study about student preferences for rental apartments involving main-effects, two-attribute, three-attribute and four-attribute interactions, which is well summarized in Table 2 of their paper. Another strand of work for the case of direct observation that incorporates three-attribute interactions using real data in a randomized clinical trial of high-risk mother-baby dyads can also be found (Shiao, Ahn, and Akazawa, 2007). Although not much attention has been given to higher-order interactions in the literature, the aforementioned works serve as a motivation for the present work when three or four of the attributes interact.

As frequently observed, in applications the choice task imposes cognitive burden when the alternatives presented are specified by too many attributes. As such the choice sets are answered anyhow which can destroy the quality of the data. In that situation, a way to simplify the choice task is to specify only a few components (attributes) of the alternatives so-called profile strength (see Großmann and Schwabe, 2015). Experiments that embody this method are known as partial profile experiments (e.g., see Green, 1974; Chrzan, 2010; Graßhoff et al., 2003). In the present setting the work or result on partial profiles is motivated by a recent study (e.g., see Großmann, 2018) in the health sector which is focused on the construction of partial profile choice design for pairs with eleven attributes, each at two levels, and where only four attributes are to be shown to respondents simultanously.

The main contribution of this thesis lies in the introduction of an appropriate model for the situation of full and partial profiles and to derive optimal designs in the presence of interactions. We consider the case when the components of the alternatives are specified by two-level and common number general-level attributes. Work on determining the structure of the optimal designs by the two-level situation has been carried out (van Berkum, 1987a; van Berkum, 1987b; Street, Bunch, and Moore, 2001) in the case of full profiles in a main-effects and first-order interactions setup, and by Schwabe et al. (2003) for partial profiles. Corresponding results when the common number of the attribute levels is larger than two have been obtained by Graßhoff et al. (2003) in a first-order interactions setup for both full and partial profiles. Here we treat the case of both second-order and third-order interactions.

This thesis is organized as follows. In Chapter 2 general linear models are introduced for paired comparisons. The fundamentals of the statistical theory of optimal experimental designs are presented in Chapter 3. Optimal design for the linear paired comparison model with emphasis on the part-worth model as well as the standard parameterization with a single attribute are briefly considered in Chapter 4, and a brief summary of optimal designs in the presence of the first-order interaction models as a motivation for the present work are presented in Chapter 5. The main part of this work on determination of optimal designs in the presence of the second-order interaction models for the case of two-level attributes with full and partial profiles are presented in Chapter 6. These results are extended to the case of common number general-level attributes in Chapter 7. In addition, optimal designs in the presence of the third-order interaction models for the case of two-level attributes with full and partial profiles are presented in Chapter 8, and the corresponding results for common number general-level attributes are presented in Chapter 9. The final Chapter 10 offers a discussion of the results and an outlook on future research.

2 Paired Comparison Models

For paired comparison models the works by van Berkum (1987b), Bradley and Terry (1952), El-Helbawy and Bradley (1978) and Großmann and Schwabe (2015) amongst others build a good introduction. The problem of paired comparison experiments in the linear paired comparisons model is well discussed in the literature (see van Berkum, 1987b; Quenouille and John, 1971; Graßhoff et al., 2003; Großmann, 2003; Graßhoff et al., 2004; Großmann and Schwabe, 2015, amongst others). The aforementioned publications serve as a motivation for the present work.

In the underlying sections we provide some basic concepts about paired comparison experiments in the linear paired comparisons model. A special case becomes the part-worth model (see Green and Srinivasan, 1978, p. 105). These models will serve as a building block to construct paired comparison designs with interactions later on after Chapter 4.

2.1 General Linear Model for Paired Comparisons

In paired comparisons the outcome of the experiment depends on some factors (attributes), say, K of influence. In this setting the dependence can be described by a functional relationship $\tilde{\mathbf{f}}$ of dimension p which quantifies the effect of the alternative $\mathbf{i}$ of the K attributes of influence. Thus, for every $k = 1, \ldots, K$ the set of all possible realizations of the i_k level of the k-th attribute can be identified with a finite set $\mathcal{I}_k \subseteq \mathbb{R}$. Hence, each alternative can be represented as $\mathbf{i} = (i_1, \ldots, i_K)$ which are elements from the set $\mathcal{I} = \mathcal{I}_1 \times \cdots \times \mathcal{I}_K$. Any observation (utility) $\tilde{Y}_{na}(\mathbf{i})$ of a single alternative $\mathbf{i} = (i_1, \ldots, i_K)$ within a pair of block alternatives $(a = 1, 2)$ is subject to a random error $\tilde{\varepsilon}_{na}$, which is assumed to be uncorrelated with constant variance and zero mean. Hence, we formalize the experimental situation by a general linear model

$$\tilde{Y}_{na}(\mathbf{i}) = \mu_n + \tilde{\mathbf{f}}(\mathbf{i})^\top \boldsymbol{\beta} + \tilde{\varepsilon}_{na}, \tag{2.1}$$

with $\mathbf{i} \in \mathcal{I}$ where the index n denotes the n-th presentation for which $\mathbf{i}$ is chosen from the set $\mathcal{I}$ of possible realizations for the alternative, $n = 1, \ldots, N$, and N is defined as the fixed sample size. If we assume that the corresponding mean response (function) $E(\tilde{Y}_{na}(\mathbf{i})) = \mu_n + \tilde{\mathbf{f}}(\mathbf{i})^\top \boldsymbol{\beta}$ can be parameterized in a linear way, then $E(\tilde{Y}_{na}(\mathbf{i}))$ is known to be a linear combination of some p known regression functions $f_i : \mathcal{I} \to \mathbb{R},\ i = 1, \ldots, p$. For notational convenience we summarize the regression functions into one regression function $\tilde{\mathbf{f}} : \mathcal{I} \to \mathbb{R}^p$ with $\tilde{\mathbf{f}}(\mathbf{i}) = (f_1(\mathbf{i}), \ldots, f_p(\mathbf{i}))^\top$ for every $\mathbf{i} \in \mathcal{I}$. Hence, the response function $E(\tilde{Y}_{na}(\mathbf{i}))$ is determined up to the unkown parameter vector $\boldsymbol{\beta} \in \mathbb{R}^p$.

By denoting $\tilde{\mathbf{Y}}_a = (\tilde{Y}_{1a}(\mathbf{i}_1), \ldots, \tilde{Y}_{Na}(\mathbf{i}_N))^\top$ as the vector of observations, $\tilde{\boldsymbol{\varepsilon}}_a = (\tilde{\varepsilon}_{1a}, \ldots, \tilde{\varepsilon}_{Na})^\top$ as the vector of errors and $\tilde{\mathbf{F}} = (\tilde{\mathbf{f}}(\mathbf{i}_1), \ldots, \tilde{\mathbf{f}}(\mathbf{i}_N))^\top$ as the matrix of dimension N, then model (2.1) can be reformulated in a vector notation as

$$\tilde{\mathbf{Y}}_a = \tilde{\mathbf{F}}\boldsymbol{\beta} + \tilde{\boldsymbol{\varepsilon}}_a. \tag{2.2}$$

We now consider paired comparison experiments. In paired comparison experiments the utilities for the alternatives are not directly measurable. Only preferences can be observed for comparing pairs of alternatives $(\mathbf{i}, \mathbf{j}) = ((i_1, \ldots, i_K), (j_1, \ldots, j_K)) \in \mathcal{I} \times \mathcal{I}$. Hence, we assume that the preference is quantified as the difference between utilities $Y_n(\mathbf{i}, \mathbf{j}) = \tilde{Y}_{n1}(\mathbf{i}) - \tilde{Y}_{n2}(\mathbf{j})$. In this case the utilities for the alternatives are properly described by the linear paired comparison model

$$Y_n(\mathbf{i}, \mathbf{j}) = (\tilde{\mathbf{f}}(\mathbf{i}) - \tilde{\mathbf{f}}(\mathbf{j}))^\top \boldsymbol{\beta} + \varepsilon_n, \tag{2.3}$$

with $(\mathbf{i}, \mathbf{j}) \in \mathcal{I} \times \mathcal{I}$. Here if we impose the concept of linear parametrization as described in (2.1) on the corresponding mean response $E(Y_n(\mathbf{i}, \mathbf{j})) = (\tilde{\mathbf{f}}(\mathbf{i}) - \tilde{\mathbf{f}}(\mathbf{j}))^\top \boldsymbol{\beta}$ for $\tilde{\mathbf{f}} : \mathcal{I} \times \mathcal{I} \to \mathbb{R}^p$ then $\tilde{\mathbf{f}}(\mathbf{i}) - \tilde{\mathbf{f}}(\mathbf{j})$ is the derived regression function and the random error $\varepsilon_n = \tilde{\varepsilon}_{n1} - \tilde{\varepsilon}_{n2}$ associated with the pairs $(\mathbf{i}, \mathbf{j}) \in \mathcal{I} \times \mathcal{I}$ is assumed to be uncorrelated with constant variance and zero mean.

By letting $\mathbf{Y} = (\tilde{Y}_{n1}(\mathbf{i}_n) - \tilde{Y}_{n2}(\mathbf{j}_n), \ldots, \tilde{Y}_{N1}(\mathbf{i}_N) - \tilde{Y}_{N2}(\mathbf{j}_N))^\top$ or $\mathbf{Y} = \tilde{\mathbf{Y}}_1 - \tilde{\mathbf{Y}}_2$ and $\boldsymbol{\varepsilon} = (\tilde{\varepsilon}_{n1} - \tilde{\varepsilon}_{n2}, \ldots, \tilde{\varepsilon}_{N1} - \tilde{\varepsilon}_{N2})^\top$ or $\boldsymbol{\varepsilon} = \tilde{\boldsymbol{\varepsilon}}_1 - \tilde{\boldsymbol{\varepsilon}}_2$ be the vectors of observations and errors of dimension N, respectively, and $\mathbf{F} = (\tilde{\mathbf{f}}(\mathbf{i}_n) - \tilde{\mathbf{f}}(\mathbf{j}_n), \ldots, \tilde{\mathbf{f}}(\mathbf{i}_N) - \tilde{\mathbf{f}}(\mathbf{j}_N))^\top$ be the matrix of dimension $N \times p$, the model (2.3) can be reformulated in a vector notation as

$$\mathbf{Y} = \mathbf{F}\boldsymbol{\beta} + \boldsymbol{\varepsilon}. \tag{2.4}$$

2.2 Part-Worth Model

In paired comparison every attribute is usually assigned with some small number of levels. For this situation the utility of every attribute-levels is equal to its part, and the overall utility of every alternative constitutes the sum of its parts (sum of so-called part-worth utilities). Among the preference models, and for the particular case of generic attributes the part-worth model approach has received wide acceptance because it is readily interpretable (Green and Srinivasan, 1978).

With this model, of course, only a finite number of levels for the K attributes will be considered for the paired comparison experiments. Thus the k-th attribute at v_k levels in the paired comparison experiments can be represented by the set $\mathcal{I}_k = \{1, \ldots, v_k\}$ for $k = 1, \ldots, K$. As pointed out by Green and Srinivasan (1978), modeling the full range of the attributes is essential to enhance the validity of part-worth values. Hence, for the multiple-attributes $k = 1, \ldots, K$ the mean response $E(\tilde{Y}_{na}(\mathbf{i}))$ of the alternative $\mathbf{i} = (i_1, \ldots, i_K)$ from the set $\mathcal{I} = \mathcal{I}_1 \times \cdots \times \mathcal{I}_K$ can be formulated as

$$E(\tilde{Y}_{na}(\mathbf{i})) = \mu_n + \sum_{k=1}^{K} \alpha_{i_k}^{(k)}, \tag{2.5}$$

where the parameter $\alpha_{i_k}^{(k)}$ is the part-worth of a particular level i_k of the attribute k as will be explored for the particular case of just a single-attribute in Section 4.2. Hence, from (2.1) the mean response $E(\tilde{Y}_{na}(\mathbf{i}))$ for every $\mathbf{i} \in \mathcal{I}$ can be defined by the vector of regression functions

$$\tilde{\mathbf{f}}(\mathbf{i}) = (\mathbf{f}_1(\mathbf{i})^\top, \ldots, \mathbf{f}_K(\mathbf{i})^\top)^\top$$

with $i_k \in \mathcal{I}_k$ where for $k = 1, \ldots, K$ each at levels $i_k = 1, \ldots, v_k - 1$ the vector of regression functions are given by

$$\mathbf{f}_k(\mathbf{i}) = \mathbf{e}_{i_k} \text{ for } i_k = 1, \ldots, v_k - 1, \text{ and } \mathbf{f}_k(\mathbf{i}) = \mathbf{1}_{v_k - 1}, \tag{2.6}$$

where $\mathbf{e}_i$ denotes the i-th unit vector of length $v_k - 1$ and $\mathbf{1}_m$ the m-dimensional vector with all entries equal to 1. The corresponding part-worth (reduced) parameter vector is given by

$$\boldsymbol{\beta} = (\alpha_1^{(1)}, \ldots, \alpha_{v_1-1}^{(1)}, \ldots, \alpha_1^{(K)}, \ldots, \alpha_{v_K-1}^{(K)})^\top, \tag{2.7}$$

which satisfies the usual identifiability condition $\sum_{i_k=1}^{v_k} \alpha_{i_k}^{(k)} = 0$ for $k = 1, \ldots, K$. As a consequence, from (2.4) the mean response $E(Y_n(\mathbf{i}, \mathbf{j}))$, $n = 1, \ldots, N$ for all $(\mathbf{i}, \mathbf{j}) \in \mathcal{I} \times \mathcal{I}$ can be reformulated by the matrix of regression functions $\mathbf{F}$ defined in (2.4).

3 Optimal Experimental Designs

The statistical theory of optimal experimental design is concerned with the allocation of treatments (combinatons of a finite number of levels of attributes) to respondents, and the choice of those values of the attributes in a linear (or non-linear) model at which observations should be taken (Smith, 1918; Atkinson, 2011). Usually, best experimental designs are assertained by a certain criteria. The modern statistical theory of optimal experimental designs can be found in a series of papers by Kiefer (see Atkinson, 2011). A comprehensive introduction to the theory of optimal experimental designs is well established in the literature (e.g., see Pázman, 1986; Pukelsheim, 1993; Schwabe, 1996; Fedorov and Hackl, 1997; Cox and Reid, 2000; Atkinson, Donev, and Tobias, 2007). A general introduction to the applications of optimal experimental designs in the area of paired comparisons (conjoint and choice experiments) can be found in the works by Kuhfeld, Tobias, and Garratt (1994), Louviere and Woodworth (1983), Louviere, Street, and Burgess (2004), Huber and Zwerina (1996) and Großmann (2003), amongst others. For more recent applications of optimal experimental designs (see e.g. Verelst et al., 2018; Luyten et al., 2015; Großmann, 2017; Großmann, 2018). The subsequent chapters draw on the concepts about optimal experimental designs in the aforementioned publications.

3.1 Preliminaries

In design of experiments several attributes may involve. For instance, in a chemical experiment possible attributes of influence on the response may be the time of reaction, pressure, temperature and the catalyst used. In psychology career preferences the possible attributes of influence may be academic, industrial, educational and clinical. It is up to the experimenter (investigator) to then select the treatments under study in order to obtain good estimates of responses. In this case, we make use of the general linear model as a function of qualitative independent K attributes. Suppose x_k is the levels of the k-th component (attribute) selected from the set $\mathcal{I}_k$, $k = 1, \dots, K$. Then the experimental setting $\mathbf{x} = (x_1, \dots, x_K)$ is a K-tuple from the experimental region $\mathcal{X}$. The general linear model is then formalized by

$$Y_n(\mathbf{x}) = \mathbf{f}(\mathbf{x})^\top \boldsymbol{\beta} + \varepsilon_n, \tag{3.1}$$

where as before $Y_n(\mathbf{x})$ denotes the n-th observation on the response at setting $\mathbf{x}$, $\mathbf{f} = (f_1, \dots, f_p)^\top : \mathcal{X} \to \mathbb{R}^p$ is a vector of p known regression functions, $\boldsymbol{\beta} \in \mathbb{R}^p$ denotes the unkown parameter vector, the experimental region $\mathcal{X}$ is assumed to be a compact set with image $\mathcal{X} \subset \mathbb{R}^p$. The observational errors ε_n are assumed to be uncorrelated with constant variance σ^2.

In the vector notation the general linear model (3.1) can be reformulated as

$$\mathbf{Y} = \mathbf{F}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{3.2}$$

where as before $\mathbf{Y} = (Y_1(\mathbf{x}_1), \ldots, Y_N(\mathbf{x}_N))^\top$ is the vector of N observations, $\mathbf{F} = (\mathbf{f}(\mathbf{x}_1), \ldots, \mathbf{f}(\mathbf{x}_N))^\top$ denotes the design matrix and $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_N)^\top$ is the vector of N observational errors. We assume that the design matrix $\mathbf{F}$ has full column rank p, $\boldsymbol{\varepsilon}$ is uncorrelated have zero mean and constant variance σ^2. With these assumptions the method of least squares or the Gauss-Markov theorem is usually used to obtain the model parameters $\boldsymbol{\beta}$.

By letting $\hat{\boldsymbol{\beta}}$ be the estimate of $\boldsymbol{\beta}$, the best linear unbiased estimator for $\boldsymbol{\beta}$ is given by the Gauss-Markov estimator

$$\hat{\boldsymbol{\beta}} = (\mathbf{F}^\top\mathbf{F})^{-1}\mathbf{F}^\top\mathbf{Y}, \tag{3.3}$$

where the covariance matrix of $\hat{\boldsymbol{\beta}}$ is given by

$$\mathrm{Cov}(\hat{\boldsymbol{\beta}}) = \sigma^2(\mathbf{F}^\top\mathbf{F})^{-1}. \tag{3.4}$$

As the theory of optimal experimental designs is focused on minimizing functions of the variances and covariances in order to obtain good parameter estimates, an experimenter has to select the experimental settings $\mathbf{x}_1, \ldots, \mathbf{x}_N$ in a suitable sense.

For convenience in notation, we point out that how the situation of paired comparison fits into the corresponding experimental setting $\mathbf{x}$ (the standard experimental situation) will be briefly introduced/discussed at the beginning of Chapter 4, because in paired comparison experiments one is interested in differences between alternatives (van Berkum, 1987b). This will further lead to a discussion in Section 4.1 about the minimal reparameterization of the corresponding models, in particular the parth-worth model (2.5).

3.2 Designs and Information

In this section we give the description of a design, which is a collection of experimental settings $\mathbf{x}_1, \ldots, \mathbf{x}_N$ (usually under the control of an experimenter) from the experimental region $\mathcal{X}$ that specifies the range of values of the experimental attributes, and optimal for the estimation of the unknown parameter $\boldsymbol{\beta}$ in the general linear model (3.1). The quality of the design may be measured by means of the covariance matrix (3.4).

An exact experimental design $\tilde{\xi}$ of size N of possible level combinations from the experimental region $\mathcal{X}$ is a vector $(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ of experimental setting $\mathbf{x}_n \in \mathcal{X},\ n = 1, \ldots, N$, which need not necessarily be distinct. Here, if we assume that the experimental settings are the same as characterized in (3.2) then the exact experimental design $\tilde{\xi}$ of size N from $\mathcal{X}$ has corresponding design matrix $\mathbf{F}$ defined in (3.2). Hence, the normalized information matrix $\mathbf{M}(\tilde{\xi})$ of the exact design $\tilde{\xi}$ of size N from $\mathcal{X}$ is defined by

$$\mathbf{M}(\tilde{\xi}) = \frac{1}{N}\mathbf{F}^\top\mathbf{F}$$

which can be obtained from the inverse of the covariance matrix (3.4) of the best linear unbiased estimator $\hat{\boldsymbol{\beta}}$ when we get rid of the constant term σ^2. The design matrix $\mathbf{F}$ as defined in (3.2) has full column rank.

The exact experimental design $\tilde{\xi}$ of size N from $\mathcal{X}$ may be alternatively represented by its M distinct settings (different level combinations) defined by $\mathbf{x}_1, \ldots, \mathbf{x}_M$ and the corresponding numbers $N_1, \ldots, N_M$ of replications with $\sum_{m=1}^{M} N_m = N$ as

$$\tilde{\xi} = \begin{pmatrix} \mathbf{x}_1 & \ldots & \mathbf{x}_M \\ N_1 & \ldots & N_M \end{pmatrix}. \tag{3.5}$$

The exact experimental design $\tilde{\xi}$ of size N from $\mathcal{X}$ can be identified with a discrete probability (design) measure

$$\xi = \sum_{m=1}^{M} w_m \epsilon_{\{\mathbf{x}_m\}} \tag{3.6}$$

where $w_m = \frac{N_m}{N}$ is the proportion of observations at the setting $\mathbf{x}_m,\ m = 1, \ldots, M$; however, the assumption of integer values for the numbers $N_1, \ldots, N_M$ of replications is dropped and only the conditions $w_m \geq 0$ and $\sum_{m=1}^{M} w_m = 1$ have to be satisfied, and $\epsilon_{\{\mathbf{x}_m\}}$ denotes the one-point measure on $\mathbf{x}_m$ (see Schwabe, 1996). The information matrix $\mathbf{M}(\xi)$ of the generalized design ξ is defined by

$$\mathbf{M}(\xi) = \sum_{m=1}^{M} w_m \mathbf{f}(\mathbf{x}_m)\mathbf{f}(\mathbf{x}_m)^\top.$$

Hence, by letting $\epsilon_{\{\mathbf{x}\}}$ denotes the one-point measure on $\mathbf{x}$ as similarly specified in (3.6) the information matrix can be alternatively written as an integral with respect to the design measure ξ as

$$\mathbf{M}(\xi) = \int_{\mathcal{X}} \mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top \xi(d\mathbf{x}).$$

For the exact experimental design $\tilde{\xi}$ the (normalized) variance of the mean response $\frac{N}{\sigma^2}\operatorname{Var}(\mathbf{f}(\mathbf{x})^\top \hat{\boldsymbol{\beta}}))$ is defined by

$$V(\mathbf{x}, \tilde{\xi}) = \mathbf{f}(\mathbf{x})^\top \mathbf{M}(\tilde{\xi})^{-1}\mathbf{f}(\mathbf{x}),$$

which is a function of both $\tilde{\xi}$ and the experimental setting $\mathbf{x}$. Moreover, for the generalized design ξ the corresponding variance function is defined by

$$V(\mathbf{x}, \xi) = \mathbf{f}(\mathbf{x})^\top \mathbf{M}(\xi)^{-1}\mathbf{f}(\mathbf{x}). \tag{3.7}$$

The departure from the exact experimental design $\tilde{\xi}$ of size N from $\mathcal{X}$ to the probability measure ξ is stimulated by the fact that it is usually convenient to apply optimization techniques to treat experimental designs incorporating probability measures instead of exact experimental designs. This concept of transition is discussed in the literature (see Kiefer, 1959; Schwabe, 1996; Atkinson and Donev, 1992, amongst others).

For technical ease, we follow Schwabe (1996, p. 7) and give a formal definition of a (generalized) experimental design ξ as a probability measure on (the experimental or design region) $\mathcal{X}$. We note that for the generalized experimental design ξ the information matrix $\mathbf{M}(\tilde{\xi}) = \mathbf{M}(\xi) = \int_{\mathcal{X}} \mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})^\top \xi(d\mathbf{x})$ is identical to the information matrix (which is positive-semidefinite) of the corresponding exact experimental design $\tilde{\xi}$ of size N from $\mathcal{X}$.

Throughout the sequel we note that every generalized experimental design ξ on $\mathcal{X}$ in (3.6) will be characterized by distinct design points $\{\mathbf{x} \in \mathcal{X} : \ \xi(\{\mathbf{x}\})\}$ where the probabilities $\xi(\{\mathbf{x}\})$ on all the distinct design points are rational numbers. In this case the probability measure $\xi = \sum_{m=1}^{M} w_m \epsilon_{\{\mathbf{x}_m\}}$ coincides with every exact experimental design $\tilde{\xi}$ with least size N characterized in (3.5). However, since in practice all designs are exact, if the weights w_m are not rational, it will not be possible to find an exact experimental design $\tilde{\xi}$ which is identical with the generalized experimental design ξ (see Atkinson and Donev, 1992, p. 94).

3.3 Optimality Criteria

In this section we describe some specific (alphabetic) design criteria of A-, D-, E- and G-optimality (e.g. see Kiefer, 1959; Atkinson and Donev, 1992; Schwabe, 1996, for detailed discussion), which have statistical interpretation in terms of the covariance matrix (3.4) of the Gauss-Markov estimator (3.3) or the information matrix $\mathbf{M}(\tilde{\xi})$ of the exact experimental design $\tilde{\xi}$ of size N and the information matrix $\mathbf{M}(\xi)$ of the generalized experimental design ξ presented in Section 3.2. Thus, if ξ_1 and ξ_2 are two experimental designs such that the difference of their informaton matrix $\mathbf{M}(\xi_1) - \mathbf{M}(\xi_2)$ is positive definite, then for the aforementioned alphabetic design criteria, ξ_1 will be better than ξ_2. Additionally, if the experimental design ξ_1 can be found such that $\mathbf{M}(\xi_1) - \mathbf{M}(\xi_2)$ is at least positive-semidefinite for all ξ_2 and positive definite for some ξ_2, then the experimental design ξ_1 is considered to be globally optimal (see Atkinson and Donev, 1992, p. 115). Analogously, if $\boldsymbol{\beta}$ is identifiable, then in this case for every linear combination $c^\top\boldsymbol{\beta}$, $c \in \mathbb{R}^p$ the (variances) for $\boldsymbol{\beta}$ as well as the covariance matrix (3.4) is simultanously minimized by ξ_1, and additionally, ξ_1 will produce the best linear unbiased estimator $\hat{\boldsymbol{\beta}}$ for $\boldsymbol{\beta}$ in the experimental situation considered in Section 3.1.

Here because of completness we consider the general linear model (3.1) with the usual assumption that the regression functions $\mathbf{f}$ of components $\mathbf{f}_1, \dots, \mathbf{f}_p$ are linearly independent on the experimental region $\mathcal{X}$. In any case the alphabetic optimality criteria can be described as functions of the information matrix $\mathbf{M}(\xi)$ with eigenvalues $\lambda_1, \dots, \lambda_p$. For instance, A-optimality $\mathrm{tr}(\mathbf{M}(\xi)^{-1})$ minimizes the mean variance of the parameter vector $\boldsymbol{\beta}$, D-optimality minimizes the generalized variance (or determinant $\det(\mathbf{M}(\xi)^{-1})$) of the parameter estimates and E-optimality minimizes the variance of the least well-estimated any linear combination $c^\top\boldsymbol{\beta}$ satisfying the constraint $c^\top c = 1$.

In the following, we formalized the (definition) of the aforementioned optimality criteria. An experimental design ξ^* is (defined) as A-optimal if $\mathrm{tr}(\mathbf{M}(\xi^*)^{-1}) \leq \mathrm{tr}(\mathbf{M}(\xi)^{-1})$ holds for every experimental design ξ with (regular) information matrix $\mathbf{M}(\xi)$. This definition is equivalent to considering the exact experimental design $\tilde{\xi}$ with regular information matrix $\mathbf{M}(\tilde{\xi})$ proportional to $\mathrm{tr}(\mathbf{M}(\tilde{\xi})^{-1})$. An experimental design ξ^* is (defined) as E-optimal if $\lambda^*_{\max}(\mathbf{M}(\xi^*)^{-1}) \leq \lambda_{\max}(\mathbf{M}(\xi)^{-1})$ holds for every experimental design ξ with regular information matrix $\mathbf{M}(\xi)$ where $\lambda_{\max}$ denotes the largest eigenvalue $\lambda_1, \dots, \lambda_p$ of $\mathbf{M}(\xi)$. This definition is equivalent to considering the exact experimental design $\tilde{\xi}$ with regular information matrix $\mathbf{M}(\tilde{\xi})$ proportional to $\lambda_{\max}(\mathbf{M}(\tilde{\xi})^{-1})$.

Further we consider the celebrated criterion of D-optimality. An experimental design ξ^* is (defined) as D-optimal if $\det(\mathbf{M}(\xi^*)^{-1}) \leq \det(\mathbf{M}(\xi)^{-1})$ or, equivalently, $\det(\mathbf{M}(\xi^*)) \geq \det(\mathbf{M}(\xi))$ holds for every experimental design ξ with regular information matrix $\mathbf{M}(\xi)$. As a result, if we assume that the observations on $Y_n(\mathbf{x})$ at the experimental setting $\mathbf{x}$ in model (3.1) are independent or uncorrelated, then this can be motivated by a confidence ellipsoid for the parameter vector $\boldsymbol{\beta}$. The volume of the ellipsoid is inversely proportional to $\det(\mathbf{M}(\xi))^{1/2}$, which the D-optimality seeks to maximizing $\det(\mathbf{M}(\xi))$ (see Silvey, 1980, p. 10). In particular, this criterion is convex and, hence convex optimation can be used. The G-optimality is another important criterion. We (define) an experimental design ξ^* to be G-optimal if it minimizes the maximum over the experimental region $\mathcal{X}$ of the variance function $V(\mathbf{x}, \xi)$ in (3.7) satisfying the property

$$\max_{\mathbf{x}\, \in \mathcal{X}} V(\mathbf{x}, \xi^*) = \min_{\xi} \max_{\mathbf{x}\, \in \mathcal{X}} V(\mathbf{x}, \xi).$$

For the particular case of generalized designs, this design measure ξ^* will also be D-optimal and $V(\mathbf{x}, \xi^*) = p$, where p is the number of model parameters in the general linear model (3.1). This equality has only to hold for $\mathbf{x}$ in the support of ξ^*, i.e. for those $\mathbf{x}$ with $\xi^*(\{\mathbf{x}\}) > 0$. This concept of equivalence of D- and G-optimality designs establishes the celebrated Kiefer and Wolfowitz (1960) equivalence theorem (see Atkinson and Donev, 1992, p. 116).

To compare two competing designs to find out which is better, it is a standard technique to rely on the normalized information matrix which expresses information on a per-obsevation basis, and then use the so-called efficiency measure. For D-optimality criterion which aims at maximizing the determinant of the normalized information matrix to facilitate the analysis of the design efficiency, we define the efficiency of every experimental design ξ relative to the D-optimal experimental design ξ^* by

$$\mathrm{eff}_D(\xi) = 100 \times \left(\frac{\det(\mathbf{M}(\xi))}{\det(\mathbf{M}(\xi^*))} \right)^{1/p}$$

where p is the number of model parameters in the general linear model (3.1). Often the D-efficiency can be interpreted as the percentage of observations that can be saved by using the optimal design ξ^* instead of ξ to obtain the same precision and the maximum value of $\mathrm{eff}_D(\xi)$ is 100.

4 Optimal Designs for Linear Paired Comparison Models

As was already pointed out in the present setting we give a brief discussion about how the corresponding standard experimental situation considered in Chapter 3 fits into the paired comparison models introduced in Chapter 2. In that case for the K attributes where $i_k = 1, \ldots, v_K$ was the levels of the k-th attribute (component) of influence from the finite set $\mathcal{I}_k$, each alternative was presented as $\mathbf{i} = (i_1, \ldots, i_K)$ from the set $\mathcal{I} = \mathcal{I}_1 \times \cdots \times \mathcal{I}_K$ as formalized in the general linear model characterized in Section 2.1. Here and, in particular for the situation of paired comparisons, we mention that the experimental setting $\mathbf{x} = (x_1, \ldots, x_K)$ of component x_k for the general linear model described in Section 3.1 consists of ordered pairs $\mathbf{x} = (\mathbf{i}, \mathbf{j}) = ((i_1, \ldots, i_K), (j_1, \ldots, j_K)) \in \mathcal{X} = \mathcal{I} \times \mathcal{I}$ of K-tuples where the component x_k is a pair $(i_k, j_k) \in \mathcal{I}_k \times \mathcal{I}_k$ of levels of the k-th attribute selected from the experimental (design) region $\mathcal{X}$. For the rest of the present work we will mainly focus on the corresponding situation in which the experimental region $\mathcal{X}$ consists of ordered pairs $(\mathbf{i}, \mathbf{j})$.

4.1 Part-Worth Model

As was already pointed out here we note that the parameter vector $\boldsymbol{\beta} = (\alpha_1^{(1)}, \ldots, \alpha_{v_1-1}^{(1)}, \ldots, \alpha_1^{(K)}, \ldots, \alpha_{v_K-1}^{(K)})^\top$ of the part-worth model described in Section 2.2 is minimal (reduced) i.e. the model (2.5) is not over-parameterized. As a consequence, by the usual identifiability condition we obtain $\sum_{i_1=1}^{v_1} \alpha_{i_1}^{(1)} = \cdots = \sum_{i_K=1}^{v_K} \alpha_{i_K}^{(K)} = 0$ for levels $i_k = 1, \ldots, v_K$ and attributes $k = 1, \ldots, K$. In what follows, here we note that instead of considering the individual effects of all the multiple-attributes $k = 1, \ldots, K$, in the next Section 4.2 we will focus on the standard parameterization of only the effects of a single-attribute (component) where $\sum_{i_k=1}^{v_k} \alpha_{i_k}^{(k)} = 0$. It is worthwhile mentioning that the effects of the single-attribute sum up to the effects of the multiple-attributes where $\sum_{k=1}^{K} \alpha_{i_k}^{(k)}$ for $i_k \in \mathcal{I}_k$ in model (2.5). Specifically, by using the single-attribute one can recover the multiple-attributes (e.g., see Graßhoff et al., 2004). It should be noted that for the case of paired comparison only differences of the effects is of interest.

Often, in applications one may only be interested in the corresponding reduced parameter vector $\boldsymbol{\beta}$. As a consequence, the concepts of optimal experimental designs discussed in Chapter 3 which embody the definition of the information matrix as well as the optimality criteria and the Gauss-Markov estimator can therefore, as here, be convieniently applied to the reduced parameter vector $\boldsymbol{\beta}$ of the underlying models.

In the following after the representation of the standard parameterization of the marginal model of the single-attribute becomes first optimal designs for the corresponding single-attribute at general-levels, which is followed by a brief summary of optimal designs for the first-order interaction models. We then introduce appropriate models for the second-order interactions, and provide results for the corresponding optimal designs for paired comparisons of full profiles (paired alternatives) involving multiple-attributes each at two-levels. This results will be generalized to the case when the optimal designs for the paired comparisons of full profiles are characterized by multiple-attributes each at common number of general-levels. Further the results involving multiple-attributes each at two-levels and common number of general-levels will be extended to the situation in which the paired alternatives are characterized by a subset of the multiple-attributes (so called partial profiles) in order to mitigate cognitive burden as frequently encountered in practice. Besides, we will also introduce appropriate models for the third-order interaction model and derive some optimality results for the situation of both full and partial profiles when each attribute has two-levels. Finally, this results will be extended to the situation of common number of general-levels.

4.2 Standard Parameterization with Single-Attribute

In the following we consider reparameterization of the part-worth model for the situation in which the alternatives are characterized by only a single-attribute so-called one-way layout in the analysis of variance model (e.g., see Scheffé, 1952). This parameterization will enhance the generation of D-optimal designs, which will then be used in the subsequent sections as components in the construction of optimal experimental designs for the situation in which the paired alternatives to be evaluated are characterized by many attributes (so-called K-way layout in the analysis of variance model) with main effects as well as interactions.

In the present setting, the effects of each single level i of a single qualitative attribute ($K = 1$) as already specified in Section 4.1 can be written in a general form as v levels where $i = 1, \ldots, v$ with corresponding parameters denoted by α_i. For illustration of the concept of paired comparison by the one-way layout at v levels (see Graßhoff et al., 2004), we adopt the standard parameterization of effects-coding

$$\tilde{Y}_{na}(i) = \mu_n + \alpha_i + \tilde{\epsilon}_{na}, \tag{4.1}$$

$i \in \mathcal{I} = \{1, \ldots, v\}$ where $\mathcal{I}$ is the set for the single-attribute, μ_n denotes the block effects in the n-th presentation and $\tilde{\varepsilon}_{na}$ is the random error, which is assumed to be uncorrelated with constant variance and zero mean. For effects-coding the regression function $\mathbf{f} = \mathbf{f}_1$ in equation (2.6) is appropriately given by $\mathbf{f}_1(i) = \mathbf{e}_i$, $i = 1, \ldots, v-1$ and $\mathbf{f}_1(i) = -\mathbf{1}_{v-1}$ for $i = v$, respectively, where $\mathbf{e}_i$ is the i-th unit vector of length $v-1$ and $\mathbf{1}_m$ denotes a vector of length m with all entries equal to 1. With this parameterization we obtain the reduced parameter vector $\boldsymbol{\beta} = (\alpha_1, \ldots, \alpha_{v-1})^\top$, which satisfies the usual identifiability condition $\sum_{i=1}^{v} \alpha_i = 0$ where the effects of the last level v can be obtained from the reduced parameter vector $\alpha_v = -\sum_{i=1}^{v-1} \alpha_i$ such that

$$\tilde{Y}_{na}(i) = \mu_n + \mathbf{f}_1(i)\boldsymbol{\beta} + \tilde{\epsilon}_{na}. \tag{4.2}$$

Hence, the reduced parameter vector $\boldsymbol{\beta} = (\alpha_1, \ldots, \alpha_{v-1})^\top$ will be used in order to avoid singularity.

For paired comparisons an observation of the effects $\alpha_i - \alpha_j$ of level i compared to level j can be characterized by the response

$$Y_n(i,j) = (\mathbf{f}_1(i) - \mathbf{f}_1(j))\boldsymbol{\beta} + \varepsilon_n = \alpha_i - \alpha_j + \varepsilon_n. \tag{4.3}$$

where $\mathbf{f}_1(i,j) = \mathbf{e}_i - \mathbf{e}_j$, $\mathbf{f}_1(i,v) = \mathbf{e}_i + \mathbf{1}_{v-1}$, $\mathbf{f}_1(v,i) = -\mathbf{f}_1(i,v)$, for $i,j = 1,\ldots,v-1$ and $\mathbf{f}_1(v,v) = \mathbf{0}$.

Now in the following we restrict our attention to the concepts of invariance (e.g., see Schwabe, 1996, Chapter 3) and show that the D-optimality is not affected by the particular choice of the parameterization in the corresponding marginal paired comparison model for the one-way layout based on symmetric properties of the underlying theorem. By Schwabe (1996, p. 27) we define G as a group of transformation (permutations) of the set $\mathcal{I} = \{1,\ldots,v\}$ and g a transformation of the set $\mathcal{I} = \{1,\ldots,v\}$. We define the orbit by $(i,j) \in \mathcal{X} = \mathcal{I} \times \mathcal{I}$ where $\mathcal{X}$ is the corresponding design region for the single-attribute and $\mathbf{f}_1$ is the regression function in (4.3). Moreover, the regression function $\mathbf{f}_1 : \mathcal{I} \times \mathcal{I} \to \mathbb{R}^{v-1}$ is defined as linearly equivariant with respect to the transformation g of the set $\mathcal{I} = \{1,\ldots,v\}$ if there exists a transformation matrix $\mathbf{Q}_g \in \mathbb{R}^{v-1}$, such that

$$\mathbf{f}_1(g(i,j)) = \mathbf{Q}_g \mathbf{f}_1(i,j), \tag{4.4}$$

for all $i,j = 1\ldots,v$. We mention that the corresponding regression function $\mathbf{f}_1 : \mathcal{I}\times\mathcal{I} \to \mathbb{R}^{v-1}$ is also linearly equivariant with respect to the group G of transformation of the set $\mathcal{I}$ if $\mathbf{f}_1$ is linearly equivariant with respect to the transformation g for every $g \in G$. Further by Theorem 3.3 and Lemma 3.4 in Schwabe (1996) we denote by ξ an invariant design which is uniform on the orbits and show that a design $\xi^*_{v,0}$ which assigns equal weights $\frac{1}{v(v-1)}$ to all pairs (i,j) with $i \neq j$ is D-optimal. Analogous results in the following theorem and corollary can be found (e.g., see Schwabe, 1996; Großmann, 2003; Graßhoff et al., 2004).

Theorem 4.1. *In the one-way layout for paired comparisons the design $\xi^*_{v,0}$ on the set $\{(i,j) \in \mathcal{I}\times\mathcal{I} : \ i \neq j\}$ is D-optimal.*

Proof. Let $\mathcal{X}_0 = \{(i,j) : i = j\}$ and $\mathcal{X}_1 = \{(i,j) : i \neq j\}$ be the two orbits with respect to G. Then the invariant design ξ is uniform on the two orbit. Because the information is zero on the orbit $\mathcal{X}_0$, this orbit does not contribute to the information matrix. Hence, the uniform design on the orbit $\mathcal{X}_1$ of all pairs with differing alternatives is optimal.

If $\xi_{v,0}^*$ on the set $\{(i,j) \in \mathcal{I} \times \mathcal{I} : \; i \neq j\}$ is invariant with respect to G then

$$\xi_{v,0}^* = \frac{1}{v(v-1)} \sum_{(i,j): i \neq j} \epsilon_{\{(i,j)\}}.$$

Here $\epsilon_{\{(i,j)\}}$ is the one-point measure in (i,j). The uniform design $\xi_{v,0}^*$ has corresponding information matrix

$$\mathbf{M}(\xi_{v,0}^*) = \frac{2}{v-1}(\mathbf{Id}_{v-1} + \mathbf{1}_{v-1}\mathbf{1}_{v-1}^\top), \tag{4.5}$$

where $\mathbf{Id}_m$ denotes the m-dimensional identity matrix. □

We note that the uniform design $\xi_{v,0}^*$ can be realized with a fixed sample size $v(v-1)$ which can be further reduced to a sample size $v(v-1)/2$ by considering experiments to those comparisons where $i < j$, which by interchanging the internal order does not affect the corresponding information matrix $\mathbf{M}(i,j) = \mathbf{M}(j,i)$. We denote the corresponding uniform design by $\xi_{v,1}^*$.

Corollary 4.1. *In the one-way layout for paired comparisons the design $\xi_{v,1}^*$ which is uniform on the set $\{(i,j) \in \mathcal{I} \times \mathcal{I} : \; i < j\}$ is D-optimal.*

Proof. The information matrix of the uniform design $\xi_{v,1}^*$ coincides with that of the D-optimal design $\xi_{v,0}^*$

$$\begin{aligned}
\mathbf{M}(\xi_{v,1}^*) &= \frac{2}{v(v-1)} \sum_{(i,j): i<j} \mathbf{f}_1(i,j)\mathbf{f}_1(i,j)^\top \\
&= \frac{1}{v(v-1)} \sum_{(i,j): i \neq j} \mathbf{f}_1(i,j)\mathbf{f}_1(i,j)^\top = \mathbf{M}(\xi_{v,0}^*).
\end{aligned}$$

Hence, $\xi_{v,1}^*$ is also D-optimal and has smaller sample size $v(v-1)/2$. □

Remark 4.1. *The corresponding information matrix* $\mathbf{M}(\xi^*_{v,0})$ *has an inverse of the form*

$$\mathbf{M}(\xi^*_{v,0})^{-1} = \frac{v-1}{2}(\mathbf{Id}_{v-1} - \frac{1}{v}\mathbf{1}_{v-1}\mathbf{1}_{v-1}^\top).$$

Further we note that by Lemma 2 *of Graßhoff et al. (2003) we obtain* $\mathbf{f}_1(i)^\top \mathbf{M}(\xi^*_{v,0})^{-1}\mathbf{f}_1(i) = \frac{(v-1)^2}{2v}$ *and* $\mathbf{f}_1(i)^\top \mathbf{M}(\xi^*_{v,0})^{-1}\mathbf{f}_1(j) = -\frac{v-1}{2v}$ *for* $i \neq j$*. Hence, for the variance function we obtain*

$$\begin{aligned}
V((i,j),\xi^*_{v,0}) &= (\mathbf{f}_1(i) - \mathbf{f}_1(j))^\top \mathbf{M}(\xi^*_{v,0})^{-1}(\mathbf{f}_1(i) - \mathbf{f}_1(j)) \\
&= \mathbf{f}_1(i)^\top \mathbf{M}(\xi^*_{v,0})^{-1}\mathbf{f}_1(i) + \mathbf{f}_1(j)^\top \mathbf{M}^{-1}\mathbf{f}_1(j) - \mathbf{f}_1(i)^\top \mathbf{M}^{-1}\mathbf{f}_1(j) - \mathbf{f}_1(j)^\top \mathbf{M}^{-1}\mathbf{f}_1(i) \\
&= 2\mathbf{f}_1(i)^\top \mathbf{M}(\xi^*_{v,0})^{-1}\mathbf{f}_1(i) - 2\mathbf{f}_1(i)^\top \mathbf{M}(\xi^*_{v,0})^{-1}\mathbf{f}_1(j) \\
&= v - 1, \qquad (4.6)
\end{aligned}$$

for $i \neq j$*, while* $V((i,i),\xi^*_{v,0}) = 0$*, which establishes the D-optimality of the uniform design* $\xi^*_{v,0}$ *in view of the celebrated Kiefer and Wolfowitz (1960) equivalence theorem.*

5 Optimal Designs for First-Order Interactions Models

As a motivation for the present work in this chapter we give a brief summary of some results in the design of paired comparison experiments for main effects and first-order interaction models. In particular, Graßhoff et al. (2004) and Graßhoff et al. (2003), respectively, considered D-optimal designs for paired comparisons in the presence of the the main effects and the first-order interaction models for the situation when the paired alternatives are described by analysis of variance model as considered by Scheffé (1952). They constructed designs for the case when each attribute has common number of general levels, and for which either full or partial profiles are presented simultaneously. In the particular case of two-level attributes the corresponding results for which only partial profiles are presented simultaneously can be found in Schwabe et al. (2003). The designs found can also be constructed by the method of van Berkum (1987b) for the case when full profiles are presented simultaneously. These results will be considered in detail in the subsequent chapters to derive D-optimal designs for paired comparisons in the presence of both the second and third-order interaction models.

In applications one may be interested in the utility estimates of both the main effects and interactions between the levels $i_k = 1, \ldots, v$ of the attributes $k = 1, \ldots, K$. However, certain level combinations of the attributes may result in higher or lower utilities. In this setting for alternatives in a choice set of size two where both $\mathbf{i} = (i_1, \ldots, i_K)$ and $\mathbf{j} = (j_1, \ldots, j_K)$ denote the first alternative and the second alternative, respectively, which are both elements of the set $\mathcal{I} = \mathcal{I}_1 \times \cdots \times \mathcal{I}_K = \{1, \ldots, v\}^K$ and where the alternatives $\mathbf{i}$ and $\mathbf{j}$ are ordered pairs which are chosen from the design region $\mathcal{X} = \mathcal{I} \times \mathcal{I}$, Graßhoff et al. (2003) obtained optimal designs in the first-order interactions setup by considering analogous main effects and first-order interactions model of the form

$$\tilde{Y}_{na}(\mathbf{i}) = \mu_n + \sum_{k=1}^{K} \alpha_{i_k}^{(k)} + \sum_{k<\ell} \alpha_{i_k i_\ell}^{(k\ell)} + \tilde{\varepsilon}_{na}, \tag{5.1}$$

with direct response (utility) $\tilde{Y}_{na}(\mathbf{i})$ where $\alpha_{i_k}^{(k)}$ is the main effects of the k-th attribute when the corresponding level is $i_k = 1, \ldots, v$ for $k = 1, \ldots, K$ in total and $\alpha_{i_k i_\ell}^{(k\ell)}$ is the first-order interaction effects of the k-th and ℓ-th attribute when the corresponding levels are $i_k = 1, \ldots, v$ and $i_\ell = 1, \ldots, v$, respectively. Hence, by the common identifiability conditions of effects-coding the following equalities hold:

$$\alpha_{i_k}^{(k)} = \beta_{i_k}^{(k)} \text{ for } i_k = 1, \ldots, v-1 \text{ and } \alpha_v^{(k)} = -\sum_{i_k=1}^{v-1} \beta_{i_k},$$

$$\alpha_{i_k i_\ell}^{(k\ell)} = \beta_{i_k i_\ell}^{(k\ell)} \text{ for } i_k, i_\ell = 1, \ldots, v-1, \; \alpha_{i_k v}^{(k\ell)} = -\sum_{i_\ell=1}^{v-1} \beta_{i_k i_\ell}^{(k\ell)}, i_k = 1, \ldots, v-1,$$

$$\alpha_{v i_\ell}^{(k\ell)} = -\sum_{i_k=1}^{v-1} \beta_{i_k i_\ell}^{(k\ell)}, i_\ell = 1, \ldots, v-1 \text{ and } \alpha_{vv}^{(k\ell)} = \sum_{i_k=1}^{v-1} \sum_{i_\ell=1}^{v-1} \beta_{i_k i_\ell}^{(k\ell)}.$$

The parameters for the main effects and the first-order interactions, respectively, can be summarized as follows

$$\boldsymbol{\beta}_k = (\beta_{i_k}^{(k)})_{i_k=1,\ldots,v-1} \text{ and } \boldsymbol{\beta}_{k\ell} = (\beta_{i_k i_\ell}^{(k\ell)})_{i_k=1,\ldots,v-1,\; i_\ell=1,\ldots,v-1}, \tag{5.2}$$

where e.g. $\boldsymbol{\beta}_k$ describes the main effects of the k-th attribute and $\boldsymbol{\beta}_{k\ell}$ describes the effect of the first-order interaction of the k-th and ℓ-th attribute. Hence, the vector of parameters of dimension $p = K(v-1) + \binom{K}{2}(v-1)^2$ can be written as $\boldsymbol{\beta} = ((\boldsymbol{\beta}_k)_{k=1,\ldots,K}^\top, (\boldsymbol{\beta}_{k\ell})_{k<\ell}^\top)^\top$. Further with the above notation the corresponding model (5.1) can be reformulated as

$$\tilde{Y}_{na}(\mathbf{i}) = \mu_n + \sum_{k=1}^{K} \mathbf{f}_1(i_k)^\top \boldsymbol{\beta}_k + \sum_{k<\ell} (\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell))^\top \boldsymbol{\beta}_{k\ell} + \tilde{\varepsilon}_{na}, \tag{5.3}$$

where $\otimes$ denotes the Kronecker product of vectors or matrices, respectively, which results in the vector $\mathbf{f}(\mathbf{i})$ having corresponding regression functions $\mathbf{f}_k = \mathbf{f}_1$ (see Section 4.2)

$$\mathbf{f}(\mathbf{i}) = (\mathbf{f}_1(i_1)^\top, \ldots, \mathbf{f}_1(i_K)^\top, \mathbf{f}_1(i_1)^\top \otimes \mathbf{f}_1(i_2)^\top, \ldots, \mathbf{f}_1(i_{K-1}) \otimes \mathbf{f}_1(i_K)^\top)^\top \tag{5.4}$$

of dimension p. Here, the first K components $\mathbf{f}_1(i_1), \ldots, \mathbf{f}_1(i_K)$ of $\mathbf{f}(\mathbf{i})$ are associated with the main effects and have $p_1 = K(v-1)$ and the remaining components $\mathbf{f}_1(i_1) \otimes \mathbf{f}_1(i_2), \ldots, \mathbf{f}_1(i_{K-1}) \otimes \mathbf{f}_1(i_K)$ of $\mathbf{f}(\mathbf{i})$ are associated with the first-order interactions and have $p_2 = (1/2)K(K-1)(v-1)^2$.

The corresponding paired comparison model is given by

$$Y_n(\mathbf{i},\mathbf{j}) = \sum_{k=1}^{K}(\mathbf{f}_1(i_k) - \mathbf{f}_1(j_k))^\top \boldsymbol{\beta}_k + \sum_{k<\ell}((\mathbf{f}_1(i_k)\otimes \mathbf{f}_1(i_\ell)) - (\mathbf{f}_1(j_k)\otimes \mathbf{f}_1(j_\ell)))^\top \boldsymbol{\beta}_{k\ell} + \varepsilon_n. \tag{5.5}$$

Due to the cognitive ability or the limited information processing capacity in applications when the alternatives involve too many attributes respondents get overloaded by the complexity of the choice task and become wear-out. The choice sets are then answered anyhow which can destroy the quality of the data as well as the estimated model parameters because respondents' decision might seem contradictory with their actual preferences.

To overcome respondent wear-out or fatigue, only partial-profiles are presented within a single paired comparison. Specifically, every choice set consists of alternatives which are described by a predefined maximal number of attributes S (so-called profile strength) with potentially different levels, while the remaining $K - S$ attributes are not shown or held constant. Here only those attributes that constitute the profile strength are shown to responsents within a single paired comparison.

For a partial profile a direct observation may be described by model (5.3) when summation is taken only over those S attributes contained in the describing subset. This requires that the profile strength S must be, at least, two in order to capture the first-order interactions. To facilitate notation we introduce an additional level $i_k = 0$ for each attribute indicating that the corresponding attribute is not present in the partial profile, and the corresponding regression functions are given by $\mathbf{f}_k(0) = \mathbf{f}_1(0) = \mathbf{0}$. With this convention a direct observation can be described by (5.3) even for a partial profile $\mathbf{i}$ from the set

$$\begin{aligned} \mathcal{I}^{(S)} = &\{\mathbf{i};\ i_k \in \{1,\ldots,v\} \text{ for } S \text{ components and} \\ &i_k = 0 \text{ for } K - S \text{ components}\}, \end{aligned} \tag{5.6}$$

of alternatives with profile strength S. In particular, $\mathcal{I}^{(K)} = \mathcal{I}^{(S)}$ in the case of full profiles ($S = K$). For general profile strength S the vector $\mathbf{f}$ of regression functions in (5.4), the paired comparison model (5.5) and the interpretation of the corresponding parameter vector $\boldsymbol{\beta}$ remain unchanged. Here we note that the comparison depth d as in the work of Graßhoff et al. (2003) describes the number of attributes in which the two alternatives in the choice sets differ satisfying the inequalities $1 \leq d \leq S \leq K$.

Hence, the paired comparison model (5.5) having corresponding design region $\mathcal{X}$ is thus restricted to those paired alternatives for which exactly S attributes are presented

$$\mathcal{X}^{(S)} = \{(\mathbf{i}, \mathbf{j});\ i_k, j_k \in \{1, \dots, v\} \text{ for } S \text{ components and } \\ i_k = j_k = 0 \text{ for exactly } K - S \text{ components}\}. \tag{5.7}$$

The design region $\mathcal{X}^{(S)}$ can be partitioned into disjoint sets such that the pairs in each set differ only in a fixed number d of the attributes. Specifically, for a comparison depth $d = 0, \dots, S$, let

$$\mathcal{X}_d^{(S)} = \{(\mathbf{i}, \mathbf{j}) \in \mathcal{X}^{(S)} : |\{k : i_k \neq j_k\}| = d\}, \tag{5.8}$$

be the set of all pairs of alternatives which differ in exactly d attributes. These sets constitute the orbits with respect to permutations. The D-criterion is invariant with respect to those permutations which induce a linear reparameterization. Specifically, we mention that the regression functions (5.4) extended to the design region $\mathcal{X}^{(S)}$ are still linearly equivariant, which means that relabeling does not affect D-optimality (and D-optimality of invariant subvectors). Hence, it is sufficient to look for optimality in the class of invariant designs which are uniform on the orbits of fixed comparison depth $d \leq S$.

Accordingly, for $N_d = \binom{K}{S}\binom{S}{d} v^S (v-1)^d$ total number of different pairs in $\mathcal{X}_d^{(S)}$ which vary in exactly d attributes and the uniform approximate design ξ_d which assigns equal weight $\xi_d(\mathbf{i}, \mathbf{j}) = 1/N_d$ to each pair in $\mathcal{X}_d^{(S)}$, Graßhoff et al. (2003) obtained the information matrix of the uniform design ξ_d on the design region $\mathcal{X}_d^{(S)}$ of fixed comparison depth d of the form

$$\mathbf{M}(\xi_d) = \begin{pmatrix} h_1(d)\mathbf{Id}_{p_1} \otimes \mathbf{M} & \mathbf{0} \\ \mathbf{0} & h_2(d)\mathbf{Id}_{p_2} \otimes \mathbf{M} \otimes \mathbf{M} \end{pmatrix},$$

where $h_1(d) = \frac{d}{K}$, $h_2(d) = \frac{d}{2vK(K-1)}(2Sv - 2S - dv - v + 2)$ and $\mathbf{M} = \frac{2}{v-1}(\mathbf{Id}_{v-1} + \mathbf{1}_{v-1}\mathbf{1}_{v-1}^\top)$ is the information matrix of the corresponding (one-way layout in (4.5)). Moreover, the information matrix for a general invariant design $\xi = \sum_{d=1}^{S} w_d \xi_d$ has a blocked diagonal information matrix of the form

$$\mathbf{M}(\xi) = \begin{pmatrix} h_1(\xi)\mathbf{Id}_{p_1} \otimes \mathbf{M} & \mathbf{0} \\ \mathbf{0} & h_2(\xi)\mathbf{Id}_{p_2} \otimes \mathbf{M} \otimes \mathbf{M} \end{pmatrix},$$

where $h_q(\xi) = \sum_{d=1}^{S} w_d h_q(d)$, $q = 1, 2$.

Here it is worthwhile mentioning that a single comparison depth d may be sufficient for non-singularity of the corresponding information matrix $\mathbf{M}(\xi_d)$, i.e. for the identifiability of all parameters for both the main effects and the first-order interactions. This can be easily seen by observing $h_q(1) > 0$, $q = 1, 2$, for $d = 1$. But this is not true for all comparison depths as for example $h_2(S) = 0$.

The corresponding invariant design ξ has a variance function of the form $V((\mathbf{i}, \mathbf{j}), \xi) = (\mathbf{f}(\mathbf{i}) - \mathbf{f}(\mathbf{j}))^\top \mathbf{M}(\xi)^{-1}(\mathbf{f}(\mathbf{i}) - \mathbf{f}(\mathbf{j}))$ where the value of the variance function for the invariant design ξ evaluated at comparison depth d is denoted as $V(d, \xi)$ where $V(d, \xi) = V((\mathbf{i}, \mathbf{j}), \xi)$ on $\mathcal{X}_d^{(S)}$:

$$V(d, \xi) = d(v-1)\left(\tfrac{1}{h_1(\xi)} + \tfrac{v-1}{4vh_2(\xi)}(2Sv - 2S - dv - v + 2)\right).$$

Accordingly, on a single comparison depth the representation of the variance function $V(d, \xi)$ simplifies

$$V(d, \xi_{d'}) = \frac{d}{d'}\left(p_1 + p_2 \tfrac{2Sv-2S-dv-v+2}{2Sv-2S-d'v-v+2}\right).$$

Note that for $d = d'$, $V(d, \xi_d) = p_1 + p_2 = p$ which shows the D-optimality of ξ_d on $\mathcal{X}_d^{(S)}$ in view of the Kiefer-Wolfowitz equivalence theorem.

It is worthwhile mentioning that the corresponding results for the first-order interactions can be used in a more general context for both the second and third-order interactions as will be seen in the underlying chapters.

6 Optimal Designs for Second-Order Interactions Two Level Models

In real life situations one may be interested in the utility estimates of both the main effects and interactions between the levels of the attributes. At this point, we mention that certain level combinations of the attributes may result in higher or lower utilities. In this setting optimal designs have been derived (see van Berkum, 1987b; Graßhoff et al., 2003, amongst others) in the first-order interactions setup as mentioned in Chapter 5. Here we focused on the second-order interactions.

For what follows, in the following unlike Section 4.2 where only a single-attribute was considered. Here we take into consideration multiple-attributes $k = 1, \ldots, K$ having levels $v_k = 2$ each that are assumed to derive the preferences for the alternatives in a paired comparison experiment. In paired comparison experiments the alternatives are represented by combinations of attribute levels. For alternatives in a choice set of size two, we denote by $\mathbf{i} = (i_1, \ldots, i_K)$ the first alternative where i_k is the component of the k-th attribute and denote the second alternative by $\mathbf{j} = (j_1, \ldots, j_K)$ which are both elements of the set $\mathcal{I} = \mathcal{I}_1 \times \cdots \times \mathcal{I}_K = \{1, -1\}^K$ where the numbers 1 and -1 represent effects-coding of the first and second level of each attribute, respectively. Specifically, the alternatives $\mathbf{i}$ and $\mathbf{j}$ are ordered pairs which are chosen from the design region $\mathcal{X} = \mathcal{I} \times \mathcal{I}$. Thus, for each attribute (component) k at levels $v_k = 2$ the corresponding marginal model coincides with that of the single-attribute with regression functions $f_k = f_1$ corresponding to the (components) of the vector of regression functions $\mathbf{f}_1$.

More formally, for the case of direct response (utility) $\tilde{Y}_{na}(\mathbf{i})$, we consider the second-order interaction model

$$\tilde{Y}_{na}(\mathbf{i}) = \mu_n + \sum_{k=1}^{K} \alpha_{i_k}^{(k)} + \sum_{k<\ell} \alpha_{i_k i_\ell}^{(k\ell)} + \sum_{k<\ell<m} \alpha_{i_k i_\ell i_m}^{(k\ell m)} + \tilde{\varepsilon}_{na}, \tag{6.1}$$

where $\alpha_{i_k}^{(k)}$ is the main effects of the k-th attribute when the corresponding level is i_k, $\alpha_{i_k i_\ell}^{(k\ell)}$ is the first-order interaction effects of the k-th and ℓ-th attriubute when the corresponding levels are i_k and i_ℓ, respectively, and $\alpha_{i_k i_\ell i_m}^{(k\ell m)}$ is the second-order interaction effects of the k-th, ℓ-th and m-th attribute when the corresponding levels are i_k, i_ℓ and i_m, respectively.

Moreover, by the common identifiability conditions of effects-coding described in Section 4.2 the following equalities hold:
For the reduced parameter $\alpha_1^{(k)} = \beta^{(k)}$ the last component is obtained by $\alpha_2^{(k)} = -\beta^{(k)}$ for the main effects. Similary, for the reduced parameter $\alpha_{11}^{(k\ell)} = \alpha_{22}^{(k\ell)} = \beta^{(k\ell)}$ the last component is obtained by $\alpha_{12}^{(k\ell)} = \alpha_{21}^{(k\ell)} = -\beta^{(k\ell)}$ for the first-order interactions, and, finally, for the reduced parameter $\alpha_{111}^{(k\ell m)} = \alpha_{122}^{(k\ell m)} = \alpha_{212}^{(k\ell m)} = \alpha_{221}^{(k\ell m)} = \beta^{(k\ell m)}$ the last

component is obtained by $\alpha_{112}^{(k\ell m)} = \alpha_{121}^{(k\ell m)} = \alpha_{211}^{(k\ell m)} = \alpha_{222}^{(k\ell m)} = -\beta^{(k\ell m)}$ for the second-order interactions where e. g., $\beta^{(k\ell m)}$ describes the effect of the second-order interaction of the k-th, ℓ-th and m-th attribute. The parameters for the main effects, the first-order interactions and the second-order interactions are given by $\beta_k = \beta^{(k)}$, $\beta_{k\ell} = \beta^{(k\ell)}$ and $\beta_{k\ell m} = \beta^{(k\ell m)}$, respectively.

With the above notation the corresponding model can be reformulated as

$$\tilde{Y}_{na}(\mathbf{i}) = \mu_n + \sum_{k=1}^{K} \beta_k i_k + \sum_{k<\ell} \beta_{k\ell} i_k i_\ell + \sum_{k<\ell<m} \beta_{k\ell m} i_k i_\ell i_m + \tilde{\varepsilon}_{na}, \tag{6.2}$$

where β_k denotes the main effect of the k-th attribute, $\beta_{k\ell}$ is the first-order interaction of the k-th and ℓ-th attribute, and $\beta_{k\ell m}$ is the second-order interaction of the k-th, ℓ-th and m-th attribute. The vectors $(\beta_k)_{1\le k\le K}$ of main effects, $(\beta_{k\ell})_{1\le k<\ell\le K}$ of first-order interactions and $(\beta_{k\ell m})_{1\le k<\ell<m\le K}$ of second-order interactions have dimensions $p_1 = K$, $p_2 = K(K-1)/2$ and $p_3 = K(K-1)(K-2)/6$, respectively. Hence, the complete parameter vector $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K, (\beta_{k\ell})_{k<\ell}^\top, (\beta_{k\ell m})_{k<\ell<m}^\top)^\top$ has dimension $p = p_1+p_2+p_3$. The corresponding p-dimensional vector $\mathbf{f}$ of regression functions is given by

$$\mathbf{f}(\mathbf{i}) = (i_1, \ldots, i_K, (i_k i_\ell)_{k<\ell}^\top, (i_k i_\ell i_m)_{k<\ell<m}^\top)^\top. \tag{6.3}$$

Also here in $\mathbf{f}(\mathbf{i})$, the first $p_1 = K$ entries $i_1, \ldots, i_K$ are associated with the main effects, the second set of p_2 entries $i_k i_\ell$, $1 \le k < \ell \le K$, are associated with the first-order interactions, and the remaining p_3 entries $i_k i_\ell i_m$, $1 \le k < \ell < m \le K$, are associated with the second-order interactions.

The corresponding paired comparison model is given by

$$Y_n(\mathbf{i}, \mathbf{j}) = \sum_{k=1}^{K} (i_k - j_k)\beta_k + \sum_{k<\ell} (i_k i_\ell - j_k j_\ell)\beta_{k\ell} + \sum_{k<\ell<m} (i_k i_\ell i_m - j_k j_\ell j_m)\beta_{k\ell m} + \varepsilon_n. \tag{6.4}$$

6.1 Designs for Full Profiles

In this section we employ the concept of invariance characterized in Section 4.2, and present D-optimal designs for the second-order interaction paired comparison model (6.4) with corresponding regression functions $\mathbf{f}(\mathbf{i})$ in (6.3). Here we note that the subsequent results can be found in Nyarko and Schwabe (2019) and the idea of considering comparison depths relies on Graßhoff et al. (2003). Define by d the comparison depth which describes the number of attributes in which the paired alternatives $(\mathbf{i}, \mathbf{j})$ differ. Thus the design region $\mathcal{X}$ can be partitioned into disjoint sets such that the pairs in each set differ only in a fixed number d of the attributes. More precisely, for a comparison depth $d = 0, \ldots, K$, let

$$\mathcal{X}_d = \{(\mathbf{i}, \mathbf{j}) \in \mathcal{X} : |\{k : i_k \neq j_k\}| = d\} \tag{6.5}$$

be the set of all pairs of alternatives which differ in exactly d attributes. These sets constitute the orbits with respect to permutations of both the levels $i_k = 1, -1$ within the attributes as well as among attributes $k = 1, \ldots, K$, themselves. Here and for the particular case of effects-coding we mention that the D-criterion is invariant with respect to those permutations (see Schwabe, 1996, p. 17). As a result, it is sufficient to look for optimality in the class of invariant designs which are uniform on the orbits of fixed comparison depth.

In the following Lemmas 6.1 and 6.2 we present the information matrices of the aforementioned invariant designs by denoting $N_d = 2^K \binom{K}{d}$ as the number of different pairs in $\mathcal{X}_d$ which vary in exactly d attributes and by ξ_d the uniform approximate design which assigns equal weights $\xi_d(\mathbf{i}, \mathbf{j}) = 1/N_d$ to each pair $(\mathbf{i}, \mathbf{j})$ in $\mathcal{X}_d$ and weight zero to all remaining pairs in $\mathcal{X}$.

Lemma 6.1. *Let d be a fixed comparison depth. The uniform design ξ_d on the set $\mathcal{X}_d$ of comparison depth d has a diagonal information matrix*

$$\mathbf{M}(\xi_d) = \begin{pmatrix} h_1(d)\mathbf{Id}_K & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & h_2(d)\mathbf{Id}_{\binom{K}{2}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & h_3(d)\mathbf{Id}_{\binom{K}{3}} \end{pmatrix},$$

where

$$h_1(d) = \frac{4d}{K}, h_2(d) = \frac{8d(K-d)}{K(K-1)} \text{ and } h_3(d) = \frac{4d(3K^2 - 6dK + 4d^2 - 3K + 2)}{K(K-1)(K-2)}.$$

Proof. First we note that for the active levels $i, j = -1, 1$; $i^2 = 1$, $ij = -1$ and $(i-j)^2 = 4$ for $i \neq j$. Given a fixed comparison depth d we obtain for the regression functions $f_k = f_1$ associated with the k-th main effect

$$\sum_{(\mathbf{i},\mathbf{j}) \in \mathcal{X}_d} (i_k - j_k)^2 = \binom{K-1}{d-1} 2^{K+2}$$

because there are $\binom{K-1}{d-1} 2^K$ pairs in $\mathcal{X}_d$ for which i_k and j_k differ. Since the number N_d of paired comparisons in $\mathcal{X}_d$ equals $N_d = \binom{K}{d} 2^K$, the corresponding diagonal entries $h_1(d)$ in the information matrix are given by

$$h_1(d) = \frac{1}{N_d} \sum_{(\mathbf{i},\mathbf{j}) \in \mathcal{X}_d} (i_k - j_k)^2 = \frac{4d}{K}. \tag{6.6}$$

For first-order interactions, we consider attributes k and ℓ, say, and distinguish between pairs in which both attributes are distinct and pairs in which only one of these attributes has distinct levels in the alternatives while the same level is presented in both alternatives for the other attribute.

In the case $i_k \neq j_k$ and $i_\ell \neq j_\ell$ we have $i_k i_\ell = j_k j_\ell$, while for $i_\ell = j_\ell$ we get $i_k i_\ell = -j_k j_\ell$. Hence

$$(i_k i_\ell - j_k j_\ell)^2 = 0 \quad \text{for} \quad i_k \neq j_k \quad \text{and} \quad i_\ell \neq j_\ell \tag{6.7}$$

and

$$(i_k i_\ell - j_k j_\ell)^2 = 4 \quad \text{for} \quad i_k \neq j_k \quad \text{and} \quad i_\ell = j_\ell, \tag{6.8}$$

respectively, where (in the latter case) the roles of the attributes k and ℓ may be interchanged.

For given attributes k and ℓ the pairs with distinct levels in both attributes occur $\binom{K-2}{d-2} 2^K$ times in $\mathcal{X}_d$ while those which differ only in one attribute occur $\binom{2}{1}\binom{K-2}{d-1} 2^K$ times. As a result, for the first-order interactions the diagonal elements $h_2(d)$ in the information matrix are given by

$$h_2(d) = \frac{1}{N_d} 2 \binom{K-2}{d-1} 2^{K+2} = \frac{8d(K-d)}{K(K-1)}. \tag{6.9}$$

Accordingly, for second-order interactions, we consider attributes k, ℓ and m, say, and distinguish between pairs in which all three attributes are distinct, pairs in which two of these attributes k and ℓ, say, have distinct levels in the alternatives while the same level is presented in both alternatives for the remaining attribute and, finally, pairs in which only one of the attributes, say, k has distinct levels in the alternatives while the same level is presented in both alternatives for the two remaining attributes. Then $i_k i_\ell i_m = -j_k j_\ell j_m$ in the first and third case, while $i_k i_\ell i_m = j_k j_\ell j_m$ in the second case. Hence,

$$(i_k i_\ell i_m - j_k j_\ell j_m)^2 = 4 \quad \text{for} \quad i_k \neq j_k, \; i_\ell \neq j_\ell \quad \text{and} \quad i_m \neq j_m, \tag{6.10}$$

$$(i_k i_\ell i_m - j_k j_\ell j_m)^2 = 0 \quad \text{for} \quad i_k \neq j_k, \; i_\ell \neq j_\ell \quad \text{and} \quad i_m = j_m \tag{6.11}$$

and

$$(i_k i_\ell i_m - j_k j_\ell j_m)^2 = 4 \quad \text{for} \quad i_k \neq j_k, \; i_\ell = j_\ell \quad \text{and} \quad i_m = j_m, \tag{6.12}$$

respectively, where again the roles of the attributes k, ℓ and m may be interchanged.

For given attributes k, ℓ and m the pairs with distinct levels in the three attributes occur $\binom{K-3}{d-3}2^K$ times in $\mathcal{X}_d$ while those which differ in two attributes occur $\binom{3}{2}\binom{K-3}{d-2}2^K$ times in $\mathcal{X}_d$, and those which differ only in one attribute occur $\binom{3}{1}\binom{K-3}{d-1}2^K$ times. As a result, for the second-order interactions the diagonal elements $h_3(d)$ in the information matrix are given by

$$\begin{aligned} h_3(d) &= \frac{1}{N_d}\left(\binom{K-3}{d-3}2^{K+2} + 3\binom{K-3}{d-1}2^{K+2}\right) \\ &= \frac{4d((d-1)(d-2)+3(K-d)(K-d-1))}{K(K-1)(K-2)} \\ &= \frac{4d(3K^2-6dK+4d^2-3K+2)}{K(K-1)(K-2)}. \end{aligned} \tag{6.13}$$

Finally, it can be noted that all off-diagonal entries in the information matrix vanish because the terms in the corresponding sums add up to zero due to the effects-type coding. □

For the particular case of first-order interactions the corresponding results of the entries $h_1(d)$ and $h_2(d)$ can be found (e.g., see van Berkum, 1987b; Graßhoff et al., 2003). It is worthwhile mentioning that for the comparison depth $d = 0$ we obtain $h_q(0) = 0$ and, hence $\mathbf{M}(\xi_0) = 0$.

Further we mention that general invariant designs ξ can be written as a convex combination $\xi = \sum_{d=1}^{K} w_d \xi_d$ of uniform designs on the comparison depth d with corresponding weights $w_d \geq 0$, $\sum_{d=1}^{K} w_d = 1$. Hence, also every invariant design has diagonal information matrix.

Lemma 6.2. *Let ξ be an invariant design on $\mathcal{X}$, i. e. $\xi = \sum_{d=1}^{K} w_d \xi_d$, then ξ has diagonal information matrix*

$$\mathbf{M}(\xi) = \begin{pmatrix} h_1(\xi)\mathbf{Id}_K & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & h_2(\xi)\mathbf{Id}_{\binom{K}{2}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & h_3(\xi)\mathbf{Id}_{\binom{K}{3}} \end{pmatrix},$$

where $h_q(\xi) = \sum_{d=1}^{K} w_d h_q(d)$, $q = 1, 2, 3$.

In the following Theorems 6.1, 6.2 and 6.3 we present optimal designs for the main effects, the first-order interaction and the second-order interaction terms separately which have corresponding entries $h_1(d)$, $h_2(d)$ and $h_3(d)$, respectively, in the information matrix. The resulting designs may also optimize every invariant criterion if the full parameter vector for the main effects, the first-order interaction and the second-order interaction is considered, satisfying the aforementioned usual identifiability conditions. With this in mind we note that also the D-optimality for the corresponding subset of the parameter vector $\boldsymbol{\beta} = ((\beta_k)_{1\leq k\leq K}^\top, (\beta_{k\ell})_{k<\ell}^\top, (\beta_{k\ell m})_{k<\ell<m}^\top)^\top$ considered separately in Theorems 6.1, 6.2 and 6.3 is invariant. To start with, we mention that Theorems 6.1 and 6.2 paraphrase theorems given in Graßhoff et al. (2003) for first-order interaction models and translate them to the present setting of second-order interaction models.

Theorem 6.1. *Let $d_1^* = K$. Then the uniform design $\xi_{d_1^*} = \xi_K$ on the largest possible comparison depth K is D-optimal for the main effects $(\beta_k)_{1\leq k\leq K}^\top$.*

This means that for the main effects only those pairs of alternatives should be used which differ in all attributes.

Theorem 6.2. *Let $d_2^* = K/2$ for K even and $d_2^* = (K-1)/2$ or $d_2^* = (K+1)/2$ for K odd, respectively. Then the uniform design $\xi_{d_2^*}$ is D-optimal for the first-order interaction effects $(\beta_{k\ell})_{k<\ell}^\top$.*

This means that for the first-order interactions only those pairs of alternatives should be used which differ in about half of the attributes.

Theorem 6.3. (a) *Let $d_3^* = 1$ or $d_3^* = 3$ for $K = 3$. Then the uniform design $\xi_{d_3^*}$ is D-optimal for the second-order interaction effects $(\beta_{k\ell m})_{k<\ell<m}^\top$.*
(b) *Let $d_3^* = K$ for $K \geq 4$. Then the uniform design $\xi_{d_3^*}$ is D-optimal for the second-order interaction effects $(\beta_{k\ell m})_{k<\ell<m}^\top$.*

This means that also for the second-order interactions only those pairs of alternatives should be used which differ in all attributes.

Proof. (a) Optimality is achieved when h_3 is maximized. For $K = 3$ we get $h_3(1) = h_3(3) = 4$ and $h_3(0) = h_3(2) = 0$ which establishes the result in this case.

(b) For $K \geq 4$ note that the function h_3 is a cubic polynomial in the comparison depth d with positive leading coefficient. Extended to a function on the real line h_3 is point symmetric with respect to $(K/2, h_3(K/2))$ and attains its local maximum and local minimum at $d_{3,\max} = K/2 - \sqrt{9K-6}/6$ and $d_{3,\min} = K/2 + \sqrt{9K-6}/6$, respectively. Now, the numerator of $h_3(d_{3,\min})$ is proportional to $d_{3,\min}^2(3K - 4d_{3,\min})$. Inserting the solution for $d_{3,\min}$ into the last factor yields $3K - 4d_{3,\min} = K - 2\sqrt{K - 2/3}$ which is equal to 0.349 for $K = 4$ and increasing in $K \geq 3$. Hence, $h_3(d_{3,\min}) > 0$ for $K \leq 4$ and, by symmetry, $h_3(d_{3,\max}) < h_3(K)$ which proves the result. $\square$

In the following Theorem 6.4 and Corollary 6.1 we present the variance function for the corresponding designs. The variance function is defined as $V((\mathbf{i},\mathbf{j}),\xi) = (\mathbf{f}(\mathbf{i}) - \mathbf{f}(\mathbf{j}))^\top \mathbf{M}(\xi)^{-1}(\mathbf{f}(\mathbf{i}) - \mathbf{f}(\mathbf{j}))$, which plays an important role for the D-criterion. As already defined, according to the Kiefer and Wolfowitz (1960) equivalence theorem a design ξ^* is D-optimal if the associated variance function is bounded by the number of model parameters, $V((\mathbf{i},\mathbf{j}),\xi^*) \leq p$. Now, for the invariant design ξ the variance function $V((\mathbf{i},\mathbf{j}),\xi)$ is also invariant with respect to permutations of levels and attributes and is, hence, constant on the orbits $\mathcal{X}_d$ of fixed comparison depth d. As a result, the value of the variance function for the invariant design ξ evaluated at comparison depth d may be denoted by $V(d,\xi)$, say, where $V(d,\xi) = V((\mathbf{i},\mathbf{j}),\xi)$ on $\mathcal{X}_d$.

Theorem 6.4. *For every invariant design ξ the variance function $V(d,\xi)$ is given by*

$$V(d,\xi) = 4d\left(\frac{1}{h_1(\xi)} + \frac{K-d}{h_2(\xi)} + \frac{3K^2 - 6dK + 4d^2 - 3K + 2}{6h_3(\xi)}\right).$$

Proof. First we note that

$$\mathbf{M}(\xi)^{-1} = \begin{pmatrix} \frac{1}{h_1(\xi)}\mathbf{Id}_K & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{h_2(\xi)}\mathbf{Id}_{\binom{K}{2}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{1}{h_3(\xi)}\mathbf{Id}_{\binom{K}{3}} \end{pmatrix},$$

for the inverse of the information matrix of the design ξ. Hence, we obtain for the variance function

$$\begin{aligned} V((\mathbf{i},\mathbf{j}),\xi) =& (\mathbf{f}(\mathbf{i}) - \mathbf{f}(\mathbf{j}))^\top \mathbf{M}(\xi)^{-1}(\mathbf{f}(\mathbf{i}) - \mathbf{f}(\mathbf{j})) \\ =& \frac{1}{h_1(\xi)}\sum_{k=1}^{K}(i_k - j_k)^2 \\ &+ \frac{1}{h_2(\xi)}\sum_{k<\ell}(i_k i_\ell - j_k j_\ell)^2 \\ &+ \frac{1}{h_3(\xi)}\sum_{k<\ell<m}(i_k i_\ell i_m - j_k j_\ell j_m)^2. \end{aligned} \tag{6.14}$$

As in the proof of Lemma 6.1 we note first that for the terms associated with the main effects we have $(i_k - j_k)^2 = 4$, when $i_k \neq j_k$, and $(i_k - j_k)^2 = 0$ otherwise. For a pair $(\mathbf{i},\mathbf{j}) \in \mathcal{X}_d$ of comparison depth d there are exactly d attributes for which i_k and j_k differ. Hence, the first sum on the right hand side of (6.14) equals $4d$.

Second, for the terms associated with the first-order interactions we have $(i_k i_\ell - j_k j_\ell)^2 = 4$, if either $i_k \neq j_k$ and $i_\ell = j_\ell$ or $i_k = j_k$ and $i_\ell \neq j_\ell$, and $(i_k i_\ell - j_k j_\ell)^2 = 0$ otherwise. For a pair $(\mathbf{i},\mathbf{j}) \in \mathcal{X}_d$ of comparison depth d there are $d(K-d)$ first-order interaction terms for which $(i_k i_\ell)$ and $(j_k j_\ell)$ differ in exactly one attribute k or ℓ. Hence, the second sum on the right hand side of (6.14) equals $4d(K-d)$.

Finally, for the terms associated with the second-order interactions we have $(i_k i_\ell i_m - j_k j_\ell j_m)^2 = 4$, if $(i_k i_\ell i_m)$ and $(j_k j_\ell j_m)$ differ either in all three attributes k, ℓ and m

or in exactly one of these attributes, and $(i_k i_\ell i_m - j_k j_\ell j_m)^2 = 0$ otherwise. For a pair $(\mathbf{i}, \mathbf{j}) \in \mathcal{X}_d$ of comparison depth d there are $\binom{d}{3}$ second-order interaction terms for which all three associated attributes differ, and there are $d\binom{K-d}{2}$ second-order interaction terms for which $(i_k i_\ell i_m)$ and $(j_k j_\ell j_m)$ differ in exactly one attribute. Hence, there are

$$\begin{aligned}\binom{d}{3} + d\binom{K-d}{2} &= d(d-1)(d-2)/6 + d(K-d)(S-d-1)/2 \\ &= d(3K^2 - 6dK + 4d^2 - 3K + 2)/6\end{aligned}$$

non-zero entries in the third sum on the right hand side of (6.14), and this sum equals $4d(3K^2 - 6dK + 4d^2 - 3K + 2)/6$.

By inserting these results into (6.14) for fixed K, we see that the value of the variance function depends on the pair $(\mathbf{i}, \mathbf{j}) \in \mathcal{X}_d$ only through its comparison depth d and obtain the formula proposed. □

It is worthwhile mentioning that for comparison depth $d = 0$ the corresponding variance function $V(0, \xi) = 0$.

If the invariant design ξ is concentrated on a single comparison depth, then the representation of the variance function $V(d, \xi)$ simplifies.

Corollary 6.1. *For a uniform design $\xi_{d'}$ on a single comparison depth d' the variance function is given by*

$$V(d, \xi_{d'}) = \frac{d}{d'}\left(p_1 + p_2\frac{K-d}{K-d'} + p_3\frac{3K^2 - 6dK + 4d^2 - 3K + 2}{3K^2 - 6d'K + 4d'^2 - 3K + 2}\right).$$

Proof. In view of Theorem 6.4 it is sufficient to note that the representation of the variance function follows immediately by inserting the values of $h_q(\xi_d)$ from Lemma 6.1 and $p_q = \binom{K}{q}$, $q = 1, 2, 3$. □

Note that for $d = d'$, we obtain $V(d, \xi_d) = p_1 + p_2 + p_3 = p$ which shows the D-optimality of ξ_d on $\mathcal{X}_d$ in view of the Kiefer and Wolfowitz (1960) equivalence theorem.

It is worth noting that a single comparison depth d may be sufficient for non-singularity of the information matrix $\mathbf{M}(\xi_d)$, i.e. for the identifiability of all parameters. This can be easily seen by observing $h_q(1) > 0$, $q = 1, 2, 3$, for $d = 1$. But this is not true for all comparison depths as $h_2(K) = 0$. Moreover, in view of Theorems 6.1, 6.2 and 6.3 no design exists which is simultaneously optimal for the main effects, the first-order interactions and the second-order interactions. As a consequence, we confine ourselves to the D-criterion to derive optimal design for the whole parameter vector. The following result gives an upper bound on the number of comparison depths required for a D-optimal design.

Theorem 6.5. *In the second-order interactions model the D-optimal design ξ^* is supported on, at most, three different comparison depths K, d^* and d^*+1, say, i.e.* $\xi^* = w_K^*\xi_K + w_{d^*}^*\xi_{d^*} + (1 - w_K^* - w_{d^*}^*)\xi_{d^*+1}$.

Proof. According to a corollary of the Kiefer-Wolfowitz equivalence theorem for the D-optimal design ξ^* the variance function $V(d, \xi^*)$ is equal to the number of parameters p for all d such that $\xi^* = \sum_{d=1}^{K} w_d^*\xi_d$ for $w_d^* > 0$. By Theorem 6.4 the variance function is a cubic polynomial in the comparison depth d with positive leading coefficient. According to the fundamental theorem of algebra the variance function $V(d, \xi^*)$ may thus be equal to p for, at most, three different values $d_1 < d_2 < d_3$ of d, say. Now, by the Kiefer-Wolfowitz equivalence theorem itself $V(d, \xi^*) \leq p$ for all $d = 0, 1, \ldots, K$. Hence, by the shape of the variance function we obtain that in the case of three different comparison depths $d_3 = K$ and $d_2 = d_1 + 1$ must hold. For two comparison depths either $d_3 = K$ or two adjacent comparison depths d_1 and $d_2 = d_1 + 1$ are included. □

Further for $K = 3$ the D-optimal design can be given explicitly. We note that for $K = 3$ the second-order interaction model is a full interaction model. As a consequence, the result follows from Theorem 4 in Graßhoff et al. (2003). Here we provide an explicit result.

Theorem 6.6. *For $K = 3$ the design $\xi^* = \frac{3}{7}\xi_1 + \frac{3}{7}\xi_2 + \frac{1}{7}\xi_3$ on all pairs with non-zero comparison depth is D-optimal in the second-order interactions model.*

Proof. We may compute the variance function by first computing $h_q(\xi^*) = 16/7$, $q = 1, 2, 3$, and then deriving $V(d, \xi^*) = 7d(d^2 - 6d + 11)/6$ which results in $V(d, \xi^*) = 7$ for $d = 1, 2, 3$. Because $p = 7$ for $K = 3$ the D-optimality of ξ^* follows from the Kiefer-Wolfowitz equivalence theorem. □

Hence, for $K = 3$ all three comparison depths are needed for D-optimality. For $K \geq 4$ numerical computations indicate that only two different comparison depths K and d^* are required. In Table 6.1 numerical solutions for the invariant designs $\xi^* = w_{d^*}^*\xi_{d^*} + (1 - w_{d^*}^*)\xi_K$ having optimal weights $w_{d^*}^*$ and $w_K^* = 1 - w_{d^*}^*$ are presented for numbers K of attributes between 4 and 10.

TABLE 6.1: Optimal Comparison Depths and Optimal Weights for K Binary Attributes

K	4	5	6	7	8	9	10
w_K^*	0.143	0.167	0.268	0.303	0.356	0.423	0.462
d^*	2	2	3	3	3	4	4
$w_{d^*}^*$	0.857	0.833	0.732	0.697	0.644	0.577	0.538

In the following we provide analytical results for fixed number K of the attributes, comparison depth d and weights w_K by direct maximization of $\ln(\det(\mathbf{M}(w_K\xi_K + (1 - w_K)\xi_d)))$ for the corresponding optimal comparison depth d^* and optimal weights w_K^* and $w_{d^*}^* = 1 - w_K^*$ in Table 6.1. The optimality of these designs has been checked numerically by virtue of the Kiefer-Wolfowitz equivalence theorem. The corresponding values of the normalized variance function $V(d, \xi^*)/p$ are recorded in Table 6.2, where maximal values less or equal to 1 establish optimality.

By Lemma 6.1 it follows that for $K = 4$ and 6 we consider the optimal intermediate comparison depth $d^* = K/2$. The entries of the information matrix $\mathbf{M}(\xi)$ are specified by

$$h_1(\xi) = w_K h_1(K) + (1 - w_K)h_1(d^*) = \frac{1 + w_K}{2},$$

$$h_2(\xi) = w_K h_2(K) + (1 - w_K)h_2(d^*) = \frac{K(1 - w_K)}{8(K - 1)},$$

and

$$h_3(\xi) = w_K h_3(K) + (1 - w_K)h_3(d^*) = \frac{w_K + 1}{32}.$$

Now, since the determinant of the information matrix $\mathbf{M}(\xi)$ is proportional to $h_1(\xi)^{p_1}$ $h_2(\xi)^{p_2}$ $h_3(\xi)^{p_3}$, we thus obtain

$$\begin{aligned}\ln\det(\mathbf{M}(\xi)) = c + K\ln(w_K + 1) + \frac{K(K-1)}{2}\ln(K - Kw_K) \\ + \frac{K(K-1)(K-2)}{6}\ln(w_K + 1),\end{aligned}$$

where c is a constant independent of the weight w_K. Taking derivatives with respect to w_K we obtain

$$\frac{\partial}{\partial w_K}\ln\det(\mathbf{M}(\xi)) = \frac{K}{w_K + 1} - \frac{K(K-1)}{2(1 - w_K)} + \frac{K(K-1)(K-2)}{6(w_K + 1)}$$

which has root

$$w_K = w_K^* = \frac{K^2 - 6K + 11}{K^2 + 5}.\text{ It follows that}$$

$$h_1(\xi_{,w_K^*}) = \frac{K^2 - 3K + 8}{K^2 + 5},\ h_2(\xi_{,w_K^*}) = \frac{3K}{4(K^2 + 5)}\text{ and } h_3(\xi_{,w_K^*}) = \frac{K^2 - 3K + 8}{16(K^2 + 5)}.$$

Inserting the expressions $h_1(\xi_{,w_K^*}), h_2(\xi_{,w_K^*})$ and $h_3(\xi_{,w_K^*})$ into the representation given by Theorem 6.4 we obtain

$$V(d^*, \xi^*) = \frac{K(K^2+5)}{6} = p.$$

Hence, the comparison depth d^* is an integer solution for the maximum of the variance function in view of the equivalence theorem by Kiefer and Wolfowitz (1960) when we consider the reduced design region $\mathcal{X}_K \cup \mathcal{X}_{d^*}$.

Further for the case, $K = 5, 7$ and 9 we consider the optimal intermediate comparison depth $d^* = (K-1)/2$. The entries of the information matrix $\mathbf{M}(\xi)$ are similarly specified as

$$h_1(\xi) = w_K h_1(K) + (1-w_K)h_1(d^*) = \frac{K + Kw_K + w_K - 1}{2K},$$

$$h_2(\xi) = w_K h_2(K) + (1-w_K)h_2(d^*) = \frac{K - Kw_K - w_K + 1}{8K},$$

and

$$\begin{aligned} h_3(\xi) &= w_K h_3(K) + (1-w_K)h_3(d^*) \\ &= \frac{K^2 + K^2 w_K - 2Kw_K - 3w_K - 2K + 3}{32K(K-2)}. \end{aligned}$$

As the determinant of the information matrix $\mathbf{M}(\xi)$ is proportional to $h_1(\xi)^{p_1}$ $h_2(\xi)^{p_2}$ $h_3(\xi)^{p_3}$, we thus obtain

$$\begin{aligned} \ln\det(\mathbf{M}(\xi)) = c + K\ln(K + Kw_K + w_K - 1) + \frac{K(K-1)}{2}\ln(K - Kw_K - w_K + 1) \\ + \frac{K(K-1)(K-2)}{6}\ln(K^2 + K^2 w_K - 2Kw_K - 3w_K - 2K + 3). \end{aligned}$$

Taking derivatives with respect to w_K we obtain

$$\begin{aligned} \frac{\partial}{\partial w}\ln\det(\mathbf{M}(\xi)) = \frac{K(K+1)}{K + Kw_K + w_K - 1} - \frac{K(K-1)(K+1)}{2(K - Kw_K - w_K + 1)} \\ + \frac{K(K-1)(K-2)(K^2 - 2K - 3)}{6(K^2 + K^2 w_K - 2Kw_K - 3w_K - 2K + 3)} \end{aligned}$$

which has root

$$w_K = w_K^* = \frac{2K^3 - 6K^2 + 7K - K\sqrt{K^6 - 12K^5 + 64K^4 - 198K^3 + 448K^2 - 636K + 369} + 15}{-K^4 + 2K^3 - 2K^2 + 10K + 15}.$$

This root w_K^* gives a maximum for the determinant. The design ξ^* is thus D-optimal when we consider the reduced design region $\mathcal{X}_K \cup \mathcal{X}_{d^*}$.

Moreover, for the case, $K = 8$ and 10 we consider the optimal intermediate comparison depth $d^* = (K/2) - 1$. The entries of the information matrix $\mathbf{M}(\xi)$ are

$$h_1(\xi) = w_K h_1(K) + (1 - w_K)h_1(d^*) = \frac{K + Kw_K + 2w_K - 2}{2K},$$

$$h_2(\xi) = w_K h_2(K) + (1 - w_K)h_2(d^*) = \frac{K^2 - K^2 w_K + 4w_K - 4}{8K(K-1)},$$

and

$$\begin{aligned} h_3(\xi) &= w_K h_3(K) + (1 - w_K)h_3(d^*) \\ &= \frac{K^2 + K^2 w_K - K w_K - K - 6w_K + 6}{32K(K-1)}. \end{aligned}$$

The determinant of the information matrix $\mathbf{M}(\xi)$ is proportional to $h_1(\xi)^{p_1}$ $h_2(\xi)^{p_2}$ $h_3(\xi)^{p_3}$, we thus obtain

$$\begin{aligned} \ln\det(\mathbf{M}(\xi)) = c + K\ln(K + Kw_K + 2w_K - 2) + \frac{K(K-1)}{2}\ln(K^2 - K^2 w_K + 4w_K - 4) \\ + \frac{K(K-1)(K-2)}{6}\ln(K^2 + K^2 w_K - K w_K - K - 6w_K + 6). \end{aligned}$$

Taking derivatives with respect to w we obtain

$$\begin{aligned} \frac{\partial}{\partial w}\ln\det(\mathbf{M}(\xi)) = \frac{K(K+2)}{K + Kw_K + 2w_K - 2} + \frac{K(K-1)(-K^2+4)}{2(K^2 - K^2 w_K + 4w_K - 4)} \\ + \frac{K(K-1)(K-2)(K^2 - K - 6)}{6(K^2 + K^2 w_K - K w_K - K - 6w_K + 6)} \end{aligned}$$

which has root

$$w_K = w_K^* = \frac{K^3 + 5K + (K - K^2)\sqrt{K^4 - 10K^3 + 37K^2 - 60K + 180} + 30}{-K^4 + K^3 + K^2 + 5K + 30}.$$

This root w_K^* gives a maximum for the determinant. The design ξ^* is thus D-optimal when we consider the reduced design region $\mathcal{X}_K \cup \mathcal{X}_{d^*}$.

Again, the comparison depth d^* is an integer solution for the maximum of the variance function which establishes the D-optimality of the design ξ^* in view of the Kiefer-Wolfowitz equivalence theorem. □

Exhibited in Table 6.2 are values of the normalized variance function $V(d, \xi^*)/p$ and comparison depth d, which show D-optimality of the design ξ^* in view of the equivalence theorem by Kiefer and Wolfowitz (1960).

TABLE 6.2: Values of the Variance Function of the D-Optimal Design ξ^* for K Binary Attributes (Boldface **1** Corresponds to the Optimal Comparison Depths d^*)

	d									
K	1	2	3	4	5	6	7	8	9	10
4	0.875	**1**	0.875	**1**						
5	0.760	**1**	0.960	0.880	**1**					
6	0.701	0.983	**1**	0.906	0.855	**1**				
7	0.615	0.917	**1**	0.956	0.879	0.863	**1**			
8	0.559	0.872	**1**	1	0.945	0.884	0.882	**1**		
9	0.504	0.811	0.962	**1**	0.969	0.910	0.868	0.883	**1**	
10	0.462	0.763	0.932	**1**	0.997	0.956	0.905	0.874	0.896	**1**

6.2 Designs for Partial Profiles

In the following we reformulate the corresponding concept about full profiles when there is a possibility of too many attributes, K to the situation when only a subset of the attributes (so-called partial profiles) are to be evaluated by respondents (see Großmann, 2018, for motivation). Analogous to Chapter 5 for a partial profile a direct observation may be described by model (6.1) when summation is taken only over those S attributes contained in the describing subset. This requires that the profile strength S must be, at least, three in order to capture the second-order interactions. To facilitate notation we introduce an additional level 0 for each attribute indicating that the corresponding attribute is not present in the partial profile, and the corresponding regression functions are given by $f_k(0) = f_1(0) = 0$. With this convention a direct observation can be described by (6.1) even for a partial profile $\mathbf{i}$ from the set

$$\begin{aligned} \mathcal{I}^{(S)} =& \{\mathbf{i};\ i_k \in \{-1, 1\} \text{ for } S \text{ components and} \\ & i_k = 0 \text{ for } K - S \text{ components}\}, \end{aligned} \tag{6.15}$$

of alternatives with profile strength S. In particular, $\mathcal{I}^{(K)} = \mathcal{I}^{(S)}$ in the case of full profiles ($S = K$). For general profile strength S the vector $\mathbf{f}$ of regression functions in (6.3), the paired comparison model in (6.4) which is given by

$$Y_n(\mathbf{i}, \mathbf{j}) = \sum_{k=1}^{K} (i_k - j_k)\beta_k + \sum_{k<\ell} (i_k i_\ell - j_k j_\ell)\beta_{k\ell} + \sum_{k<\ell<m} (i_k i_\ell i_m - j_k j_\ell j_m)\beta_{k\ell m} + \varepsilon_n, \tag{6.16}$$

and the interpretation of the corresponding parameter vector $\boldsymbol{\beta}$ remain unchanged. Here we note that the comparison depth d as in the work of Graßhoff et al. (2003) describes the number of attributes in which the two alternatives in the choice sets differ satisfying the inequalities $1 \leq d \leq S < K$. This is the approach taken by Großmann (2018) for a study in the health sector where only 4 out of a total number of 11 attributes in a choice set of size two were presented to respondents simultaneously.

Hence, the paired comparison model (6.4) having corresponding design region $\mathcal{X}$ in (6.5) is thus restricted to those paired alternatives for which exactly S attributes are presented

$$\begin{aligned} \mathcal{X}^{(S)} = \{(\mathbf{i}, \mathbf{j});\ & i_k, j_k \in \{1, -1\} \text{ for } S \text{ components and} \\ & i_k = j_k = 0 \text{ for exactly } K - S \text{ components}\}. \end{aligned} \tag{6.17}$$

As already noted the design region $\mathcal{X}^{(S)}$ can be also partitioned into disjoint sets such that the pairs in each set differ only in a fixed number d of the attributes. Specifically, for a comparison depth $d = 0, \ldots, S$, let

$$\mathcal{X}_d^{(S)} = \{(\mathbf{i}, \mathbf{j}) \in \mathcal{X}^{(S)} : |\{k : i_k \neq j_k\}| = d\}, \tag{6.18}$$

be the set of all pairs of alternatives which differ in exactly d attributes. These sets also constitute the orbits with respect to permutations. The D-criterion is also invariant with respect to those permutations which induce a linear reparameterization. Specifically, we mention that the regression functions (6.3) extended to the design region $\mathcal{X}^{(S)}$ are still linearly equivariant i. e. that also here relabeling does not affect D-optimality (and D-optimality of invariant subvectors). Hence, as already specified it is sufficient to look for optimality in the class of invariant designs which are uniform on the orbits of fixed comparison depth $d \leq S$. These designs can be found in the work of Nyarko and Schwabe (2019).

Further let $N_d = 2^S \binom{K}{S} \binom{S}{d}$ be the number of different pairs in $\mathcal{X}_d^{(S)}$ which vary in exactly d attributes and denote by ξ_d the uniform approximate design which assigns equal weight $\xi_d(\mathbf{i}, \mathbf{j}) = 1/N_d$ to each pair in $\mathcal{X}_d^{(S)}$. The corresponding information matrices for the invariant designs are presented in the following Lemmas 6.3 and 6.4. We note that results for full profiles can be obtained as special cases by letting $S = K$.

Lemma 6.3. *Let $d \in \{0, \ldots, S\}$. The uniform design ξ_d on the set $\mathcal{X}_d^{(S)}$ of comparison depth d has a diagonal information matrix*

$$\mathbf{M}(\xi_d) = \begin{pmatrix} h_1(d)\mathbf{Id}_K & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & h_2(d)\mathbf{Id}_{\binom{K}{2}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & h_3(d)\mathbf{Id}_{\binom{K}{3}} \end{pmatrix},$$

where $h_1(d) = \frac{4d}{K}$, $h_2(d) = \frac{8d(S-d)}{K(K-1)}$ *and* $h_3(d) = \frac{4d(3S^2-6dS+4d^2-3S+2)}{K(K-1)(K-2)}$.

Proof. From Lemma 6.1 the regression functions $f_k = f_1$ associated with the k-th main effects are given by

$$\begin{aligned}\sum_{(\mathbf{i},\mathbf{j}) \in \mathcal{X}_d^{(S)}} (i_k - j_k)^2 &= \binom{K-1}{S-1}\binom{S-1}{d-1} 2^S \sum_{i_k \neq j_k} (i_k - j_k)^2 \\ &= \binom{K-1}{S-1}\binom{S-1}{d-1} 2^{S+2},\end{aligned}$$

as there are $\binom{K-1}{S-1}\binom{S-1}{d-1}2^S$ pairs in $\mathcal{X}_d^{(S)}$ for which i_k and j_k differ.

The number N_d of paired comparisons in $\mathcal{X}_d^{(S)}$ with comparison depth d equals $N_d = \binom{K}{S}\binom{S}{d}2^S$. Hence the corresponding diagonal elements $h_1(d)$ are given by

$$h_1(d) = \frac{1}{N_d}\sum_{(\mathbf{i},\mathbf{j})\in\mathcal{X}_d^{(S)}}(i_k - j_k))^2 = \frac{1}{N_d}\binom{K-1}{S-1}\binom{S-1}{d-1}2^{S+2} = \frac{4d}{K} \tag{6.19}$$

in the information matrix.

Further for the first-order interactions and the given attributes k and ℓ the pairs with distinct levels in both attributes occur $\binom{K-2}{S-2}\binom{S-2}{d-2}2^S$ times in $\mathcal{X}_d^{(S)}$, while those which differ only in one attribute occur $\binom{2}{1}\binom{K-2}{S-2}\binom{S-2}{d-1}2^S$ times. As a result, from (6.7)–(6.8) the diagonal elements $h_2(d)$ in the corresponding information matrix are given by

$$h_2(d) = \frac{1}{N_d}\binom{K-2}{S-2}\binom{S-2}{d-1}2^{S+3} = \frac{8d(S-d)}{K(K-1)}. \tag{6.20}$$

Accordingly, for the second-order interactions, and for the given attributes k, ℓ and m the pairs with distinct levels in the three attributes occur $\binom{K-3}{S-3}\binom{S-3}{d-3}2^S$ times in $\mathcal{X}_d^{(S)}$, while those which differ in the two attributes occur $\binom{3}{2}\binom{K-3}{S-3}\binom{S-3}{d-2}2^S$ times in $\mathcal{X}_d^{(S)}$, and those which differ only in the one attribute occur $\binom{3}{1}\binom{K-3}{S-3}\binom{S-3}{d-1}2^S$ times in $\mathcal{X}_d^{(S)}$. Hence, from (6.10)–(6.12) the diagonal elements $h_3(d)$ are given by

$$\begin{aligned} h_3(d) &= \frac{1}{N_d}\binom{K-3}{S-3}\left(\binom{S-3}{d-3}2^{S+2} + 3\binom{S-3}{d-1}2^{S+2}\right) \\ &= \frac{4d(d-1)(d-2)}{K(K-1)(K-2)} + \frac{12(S-d)(S-d-1)d}{K(K-1)(K-2)} \\ &= \frac{4d}{K(K-1)(K-2)}\Big((d-1)(d-2) + 3(S-d)(S-d-1)\Big) \\ &= \frac{4d}{K(K-1)(K-2)}(d^2 - 3d + 2 + 3S^2 - 6Sd - 3S + 3d^2 + 3d) \\ &= \frac{4d}{K(K-1)(K-2)}(3S^2 - 6Sd + 4d^2 - 3S + 2). \end{aligned} \tag{6.21}$$

The off-diagonal elements all vanish because the terms in the corresponding entries sum up to zero due to the effects-type coding. □

For the particular case of first-order interactions the corresponding results of the entries $h_1(d)$ and $h_2(d)$ can be found (e.g., see Schwabe et al., 2003; Graßhoff et al., 2003). It is worth mentioning that the corresponding functions $h_q(0) = 0$ for $q = 1, 2, 3$.

General invariant designs ξ as already pointed out can be written as a convex combination $\xi = \sum_{d=1}^{S} w_d\xi_d$ of the corresponding uniform designs on the comparison depth d with corresponding weights $w_d \geq 0$, $\sum_{d=1}^{S} w_d = 1$. Hence, also every invariant design has diagonal information matrix.

Lemma 6.4. *Let ξ be an invariant design on $\mathcal{X}^{(S)}$, i. e. $\xi = \sum_{d=1}^{S} w_d \xi_d$, then ξ has diagonal information matrix*

$$\mathbf{M}(\xi) = \begin{pmatrix} h_1(\xi)\mathbf{Id}_K & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & h_2(\xi)\mathbf{Id}_{\binom{K}{2}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & h_3(\xi)\mathbf{Id}_{\binom{K}{3}} \end{pmatrix},$$

where $h_q(\xi) = \sum_{d=1}^{S} w_d h_q(d)$, $q = 1, 2, 3$.

In the following we consider optimal designs for the main effects, the first-order interaction and the second-order interaction terms having corresponding entries $h_1(d)$, $h_2(d)$ and $h_3(d)$, respectively, in the information matrix $\mathbf{M}(\xi_d)$ of Lemma 6.3. To start with, we mention that the following Theorems 6.7 and 6.8 paraphrase theorems given in Schwabe et al. (2003) and Graßhoff et al. (2003) for first-order interaction models and translate them to the present setting of second-order interaction models.

Theorem 6.7. *Let $d_1^* = S$. Then the uniform design $\xi_{d_1^*} = \xi_S$ on the largest possible comparison depth S is D-optimal for the main effects $(\beta_k)_{1\leq k\leq K}^{\top}$.*

This means that for the main effects only those pairs of alternatives should be used which differ in all attributes subject to the profile strength S.

Theorem 6.8. *Let $d_2^* = S/2$ for S even and $d_2^* = (S-1)/2$ or $d_2^* = (S+1)/2$ for S odd, respectively. Then the uniform design $\xi_{d_2^*}$ on $\mathcal{X}_d^{(S)}$ is D-optimal for the first-order interaction effects $(\beta_{k\ell})_{k<\ell}^{\top}$.*

This means that for the first-order interactions only those pairs of alternatives should be used which differ in about half of the attributes subject to the profile strength S.

Theorem 6.9. *Let $d_3^* = 1$ or $d_3^* = 3$ for $S = 3$ and $d_3^* = S$ for $S \geq 4$, respectively. Then the uniform design $\xi_{d_3^*}$ on $\mathcal{X}_d^{(S)}$ is D-optimal for the second-order interaction effects $(\beta_{k\ell m})_{k<\ell<m}^{\top}$.*

This means that also for the second-order interactions it is sufficient to use those pairs of alternatives that differ in all the profile strength S.

Proof. We first note that optimality is achieved when h_3 is maximized. Hence, for $S = 3$ we get $h_3(1) = h_3(3) = 4$ and $h_3(2) = h_3(0) = 0$ which proves the result in this case. Moreover, for the proof of $S \geq 4$ it follows directly from Theorem 6.3 that h_3 attains its local maximum and local minimum at $d_{3,\max} = S/2 - \sqrt{9S-6}/6$ and $d_{3,\min} = S/2 + \sqrt{9S-6}/6$, respectively. $\square$

As already defined in connection with Theorem 6.4 here the variance function $V((\mathbf{i},\mathbf{j}),\xi)=(\mathbf{f}(\mathbf{i})-\mathbf{f}(\mathbf{j}))^\top\mathbf{M}(\xi)^{-1}(\mathbf{f}(\mathbf{i})-\mathbf{f}(\mathbf{j}))$ is also invariant with respect to permutation of levels and attributes and, also constant on the orbits $\mathcal{X}_d^{(S)}$ of fixed comparison depth d. We define by $V(d,\xi)$, $V(d,\xi)=V((\mathbf{i},\mathbf{j}),\xi)$ on $\mathcal{X}_d^{(S)}$ the value of the variance function for the corresponding invariant design evaluated at comparison depth d. Hence we obtain the following results.

Theorem 6.10. *For every invariant design ξ the variance function $V(d,\xi)$ on $\mathcal{X}_d^{(S)}$ is given by*

$$V(d,\xi)=4d\left(\frac{1}{h_1(\xi)}+\frac{S-d}{h_2(\xi)}+\frac{3S^2-6dS+4d^2-3S+2}{6h_3(\xi)}\right).$$

Proof. First we note that for the design ξ on $\mathcal{X}_d^{(S)}$ the inverse of the corresponding information matrix $\mathbf{M}(\xi)$ is given by

$$\mathbf{M}(\xi)^{-1}=\begin{pmatrix}\frac{1}{h_1(\xi)}\mathbf{Id}_K & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \frac{1}{h_2(\xi)}\mathbf{Id}_{\binom{K}{2}} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \frac{1}{h_3(\xi)}\mathbf{Id}_{\binom{K}{3}}\end{pmatrix}.$$

Hence, we obtain for the variance function

$$\begin{aligned}V((\mathbf{i},\mathbf{j}),\xi)=&(\mathbf{f}(\mathbf{i})-\mathbf{f}(\mathbf{j}))^\top\mathbf{M}(\xi)^{-1}(\mathbf{f}(\mathbf{i})-\mathbf{f}(\mathbf{j}))\\ =&\frac{1}{h_1(\xi)}\sum_{k=1}^{K}(i_k-j_k)^2\\ &+\frac{1}{h_2(\xi)}\sum_{k<\ell}(i_ki_\ell-j_kj_\ell)^2\\ &+\frac{1}{h_3(\xi)}\sum_{k<\ell<m}(i_ki_\ell i_m-j_kj_\ell j_m)^2.\end{aligned}\tag{6.22}$$

In view of Theorem 6.4 and Lemma 6.3 it is sufficient to note that

For a pair $(\mathbf{i},\mathbf{j})\in\mathcal{X}_d^{(S)}$ of comparison depth d there are exactly d attributes of the main effects for which i_k and j_k differ. Hence, the first sum on the right hand side of (6.22) equals $4d$.

Again, for a pair $(\mathbf{i},\mathbf{j})\in\mathcal{X}_d^{(S)}$ of comparison depth d there are $d(S-d)$ first-order interaction terms for which (i_ki_ℓ) and (j_kj_ℓ) differ in exactly one attribute k or ℓ. Hence, the second sum on the right hand side of (6.22) equals $4d(S-d)$.

Finally, for a pair $(\mathbf{i},\mathbf{j})\in\mathcal{X}_d^{(S)}$ of comparison depth d there are $\binom{d}{3}$ second-order interaction terms for which $(i_ki_\ell i_m)$ and $(j_kj_\ell j_m)$ of all the associated three attributes k,ℓ and m differ, and there are $d\binom{S-d}{2}$ second-order interaction terms for which $(i_ki_\ell i_m)$

and $(j_k j_\ell j_m)$ differ in exactly one attribute. Hence, there are

$$\begin{aligned}\binom{d}{3} + d\binom{S-d}{2} &= d(d-1)(d-2)/6 + d(S-d)(S-d-1)/2 \\ &= d(3S^2 - 6dS + 4d^2 - 3S + 2)/6\end{aligned}$$

non-zero entries in the third sum on the right hand side of (6.22), and this sum equals $4d(3S^2 - 6dS + 4d^2 - 3S + 2)/6$.

By inserting these results into (6.22) for fixed S we see that the value of the variance function depends on the pair $(\mathbf{i}, \mathbf{j}) \in \mathcal{X}_d^{(S)}$ only through its comparison depth d which proofs the representation of the variance function. □

As was already pointed out here we similarly note that for comparison depth $d = 0$ the corresponding variance function $V(0, \xi) = 0$.

The corresponding variance function simplifies if the invariant design ξ is concentrated on a single comparison depth.

Corollary 6.2. *For a uniform design $\xi_{d'}$ on a single comparison depth d' the variance function is given by*

$$V(d, \xi_{d'}) = \frac{d}{d'}\left(p_1 + p_2\frac{S-d}{S-d'} + p_3\frac{3S^2 - 6dS + 4d^2 - 3S + 2}{3S^2 - 6d'S + 4d'^2 - 3S + 2}\right).$$

Proof. In view of Theorem 6.10 it is sufficient to note that the representation of the variance function follows immediately by inserting the values of $h_q(\xi_d)$ from Lemma 6.3 and $p_q = \binom{K}{q}$, $q = 1, 2, 3$. □

We note that if $d = d'$ then the varaince $V(d, \xi_d) = p_1 + p_2 + p_3 = p$, which shows the D-optimality of ξ_d on $\mathcal{X}_d^{(S)}$ in view of the Kiefer and Wolfowitz (1960) equivalence theorem.

Obviously, no optimal design exists for the whole parameter vector of Theorems 6.7, 6.8 and 6.9. Hence, in view of Kiefer and Wolfowitz (1960) equivalence theorem we similarly focus on the D-criterion to derive optimal design for the whole parameter vector of the following theorems.

Theorem 6.11. *In the second-order interactions model the D-optimal design ξ^* is supported on at most, three different comparison depths S, d^* and $d^* + 1$.*

Proof. According to the Kiefer-Wolfowitz equivalence theorem for an invariant D-optimal design ξ^* with weights w_d^* on the comparison depths d the variance function $V(d, \xi^*)$ is equal to the number of parameters p for all d such that $\xi^* = \sum_{d=1}^{S} w_d^* \xi_d$ for $w_d^* > 0$. By Theorem 6.10 the variance function is a cubic polynomial in the comparison depth d with positive leading coefficient. According to the fundamental theorem of algebra the variance function $V(d, \xi^*)$ may thus be equal to p for, at most, three different values of d. Now, by the Kiefer-Wolfowitz equivalence theorem itself $V(d, \xi^*) \leq p$ for all $d = 0, 1, \ldots, S$. Hence, by the shape of the variance function we obtain that $V(d, \xi^*) = p$ may occur only at the maximal comparison depth $d = S$ allowed or at, at most, two adjacent comparison depths d^* and $d^* + 1$, say, in the interior. □

Also for the case $S = 3 < K$ the corresponding D-optimal design ξ^* can be given in a more intuitive manner.

Theorem 6.12. (a) *If $S = 3$ and $K = 4$, then the design $\xi^* = \frac{9}{10}\xi_1 + \frac{1}{10}\xi_3$ is D-optimal in the second-order interactions model.*
(b) *If $S = 3$ and $K > 4$, then the design $\xi^* = \xi_1$ is D-optimal in the second-order interactions model.*

Proof. (a) For the design ξ^* we obtain $h_1(\xi^*) = h_2(\xi^*) = 6/5$ and $h_3(\xi^*) = 1$. Inserting this into the variance function of Theorem 6.10 yields $V(d, \xi^*) = 2d(4d^2 - 23d + 40)/3$ which results in $V(1, \xi^*) = V(3, \xi^*) = 14$ and $V(2, \xi^*) = 40/3 < 14$. Hence, the variance function is bounded by the number of parameters $p = 14$ which establishes the D-optimality of ξ^* in view of the Kiefer-Wolfowitz equivalence theorem.

(b) From Corollary 6.2 we derive the variance function

$$V(d, \xi_1) = dp_1 + dp_2(3 - d)/2 + dp_3(2d^2 - 9d + 10)/3$$

which results in $V(1, \xi^*) = p_1 + p_2 + p_3 = p$, $V(2, \xi^*) = 2p_1 + p_2$ and $V(3, \xi^*) = 3p_1 + p_3$. Because $p_1 = K \leq K(K-1)(K-2)/6 = p_3$ for $K \geq 5$ and $2p_1 = 2K \leq K(K-1)/2 = p_2$ for $K \geq 5$ the variance function is bounded by the number of parameters p, and the D-optimality of ξ^* follows from the Kiefer-Wolfowitz equivalence theorem. □

Hence, for $S = 3$ and $K = 4$ two comparison depths are needed for D-optimality while for $S = 3$ and $K > 4$ one comparison depth is sufficient. Moreover, for $S \geq 4$ numerical computations indicate that only two different comparison depths S and d^* are required for D-optimality. Exhibited in Table 6.3 are numerical solutions for the invariant designs $\xi^* = w^*_{d^*}\xi_{d^*} + (1 - w^*_{d^*})\xi_S$ having optimal weights $w^*_{d^*}$ and optimal comparison depth d^* with entries of the form $(d^*, w^*_{d^*})$ for various choices of profile strengths S and attributes K between 3 and 10. It is worthwhile mentioning that the results for full profiles presented in Table 6.1 can be recovered in Table 6.3 for the (row) entries when $S = K$.

Table 6.3: Optimal Designs with Intermediate Comparison Depths d^* and Optimal Weights $w^*_{d^*}$ of the form $(d^*,\ w^*_{d^*})$ for $S \le K$ Binary Attributes

	S							
K	3	4	5	6	7	8	9	10
4	(1, 0.900)	(2, 0.857)						
5	(1,1)	(2, 0.800)	(2, 0.833)					
6	(1,1)	(2, 0.732)	(2, 0.802)	(3, 0.732)				
7	(1,1)	(1, 0.836)	(2, 0.756)	(2, 0.728)	(3, 0.697)			
8	(1,1)	(1, 0.832)	(2, 0.707)	(2, 0.687)	(3, 0.643)	(3, 0.644)		
9	(1,1)	(1, 0.819)	(2, 0.659)	(2, 0.645)	(3, 0.594)	(3, 0.598)	(4, 0.577)	
10	(1,1)	(1, 0.800)	(2, 0.615)	(2, 0.604)	(3, 0.551)	(3, 0.556)	(4, 0.533)	(4, 0.538)

For fixed number $S \le K$ of attributes and intermediate comparison depth d the optimal weights w^*_S and $w^*_{d^*} = 1 - w^*_S$ have been determined analytically by direct maximization of $\ln(\det(\mathbf{M}(w\xi_S + (1-w_S)\xi_d)))$. In particular, by Lemma 6.3 the optimal weights w^*_S can be obtained analytically:

First we note that for the case where $S = 3$ and $K = 4$ we consider the optimal intermediate comparison depth $d^* = (S-1)/2$ and corresponding optimal weight $w^*_S = \frac{2S-K-1}{KS-K+S-1}$. Hence, the entries of the information matrix $\mathbf{M}(\xi)$ are specified as

$$h_1(\xi_{,w^*_S}) = \frac{S}{2K+2}, h_2(\xi_{,w^*_S}) = \frac{KS-S}{(K+1)(S-1)} \text{ and } h_3(\xi_{,w^*_S}) = \frac{1}{64}.$$

Inserting the expressions $h_1(\xi_{,w^*_S})$, $h_2(\xi_{,w^*_S})$ and $h_3(\xi_{,w^*_S})$ into the representation of the variance function in Theorem 6.10 we obtain

$$\begin{aligned} V(d^*, \xi^*) &= \frac{2K+2}{S} + \frac{3(K+1)(S-1)^2}{(K-1)S} + 2S^2 - 6S + 4 \\ &= \frac{2K^2 + 3K + 2KS^3 - 3KS^2 - 2KS - 2S^3 + 9S^2 - 10S + 1}{KS - S} \\ &= p. \end{aligned}$$

Hence, the comparison depth d^* is an integer solution for the maximum of the variance function in view of the equivalence theorem by Kiefer and Wolfowitz (1960). Moreover, for the case where $S = 5, 7$ or 9, and $K \ge S$ the entries of the information matrix $\mathbf{M}(\xi)$ are specified as

$$h_1(\xi) = wh_1(S) + (1-w)h_1(d^*) = \frac{S + Sw + w - 1}{2K},$$

$$h_2(\xi) = wh_2(S) + (1-w)h_2(d^*) = \frac{S^2 - S^2w + w - 1}{8K(K-1)},$$

and

$$\begin{aligned} h_3(\xi) &= wh_3(S) + (1-w)h_3(d^*) \\ &= \frac{S^3 + S^3w - 3S^2w - 3S^2 - Sw + 5S + 3w - 3}{32K(K-1)(K-2)}. \end{aligned}$$

Now as the determinant of the information matrix $\mathbf{M}(\xi)$ is proportional to $h_1(\xi)^{p_1}$ $h_2(\xi)^{p_2} h_3(\xi)^{p_3}$, we obtain

$$\begin{aligned} \ln\det(\mathbf{M}(\xi)) = c &+ p_1 \ln(S + Sw + w - 1) + p_2 \ln(S^2 - S^2w + w - 1) \\ &+ p_3 \ln(S^3 + S^3w - 3S^2w - 3S^2 - Sw + 5S + 3w - 3), \end{aligned}$$

where c is a constant independent of the weight w. Taking derivatives with respect to w we obtain

$$\begin{aligned} \frac{\partial}{\partial w} \ln\det(\mathbf{M}(\xi)) = &\frac{p_1(S+1)}{S + Sw + w - 1} + \frac{p_2(-S^2+1)}{S^2 - S^2w + w - 1} \\ &+ \frac{p_3(S^3 - 3S^2 - S + 3)}{S^3 + S^3w - 3S^2w - 3S^2 - Sw + 5S + 3w - 3} \end{aligned}$$

which has root

$$w = w_S^* = \frac{K^2S - 3KS^2 - 3K^2 + 6KS + 3S^2 + \sqrt{\lambda} - 7S - 15}{K^2S^2 - 2K^2S - 3K^2 + 5S^2 - 10S - 15}$$

where

$$\begin{aligned} \lambda = K^4S^4 &- 6K^4S^3 - 6K^3S^4 + 9K^4S^2 + 30K^3S^3 + 25K^2S^4 - 36K^3S^2 - 114K^2S^3 \\ &- 48KS^4 + 126K^2S^2 + 258KS^3 + 64S^4 - 324KS^2 - 312S^3 + 369S^2. \end{aligned}$$

This root w_S^* gives a maximum for the determinant. The design ξ^* is thus D-optimal when we consider the reduced design region $\mathcal{X}_S^{(S)} \cup \mathcal{X}_{d^*}^{(S)}$.

Further for the case where $S = 4, 6, 8$ or 10, and $K \geq S$ we consider the optimal intermediate comparison depth $d^* = (S/2) - 1$. The entries of the information matrix $\mathbf{M}(\xi)$ are specified as

$$h_1(\xi) = wh_1(S) + (1-w)h_1(d^*) = \frac{S + Sw + 2w - 2}{2K},$$

$$h_2(\xi) = wh_2(S) + (1-w)h_2(d^*) = \frac{S^2 - S^2w + 4w - 4}{8K(K-1)},$$

and

$$\begin{aligned} h_3(\xi) &= wh_3(S) + (1-w)h_3(d^*) \\ &= \frac{S^3 + S^3w - 3S^2 - 3S^2w + 8S - 4Sw + 12w - 12}{32K(K-1)(K-2)}. \end{aligned}$$

Now the determinant of the information matrix $\mathbf{M}(\xi)$ is given by

$$\begin{aligned} \ln\det(\mathbf{M}(\xi^*)) = c &+ p_1 \ln(S + Sw + 2w - 2) + p_2 \ln(S^2 - S^2w + 4w - 4) \\ &+ p_3 \ln(S^3 + S^3w - 3S^2 - 3S^2w + 8S - 4Sw + 12w - 12). \end{aligned}$$

Taking derivatives with respect to w we obtain

$$\begin{aligned} \frac{\partial}{\partial w} \ln\det(\mathbf{M}(\xi)) &= \frac{p_1(S+2)}{S + Sw + 2w - 2} + \frac{p_2(-S^2+4)}{S^2 - S^2w + 4w - 4} \\ &+ \frac{p_3(S^3 - 3S^2 - 4S + 12)}{S^3 + S^3w - 3S^2 - 3S^2w + 8S - 4Sw + 12w - 12} \end{aligned}$$

which has root

$$w = w_S^* = \frac{2K^2S - 3KS^2 - 6K^2 + 3KS + 3S^2 + \sqrt{\lambda} - 5S - 30}{K^2S^2 - K^2S - 6K^2 + 5S^2 - 5S - 30}$$

where

$$\begin{aligned} \lambda = K^4S^4 &- 6K^4S^3 - 6K^3S^4 + 9K^4S^2 + 24K^3S^3 + 25K^2S^4 - 18K^3S^2 - 78K^2S^3 \\ &- 48KS^4 + 45K^2S^2 + 228KS^3 + 64S^4 - 180KS^2 - 240S^3 + 180S^2. \end{aligned}$$

Finally, for the case where $S = 4$ for $K = 5$ or 6 we consider the optimal intermediate comparison depth $d^* = S/2$. The entries of the information matrix $\mathbf{M}(\xi)$ are specified as

$$h_1(\xi) = wh_1(S) + (1-w)h_1(d^*) = \frac{S + Sw}{2K},$$

$$h_2(\xi) = wh_2(S) + (1-w)h_2(d^*) = \frac{S^2 - S^2w}{8K(K-1)},$$

and

$$h_3(\xi) = wh_3(S) + (1-w)h_3(d^*) = \frac{S^3 - 3S^2 + 2S + S^3w - 3S^2w + 2Sw}{32K(K-1)(K-2)}.$$

The determinant of the information matrix $\mathbf{M}(\xi)$ is given by

$$\ln\det(\mathbf{M}(\xi)) = c + p_1 \ln(S + Sw) + p_2 \ln(S^2 - S^2 w) + p_3 \ln(S^3 - 3S^2 + 2S + S^3 w - 3S^2 w + 2Sw).$$

Taking derivatives with respect to w we obtain

$$\frac{\partial}{\partial w}\ln\det(\mathbf{M}(\xi)) = \frac{p_1 S}{S + Sw} + \frac{p_2(-S^2)}{S^2 - S^2 w} + \frac{p_3(S^3 - 3S^2 + 2S)}{S^3 - 3S^2 + 2S + S^3 w - 3S^2 w + 2Sw}$$

which has root

$$w = w_S^* = \frac{K^2S^3 - 6KS^3 - 3K^2S^2 + 11S^3 + 18KS^2 + 2K^2S - 33S^2 - 12KS + 22S}{K^2S^3 - 3K^2S^2 + 5S^3 + 2K^2S - 15S^2 + 10S}.$$

This root w_S^* gives a maximum for the determinant. The design ξ^* is thus D-optimal when we consider the reduced design region $\mathcal{X}_S^{(S)} \cup \mathcal{X}_{d^*}^{(S)}$.

Again, the comparison depth d^* is an integer solution for the maximum of the variance function which proofs the D-optimality of the design ξ^* in view of the Kiefer and Wolfowitz (1960) equivalence theorem. □

Values of the normalized variance function $V(d, \xi^*)/p$, which show D-optimality of the design ξ^* in view of the equivalence theorem by Kiefer and Wolfowitz (1960) are presented in the following Table 6.4.

TABLE 6.4: Values of the Variance Function of the D-Optimal Design ξ^* for $S = K - 1$ Binary Attributes (Boldface **1** Corresponds to the Optimal Comparison Depths d^*)

		d								
K	S	1	2	3	4	5	6	7	8	9
4	3	**1**	0.952	**1**						
5	4	0.958	**1**	0.792	**1**					
6	5	0.791	**1**	0.913	0.817	**1**				
7	6	0.701	**1**	1	0.915	0.858	**1**			
8	7	0.626	0.926	**1**	0.946	0.862	0.847	**1**		
9	8	0.565	0.877	**1**	0.998	0.933	0.869	0.870	**1**	
10	9	0.508	0.815	0.964	**1**	0.965	0.904	0.860	0.878	**1**

7 Optimal Designs for Second-Order Interactions General Level Models

For the present setting of common number general-level attributes results for main effects and first-order interactions have been derived by Graßhoff et al. (2003) as pointed out in Chapter 5. Here we focus on second-order interactions where the attributes $k = 1, \ldots, K$ having levels $i_k = 1, \ldots, v$ each are assumed to derive the preferences for the alternatives in a paired comparison experiment. For alternatives in a choice set of size two, as before we denote by $\mathbf{i} = (i_1, \ldots, i_K)$ the first alternative and the second alternative by $\mathbf{j} = (j_1, \ldots, j_K)$ which are both elements of the set $\mathcal{I} = \mathcal{I}_1 \times \cdots \times \mathcal{I}_K = \{1, \ldots, v\}^K$ where the numbers 1 and v represent the first and last level of each attribute. The alternatives $\mathbf{i}$ and $\mathbf{j}$ are ordered pairs which are chosen from the design region $\mathcal{X} = \mathcal{I} \times \mathcal{I}$. For each attribute (component) k the corresponding marginal model coincides with that of the one-way layout with regression functions $\mathbf{f}_k = \mathbf{f}_1$ defined in Section 4.2.

Now the second-order interaction model (6.1) can be similarly formulated as

$$\tilde{Y}_{na}(\mathbf{i}) = \mu_n + \sum_{k=1}^{K} \alpha_{i_k}^{(k)} + \sum_{k<\ell} \alpha_{i_k i_\ell}^{(k\ell)} + \sum_{k<\ell<m} \alpha_{i_k i_\ell i_m}^{(k\ell m)} + \tilde{\varepsilon}_{na}, \tag{7.1}$$

with direct response (utility) $\tilde{Y}_{na}(\mathbf{i})$ where $\alpha_{i_k}^{(k)}$ is the main effects of the k-th attribute when the corresponding level is $i_k = 1, \ldots, v$ for $k = 1, \ldots, K$ in total, $\alpha_{i_k i_\ell}^{(k\ell)}$ is the first-order interaction effects of the k-th and ℓ-th attribute when the corresponding levels are $i_k = 1, \ldots, v$ and $i_\ell = 1, \ldots, v$, respectively, and $\alpha_{i_k i_\ell i_m}^{(k\ell m)}$ is the second-order interaction effects of the k-th, ℓ-th and m-th attribute when the corresponding levels are $i_k = 1, \ldots, v, i_\ell = 1, \ldots, v$ and $i_m = 1, \ldots, v$, respectively. As a result, by the common identifiability conditions of effects-coding, in particular for the second-order interactions effects $\alpha_{i_k i_\ell i_m}^{(k\ell m)} = \beta_{i_k i_\ell i_m}^{(k\ell m)}$ for $i_k, i_\ell, i_m = 1, \ldots, v-1$ the following equalities hold:

$$\alpha_{i_k i_\ell v}^{(k\ell m)} = -\sum_{i_m=1}^{v-1} \beta_{i_k i_\ell i_m}^{(k\ell m)}, i_k, i_\ell = 1, \ldots, v-1,\ \alpha_{i_k v i_m}^{(k\ell m)} = -\sum_{i_\ell=1}^{v-1} \beta_{i_k i_\ell i_m}^{(k\ell m)}, i_k, i_m = 1, \ldots, v-1,$$

$$\alpha_{v i_\ell i_m}^{(k\ell m)} = -\sum_{i_k=1}^{v-1} \beta_{i_k i_\ell i_m}^{(k\ell m)}, i_\ell, i_m = 1, \ldots, v-1,\ \alpha_{i_k v v}^{(k\ell m)} = \sum_{i_\ell=1}^{v-1} \sum_{i_m=1}^{v-1} \beta_{i_k i_\ell i_m}^{(k\ell m)}, i_k = 1, \ldots, v-1,$$

$$\alpha_{v i_\ell v}^{(k\ell m)} = \sum_{i_k=1}^{v-1} \sum_{i_m=1}^{v-1} \beta_{i_k i_\ell i_m}^{(k\ell m)}, i_\ell = 1, \ldots, v-1, \alpha_{v v i_m}^{(k\ell m)} = \sum_{i_k=1}^{v-1} \sum_{i_\ell=1}^{v-1} \beta_{i_k i_\ell i_m}^{(k\ell m)}, i_m = 1, \ldots, v-1, \text{and}$$

$$\alpha_{vvv}^{(k\ell m)} = -\sum_{i_k=1}^{v-1} \sum_{i_\ell=1}^{v-1} \sum_{i_m=1}^{v-1} \beta_{i_k i_\ell i_m}^{(k\ell m)}.$$

The parameters for the main effects, the first-order interactions and the second-order interactions, respectively, can be summarized as follows

$$\boldsymbol{\beta}_k = (\beta^{(k)}_{i_k})_{i_k=1,\ldots,v-1},\ \boldsymbol{\beta}_{k\ell} = (\beta^{(k\ell)}_{i_k i_\ell})_{i_k=1,\ldots,v-1,\ i_\ell=1,\ldots,v-1} \text{ and}$$
$$\boldsymbol{\beta}_{k\ell m} = (\beta^{(k\ell m)}_{i_k i_\ell i_m})_{i_k=1,\ldots,v-1,\ i_\ell=1,\ldots,v-1,\ i_m=1,\ldots,v-1},$$

where e. g. $\boldsymbol{\beta}_{k\ell m}$ describes the effect of the second-order interaction of the k-th, ℓ-th and m-th attribute. Then the vector of parameters of dimension $p = K(v-1) + \binom{K}{2}(v-1)^2 + \binom{K}{3}(v-1)^3$ is given by

$$\boldsymbol{\beta} = ((\boldsymbol{\beta}_k)^\top_{k=1,\ldots,K}, (\boldsymbol{\beta}_{k\ell})^\top_{k<\ell}, (\boldsymbol{\beta}_{k\ell m})^\top_{k<\ell<m})^\top. \tag{7.2}$$

With the above notation the model (6.2) can be rewritten as

$$\begin{aligned}\tilde{Y}_{na}(\mathbf{i}) = \mu_n + \sum_{k=1}^{K} \mathbf{f}_1(i_k)^\top \boldsymbol{\beta}_k + \sum_{k<\ell} (\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell))^\top \boldsymbol{\beta}_{k\ell} \\ + \sum_{k<\ell<m} (\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) \otimes \mathbf{f}_1(i_m))^\top \boldsymbol{\beta}_{k\ell m} + \tilde{\varepsilon}_{na},\end{aligned} \tag{7.3}$$

which results in the vector

$$\begin{aligned}\mathbf{f}(\mathbf{i}) = (\mathbf{f}_1(i_1)^\top, \ldots, \mathbf{f}_1(i_K)^\top, \mathbf{f}_1(i_1)^\top \otimes \mathbf{f}_1(i_2)^\top, \ldots, \mathbf{f}_1(i_{K-1}) \otimes \mathbf{f}_1(i_K)^\top, \\ \mathbf{f}_1(i_1)^\top \otimes \mathbf{f}_1(i_2)^\top \otimes \mathbf{f}_1(i_3)^\top, \ldots, \mathbf{f}_1(i_{K-2})^\top \otimes \mathbf{f}_1(i_{K-1})^\top \otimes \mathbf{f}_1(i_K)^\top)^\top\end{aligned} \tag{7.4}$$

of dimension p. Here the first K components $\mathbf{f}_1(i_1), \ldots, \mathbf{f}_1(i_K)$ of $\mathbf{f}(\mathbf{i})$ are associated with the main effects and have $p_1 = K(v-1)$, the second components $\mathbf{f}_1(i_1) \otimes \mathbf{f}_1(i_2), \ldots, \mathbf{f}_1(i_{K-1}) \otimes \mathbf{f}_1(i_K)$ of $\mathbf{f}(\mathbf{i})$ are associated with the first-order interactions and have $p_2 = (1/2)K(K-1)(v-1)^2$, and the remaining components $\mathbf{f}_1(i_1) \otimes \mathbf{f}_1(i_2) \otimes \mathbf{f}_3(i_3), \ldots, \mathbf{f}_1(i_{K-2}) \otimes \mathbf{f}_1(i_{K-1}) \otimes \mathbf{f}_1(i_K)$ of $\mathbf{f}(\mathbf{i})$ are associated with the second-order interactions and have $p_3 = (1/6)K(K-1)(K-2)(v-1)^3$.

The corresponding paired comparison model is given by

$$\begin{aligned}Y_n(\mathbf{i},\mathbf{j}) = \sum_{k=1}^{K} (\mathbf{f}_1(i_k) - \mathbf{f}_1(j_k))^\top \boldsymbol{\beta}_k + \sum_{k<\ell} ((\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell)) - (\mathbf{f}_1(j_k) \otimes \mathbf{f}_1(j_\ell)))^\top \boldsymbol{\beta}_{k\ell} \\ + \sum_{k<\ell<m} ((\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) \otimes \mathbf{f}_1(i_m)) - (\mathbf{f}_1(j_k) \otimes \mathbf{f}_1(j_\ell) \otimes \mathbf{f}_1(j_m)))^\top \boldsymbol{\beta}_{k\ell m} + \varepsilon_n.\end{aligned} \tag{7.5}$$

7.1 Designs for Full Profiles

In the following we consider the second-order interaction paired comparison model (7.5) with corresponding regression functions $\mathbf{f}(\mathbf{i})$ in (7.4), and present some optimality results in the class of invariant designs which are uniform on their orbits of fixed comparison depth d.

For paired alternatives $\mathbf{i} = (i_1, \ldots, i_K)$ and $\mathbf{j} = (j_1, \ldots, j_K)$ the design region $\mathcal{X}$ is defined as

$$\mathcal{X} = \{(\mathbf{i}, \mathbf{j}); i_k, j_k \in \{1, \ldots, v\} \text{ for } k = 1, \ldots, K \text{ attributes}\}, \tag{7.6}$$

which can be partitioned into disjoint sets such that the pairs in each set differ only in a specified number of the attributes. As already noted these sets constitute the orbits with respect to permutations of both the levels $i_k, j_k = 1, \ldots, v$ within the attributes as well as among attributes $k = 1, \ldots, K$, themselves. Hence for the comparison depth $d = 0, \ldots, K$, we similarly formulate the set $\mathcal{X}_d$ in (6.5) as

$$\mathcal{X}_d = \{(\mathbf{i}, \mathbf{j}) \in \mathcal{X} : |\{k : i_k \neq j_k\}| = d\} \tag{7.7}$$

with a total number of $N_d = \binom{K}{d} v^K (v-1)^d$ pairs which vary in exactly d attributes and denote by ξ_d the uniform approximate design which assigns equal weight $\xi_d(\mathbf{i}, \mathbf{j}) = 1/N_d$ to each pair in $\mathcal{X}_d$ and weight zero to all remaining pairs in $\mathcal{X}$.

We next derive the information matrices for the aforementioned invariant designs. To begin with, we note that $\mathbf{M} = \frac{2}{v-1}(\mathbf{Id}_{v-1} + \mathbf{1}_{v-1}\mathbf{1}_{v-1}^\top)$ is the information matrix of the corresponding one-way layout in (4.5).

Lemma 7.1. *Let d be a fixed comparison depth. The uniform design ξ_d on the set $\mathcal{X}_d$ of comparison depth d has block diagonal information matrix*

$$\mathbf{M}(\xi_d) = \begin{pmatrix} h_1(d)\mathbf{Id}_{p_1} \otimes \mathbf{M} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & h_2(d)\mathbf{Id}_{p_2} \otimes \mathbf{M} \otimes \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & h_3(d)\mathbf{Id}_{p_3} \otimes \mathbf{M} \otimes \mathbf{M} \otimes \mathbf{M} \end{pmatrix},$$

where

$$h_1(d) = \frac{d}{K}, \; h_2(d) = \frac{d}{2vK(K-1)}(2Kv - 2K - dv - v + 2) \text{ and}$$

$$h_3(d) = \frac{d}{4v^2K(K-1)(K-2)}(3K^2 + 3K^2v^2 - 6K^2v - 3Kdv^2 + 3Kdv - 6Kv^2 + 15Kv - 9K + d^2v^2 + 3dv^2 - 6dv + 2v^2 - 6v + 6).$$

We note that the results for $v = 2$ in Lemma 6.1 can be obtained as a special case by replacing $h_q(d)$ for $q = 1, 2, 3$ by $4^q h_q(d)$ because $\mathbf{M} = 4$.

Proof. In view of Lemma 1 of Graßhoff et al. (2003) it follows that $\sum_{i=1}^{v} \mathbf{f}_1(i)\mathbf{f}_1(i)^\top = \frac{v-1}{2}\mathbf{M}$ and $\sum_{i \neq j} \mathbf{f}_1(i)\mathbf{f}_1(j)^\top = -\frac{v-1}{2}\mathbf{M}$. Hence

$$\begin{aligned}
&\sum_{i \neq j} (\mathbf{f}_1(i) - \mathbf{f}_1(j))(\mathbf{f}_1(i) - \mathbf{f}_1(j))^\top \\
&\quad = \sum_{i=1}^{v} \sum_{j=1, j \neq i}^{v} (\mathbf{f}_1(i)\mathbf{f}_1(i)^\top + \mathbf{f}_1(j)\mathbf{f}_1(j)^\top - \mathbf{f}_1(i)\mathbf{f}_1(j)^\top - \mathbf{f}_1(j)\mathbf{f}_1(i)^\top) \\
&\quad = (v-1)(v-1)\mathbf{M} + (v-1)\mathbf{M} \\
&\quad = v(v-1)\mathbf{M}.
\end{aligned}$$

For the regression functions $\mathbf{f}_k = \mathbf{f}_1$ associated with the k-th main effect we have

$$\begin{aligned}
&\sum_{(\mathbf{i},\mathbf{j}) \in \mathcal{X}_d} (\mathbf{f}_k(i_k) - \mathbf{f}_k(j_k))(\mathbf{f}_k(i_k) - \mathbf{f}_k(j_k))^\top \\
&\quad = \binom{K-1}{d-1} v^{K-1}(v-1)^{d-1} \sum_{i_k \neq j_k} (\mathbf{f}_1(i_k) - \mathbf{f}_1(j_k))(\mathbf{f}_1(i_k) - \mathbf{f}_1(j_k))^\top \\
&\quad = \binom{K-1}{d-1} v^{K}(v-1)^{d}\mathbf{M},
\end{aligned}$$

because there are $\binom{K-1}{d-1} v^{K-1}(v-1)^{d-1}$ pairs in $\mathcal{X}_d$ for which i_k and j_k differ.

Now as the number N_d of paired comparisons in $\mathcal{X}_d$ with comparison depth d equals $N_d = \binom{K}{d} v^K (v-1)^d$ for every attribute k, the corresponding diagonal elements $h_1(d)$ are given by

$$\begin{aligned}
h_1(d) &= \frac{1}{N_d} \sum_{(\mathbf{i},\mathbf{j}) \in \mathcal{X}_d} (\mathbf{f}_k(i_k) - \mathbf{f}_k(j_k))(\mathbf{f}_k(i_k) - \mathbf{f}_k(j_k))^\top \\
&= \frac{1}{N_d} \binom{K-1}{d-1} v^K (v-1)^d \mathbf{M} = \frac{d}{K}\mathbf{M} \qquad (7.8)
\end{aligned}$$

in the information matrix.

For first-order interactions, we similarly consider attributes k and ℓ, say, and distinguish between pairs in which both attributes are distinct and pairs in which only one of these attributes has distinct levels in the alternatives while the same level is presented in both alternatives for the other attribute:

$$\begin{aligned}
&\sum_{i_k \neq j_k} \sum_{i_\ell \neq j_\ell} (\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) - \mathbf{f}_1(j_k) \otimes \mathbf{f}_1(j_\ell))(\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) - \mathbf{f}_1(j_k) \otimes \mathbf{f}_1(j_\ell))^\top \\
&\quad = \sum_{i_k=1}^{v} \sum_{j_k \neq i_k} \sum_{i_\ell=1}^{v} \sum_{j_\ell \neq i_\ell} (\mathbf{f}_1(i_k)\mathbf{f}_1(i_k)^\top \otimes \mathbf{f}_1(i_\ell)\mathbf{f}_1(i_\ell)^\top + \mathbf{f}_1(j_k)\mathbf{f}_1(j_k)^\top \otimes \mathbf{f}_1(j_\ell)\mathbf{f}_1(j_\ell)^\top \\
&\qquad\qquad - \mathbf{f}_1(i_k)\mathbf{f}_1(j_k)^\top \otimes \mathbf{f}_1(i_\ell)\mathbf{f}_1(j_\ell)^\top - \mathbf{f}_1(j_k)\mathbf{f}_1(i_k)^\top \otimes \mathbf{f}_1(j_\ell)\mathbf{f}_1(i_\ell)^\top) \\
&\quad = 2(v-1)^2 \sum_{i_k=1}^{v} \mathbf{f}_1(i_k)\mathbf{f}_1(i_k)^\top \otimes \sum_{i_\ell=1}^{v} \mathbf{f}_1(i_\ell)\mathbf{f}_1(i_\ell)^\top \\
&\qquad\qquad - 2 \sum_{i_k \neq j_k} \mathbf{f}_1(i_k)\mathbf{f}_1(j_k)^\top \otimes \sum_{i_\ell \neq j_\ell} \mathbf{f}_1(i_\ell)\mathbf{f}_1(j_\ell)^\top \\
&\quad = \frac{1}{2} v(v-1)^2(v-2)\mathbf{M} \otimes \mathbf{M}
\end{aligned} \tag{7.9}$$

and

$$\begin{aligned}
&\sum_{i_k \neq j_k} \sum_{i_\ell = j_\ell} (\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) - \mathbf{f}_1(j_k) \otimes \mathbf{f}_1(j_\ell))(\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) - \mathbf{f}_1(j_k) \otimes \mathbf{f}_1(j_\ell))^\top \\
&= \sum_{i_k=1}^{v} \sum_{j_k \neq i_k} \sum_{i_\ell=1}^{v} \sum_{j_\ell = i_\ell} (\mathbf{f}_1(i_k)\mathbf{f}_1(i_k)^\top \otimes \mathbf{f}_1(i_\ell)\mathbf{f}_1(i_\ell)^\top + \mathbf{f}_1(j_k)\mathbf{f}_1(j_k)^\top \otimes \mathbf{f}_1(j_\ell)\mathbf{f}_1(j_\ell)^\top \\
&\qquad\qquad - \mathbf{f}_1(i_k)\mathbf{f}_1(j_k)^\top \otimes \mathbf{f}_1(i_\ell)\mathbf{f}_1(j_\ell)^\top - \mathbf{f}_1(j_k)\mathbf{f}_1(i_k)^\top \otimes \mathbf{f}_1(j_\ell)\mathbf{f}_1(i_\ell)^\top) \\
&= 2(v-1) \sum_{i_k=1}^{v} \mathbf{f}_1(i_k)\mathbf{f}_1(i_k)^\top \otimes \sum_{i_\ell=1}^{v} \mathbf{f}_1(i_\ell)\mathbf{f}_1(i_\ell)^\top - 2 \sum_{i_k \neq j_k} \mathbf{f}_1(i_k)\mathbf{f}_1(j_k)^\top \otimes \sum_{i_\ell = j_\ell} \mathbf{f}_1(i_\ell)\mathbf{f}_1(j_\ell)^\top \\
&= \frac{1}{2} v(v-1)^2 \mathbf{M} \otimes \mathbf{M},
\end{aligned} \tag{7.10}$$

respectively.

For given attributes k and ℓ the pairs with distinct levels in both attributes occur $\binom{K-2}{d-2} v^{K-2}(v-1)^{d-2}$ times in $\mathcal{X}_d$, while those which differ only in one attribute occur $\binom{2}{1}\binom{K-2}{d-1} v^{K-2}(v-1)^{d-1}$ times in $\mathcal{X}_d$. As a result, for the first-order interactions the diagonal elements $h_2(d)$ in the information matrix are given by

$$\begin{aligned}
h_2(d) &= \frac{1}{N_d}\left(\frac{1}{2}\binom{K-2}{d-2} v(v-1)^2(v-2)v^{K-2}(v-1)^{d-2}\mathbf{M}\otimes\mathbf{M}\right.\\
&\qquad\left. + \binom{K-2}{d-1} v(v-1)^2 v^{K-2}(v-1)^{d-1}\mathbf{M}\otimes\mathbf{M}\right)\\
&= \left(\tfrac{d(d-1)}{2vK(K-1)}(v-2) + \tfrac{d(K-d)}{vK(K-1)}(v-1)\right)\mathbf{M}\otimes\mathbf{M}\\
&= \frac{d}{2vK(K-1)}\Big((v-2)(d-1) + 2(K-d)(v-1)\Big)\mathbf{M}\otimes\mathbf{M}\\
&= \frac{d}{2vK(K-1)}(2Kv - 2K - dv - v + 2)\mathbf{M}\otimes\mathbf{M}. \qquad (7.11)
\end{aligned}$$

Further for the second-order interactions, we similarly consider attributes k, ℓ and m, say, and distinguish between pairs in which all three attributes are distinct, pairs in which two of these attributes k and ℓ, say, have distinct levels in the alternatives while the same level is presented in both alternatives for the remaining attribute and, finally, pairs in which only one of the attributes, say, k has distinct levels in the alternatives while the same level is presented in both alternatives for the two remaining attributes:

$$\begin{aligned}
&\sum_{i_k\neq j_k}\sum_{i_\ell\neq j_\ell}\sum_{i_m\neq j_m}(\mathbf{f}_1(i_k)\otimes\mathbf{f}_1(i_\ell)\otimes\mathbf{f}_1(i_m) - \mathbf{f}_1(j_k)\otimes\mathbf{f}_1(j_\ell)\otimes\mathbf{f}_1(j_m))\\
&\qquad\qquad\cdot(\mathbf{f}_1(i_k)\otimes\mathbf{f}_1(i_\ell)\otimes\mathbf{f}_1(i_m) - \mathbf{f}_1(j_k)\otimes\mathbf{f}_1(j_\ell)\otimes\mathbf{f}_1(j_m))^\top\\
&= \sum_{i_k=1}^{v}\sum_{j_k\neq i_k}\sum_{i_\ell=1}^{v}\sum_{j_\ell\neq i_\ell}\sum_{i_m=1}^{v}\sum_{j_m\neq i_m}(\mathbf{f}_1(i_k)\mathbf{f}_1(i_k)^\top\otimes\mathbf{f}_1(i_\ell)\mathbf{f}_1(i_\ell)^\top\otimes\mathbf{f}_1(i_m)\mathbf{f}_1(i_m)^\top\\
&\quad + \mathbf{f}_1(j_k)\mathbf{f}_1(j_k)^\top\otimes\mathbf{f}_1(j_\ell)\mathbf{f}_1(j_\ell)^\top\otimes\mathbf{f}_1(j_m)\mathbf{f}_1(j_m)^\top\\
&\quad - \mathbf{f}_1(i_k)\mathbf{f}_1(j_k)^\top\otimes\mathbf{f}_1(i_\ell)\mathbf{f}_1(j_\ell)^\top\otimes\mathbf{f}_1(i_m)\mathbf{f}_1(j_m)^\top\\
&\quad - \mathbf{f}_1(j_k)\mathbf{f}_1(i_k)^\top\otimes\mathbf{f}_1(j_\ell)\mathbf{f}_1(i_\ell)^\top\otimes\mathbf{f}_1(j_m)\mathbf{f}_1(i_m)^\top)\\
&= 2(v-1)^3\sum_{i_k=1}^{v}\mathbf{f}_1(i_k)\mathbf{f}_1(i_k)^\top\otimes\sum_{i_\ell=1}^{v}\mathbf{f}_1(i_\ell)\mathbf{f}_1(i_\ell)^\top\otimes\sum_{i_m=1}^{v}\mathbf{f}_1(i_m)\mathbf{f}_1(i_m)^\top\\
&\quad - 2\sum_{i_k\neq j_k}\mathbf{f}_1(i_k)\mathbf{f}_1(j_k)^\top\otimes\sum_{i_\ell\neq j_\ell}\mathbf{f}_1(i_\ell)\mathbf{f}_1(j_\ell)^\top\otimes\sum_{i_m\neq j_m}\mathbf{f}_1(i_m)\mathbf{f}_1(j_m)^\top\\
&= \frac{1}{4}v(v-1)^3(v^2-3v+3)\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M}, \qquad (7.12)
\end{aligned}$$

also

$$
\begin{aligned}
&\sum_{i_k \neq j_k} \sum_{i_\ell \neq j_\ell} \sum_{i_m = j_m} (\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) \otimes \mathbf{f}_1(i_m) - \mathbf{f}_1(j_k) \otimes \mathbf{f}_1(j_\ell) \otimes \mathbf{f}_1(j_m)) \\
&\qquad\qquad\qquad \cdot (\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) \otimes \mathbf{f}_1(i_m) - \mathbf{f}_1(j_k) \otimes \mathbf{f}_1(j_\ell) \otimes \mathbf{f}_1(j_m))^\top \\
&= \sum_{i_k=1}^{v} \sum_{j_k \neq i_k} \sum_{i_\ell=1}^{v} \sum_{j_\ell \neq i_\ell} \sum_{i_m=1}^{v} \sum_{j_m = i_m} (\mathbf{f}_1(i_k)\mathbf{f}_1(i_k)^\top \otimes \mathbf{f}_1(i_\ell)\mathbf{f}_1(i_\ell)^\top \otimes \mathbf{f}_1(i_m)\mathbf{f}_1(i_m)^\top \\
&\qquad + \mathbf{f}_1(j_k)\mathbf{f}_1(j_k)^\top \otimes \mathbf{f}_1(j_\ell)\mathbf{f}_1(j_\ell)^\top \otimes \mathbf{f}_1(j_m)\mathbf{f}_1(j_m)^\top \\
&\qquad - \mathbf{f}_1(i_k)\mathbf{f}_1(j_k)^\top \otimes \mathbf{f}_1(i_\ell)\mathbf{f}_1(j_\ell)^\top \otimes \mathbf{f}_1(i_m)\mathbf{f}_1(j_m)^\top \\
&\qquad - \mathbf{f}_1(j_k)\mathbf{f}_1(i_k)^\top \otimes \mathbf{f}_1(j_\ell)\mathbf{f}_1(i_\ell)^\top \otimes \mathbf{f}_1(j_m)\mathbf{f}_1(i_m)^\top) \\
&= 2(v-1)^2 \sum_{i_k=1}^{v} \mathbf{f}_1(i_k)\mathbf{f}_1(i_k)^\top \otimes \sum_{i_\ell=1}^{v} \mathbf{f}_1(i_\ell)\mathbf{f}_1(i_\ell)^\top \otimes \sum_{i_m=1}^{v} \mathbf{f}_1(i_m)\mathbf{f}_1(i_m)^\top \\
&\qquad - 2 \sum_{i_k \neq j_k} \mathbf{f}_1(i_k)\mathbf{f}_1(j_k)^\top \otimes \sum_{i_\ell \neq j_\ell} \mathbf{f}_1(i_\ell)\mathbf{f}_1(j_\ell)^\top \otimes \sum_{i_m = j_m} \mathbf{f}_1(i_m)\mathbf{f}_1(j_m)^\top \\
&= \frac{1}{4} v(v-1)^3(v-2)\mathbf{M} \otimes \mathbf{M} \otimes \mathbf{M},
\end{aligned} \tag{7.13}
$$

and

$$
\begin{aligned}
&\sum_{i_k \neq j_k} \sum_{i_\ell = j_\ell} \sum_{i_m = j_m} (\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) \otimes \mathbf{f}_1(i_m) - \mathbf{f}_1(j_k) \otimes \mathbf{f}_1(j_\ell) \otimes \mathbf{g}(j_m)) \\
&\qquad\qquad\qquad \cdot (\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) \otimes \mathbf{f}_1(i_m) - \mathbf{f}_1(j_k) \otimes \mathbf{f}_1(j_\ell) \otimes \mathbf{f}_1(j_m))^\top \\
&= \sum_{i_k=1}^{v} \sum_{j_k \neq i_k} \sum_{i_\ell=1}^{v} \sum_{j_\ell = i_\ell} \sum_{i_m=1}^{v} \sum_{j_m = i_m} (\mathbf{f}_1(i_k)\mathbf{f}_1(i_k)^\top \otimes \mathbf{f}_1(i_\ell)\mathbf{f}_1(i_\ell)^\top \otimes \mathbf{f}_1(i_m)\mathbf{f}_1(i_m)^\top \\
&\qquad + \mathbf{f}_1(j_k)\mathbf{f}_1(j_k)^\top \otimes \mathbf{f}_1(j_\ell)\mathbf{f}_1(j_\ell)^\top \otimes \mathbf{f}_1(j_m)\mathbf{f}_1(j_m)^\top \\
&\qquad - \mathbf{f}_1(i_k)\mathbf{f}_1(j_k)^\top \otimes \mathbf{f}_1(i_\ell)\mathbf{f}_1(j_\ell)^\top \otimes \mathbf{f}_1(i_m)\mathbf{f}_1(j_m)^\top \\
&\qquad - \mathbf{f}_1(j_k)\mathbf{f}_1(i_k)^\top \otimes \mathbf{f}_1(j_\ell)\mathbf{f}_1(i_\ell)^\top \otimes \mathbf{f}_1(j_m)\mathbf{f}_1(i_m)^\top) \\
&= 2(v-1) \sum_{i_k=1}^{v} \mathbf{f}_1(i_k)\mathbf{f}_1(i_k)^\top \otimes \sum_{i_\ell=1}^{v} \mathbf{f}_1(i_\ell)\mathbf{f}_1(i_\ell)^\top \otimes \sum_{i_m=1}^{v} \mathbf{f}_1(i_m)\mathbf{f}_1(i_m)^\top \\
&\qquad - 2 \sum_{i_k \neq j_k} \mathbf{f}_1(i_k)\mathbf{f}_1(j_k)^\top \otimes \sum_{i_\ell = j_\ell} \mathbf{f}_1(i_\ell)\mathbf{f}_1(j_\ell)^\top \otimes \sum_{i_m = j_m} \mathbf{f}_1(i_m)\mathbf{f}_1(j_m)^\top \\
&= \frac{1}{4} v(v-1)^3\mathbf{M} \otimes \mathbf{M} \otimes \mathbf{M},
\end{aligned} \tag{7.14}
$$

respectively.

For given attributes k, ℓ and m the pairs with distinct levels in the three attributes occur $\binom{K-3}{d-3} v^{K-3}(v-1)^{d-3}$ times in $\mathcal{X}_d$, while those which differ in two attributes occur $\binom{3}{2}\binom{K-3}{d-2} v^{K-3}(v-1)^{d-2}$ times in $\mathcal{X}_d$ and, finally, those which differ only in one attribute occur $\binom{3}{1}\binom{K-3}{d-1} v^{K-3}(v-1)^{d-1}$ times in $\mathcal{X}_d$. As a result, for the second-order interactions the diagonal elements $h_3(d)$ in the information matrix are given by

$$
\begin{aligned}
h_3(d) = \frac{1}{N_d}\Bigg(&\frac{1}{4}\binom{K-3}{d-3} v(v-1)^3(v^2-3v+3)v^{K-3}(v-1)^{d-3}\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M} \\
&+\frac{3}{4}\binom{K-3}{d-2} v(v-1)^3(v-2)v^{K-3}(v-1)^{d-2}\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M} \\
&+\frac{3}{4}\binom{K-3}{d-1} v(v-1)^3 v^{K-3}(v-1)^{d-1}\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M}\Bigg) \\
= \Bigg(&\frac{d(d-1)(d-2)}{4v^2K(K-1)(K-2)}(v^2-3v+3) + \frac{3(K-d)d(d-1)}{4v^2K(K-1)(K-2)}(v-1)(v-2) \\
&+\frac{3(K-d)(K-d-1)d}{4v^2K(K-1)(K-2)}(v-1)^2\Bigg)\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M} \\
= &\frac{d}{4v^2K(K-1)(K-2)}(3K^2+3K^2v^2-6K^2v-3Kdv^2+3Kdv-6Kv^2 \\
&+15Kv-9K+d^2v^2+3dv^2-6dv+2v^2-6v+6)\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M}.
\end{aligned}
\tag{7.15}
$$

Finally, it can be noted that all off-diagonal entries in the information matrix vanish because the terms in the corresponding sums add up to zero due to the effects-type coding. □

Here we similarly note that the corresponding functions $h_q(0) = 0$ for $q = 1, 2, 3$. Further we mention that for the situation of first-order interactions the corresponding results of $h_1(d)$ and $h_2(d)$ can be found in Chapter 5.

In the following we give the information matrix for a general invariant design ξ. As already mentioned the invariant design ξ can be written as convex combination $\xi = \sum_{d=1}^{K} w_d\xi_d$ of uniform designs on the comparison depth d with corresponding weights $w_d \geq 0$, $\sum_{d=1}^{K} w_d = 1$.

Lemma 7.2. *Let ξ be an invariant design on $\mathcal{X}$, i. e. $\xi = \sum_{d=1}^{K} w_d\xi_d$, then ξ has block diagonal information matrix*

$$
\mathbf{M}(\xi) = \begin{pmatrix} h_1(\xi)\mathbf{Id}_{p_1}\otimes\mathbf{M} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & h_2(\xi)\mathbf{Id}_{p_2}\otimes\mathbf{M}\otimes\mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & h_3(\xi)\mathbf{Id}_{p_3}\otimes\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M} \end{pmatrix},
$$

where $h_q(\xi) = \sum_{d=1}^{K} w_d h_q(d)$, $q = 1, 2, 3$.

Now in the following Theorem 7.1, Theorem 7.2 and Remark 7.1 we consider optimal designs for the main effects, the first-order interaction and the second-order interaction terms having corresponding entries $h_1(d)$, $h_2(d)$ and $h_3(d)$, respectively, in the information matrix $\mathbf{M}(\xi_d)$ in Lemma 7.1. As already mentioned the resulting designs may optimize every design criterion which is invariant with respect to both permutations of the levels and permutations of the attributes subject to the full parameter vector, satisfying the corresponding identifiabity conditions. As a consequence, the reduced parameter vector $\boldsymbol{\beta} = ((\boldsymbol{\beta}_k)^\top_{1\leq k\leq K}, (\boldsymbol{\beta}_{k\ell})^\top_{k<\ell}, (\boldsymbol{\beta}_{k\ell m})^\top_{k<\ell<m})^\top$ in (7.2) considered separately in the underlying theorems are also invariant with respect to the D-criterion. Here we mention that Theorems 7.1 and 7.2 paraphrase theorems given in Graßhoff et al. (2003) for first-order interaction models and translate them to the present setting of second-order interaction models.

Theorem 7.1. *Let $d_1^* = K$ for $v \geq 2$. Then the uniform design $\xi_{d_1^*} = \xi_K$ on the largest possible comparison depth K is D-optimal for the main effects $(\boldsymbol{\beta}_k)^\top_{1\leq k\leq K}$.*

This means that for the main effects it is sufficient to use only those pairs of alternatives which differ in all attributes.

Moreover, for the first-order interactions attributes with distinct levels does not provide enough information. Hence, we consider the intermediate comparison depth $d^* = K - 1 - \left[\frac{K-2}{v}\right]$ where $[u]$ denotes the integer part of the decimal expansion for u satisfying $[u] \leq u < [u] + 1$.

Theorem 7.2. *Let $d_2^* = K - 1 - \left[\frac{K-2}{v}\right]$ for $v \geq 2$. Then the uniform design $\xi_{d_2^*}$ is D-optimal for the first-order interaction effects $(\boldsymbol{\beta}_{k\ell})^\top_{k<\ell}$.*

This means that for the first-order interactions it is sufficient to use only those pairs of alternatives which differ in approximately half of the attributes.

Remark 7.1. *There exists a single comparison depth d_3^* such that the uniform design $\xi_{d_3^*}$ is D-optimal for the second-order interaction effects $(\boldsymbol{\beta}_{k\ell m})^\top_{k<\ell<m}$.*

In the following Table 7.1 we note that the corresponding values of d_3^* were obtained by first calculating the values of $h_3(d)$ and determining the maximum. It is worthwhile mentioning that generally for moderate values of v the optimal comparison depth $d_3^* = K$ but this is not true for $K = 3$. Moreover, the optimal comparison depth $d_3^* = K - 2$ for sufficiently large values of v.

Table 7.1: Values of the Optimal Comparison Depths d_3^* of the D-Optimal Uniform Designs $\xi_{d_3^*}$ for the Second-Order Interactions with K Attributes and v-Levels

	v									
K	2	3	4	5	6	7	8	9	10	20
3	1	1	1	1	1	1	1	1	1	1
4	4	1	2	2	2	2	2	2	2	2
5	5	2	2	3	3	3	3	3	3	3
6	6	6	3	3	3	4	4	4	4	4
7	7	7	7	4	4	4	4	5	5	5
8	8	8	8	5	5	5	5	5	6	6
9	9	9	9	9	6	6	6	6	6	7
10	10	10	10	10	6	7	7	7	7	8

In the following Theorem 7.3 and Corollary 7.1 we present the variance function for the corresponding invariant designs. We define the variance function for the invariant designs ξ as $V((\mathbf{i},\mathbf{j}),\xi) = (\mathbf{f}(\mathbf{i})-\mathbf{f}(\mathbf{j}))^\top \mathbf{M}(\xi)^{-1}(\mathbf{f}(\mathbf{i})-\mathbf{f}(\mathbf{j}))$, which as similarly characterized in Theorem 6.4 is also invariant with respect to permutation of levels and attributes and, hence, constant on the orbits $\mathcal{X}_d$ of fixed comparison depth d. The value of the variance function for the invariant design ξ evaluated at comparison depth d can also be denoted as $V(d,\xi)$ where $V(d,\xi) = V((\mathbf{i},\mathbf{j}),\xi)$ on $\mathcal{X}_d$.

Theorem 7.3. *For every invariant design ξ the variance function $V(d,\xi)$ is given by*

$$V(d,\xi) = d(v-1)\left(\frac{1}{h_1(\xi)} + \frac{v-1}{4vh_2(\xi)}(2Kv-2K-dv-v+2) + \frac{(v-1)^2}{24v^2h_3(\xi)}\lambda(d)\right),$$

where

$$\lambda(d) = 3K^2 + 3K^2v^2 - 6K^2v - 3Kdv^2 + 3Kdv - 6Kv^2 + 15Kv \\ - 9K + d^2v^2 + 3dv^2 - 6dv + 2v^2 - 6v + 6.$$

It should be noted that the results for $v = 2$ in Theorem 6.4 can be obtained as a special case.

Proof. In view of Remark 4.1 it is sufficient to note that for the k-th main effects the variance function is given by

$$\begin{aligned}
&(\mathbf{f}_1(i_k) - \mathbf{f}_1(j_k))^\top \mathbf{M}^{-1}(\mathbf{f}_1(i_k) - \mathbf{f}_1(j_k)) \\
&\quad = \mathbf{f}_1(i_k)^\top \mathbf{M}^{-1}\mathbf{f}_1(i_k) + \mathbf{f}_1(j_k)^\top \mathbf{M}^{-1}\mathbf{f}_1(j_k) - \mathbf{f}_1(i_k)^\top \mathbf{M}^{-1}\mathbf{f}_1(j_k) - \mathbf{f}_1(j_k)^\top \mathbf{M}^{-1}\mathbf{f}_1(i_k) \\
&\quad = \frac{(v-1)^2}{v} + \frac{(v-1)}{v} \\
&\quad = v - 1. \qquad (7.16)
\end{aligned}$$

Further for the regression function associated with the first-order interactions of the attributes k and ℓ, say, we obtain

$$\begin{aligned}
&(\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) - \mathbf{f}_1(j_k) \otimes \mathbf{f}_1(j_\ell))^\top \mathbf{M}^{-1} \otimes \mathbf{M}^{-1}(\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) - \mathbf{f}_1(j_k) \otimes \mathbf{f}_1(j_\ell)) \\
&\quad = \mathbf{f}_1(i_k)^\top \mathbf{M}^{-1}\mathbf{f}_1(i_k) \cdot \mathbf{f}_1(i_\ell)^\top \mathbf{M}^{-1}\mathbf{f}_1(i_\ell) + \mathbf{f}_1(j_k)^\top \mathbf{M}^{-1}\mathbf{f}_1(j_k) \cdot \mathbf{f}_1(j_\ell)^\top \mathbf{M}^{-1}\mathbf{f}_1(j_\ell) \\
&\qquad - \mathbf{f}_1(i_k)^\top \mathbf{M}^{-1}\mathbf{f}_1(j_k) \cdot \mathbf{f}_1(i_\ell)^\top \mathbf{M}^{-1}\mathbf{f}_1(j_\ell) - \mathbf{f}_1(j_k)^\top \mathbf{M}^{-1}\mathbf{f}_1(i_k) \cdot \mathbf{f}_1(j_\ell)^\top \mathbf{M}^{-1}\mathbf{f}_1(i_\ell) \\
&\quad = \begin{cases} \dfrac{(v-1)^2(v-2)}{2v} & \text{for } i_k \neq j_k, i_\ell \neq j_\ell \\ \dfrac{(v-1)^3}{2v} & \text{for } i_k \neq j_k, i_\ell = j_\ell \text{ or } i_k = j_k, i_\ell \neq j_\ell. \end{cases} \qquad (7.17)
\end{aligned}$$

Accordingly, for the regression function associated with the interaction of the attributes k, ℓ and m, say, we obtain

$$\begin{aligned}
&(\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) \otimes \mathbf{f}_1(i_m) - \mathbf{f}_1(j_k) \otimes \mathbf{f}_1(j_\ell) \otimes \mathbf{f}_1(j_m))^\top \mathbf{M}^{-1} \otimes \mathbf{M}^{-1} \otimes \mathbf{M}^{-1} \\
&\qquad \cdot (\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) \otimes \mathbf{f}_1(i_m) - \mathbf{f}_1(j_k) \otimes \mathbf{f}_1(j_\ell) \otimes \mathbf{f}_1(j_m)) \\
&\quad = \mathbf{f}_1(i_k)^\top \mathbf{M}^{-1}\mathbf{f}_1(i_k) \cdot \mathbf{f}_1(i_\ell)^\top \mathbf{M}^{-1}\mathbf{f}_1(i_\ell) \cdot \mathbf{f}_1(i_m)^\top \mathbf{M}^{-1}\mathbf{f}_1(i_m) \\
&\qquad + \mathbf{f}_1(j_k)^\top \mathbf{M}^{-1}\mathbf{f}_1(j_k) \cdot \mathbf{f}_1(j_\ell)^\top \mathbf{M}^{-1}\mathbf{f}_1(j_\ell) \cdot \mathbf{f}_1(j_m)^\top \mathbf{M}^{-1}\mathbf{f}_1(j_m) \\
&\qquad - \mathbf{f}_1(i_k)^\top \mathbf{M}^{-1}\mathbf{f}_1(j_k) \cdot \mathbf{f}_1(i_\ell)^\top \mathbf{M}^{-1}\mathbf{f}_1(j_\ell) \cdot \mathbf{f}_1(i_m)^\top \mathbf{M}^{-1}\mathbf{f}_1(j_m) \\
&\qquad - \mathbf{f}_1(j_k)^\top \mathbf{M}^{-1}\mathbf{g}(i_k) \cdot \mathbf{f}_1(j_\ell)^\top \mathbf{M}^{-1}\mathbf{f}_1(i_\ell) \cdot \mathbf{f}_1(j_m)^\top \mathbf{M}^{-1}\mathbf{f}_1(i_m) \\
&\quad = \begin{cases} \dfrac{(v-1)^3(v^2-3v+3)}{4v^2} & \text{for } i_k \neq j_k, i_\ell \neq j_\ell, i_m \neq j_m \\ \dfrac{(v-1)^4(v-2)}{4v^2} & \text{for } i_k \neq j_k, i_\ell \neq j_\ell, i_m = j_m \\ \dfrac{(v-1)^5}{4v^2} & \text{for } i_k \neq j_k, i_\ell = j_\ell, i_m = j_m. \end{cases} \qquad (7.18)
\end{aligned}$$

Now for a pair of alternatives $(\mathbf{i}, \mathbf{j}) \in \mathcal{X}_d$ of comparison depth d: there are exactly d attributes of the main effects for which i_k and j_k differ, there are $\frac{1}{2}d(d-1)$ first-order interaction terms for which $(i_k i_\ell)$ and $(j_k j_\ell)$ differ in all two attributes k and ℓ, there are $d(K-d)$ first-order interaction terms for which $(i_k i_\ell)$ and $(j_k j_\ell)$ differ in exactly one attribute k or ℓ, there are $\frac{1}{6}d(d-1)(d-2)$ second-order interaction terms for which $(i_k i_\ell i_m)$ and $(j_k j_\ell j_m)$ differ in all three attributes k, ℓ and m, there are $\frac{1}{2}(K-d)d(d-1)$ second-order interaction terms for which $(i_k i_\ell i_m)$ and $(j_k j_\ell j_m)$ differ in exactly two of the associated three attributes and finally, there are $\frac{1}{2}(K-d)(K-d-1)d$ second-order interaction terms for which $(i_k i_\ell i_m)$ and $(j_k j_\ell j_m)$ differ in exactly one of the associated three attributes. Hence from (7.16)–(7.18) and Lemma 7.2 we obtain

$$
\begin{aligned}
V(d,\xi) &= (\mathbf{f}(\mathbf{i})-\mathbf{f}(\mathbf{j}))^\top \mathbf{M}(\xi)^{-1}(\mathbf{f}(\mathbf{i})-\mathbf{f}(\mathbf{j})) \\
&= \frac{d(v-1)}{h_1(\xi)} + \frac{d(d-1)}{2}\frac{(v-1)^2(v-2)}{2vh_2(\xi)} + d(K-d)\frac{(v-1)^3}{2vh_2(\xi)} \\
&\quad + \frac{d(d-1)(d-2)}{6}\frac{(v-1)^3(v^2-3v+3)}{4v^2h_3(\xi)} + \frac{(K-d)d(d-1)}{2}\frac{(v-1)^4(v-2)}{4v^2h_3(\xi)} \\
&\quad + \frac{(K-d)(K-d-1)d}{2}\frac{(v-1)^5}{4v^2h_3(\xi)} \\
&= \frac{d(v-1)}{h_1(\xi)} + \frac{d(v-1)^2}{4vh_2(\xi)}\Big((d-1)(v-2)+2(K-d)(v-1)\Big) \\
&\quad + \frac{d(v-1)^3}{24v^2h_3(\xi)}\Big((d-1)(d-2)(v^2-3v+3)+3(K-d)(d-1)(v-1)(v-2) \\
&\qquad\qquad + 3(K-d)(K-d-1)(v-1)^2\Big) \\
&= \frac{d(v-1)}{h_1(\xi)} + \frac{d(v-1)^2}{4vh_2(\xi)}(2Kv-2K-dv-v+2) \\
&\quad + \frac{d(v-1)^3}{24v^2h_3(\xi)}\big(3K^2v^2-6K^2v-6Kv^2+3K^2-3Kdv^2+3Kdv \\
&\qquad\qquad + 3dv^2+15Kv-9K+d^2v^2-6dv+2v^2-6v+6\big),
\end{aligned}
$$

for $(\mathbf{i},\mathbf{j}) \in \mathcal{X}_d$ which proofs the proposed formula. □

We note that for comparison depth $d=0$ the corresponding variance function $V(0,\xi)=0$.

Accordingly, if the general invariant design ξ is concentrated on a single comparison depth, then the representation of the variance function $V(d,\xi)$ simplifies

Corollary 7.1. *For a uniform design $\xi_{d'}$ on a single comparison depth d' the variance function is given by*

$$V(d,\xi_{d'}) = \frac{d}{d'}\left(p_1 + p_2\tfrac{2Kv-2K-dv-v+2}{2Kv-2K-d'v-v+2} + p_3\tfrac{\lambda(d)}{\lambda(d')}\right),$$

where

$$
\begin{aligned}
\lambda(d) = 3K^2 + 3K^2v^2 - 6K^2v - 3Kdv^2 + 3Kdv - 6Kv^2 + 15Kv \\
- 9K + d^2v^2 + 3dv^2 - 6dv + 2v^2 - 6v + 6.
\end{aligned}
$$

Proof. In view of Theorem 7.3 it is sufficient to note that the representation of the variance function follows immediately by inserting the values of $h_q(\xi_d)$ from Lemma 7.1 and $p_q = \binom{K}{q}(v-1)^q$, $q = 1, 2, 3$. $\square$

Note that for $d = d'$, $V(d, \xi_d) = p_1 + p_2 + p_3 = p$ which shows the D-optimality of ξ_d on $\mathcal{X}_d$ in view of the Kiefer-Wolfowitz equivalence theorem.

In view of Theorems 7.1, 7.2 and 7.1 we similarly note that no design exists which is simultaneously optimal for the main effects, the first-order interactions and the second-order interactions. As a consequence, we confine ourselves to the D-criterion to derive optimal design for the whole parameter vector. As before we mention that generally a single comparison depth d may be sufficient for non-singularity of the information matrix $\mathbf{M}(\xi_d)$, i. e. for the identifiability of all model parameters. The following theorem gives an upper bound on the number of comparison depths required for a D-optimal design.

Theorem 7.4. *In the second-order interactions model the D-optimal design ξ^* is supported on, at most, three different comparison depths K, d^* and $d^* + 1$, say, i.e. $\xi^* = w_K^* \xi_K + w_{d^*}^* \xi_{d^*} + (1 - w_K^* - w_{d^*}^*)\xi_{d^*+1}$.*

Proof. The proof follows directly from Theorem 7.3 by using similar arguments in Theorem 6.5. $\square$

For fixed number K of the attributes at v levels, intermediate comparison depth d and weights w_K the numerical results were obtained by direct maximization of $\ln(\det(\mathbf{M}(w_K\xi_K + (1 - w_K)\xi_d)))$ for the corresponding optimal comparison depth d^* and optimal weights w_K^* where $w_{d^*}^* = 1 - w_K^*$ in Table 7.2. In particular, the numerical computations indicate that at most two different comparison depths K and d^* may be required as presented in Table 7.2 for fixed K attributes between 4 and 10 and levels $v = 2, \ldots, 8$. Here entries of the form $(d^*, w_{d^*}^*)$ indicate that invariant designs $\xi^* = w_{d^*}^* \xi_{d^*} + (1 - w_{d^*}^*)\xi_K$ have to be considered, while for single entries d^* the optimal design $\xi^* = \xi_{d^*}$ has to be considered which is uniform on the optimal comparison depth d^* (in boldface).

TABLE 7.2: Optimal Designs with Intermediate Comparison Depths d^* in Boldface and Optimal Weights $w^*_{d^*}$ of the form $(d^*,\ w^*_{d^*})$ for K Attributes and v-Levels

	v						
K	2	3	4	5	6	7	8
4	(**2**, 0.857)	**2**	**2**	**2**	**2**	**2**	**2**
5	(**2**, 0.833)	(**2**, 0.667)	**3**	**3**	**3**	**3**	**3**
6	(**3**, 0.732)	(**3**, 0.789)	**3**	**4**	**4**	**4**	**4**
7	(**3**, 0.697)	(**4**, 0.322)	**4**	**4**	**4**	**5**	**5**
8	(**3**, 0.644)	**4**	(**5**, 0.425)	**5**	**5**	**5**	**5**
9	(**4**, 0.577)	**5**	**5**	**6**	**6**	**6**	**6**
10	(**4**, 0.538)	**5**	**6**	**6**	**7**	**7**	**7**

In the corresponding Table 7.2 it is worth noting that for the particular case $K = 3$ and $v = 2$ one can recover the corresponding result in Theorem 6.6 where the design ξ^* is uniform on all pairs with non-zero comparison depth. Accordingly, a similar result also holds for arbitrary v (see Graßhoff et al., 2003). For $K = 9, 10$ and $v = 3, 4$, the corresponding intermediate comparison depths are candidates to obtain optimal designs.

Accordingly, in the following Table 7.3 we present the values of the normalized variance function $V(d, \xi^*)/p$, which shows D-optimality of the design ξ^* in view of the equivalence theorem by Kiefer and Wolfowitz (1960) for attributes $K = 4, \ldots, 10$ and levels $v = 2, \ldots, 8$. We note that for the corresponding particular case $K = 3$ and $v = 2$ the values of the normalized variance function $V(d, \xi^*)/p$ indicate that all three comparison depths are needed for D-optimality, and that one can obtain the corresponding results in Theorem 6.6. In this case similar result also holds for arbitrary v (see Graßhoff et al., 2003). Moreover, for $K = 4, \ldots, 10$ and $v = 2$ the corresponding results in Table 6.1 can also be recovered.

TABLE 7.3: Values of the Variance Function of the D-Optimal Design ξ^* for K Attributes and v-Levels (Boldface **1** Corresponds to the Optimal Comparison Depths d^*)

		d									
K	v	1	2	3	4	5	6	7	8	9	10
4	2	0.875	**1**	0.875	**1**						
	3	0.813	**1**	0.938	1						
	4	0.793	**1**	0.953	0.983						
	5	0.783	**1**	0.962	0.980						
	6	0.777	**1**	0.968	0.980						
	7	0.773	**1**	0.973	0.981						
	8	0.770	**1**	0.976	0.982						
5	2	0.760	**1**	0.960	0.880	**1**					
	3	0.723	**1**	1	0.954	**1**					
	4	0.689	0.967	**1**	0.952	0.987					
	5	0.666	0.951	**1**	0.961	0.981					
	6	0.653	0.941	**1**	0.968	0.980					
	7	0.644	0.934	**1**	0.972	0.981					
	8	0.638	0.929	**1**	0.976	0.982					
6	2	0.701	0.983	**1**	0.906	0.855	**1**				
	3	0.624	0.921	**1**	0.968	0.932	**1**				
	4	0.591	0.895	**1**	0.993	0.963	0.997				
	5	0.576	0.882	0.997	**1**	0.972	0.992				
	6	0.560	0.865	0.987	**1**	0.976	0.989				
	7	0.550	0.854	0.981	**1**	0.979	0.988				
	8	0.543	0.846	0.977	**1**	0.982	0.988				
7	2	0.615	0.917	**1**	0.956	0.879	0.863	**1**			
	3	0.553	0.860	0.988	**1**	0.963	0.941	**1**			
	4	0.519	0.822	0.965	**1**	0.981	0.962	0.997			
	5	0.498	0.800	0.952	**1**	0.992	0.974	0.993			
	6	0.487	0.787	0.944	**1**	0.999	0.983	0.995			
	7	0.479	0.777	0.937	0.997	**1**	0.985	0.994			
	8	0.471	0.768	0.929	0.994	**1**	0.987	0.993			
8	2	0.559	0.872	**1**	1	0.945	0.884	0.884	**1**		
	3	0.490	0.792	0.948	**1**	0.990	0.958	0.948	1		
	4	0.462	0.759	0.924	0.993	**1**	0.980	0.969	**1**		
	5	0.442	0.732	0.902	0.981	**1**	0.988	0.977	0.995		
	6	0.429	0.716	0.889	0.974	**1**	0.994	0.982	0.994		
	7	0.421	0.706	0.880	0.970	**1**	0.997	0.987	0.995		
	8	0.415	0.698	0.874	0.960	**1**	1	0.991	0.996		

(To be continued)

Table 7.3 (continued)

		d									
K	v	1	2	3	4	5	6	7	8	9	10
9	2	0.504	0.811	0.962	**1**	0.969	0.910	0.868	0.883	**1**	
	3	0.437	0.726	0.894	0.972	**1**	0.969	0.946	0.946	1	
	4	0.414	0.696	0.872	0.965	**1**	0.994	0.977	0.971	1	
	5	0.397	0.674	0.853	0.953	0.995	**1**	0.989	0.981	1	
	6	0.384	0.657	0.836	0.940	0.989	**1**	0.992	0.985	0.996	
	7	0.376	0.645	0.825	0.932	0.985	**1**	0.995	0.988	0.995	
	8	0.370	0.637	0.817	0.927	0.982	**1**	0.997	0.990	0.996	
10	2	0.462	0.763	0.932	**1**	0.997	0.956	0.905	0.874	0.896	**1**
	3	0.395	0.669	0.843	0.938	**1**	0.972	0.953	0.938	0.947	1
	4	0.374	0.642	0.822	0.929	0.981	**1**	0.987	0.974	0.972	1
	5	0.359	0.622	0.803	0.917	0.977	**1**	1	0.989	0.985	1
	6	0.348	0.606	0.786	0.903	0.968	0.996	**1**	0.993	0.988	0.998
	7	0.340	0.594	0.774	0.892	0.961	0.993	**1**	0.995	0.990	0.997
	8	0.335	0.586	0.765	0.885	0.956	0.990	**1**	0.996	0.991	0.996

7.2 Designs for Partial Profiles

In this section we generalize the results of full profiles presented in Section 7.1 to partial profiles, which is also a generalization of the corresponding results of partial profiles for the case of two-level attributes in Section 6.2 to common number of general levels $i_k = 1, \ldots, v$ for each attribute $k = 1, \ldots, K$.

As similarly noted for partial profiles a direct observation (utility) may be described by model (7.3) when summation is taken only over those S attributes contained in the describing subset. This requires that a profile strength $S \geq 3$ must be considered to ensure identifiability of the second-order interactions. To facilitate notation we similarly introduce an additional level 0 as before for each attribute indicating that the corresponding k-th attribute having level $i_k = 0$ is not present in the partial profile, and the corresponding regression functions are given by $\mathbf{f}_k(0) = \mathbf{f}_1(0) = \mathbf{0}$. With this convention a direct observation can be described by (7.3) even for a partial profile $\mathbf{i}$ from the set

$$\begin{aligned}\mathcal{I}^{(S)} =&\{\mathbf{i};\ i_k \in \{1, \ldots, v\} \text{ for } S \text{ components and} \\ &i_k = 0 \text{ for } K - S \text{ components}\},\end{aligned} \tag{7.19}$$

of alternatives with profile strength S. It should be noted that for general profile strength S the vector $\mathbf{f}$ of regression functions in (7.4) which is similarly given by

$$\begin{aligned}\mathbf{f}(\mathbf{i}) = (&\mathbf{f}_1(i_1)^\top, \ldots, \mathbf{f}_1(i_K)^\top, \mathbf{f}_1(i_1)^\top \otimes \mathbf{f}_1(i_2)^\top, \ldots, \mathbf{f}_1(i_{K-1}) \otimes \mathbf{f}_1(i_K)^\top, \\ &\mathbf{f}_1(i_1)^\top \otimes \mathbf{f}_1(i_2)^\top \otimes \mathbf{f}_1(i_3)^\top, \ldots, \mathbf{f}_1(i_{K-2})^\top \otimes \mathbf{f}_1(i_{K-1})^\top \otimes \mathbf{f}_1(i_K)^\top)^\top,\end{aligned} \tag{7.20}$$

of dimension p, the paired comparison model in (7.5) given by

$$\begin{aligned}Y_n(\mathbf{i}, \mathbf{j}) = &\sum_{k=1}^{K} (\mathbf{f}_1(i_k) - \mathbf{f}_1(j_k))^\top \boldsymbol{\beta}_k + \sum_{k<\ell} ((\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell)) - (\mathbf{f}_1(j_k) \otimes \mathbf{f}_1(j_\ell)))^\top \boldsymbol{\beta}_{k\ell} \\ &+ \sum_{k<\ell<m} ((\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) \otimes \mathbf{f}_1(i_m)) - (\mathbf{f}_1(j_k) \otimes \mathbf{f}_1(j_\ell) \otimes \mathbf{f}_1(j_m)))^\top \boldsymbol{\beta}_{k\ell m} + \varepsilon_n,\end{aligned} \tag{7.21}$$

and the interpretation of the corresponding parameter vector $\boldsymbol{\beta}$ in (7.2) given by

$$\boldsymbol{\beta} = ((\boldsymbol{\beta}_k)^\top_{k=1,\ldots,K}, (\boldsymbol{\beta}_{k\ell})^\top_{k<\ell}, (\boldsymbol{\beta}_{k\ell m})^\top_{k<\ell<m})^\top, \tag{7.22}$$

remain unchanged. Also, it should be noted that the comparison depth d describes the number of attributes in which the two alternatives in the choice sets differ satisfying the inequalities $1 \leq d \leq S < K$ (e.g. see Großmann, 2017, p. 239).

Hence, the paired comparison model (7.21) is thus restricted to those paired alternatives for which exactly S attributes are presented

$$\begin{aligned}\mathcal{X}^{(S)} = \{(\mathbf{i},\mathbf{j});\ & i_k, j_k \in \{1,\ldots,v\} \text{ for } S \text{ components and} \\ & i_k = j_k = 0 \text{ for exactly } K-S \text{ components}\}.\end{aligned} \tag{7.23}$$

As already noted the design region $\mathcal{X}^{(S)}$ can be partitioned into disjoint sets such that the pairs in each set differ only in a fixed number d of the attributes. Specifically, for a comparison depth $d = 0,\ldots,S$, let

$$\mathcal{X}_d^{(S)} = \{(\mathbf{i},\mathbf{j}) \in \mathcal{X}^{(S)} : |\{k : i_k \neq j_k\}| = d\}, \tag{7.24}$$

be the set of all pairs of alternatives which differ in exactly d attributes. These sets also constitute the orbits with respect to permutations. The D-criterion is also invariant with respect to those permutations. In particular, the regression functions $\mathbf{f}$ in (7.20) extended to the design region $\mathcal{X}^{(S)}$ are still linearly equivariant i. e. also here relabeling does not affect D-optimality as well as D-optimality of invariant subvectors $\boldsymbol{\beta}$ in (7.22). Hence, it is sufficient to look for optimality in the class of invariant designs which are uniform on the orbits of fixed comparison depth $d \leq S$.

Further let $N_d = \binom{K}{S}\binom{S}{d}v^S(v-1)^d$ be the number of different pairs in $\mathcal{X}_d^{(S)}$ which vary in exactly d attributes and denote by ξ_d the uniform approximate design which assigns equal weight $\xi_d(\mathbf{i},\mathbf{j}) = 1/N_d$ to each pair in $\mathcal{X}_d^{(S)}$ and weight zero to all remaining pairs in $\mathcal{X}^{(S)}$. The corresponding information matrices for the invariant designs are presented in the following Lemma 7.3 and Lemma 7.4. It is worthwhile mentioning that the results for full profiles presented in Section 7.1 can be obtained as special cases by letting $S = K$.

Lemma 7.3. *Let d be a fixed comparison depth. The uniform design ξ_d on the set $\mathcal{X}_d^{(S)}$ of comparison depth d has block diagonal information matrix*

$$\mathbf{M}(\xi_d) = \begin{pmatrix} h_1(d)\mathbf{Id}_{p_1} \otimes \mathbf{M} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & h_2(d)\mathbf{Id}_{p_2} \otimes \mathbf{M} \otimes \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & h_3(d)\mathbf{Id}_{p_3} \otimes \mathbf{M} \otimes \mathbf{M} \otimes \mathbf{M} \end{pmatrix},$$

where

$$h_1(d) = \frac{d}{K},\ h_2(d) = \frac{d}{2vK(K-1)}(2Sv - 2S - dv - v + 2) \text{ and}$$

$$\begin{aligned}h_3(d) = \frac{d}{4v^2K(K-1)(K-2)}(&3S^2 + 3S^2v^2 - 6S^2v - 3Sdv^2 + 3Sdv - 6Sv^2 + 15Sv \\ &- 9S + d^2v^2 + 3dv^2 - 6dv + 2v^2 - 6v + 6).\end{aligned}$$

It should be noted that the results for $v = 2$ in Lemma 6.3 can be obtained as a special case by replacing $h_q(d)$ for $q = 1, 2, 3$ by $4^q h_q(d)$ because $\mathbf{M} = 4$.

Proof. In view of Lemma 7.1 and for the regression functions $\mathbf{f}_k = \mathbf{f}_1$ associated with the k-th main effects we obtain

$$\begin{aligned}
&\sum_{(\mathbf{i},\mathbf{j})\in\mathcal{X}_d^{(S)}} (\mathbf{f}_k(i_k) - \mathbf{f}_k(j_k))(\mathbf{f}_k(i_k) - \mathbf{f}_k(j_k))^\top \\
&\quad = \tbinom{K-1}{S-1}\tbinom{S-1}{d-1} v^{S-1}(v-1)^{d-1} \sum_{i_k \neq j_k} (\mathbf{f}_1(i_k) - \mathbf{f}_1(j_k))(\mathbf{f}_1(i_k) - \mathbf{f}_1(j_k))^\top \\
&\quad = \tbinom{K-1}{S-1}\tbinom{S-1}{d-1} v^{S}(v-1)^{d}\mathbf{M},
\end{aligned}$$

because there are $\tbinom{K-1}{S-1}\tbinom{S-1}{d-1} v^{S-1}(v-1)^{d-1}$ pairs in $\mathcal{X}_d^{(S)}$ for which i_k and j_k differ.

As the number N_d of paired comparisons in $\mathcal{X}_d^{(S)}$ with comparison depth d equals $N_d = \tbinom{K}{S}\tbinom{S}{d} v^S(v-1)^d$, the corresponding diagonal elements $h_1(d)$ in the information matrix are given by

$$\begin{aligned}
h_1(d) &= \frac{1}{N_d} \sum_{(\mathbf{i},\mathbf{j})\in\mathcal{X}_d^{(S)}} (\mathbf{f}_k(i_k) - \mathbf{f}_k(j_k))(\mathbf{f}_k(i_k) - \mathbf{f}_k(j_k))^\top \\
&= \frac{d}{K}\mathbf{M}.
\end{aligned}$$

Now for the given attributes k and ℓ the pairs with distinct levels in both attributes occur $\tbinom{K-2}{S-2}\tbinom{S-2}{d-2} v^{S-2}(v-1)^{d-2}$ times in $\mathcal{X}_d^{(S)}$, while those which differ only in one attribute occur $\tbinom{2}{1}\tbinom{K-2}{S-2}\tbinom{S-2}{d-1} v^{S-2}(v-1)^{d-1}$ times in $\mathcal{X}_d^{(S)}$. Hence, from (7.9)–(7.10) the diagonal elements $h_2(d)$ for the first-order interactions are given by

$$\begin{aligned}
h_2(d) &= \frac{1}{N_d}\Bigg(\frac{1}{2}\binom{K-2}{S-2}\binom{S-2}{d-2} v(v-1)^2(v-2)v^{S-2}(v-1)^{d-2}\mathbf{M}\otimes\mathbf{M} \\
&\qquad + \binom{K-2}{S-2}\binom{S-2}{d-1} v(v-1)^2 v^{S-2}(v-1)^{d-1}\mathbf{M}\otimes\mathbf{M}\Bigg) \\
&= \left(\tfrac{d(d-1)}{2vK(K-1)}(v-2) + \tfrac{d(S-d)}{vK(K-1)}(v-1)\right)\mathbf{M}\otimes\mathbf{M} \\
&= \frac{d}{2vK(K-1)}\Big((v-2)(d-1) + 2(S-d)(v-1)\Big)\mathbf{M}\otimes\mathbf{M} \\
&= \frac{d}{2vK(K-1)}(2Sv - 2S - dv - v + 2)\mathbf{M}\otimes\mathbf{M}.
\end{aligned}$$

Futher for the given attributes k, ℓ and m the pairs with distinct levels in the three attributes occur $\tbinom{K-3}{S-3}\tbinom{S-3}{d-3} v^{S-3}(v-1)^{d-3}$ times in $\mathcal{X}_d^{(S)}$, while those which differ in two attributes occur $\tbinom{3}{2}\tbinom{K-3}{S-3}\tbinom{S-3}{d-2} v^{S-3}(v-1)^{d-2}$ times in $\mathcal{X}_d^{(S)}$. Finally, those which differ only in one attribute occur $\tbinom{3}{1}\tbinom{K-3}{S-3}\tbinom{S-3}{d-1} v^{S-3}(v-1)^{d-1}$ times in $\mathcal{X}_d^{(S)}$.

Hence, from (9.8)–(9.10) the diagonal elements $h_3(d)$ for the second-order interactions are given by

$$\begin{aligned}
h_3(d) =& \frac{1}{N_d}\Bigg(\frac{1}{4}\binom{K-3}{S-3}\binom{S-3}{d-3} v(v-1)^3(v^2-3v+3)v^{S-3}(v-1)^{d-3}\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M} \\
&+\frac{3}{4}\binom{K-3}{S-3}\binom{S-3}{d-2} v(v-1)^3(v-2)v^{S-3}(v-1)^{d-2}\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M} \\
&+\frac{3}{4}\binom{K-3}{S-3}\binom{S-3}{d-1} v(v-1)^3v^{S-3}(v-1)^{d-1}\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M}\Bigg) \\
=& \Bigg(\frac{d(d-1)(d-2)}{4v^2K(K-1)(K-2)}(v^2-3v+3) + \frac{3(S-d)d(d-1)}{4v^2K(K-1)(K-2)}(v-1)(v-2) \\
&+\frac{3(S-d)(S-d-1)d}{4v^2K(K-1)(K-2)}(v-1)^2\Bigg)\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M} \\
=& \frac{d}{4v^2K(K-1)(K-2)}(3S^2+3S^2v^2-6S^2v-3Sdv^2+3Sdv-6Sv^2+15Sv \\
&-9S+d^2v^2+3dv^2-6dv+2v^2-6v+6)\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M},
\end{aligned}$$

in the information matrix.

It should be noted that the off-diagonal elements all vanish because the terms in the corresponding entries sum up to zero due to the effects-type coding. □

We note that for comparison depth $d = 0$ the corresponding functions $h_q(0) = 0$ for $q = 1, 2, 3$.

The information matrix for the general invariant design ξ which result from convex combinations of uniform designs on the comparison depth d with weight w_d, $\xi = \sum_{d=1}^{S} w_d\xi_d$, $w_d \geq 0, \sum_{d=1}^{S} w_d = 1$ also has block diagonal information matrix:

Lemma 7.4. *Let ξ be an invariant design on $\mathcal{X}^{(S)}$ then ξ has block diagonal information matrix*

$$\mathbf{M}(\xi) = \begin{pmatrix} h_1(\xi)\mathbf{Id}_{p_1}\otimes\mathbf{M} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & h_2(\xi)\mathbf{Id}_{p_2}\otimes\mathbf{M}\otimes\mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & h_3(\xi)\mathbf{Id}_{p_3}\otimes\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M}\end{pmatrix},$$

where $h_q(\xi) = \sum_{d=1}^{S} w_d h_q(d)$, $q = 1, 2, 3$.

In the following Theorems and remark we consider optimal designs for the main effects, the first-order interaction and the second-order interaction terms separately having entries $h_q(d)$ for $q = 1, 2, 3$ in the corresponding information matrix $\mathbf{M}(\xi_d)$. Analogously, the resulting designs may optimize every design criterion which is invariant with respect to both permutations of the levels and permutations of the attributes if one considers the full parameter vector, satisfying the aforementioned identifiabity conditions. Hence, the reduced parameter vector $\boldsymbol{\beta} = ((\boldsymbol{\beta}_k)^\top_{k=1,\dots,K}, (\boldsymbol{\beta}_{k\ell})^\top_{k<\ell}, (\boldsymbol{\beta}_{k\ell m})^\top_{k<\ell<m})^\top$ in (7.22) considered separately in the underlying Theorems 7.5, 7.6 and 7.2 are also invariant in particular, with the D-criterion.

Theorem 7.5. *Let $d_1^* = S$. Then the uniform design $\xi_{d_1^*} = \xi_S$ on the largest possible comparison depth S is D-optimal for the main effects $(\boldsymbol{\beta}_k)^\top_{1\leq k\leq K}$.*

This means that for the main effects it is sufficient to use only those pairs of alternatives which differ in all the profile strength.

For the first-order interactions the number of the attributes subject to the profile strength with distinct levels does not provide enough information. Hence, only those pairs of alternatives should be used which differ in approximately half of the profile strength S. In particular, and as before one has to consider the intermediate comparison depth $d^* = S - 1 - \left[\frac{S-2}{v}\right]$ where $[u]$ denotes the integer part of the decimal expansion for u, satisfying $[u] \leq u < [u] + 1$.

Theorem 7.6. *Let $d_2^* = S - 1 - \left[\frac{S-2}{v}\right]$. Then the uniform design $\xi_{d_2^*}$ is D-optimal for the first-order interaction effects $(\boldsymbol{\beta}_{k\ell})^\top_{k<\ell}$.*

It is worthwhile mentioning that the corresponding Theorems 7.5 and 7.6 paraphrase theorems given in Graßhoff et al. (2003) for first-order interaction models and translate them to the present setting of second-order interaction models.

Remark 7.2. *There exists a single comparison depth d_3^* subject to the profile strength S such that the uniform design $\xi_{d_3^*}$ is D-optimal for the second-order interaction effects $(\boldsymbol{\beta}_{k\ell m})^\top_{k<\ell<m}$.*

In the following Table 7.4 we note that the corresponding values of d_3^* were obtained by first calculating the values of $h_3(d)$ and determining the maximum. It is worthwhile mentioning that generally for moderate values of v the optimal comparison depth $d_3^* = S$ but this is not true for the case when $S = 3$ and $K = 4$. Moreover, the optimal comparison depth $d_3^* = S - 2$ for sufficiently large values of v. We further note that for the situation of full profiles ($S = K$) the corresponding results presented in Table 7.1 can also be recovered.

TABLE 7.4: Values of the Optimal Comparison Depths d_3^* of the D-Optimal Uniform Designs $\xi_{d_3^*}$ for the Second-Order Interactions with $S = K - 1$ Attributes and v-Levels

		v									
K	S	2	3	4	5	6	7	8	9	10	20
4	3	1	1	1	1	1	1	1	1	1	1
5	4	4	1	2	2	2	2	2	2	2	2
6	5	5	2	2	3	3	3	3	3	3	3
7	6	6	6	3	3	3	4	4	4	4	4
8	7	7	7	7	4	4	4	4	5	5	5
9	8	8	8	8	5	5	5	5	5	6	6
10	9	9	9	9	9	6	6	6	6	6	7

As before we mention that the value of the variance function for the corresponding invariant ξ evaluated at comparison depth d is given by $V(d,\xi)$ where $V(d,\xi) = V((\mathbf{i},\mathbf{j}),\xi) = (\mathbf{f}(\mathbf{i}) - \mathbf{f}(\mathbf{j}))^\top \mathbf{M}(\xi)^{-1}(\mathbf{f}(\mathbf{i}) - \mathbf{f}(\mathbf{j}))$ for the pair $(\mathbf{i},\mathbf{j})$ on the orbits $\mathcal{X}_d^{(S)}$ of fixed comparison depth d. The variance function $V(d,\xi)$ for the invariant design ξ is given in the following Theorem 7.7.

Theorem 7.7. *For every invariant design ξ the variance function $V(d,\xi)$ is given by*

$$V(d,\xi) = d(v-1)\left(\frac{1}{h_1(\xi)} + \frac{v-1}{4vh_2(\xi)}\Big(2Sv - 2S - dv - v + 2\Big) + \frac{(v-1)^2}{24v^2h_3(\xi)}\lambda(d)\right),$$

where

$$\begin{aligned}\lambda(d) = {} & 3S^2 + 3S^2v^2 - 6S^2v - 3Sdv^2 + 3Sdv - 6Sv^2 + 15Sv - 9S + d^2v^2 + 3dv^2 \\ & - 6dv + 2v^2 - 6v + 6.\end{aligned}$$

Proof. First we note that

$$\mathbf{M}(\xi)^{-1} = \begin{pmatrix} \frac{1}{h_1(\xi)}\mathbf{Id}_{p_1} \otimes \mathbf{M} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{h_2(\xi)}\mathbf{Id}_{p_2} \otimes \mathbf{M} \otimes \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{1}{h_3(\xi)}\mathbf{Id}_{p_3} \otimes \mathbf{M} \otimes \mathbf{M} \otimes \mathbf{M} \end{pmatrix},$$

for the inverse of the information matrix of the design ξ. Hence, we obtain for the variance function

$$
\begin{aligned}
V((\mathbf{i},\mathbf{j}),\xi) =& (\mathbf{f}(\mathbf{i})-\mathbf{f}(\mathbf{j}))^\top \mathbf{M}(\xi)^{-1}(\mathbf{f}(\mathbf{i})-\mathbf{f}(\mathbf{j})) \\
=& \frac{1}{h_1(\xi)}\sum_{k=1}^{K}(\mathbf{f}_1(i_k)-\mathbf{f}_1(j_k))(\mathbf{f}_1(i_k)-\mathbf{f}_1(j_k))^\top \\
&+ \frac{1}{h_2(\xi)}\sum_{k<\ell}(\mathbf{f}_1(i_k)\otimes\mathbf{f}_1(i_\ell)-\mathbf{f}_1(j_k)\otimes\mathbf{f}_1(j_\ell))(\mathbf{f}_1(i_k)\otimes\mathbf{f}_1(i_\ell)-\mathbf{f}_1(j_k)\otimes\mathbf{f}_1(j_\ell))^\top \\
&+ \frac{1}{h_3(\xi)}\sum_{k<\ell<m}(\mathbf{f}_1(i_k)\otimes\mathbf{f}_1(i_\ell)\otimes\mathbf{f}_1(i_m)-\mathbf{f}_1(j_k)\otimes\mathbf{f}_1(j_\ell)\otimes\mathbf{f}_1(j_m)) \\
&\qquad\cdot(\mathbf{f}_1(i_k)\otimes\mathbf{f}_1(i_\ell)\otimes\mathbf{f}_1(i_m)-\mathbf{f}_1(j_k)\otimes\mathbf{f}_1(j_\ell)\otimes\mathbf{f}_1(j_m))^\top .
\end{aligned}
\tag{7.25}
$$

Now we note that for a pair of alternatives $(\mathbf{i},\mathbf{j})\in\mathcal{X}_d^{(S)}$ of comparison depth d, there are exactly d attributes of the main effects for which i_k and j_k differ, there are $\frac{1}{2}d(d-1)$ first-order interaction terms for which $(i_k i_\ell)$ and $(j_k j_\ell)$ differ in all two attributes k and ℓ, there are $d(S-d)$ first-order interaction terms for which $(i_k i_\ell)$ and $(j_k j_\ell)$ differ in exactly one attribute k or ℓ, there are $\frac{1}{6}d(d-1)(d-2)$ second-order interaction terms for which $(i_k i_\ell i_m)$ and $(j_k j_\ell j_m)$ differ in all three attributes k, ℓ and m, there are $\frac{1}{2}(S-d)d(d-1)$ second-order interaction terms for which $(i_k i_\ell i_m)$ and $(j_k j_\ell j_m)$ differ in exactly two of the associated three attributes and finally, there are $\frac{1}{2}(S-d)(S-d-1)d$ second-order interaction terms for which $(i_k i_\ell i_m)$ and $(j_k j_\ell j_m)$ differ in exactly one of the associated three attributes. Hence from (7.16)–(7.18) and (7.25) we obtain

$$
\begin{aligned}
V(d,\xi) &= (\mathbf{f}(\mathbf{i})-\mathbf{f}(\mathbf{j}))^\top \mathbf{M}(\xi)^{-1}(\mathbf{f}(\mathbf{i})-\mathbf{f}(\mathbf{j})) \\
&= \frac{d(v-1)}{h_1(\xi)} + \frac{d(d-1)}{2}\frac{(v-1)^2(v-2)}{2vh_2(\xi)} + d(S-d)\frac{(v-1)^3}{2vh_2(\xi)} \\
&\quad + \frac{d(d-1)(d-2)}{6}\frac{(v-1)^3(v^2-3v+3)}{4v^2h_3(\xi)} + \frac{(S-d)d(d-1)}{2}\frac{(v-1)^4(v-2)}{4v^2h_3(\xi)} \\
&\quad + \frac{(S-d)(S-d-1)d}{2}\frac{(v-1)^5}{4v^2h_3(\xi)} \\
&= \frac{d(v-1)}{h_1(\xi)} + \frac{d(v-1)^2}{4vh_2(\xi)}\Big((d-1)(v-2)+2(S-d)(v-1)\Big) \\
&\quad + \frac{d(v-1)^3}{24v^2h_3(\xi)}\Big((d-1)(d-2)(v^2-3v+3)+3(S-d)(d-1)(v-1)(v-2) \\
&\qquad\qquad + 3(S-d)(S-d-1)(v-1)^2\Big) \\
&= \frac{d(v-1)}{h_1(\xi)} + \frac{d(v-1)^2}{4vh_2(\xi)}\Big(2Sv-2S-dv-v+2\Big) \\
&\quad + \frac{d(v-1)^3}{24v^2h_3(\xi)}\big(3S^2v^2-6S^2v-6Sv^2+3S^2-3Sdv^2+3Sdv \\
&\qquad\qquad + 3dv^2+15Sv-9S+d^2v^2-6dv+2v^2-6v+6\big),
\end{aligned}
$$

for $(\mathbf{i},\mathbf{j}) \in \mathcal{X}_d^{(S)}$. ☐

It is worthwhile mentioning that for the case when the comparison depth $d=0$ the corresponding variance function $V(0,\xi)=0$.

Further we mention that the corresponding representation of the variance function $V(d,\xi)$ in Theorem 7.7 simplifies if the general invariant design ξ is concentrated on a single comparison depth d':

Corollary 7.2. *For a uniform design $\xi_{d'}$ on a single comparison depth d' the variance function is given by*

$$V(d,\xi_{d'}) = \frac{d}{d'}\left(p_1 + p_2\tfrac{2Sv-2S-dv-v+2}{2Sv-2S-d'v-v+2} + p_3\tfrac{\lambda(d)}{\lambda(d')}\right),$$

where

$$
\begin{aligned}
\lambda(d) &= 3S^2+3S^2v^2-6S^2v-3Sdv^2+3Sdv-6Sv^2+15Sv-9S+d^2v^2+3dv^2 \\
&\quad -6dv+2v^2-6v+6.
\end{aligned}
$$

Proof. In view of Theorem 7.7 it is sufficient to note that the representation of the variance function follows immediately by inserting the values of $h_q(\xi_d)$ from Lemma 7.3 and $p_q = \binom{K}{q}(v-1)^q$, $q=1,2,3$. ☐

For $d = d'$ we obtain $V(d, \xi_d) = p_1 + p_2 + p_3 = p$ which shows the D-optimality of ξ_d on $\mathcal{X}_d^{(S)}$ in view of the Kiefer-Wolfowitz equivalence theorem.

In the following theorem we employ the D-criterion to derive optimal design which incorporates the main effects, the first-order interactions and the second-order interactions since no design exists which is simultaneously optimal for the corresponding Theorem 7.5, Theorem 7.6 and Remark 7.2. As already mentioned a single comparison depth d may be sufficient for non-singularity of the information matrix $\mathbf{M}(\xi_d)$, i. e. for the identifiability of all model parameters. The following theorem gives an upper bound on the number of comparison depths required for a D-optimal design subject to the profile strength S.

Theorem 7.8. *In the second-order interactions model the D-optimal design ξ^* is supported on, at most, three different comparison depths S, d^* and $d^* + 1$, say, i.e.* $\xi^* = w_S^* \xi_S + w_{d^*}^* \xi_{d^*} + (1 - w_S^* - w_{d^*}^*)\xi_{d^*+1}$.

Proof. By Theorem 7.7 the proof follows directly by using analogous arguments in Theorem 6.11. □

The following Table 7.5 shows the corresponding optimal designs with their optimal comparison depths d^* in boldface and their corresponding weights $w_{d^*}^*$ for various choices of attributes K between 4 and 10, profile strength S and levels $v = 2, \dots, 8$. Entries of the form $(d^*, w_{d^*}^*)$ indicate that invariant designs $\xi^* = w_{d^*}^* \xi_{d^*} + (1 - w_{d^*}^*)\xi_S$ have to be considered, while for single entries d^* the optimal design $\xi^* = \xi_{d^*}$ has to be considered which is uniform on the optimal comparison depth d^* (in boldface). In particular, the numerical results presented in Table 7.5 indicate that at most two different comparison depths S and d^* may be required for D-optimality. It is worthwhile mentioning that the results for full profiles presented in Table 7.2 can be recovered in Table 7.5 for the (row) entries when $S = K$. Moreover, in Table 7.5 we note that for the particular case $S = 3$, $K \geq 4$ and $v = 2$ one can recover the corresponding results in Theorem 6.12 where the design ξ^* is uniform on two different comparison depths $d^* = 1$ and $S = 3$. Further the corresponding values of the normalized variance function $V(d, \xi^*)/p$ which shows D-optimality of the design ξ^* in view of the Kiefer and Wolfowitz (1960) equivalence theorem is exhibited in Table 7.6. We note that for the situation when $S = K - 1$, $K = 4, \dots, 10$ and $v = 2$ the corresponding results in Table 6.4 can be obtained.

Table 7.5: Optimal Designs with Intermediate Comparison Depths d^* in Boldface and Optimal Weights $w^*_{d^*}$ of the form (d^*, $w^*_{d^*}$) for $S \leq K$ Attributes and v-Levels

		v						
K	S	2	3	4	5	6	7	8
4	3	(**1**, 0.900)	(**1**, 0.937)	**1**	**1**	**1**	**1**	**1**
	4	(**2**, 0.857)	**2**	**2**	**2**	**2**	**2**	**2**
5	3	**1**	**1**	**1**	**1**	**1**	**1**	**1**
	4	(**2**, 0.800)	**2**	**2**	**2**	**2**	**2**	**2**
	5	(**2**, 0.833)	(**2**, 0.667)	**3**	**3**	**3**	**3**	**3**
6	3	**1**	**1**	**1**	**1**	**1**	**1**	**1**
	4	(**2**, 0.732)	**2**	**2**	**2**	**2**	**2**	**2**
	5	(**2**, 0.802)	(**2**, 0.832)	**3**	**3**	**3**	**3**	**3**
	6	(**3**, 0.732)	(**3**, 0.789)	**3**	**4**	**4**	**4**	**4**
7	3	**1**	**1**	**1**	**1**	**1**	**1**	**1**
	4	(**1**, 0.836)	**2**	**2**	**2**	**2**	**2**	**2**
	5	(**2**, 0.756)	(**2**, 0.952)	**3**	**3**	**3**	**3**	**3**
	6	(**2**, 0.728)	(**3**, 0.755)	**3**	**3**	**4**	**4**	**4**
	7	(**3**, 0.697)	(**4**, 0.322)	**4**	**4**	**4**	**5**	**5**
8	3	**1**	**1**	**1**	**1**	**1**	**1**	**1**
	4	(**1**, 0.832)	**2**	**2**	**2**	**2**	**2**	**2**
	5	(**2**, 0.707)	**2**	**3**	**3**	**3**	**3**	**3**
	6	(**2**, 0.687)	(**3**, 0.675)	**3**	**3**	**4**	**4**	**4**
	7	(**3**, 0.643)	(**4**, 0.105)	**4**	**4**	**4**	**5**	**5**
	8	(**3**, 0.644)	**4**	(**5**, 0.425)	**5**	**5**	**5**	**5**
9	3	**1**	**1**	**1**	**1**	**1**	**1**	**1**
	4	(**1**, 0.819)	**2**	**2**	**2**	**2**	**2**	**2**
	5	(**2**, 0.659)	**2**	(**2**, 0.999)	**3**	**3**	**3**	**3**
	6	(**2**, 0.645)	(**3**, 0.559)	**3**	**3**	**4**	**4**	**4**
	7	(**3**, 0.594)	**4**	**4**	**4**	**4**	**5**	**5**
	8	(**3**, 0.598)	**4**	(**5**, 0.113)	**5**	**5**	**5**	**5**
	9	(**4**, 0.577)	**5**	**5**	**6**	**6**	**6**	**6**
10	3	**1**	**1**	**1**	**1**	**1**	**1**	**1**
	4	(**1**, 0.800)	**2**	**2**	**2**	**2**	**2**	**2**
	5	(**2**, 0.615)	**2**	(**2**, 0.997)	**3**	**3**	**3**	**3**
	6	(**2**, 0.604)	(**3**, 0.418)	**3**	**3**	**4**	**4**	**4**
	7	(**3**, 0.551)	**4**	**4**	**4**	**4**	(**4**, 0.996)	**5**
	8	(**3**, 0.556)	**4**	**5**	**5**	**5**	**5**	**5**
	9	(**4**, 0.533)	**5**	**5**	**6**	**6**	**6**	**6**
	10	(**4**, 0.538)	**5**	**6**	**6**	**7**	**7**	**7**

TABLE 7.6: Values of the Variance Function of the D-Optimal Design ξ^* for $S = K - 1$ Attributes and v-Levels (Boldface **1** Corresponds to the Optimal Comparison Depths d^*)

			d								
K	S	v	1	2	3	4	5	6	7	8	9
4	3	2	**1**	0.952	**1**						
		3	**1**	0.952	**1**						
		4	**1**	0.966	1						
		5	**1**	0.967	0.989						
		6	**1**	0.970	0.985						
		7	**1**	0.973	0.984						
		8	**1**	0.975	0.983						
5	4	2	0.958	**1**	0.792	**1**					
		3	0.859	**1**	0.885	0.974					
		4	0.822	**1**	0.917	0.958					
		5	0.804	**1**	0.935	0.958					
		6	0.794	**1**	0.947	0.962					
		7	0.787	**1**	0.955	0.965					
		8	0.781	**1**	0.961	0.968					
6	5	2	0.791	**1**	0.913	0.817	**1**				
		3	0.738	**1**	0.993	0.924	**1**				
		4	0.707	0.981	**1**	0.942	0.985				
		5	0.679	0.960	**1**	0.954	0.978				
		6	0.663	0.948	**1**	0.962	0.977				
		7	0.652	0.940	**1**	0.967	0.978				
		8	0.644	0.934	**1**	0.972	0.979				
7	6	2	0.714	**1**	1	0.915	0.858	**1**			
		3	0.637	0.932	**1**	0.957	0.919	**1**			
		4	0.599	0.902	**1**	0.986	0.953	0.992			
		5	0.580	0.887	**1**	1	0.973	0.995			
		6	0.566	0.871	0.991	**1**	0.974	0.989			
		7	0.555	0.859	0.984	**1**	0.978	0.988			
		8	0.546	0.850	0.979	**1**	0.980	0.988			

(To be continued)

Table 7.6 (continued)

			d								
K	S	v	1	2	3	4	5	6	7	8	9
8	7	2	0.626	0.926	**1**	0.946	0.862	0.847	**1**		
		3	0.559	0.867	0.992	**1**	0.959	0.937	**1**		
		4	0.524	0.828	0.969	**1**	0.978	0.958	0.996		
		5	0.502	0.804	0.954	**1**	0.989	0.971	0.991		
		6	0.489	0.790	0.946	**1**	0.997	0.980	0.993		
		7	0.481	0.781	0.939	0.999	**1**	0.985	0.994		
		8	0.474	0.771	0.932	0.995	**1**	0.986	0.993		
9	8	2	0.565	0.877	**1**	0.998	0.933	0.869	0.870	**1**	
		3	0.492	0.793	0.945	**1**	0.980	0.947	0.940	1	
		4	0.467	0.764	0.929	0.997	**1**	0.981	0.970	**1**	
		5	0.444	0.736	0.905	0.983	**1**	0.987	0.975	0.995	
		6	0.431	0.719	0.891	0.975	**1**	0.993	0.981	0.993	
		7	0.423	0.708	0.882	0.970	**1**	0.996	0.986	0.994	
		8	0.417	0.700	0.875	0.967	**1**	0.999	0.990	0.996	
10	9	2	0.509	0.816	0.965	**1**	0.966	0.905	0.861	0.878	**1**
		3	0.439	0.727	0.893	0.968	**1**	0.960	0.937	0.941	1
		4	0.415	0.699	0.874	0.965	**1**	0.992	0.975	0.969	1
		5	0.399	0.677	0.855	0.954	0.996	**1**	0.988	0.981	0.999
		6	0.386	0.659	0.838	0.942	0.990	**1**	0.992	0.984	0.996
		7	0.377	0.647	0.826	0.933	0.985	**1**	0.995	0.987	0.995
		8	0.371	0.639	0.818	0.927	0.982	**1**	0.997	0.990	0.996

8 Optimal Designs for Third-Order Interactions Two Level Models

In this chapter we consider the case where the paired comparisons are characterized by both full and partial profiles. For this situation we introduce an appropriate model and derive optimal designs in the presence of third-order interactions when all the attributes $k = 1, \ldots, K$ have two-levels each i. e. $i_k = 1, -1$.

For the direct observation (utility) $\tilde{Y}_{na}(\mathbf{i})$ and in the present setting of third-order interactions, we introduce additional 4-th attribute of influence, and reformulate model (6.2) as

$$\tilde{Y}_{na}(\mathbf{i}) = \mu_n + \sum_{k=1}^{K} \beta_k i_k + \sum_{k<\ell} \beta_{k\ell} i_k i_\ell + \sum_{k<\ell<m} \beta_{k\ell m} i_k i_\ell i_m + \sum_{k<\ell<m<r} \beta_{k,\ell,m,r} i_k i_\ell i_m i_r + \tilde{\varepsilon}_{na}, \tag{8.1}$$

where as before β_k denotes the main effect of the k-th attribute, $\beta_{k\ell}$ is the first-order interaction of the k-th and ℓ-th attribute, $\beta_{k\ell m}$ is the second-order interaction of the k-th, ℓ-th and m-th attribute and $\beta_{k\ell mr}$ is the third-order interaction of the k-th, ℓ-th, m-th and r-th attribute. The vectors $(\beta_k)_{1\leq k\leq K}$ of main effects, $(\beta_{k\ell})_{1\leq k<\ell\leq K}$ of first-order interactions, $(\beta_{k\ell m})_{1\leq k<\ell<m\leq K}$ of second-order interactions and $(\beta_{k\ell mr})_{1\leq k<\ell<m<r\leq K}$ of third-order interactions have dimensions $p_1 = K$, $p_2 = K(K-1)/2$, $p_3 = K(K-1)(K-2)/6$ and $p_4 = (1/24)K(K-1)(K-2)(K-3)$, respectively. Hence the reduced parameter vector $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K, (\beta_{k\ell})_{k<\ell}^\top, (\beta_{k\ell m})_{k<\ell<m}^\top, (\beta_{k\ell mr})_{k<\ell<m<r}^\top)^\top$ has dimension $p = p_1 + p_2 + p_3 + p_4$. The corresponding p-dimensional vector $\mathbf{f}$ of regression functions is given by

$$\mathbf{f}(\mathbf{i}) = (i_1, \ldots, i_K, (i_k i_\ell)_{k<\ell}^\top, (i_k i_\ell i_m)_{k<\ell<m}^\top, (i_k i_\ell i_m i_r)_{k<\ell<m<r}^\top)^\top. \tag{8.2}$$

As already defined here in $\mathbf{f}(\mathbf{i})$, the first $p_1 = K$ components $i_1, \ldots, i_K$ are associated with the main effects, the second set of p_2 components $i_k i_\ell$, $1 \leq k < \ell \leq K$, are associated with the first-order interactions, the third set of p_3 components $i_k i_\ell i_m$, $1 \leq k < \ell < m \leq K$, are associated with the second-order interactions, and the remaining p_4 components $i_k i_\ell i_m i_r$, $1 \leq k < \ell < m < r \leq K$, are associated with the third-order interactions.

The corresponding paired comparison model is given by

$$\begin{aligned} Y_n(\mathbf{i}, \mathbf{j}) = &\sum_{k=1}^{K} (i_k - j_k)\beta_k + \sum_{k<\ell} (i_k i_\ell - j_k j_\ell)\beta_{k\ell} + \sum_{k<\ell<m} (i_k i_\ell i_m - j_k j_\ell j_m)\beta_{k\ell m} \\ &+ \sum_{k<\ell<m<r} (i_k i_\ell i_m i_r - j_k j_\ell j_m j_r)\beta_{k\ell mr} + \varepsilon_n. \end{aligned} \tag{8.3}$$

We similarly note that the comparison depth d describes the number of attributes in which the two alternatives in the choice sets differ satisfying the inequalities $1 \leq d \leq S \leq K$. Hence, the paired comparison model (8.3) is restricted to those paired alternatives for which exactly $S = K$ attributes are presented

$$\begin{aligned} \mathcal{X}^{(S)} = \{(\mathbf{i},\mathbf{j});\ & i_k, j_k \in \{1,-1\} \text{ for } S \text{ components and} \\ & i_k = j_k = 0 \text{ for exactly } K-S \text{ components}\}. \end{aligned} \tag{8.4}$$

Analogously, the design region $\mathcal{X}^{(S)}$ can be also partitioned into disjoint sets such that the pairs in each set differ only in a fixed number d of the attributes. Specifically, for a comparison depth $d = 0, \ldots, S$, let

$$\mathcal{X}_d^{(S)} = \{(\mathbf{i},\mathbf{j}) \in \mathcal{X}^{(S)} : |\{k : i_k \neq j_k\}| = d\}, \tag{8.5}$$

be the set of all pairs of alternatives which differ in exactly d attributes. As before these sets also constitute the orbits with respect to permutations. The D-criterion is also invariant with respect to those permutations, and that the corresponding regression functions (8.2) extended to the design region $\mathcal{X}^{(S)}$ are still linearly equivariant. Hence, one can consider optimality for the full parameter vector (and D-optimality of invariant subvectors). As a consequence, it is sufficient to look for optimality in the class of invariant designs which are uniform on the orbits of fixed comparison depth. In what follows, we denote by $N_d = 2^S \binom{K}{S}\binom{S}{d}$ the number of different pairs in $\mathcal{X}_d^{(S)}$ which vary in exactly d attributes and let ξ_d denotes the uniform approximate design which assigns equal weight $\xi_d(\mathbf{i},\mathbf{j}) = 1/N_d$ to each pair in $\mathcal{X}_d^{(S)}$. In the following Lemma 8.1 and 8.2 we present the information matrices for the corresponding invariant designs.

Lemma 8.1. *Let $d \in \{1, \ldots, S\}$. The uniform design ξ_d on the set $\mathcal{X}_d^{(S)}$ has block diagonal information matrix*

$$\mathbf{M}(\xi_d) = \begin{pmatrix} h_1(d)\mathbf{Id}_K & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & h_2(d)\mathbf{Id}_{\binom{K}{2}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & h_3(d)\mathbf{Id}_{\binom{K}{3}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & h_4(d)\mathbf{Id}_{\binom{K}{4}} \end{pmatrix},$$

where $h_1(d) = \dfrac{4d}{K}$, $h_2(d) = \dfrac{8d(S-d)}{K(K-1)}$, $h_3(d) = \dfrac{4d(3S^2 - 6Sd + 4d^2 - 3S + 2)}{K(K-1)(K-2)}$ *and* $h_4(d) = \dfrac{16d(S-d)(2d^2 - 2Sd + S^2 - 3S + 4)}{K(K-1)(K-2)(K-3)}$.

Proof. We first note that the entries $h_q(d)$, $q = 1, 2, 3$ are the same as in Lemma 6.3 for the second-order interaction models. Now for the third-order interactions we similarly consider attributes k, ℓ, m and r, say, and distinguish between pairs in which all four attributes are distinct, pairs in which three of these attributes k, ℓ and m, say, have distinct levels in the alternatives while the same level is presented in both alternatives for the remaining attribute, pairs in which two of these attributes k, ℓ, say, have distinct levels in the alternatives while the same level is presented in both alternatives for the remaining attribute two attributes and, finally, pairs in which only one of the attributes, say, k has distinct levels in the alternatives while the same level is presented in both alternatives for the three remaining attributes. Then $i_k i_\ell i_m i_r = j_k j_\ell j_m j_r$ in the first and third case, while $i_k i_\ell i_m i_r = -j_k j_\ell j_m j_r$ in the second and last case. Hence,

$$(i_k i_\ell i_m i_r - j_k j_\ell j_m j_r)^2 = 0 \quad \text{for} \quad i_k \neq j_k,\ i_\ell \neq j_\ell,\ i_m \neq j_m \quad \text{and} \quad i_r \neq j_r, \tag{8.6}$$

$$(i_k i_\ell i_m i_r - j_k j_\ell j_m j_r)^2 = 4 \quad \text{for} \quad i_k \neq j_k,\ i_\ell \neq j_\ell,\ i_m \neq j_m \quad \text{and} \quad i_r = j_r, \tag{8.7}$$

$$(i_k i_\ell i_m i_r - j_k j_\ell j_m j_r)^2 = 0 \quad \text{for} \quad i_k \neq j_k,\ i_\ell \neq j_\ell,\ i_m = j_m \quad \text{and} \quad i_r = j_r, \tag{8.8}$$

and

$$(i_k i_\ell i_m i_r - j_k j_\ell j_m j_r)^2 = 4 \quad \text{for} \quad i_k \neq j_k,\ i_\ell = j_\ell,\ i_m = j_m \quad \text{and} \quad i_r = j_r, \tag{8.9}$$

respectively, where the roles of the attributes k, ℓ, m and r may be interchanged.

For given attributes k, ℓ, m and r the pairs with distinct levels in the four attributes occur

$$\tbinom{K-4}{S-4} \tbinom{S-4}{d-4} 2^S$$

times in $\mathcal{X}_d^{(S)}$ with corresponding number of paired comparisons

$$N_d = \tbinom{K-4}{S-4} \tbinom{S-4}{d-4} \frac{K(K-1)(K-2)(K-3)}{S(S-1)(S-2)(S-3)} \frac{S(S-1)(S-2)(S-3)}{d(d-1)(d-2)(d-3)} 2^S,$$

while those which differ in the three attributes occur

$$\tbinom{4}{3} \tbinom{K-4}{S-4} \tbinom{S-4}{d-3} 2^S$$

times in $\mathcal{X}_d^{(S)}$ with corresponding number of paired comparisons

$$N_d = \tbinom{K-4}{S-4} \tbinom{S-4}{d-3} \frac{K(K-1)(K-2)(K-3)}{S(S-1)(S-2)(S-3)} \frac{S(S-1)(S-2)(S-3)}{(S-d)d(d-1)(d-2)} 2^S,$$

while those which differ in the two attributes occur

$$\tbinom{4}{2} \tbinom{K-4}{S-4} \tbinom{S-4}{d-2} 2^S$$

times in $\mathcal{X}_d^{(S)}$ with corresponding number of paired comparisons

$$N_d = \binom{K-4}{S-4}\binom{S-4}{d-2}\frac{K(K-1)(K-2)(K-3)}{S(S-1)(S-2)(S-3)}\frac{S(S-1)(S-2)(S-3)}{(S-d)(S-d-1)d(d-1)}2^S,$$

and, finally, those which differ only in one attribute occur

$$\binom{4}{1}\binom{K-4}{S-4}\binom{S-4}{d-1}2^S$$

times in $\mathcal{X}_d^{(S)}$ with corresponding number of paired comparisons

$$N_d = \binom{K-4}{S-4}\binom{S-4}{d-1}\frac{K(K-1)(K-2)(K-3)}{S(S-1)(S-2)(S-3)}\frac{S(S-1)(S-2)(S-3)}{(S-d)(S-d-1)(S-d-2)d}2^S.$$

Hence, the diagonal elements $h_4(d)$ are given by

$$\begin{aligned}
h_4(d) &= \frac{1}{N_d}\binom{K-4}{S-4}\left(\binom{S-4}{d-3}2^{S+4} + \binom{S-4}{d-1}2^{S+4}\right) \\
&= \frac{16(S-d)d(d-1)(d-2)}{K(K-1)(K-2)(K-3)} + \frac{16(S-d)(S-d-1)(S-d-2)d}{K(K-1)(K-2)(K-3)} \\
&= \frac{16d(S-d)(2d^2-2Sd+S^2-3S+4)}{K(K-1)(K-2)(K-3)}. \qquad (8.10)
\end{aligned}$$

□

We note that for comparison depth $d = 0$ the corresponding function $h_4(0) = 0$.

As general invariant designs ξ result from convex combinations of uniform designs on the comparison depth d with weight w_d, $\xi = \sum_{d=1}^{S} w_d\xi_d$, $w_d \geq 0, \sum_{d=1}^{S} w_d = 1$, then also the invariant designs ξ have block diagonal information matrix:

Lemma 8.2. *Every invariant design $\xi = \sum_{d=1}^{S} w_d\xi_d$ on the set $\mathcal{X}^{(S)}$ has block diagonal information matrix*

$$\mathbf{M}(\xi) = \begin{pmatrix} h_1(\xi)\mathbf{Id}_K & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & h_2(\xi)\mathbf{Id}_{\binom{K}{2}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & h_3(\xi)\mathbf{Id}_{\binom{K}{3}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & h_4(\xi)\mathbf{Id}_{\binom{K}{4}} \end{pmatrix},$$

where $h_q(\xi) = \sum_{d=1}^{S} w_d h_q(d)$, $q = 1, 2, 3, 4$.

It is worthwhile mentioning that the corresponding Theorems 6.1, 6.2 and 6.3 also remain optimal for main effects and lower order interactions in the present third-order interaction model.

The following Remark 8.1 and the numerical results presented in Table 8.1 show the optimality of the corresponding third-order interaction term $h_4(d)$ presented in Lemma 8.1. We note that h_4 is symmetric with respect to $S/2$. Therefore there are, at least two maxima d^* and $S - d^*$ symmetric with respect to $S/2$.

Remark 8.1. *There exists a single comparison depth d^* subject to the profile strength S such that the uniform design ξ_{d^*} is D-optimal for the third-order interaction effects $(\beta_{k\ell mr})^\top_{k<\ell<m<r}$.*

We note that the corresponding values of d^* presented in Table 8.1 were obtained by first calculating the values of $h_4(d)$ and determining the maximum.

TABLE 8.1: Values of the Optimal Comparison Depths d^* of the D-Optimal Uniform Designs ξ_{d^*} for the Third-Order Interactions with $S \leq K$ Binary Attributes

S	4	5	6	7	8	9	10	11	12
d^*	1	1	1	1	2	2	2	3	3

We mention that the corresponding invariant design ξ has a variance function of the form $V((\mathbf{i},\mathbf{j}),\xi) = (\mathbf{f}(\mathbf{i}) - \mathbf{f}(\mathbf{j}))^\top \mathbf{M}(\xi)^{-1}(\mathbf{f}(\mathbf{i}) - \mathbf{f}(\mathbf{j}))$, which as before is also invariant with respect to permutation of levels and attributes and, hence, constant on the orbits $\mathcal{X}_d^{(S)}$ of fixed comparison depth d. The value of the variance function for the invariant design ξ evaluated at comparison depth d is denoted as $V(d,\xi)$, $V(d,\xi) = V((\mathbf{i},\mathbf{j}),\xi)$ for all $(\mathbf{i},\mathbf{j})$ on $\mathcal{X}_d^{(S)}$.

Theorem 8.1. *For every invariant design ξ the variance function $V(d,\xi)$ on $\mathcal{X}_d^{(S)}$ is given by*

$$V(d,\xi) = 4d\left(\tfrac{1}{h_1(\xi)} + \tfrac{S-d}{h_2(\xi)} + \tfrac{3S^2-6dS+4d^2-3S+2}{6h_3(\xi)} + \tfrac{(S-d)(2d^2-2Sd+S^2-3S+4)}{6h_4(\xi)}\right).$$

Proof. From Lemma 8.2 the inverse of the corresponding information matrix $\mathbf{M}(\xi)$ is given by

$$\mathbf{M}(\xi)^{-1} = \begin{pmatrix} \frac{1}{h_1(\xi)}\mathbf{Id}_K & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{1}{h_2(\xi)}\mathbf{Id}_{\binom{K}{2}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{1}{h_3(\xi)}\mathbf{Id}_{\binom{K}{3}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{1}{h_4(\xi)}\mathbf{Id}_{\binom{K}{4}} \end{pmatrix}.$$

Hence, we obtain for the variance function

$$
\begin{aligned}
V((\mathbf{i},\mathbf{j}),\xi) =& (\mathbf{f}(\mathbf{i})-\mathbf{f}(\mathbf{j}))^\top \mathbf{M}(\xi)^{-1}(\mathbf{f}(\mathbf{i})-\mathbf{f}(\mathbf{j})) \\
=& \frac{1}{h_1(\xi)}\sum_{k=1}^{K}(i_k-j_k)^2 \\
&+\frac{1}{h_2(\xi)}\sum_{k<\ell}(i_k i_\ell - j_k j_\ell)^2 \\
&+\frac{1}{h_3(\xi)}\sum_{k<\ell<m}(i_k i_\ell i_m - j_k j_\ell j_m)^2 \\
&+\frac{1}{h_4(\xi)}\sum_{k<\ell<m<r}(i_k i_\ell i_m i_r - j_k j_\ell j_m j_r)^2.
\end{aligned} \tag{8.11}
$$

For a pair $(\mathbf{i},\mathbf{j}) \in \mathcal{X}_d^{(S)}$ of comparison depth d there are $(S-d)\binom{d}{3}$ third-order interaction terms for which $(i_k i_\ell i_m i_r)$ and $(j_k j_\ell j_m j_r)$ of the associated four attributes k,ℓ,m and r differ in exactly three of the attributes, and there are $d\binom{S-d}{3}$ third-order interaction terms for which $(i_k i_\ell i_m i_r)$ and $(j_k j_\ell j_m j_r)$ differ in exactly one attribute. As a result, there are

$$
\begin{aligned}
(S-d)\binom{d}{3}+d\binom{S-d}{3} &= (S-d)d(d-1)(d-2)/6 + d(S-d)(S-d-1)(S-d-2)/6 \\
&= d((S-d)(2d^2-2Sd+S^2-3S+4))/6
\end{aligned}
$$

non-zero entries in the fourth sum on the right hand side of (8.11), and this sum equals $4d((S-d)(2d^2-2Sd+S^2-3S+4))/6$.

Now by substituting the corresponding results into (8.11) for fixed S we see that the value of the variance function depends on the pair $(\mathbf{i},\mathbf{j}) \in \mathcal{X}_d^{(S)}$ only through its comparison depth d and obtain the representation of the variance function. □

We mention that for comparison depth $d=0$ the corresponding variance function $V(0,\xi)=0$.

The variance function $V(d,\xi)$ simplifies if the invariant design ξ is concentrated on a single comparison depth.

Corollary 8.1. *For a uniform design $\xi_{d'}$ on a single comparison depth d' the variance function is given by*

$$
V(d,\xi_{d'}) = \frac{d}{d'}\left(p_1 + p_2\tfrac{S-d}{S-d'} + p_3\tfrac{3S^2-6dS+4d^2-3S+2}{3S^2-6d'S+4d'^2-3S+2} + p_4\tfrac{(S-d)(2d^2-2Sd+S^2-3S+4)}{(S-d')(2d'^2-2Sd'+S^2-3S+4)}\right).
$$

Proof. In view of Theorem 8.1 it is sufficient to note that the representation of the variance function follows immediately by inserting the values of $h_q(\xi_d)$ from Lemma 8.1 and $p_q=\binom{K}{q}$, $q=1,2,3,4$. □

We note that for $d = d'$ the varaince $V(d, \xi_d) = p_1 + p_2 + p_3 + p_4 = p$, which shows the D-optimality of ξ_d by the Kiefer and Wolfowitz (1960) equivalence theorem.

In the following theorem we employ the D-criterion to derive optimal design for the main effects, the first-order interactions, the second-order interactions and the third-order interactions. We similarly note that a single comparison depth d may be sufficient for the identifiability of all model parameters. Theorem 8.2 gives an upper bound on the number of comparison depths required for a D-optimal design.

Theorem 8.2. *In the third-order interactions model the D-optimal design is supported on at most four orbits $d^*, d_1^*, d^* + 1$ and $d_1^* + 1$.*

Proof. Let ξ^* be an invariant D-optimal design with weights w_d^* on the comparison depths d for which the variance function $V(d, \xi^*)$ is equal to the number of parameters p for all d such that $w_d^* > 0$. By Theorem 8.1 the variance function is a polynomial of degree 4 in the comparison depth d with negative leading coefficient. For integer d the variance function $V(d, \xi^*)$ may thus be equal to p for, at most, four different values of d. Now, by the Kiefer and Wolfowitz (1960) equivalence theorem itself $V(d, \xi^*) \leq p$ for all $d = 0, 1, \ldots, S$. Hence, by the shape of the variance function we obtain that $V(d, \xi^*) = p$ may occur only at, at most two adjacent pairs $d^*, d^* + 1$ and $d_1^*, d_1^* + 1$, say. □

Further for the case $S = K = 4$ of full profiles the D-optimal design can be given explicitly. It is worth mentioning that this situation of $S = K = 4$ of full profiles can also be regarded as complete interactions. Analogous result can be found in Theorem 4 of Graßhoff et al. (2003). Here we show that the corresponding result can be obtained explicitly.

Theorem 8.3. *If $S = K = 4$ then the design $\xi^* = \frac{4}{15}\xi_1 + \frac{2}{5}\xi_2 + \frac{4}{15}\xi_3 + \frac{1}{15}\xi_4$ which is uniform on all pairs with non-zero comparison depth is D-optimal in the third-order interactions model.*

Proof. For the design ξ^* we obtain $h_1(\xi^*) = 8/15$, $h_2(\xi^*) = 2/15$, $h_3(\xi^*) = 1/30$ and $h_4(\xi^*) = 1/120$. Inserting this into the variance function of Theorem 8.1 yields $V(d, \xi^*) = 5d(-1/2d^3 + 5d^2 - 35/2d + 25)/4$ which results in $V(1, \xi^*) = V(2, \xi^*) = V(3, \xi^*) = V(4, \xi^*) = 15$. Hence, the variance function is bounded by the number of parameters $p = 15$ which establishes the D-optimality of ξ^* by virtue of the Kiefer-Wolfowitz equivalence theorem. □

It should be noted that in this case $d^* = 1$ and $d_1^* = 3$ in Theorem 8.2. We further note that for $S = K = 4$ all four comparison depths are needed for D-optimality.

Moreover, for full profiles $S = K$ between 5 and 12, intermediate comparison depths d, d_1 and weights w_d, w_{d_1} the numerical results presented in Table 8.2 were obtained by direct maximization of $\ln(\det(\mathbf{M}(w_d\xi_d + (1-w_d)\xi_{d_1})))$ for the corresponding optimal comparison depth d^* and optimal weights $w^*_{d^*}$ where $1 - w^*_{d^*} = w^*_{d^*_1}$. In particular, by considering the invariant designs $\xi^* = w^*_{d^*}\xi_{d^*} + (1-w^*_{d^*})\xi_{d^*_1}$ the numerical results show that two different comparison depths d^* and d^*_1 may be needed for D-optimality, which is verified by the Kiefer and Wolfowitz (1960) equivalence theorem in Table 8.3. Specifically, for the various choices of profile strengths S between 5 and 12 the corresponding optimal comparison depths $d^* = (S-1)/2$ for $S = 5$, $d^* = (S-2)/2$ for $S = 6, 8$, $d^* = (S-3)/2$ for $S = 7, 9, 11$ and $d^* = (S-4)/2$ for $S = 10, 12$.

TABLE 8.2: Optimal Designs with Intermediate Comparison Depths d^* and Optimal Weights $w^*_{d^*}$ for $S = K$ Binary Attributes

	S							
	5	6	7	8	9	10	11	12
d^*	2	2	2	3	3	3	4	4
$w^*_{d^*}$	0.667	0.714	0.750	0.667	0.700	0.727	0.667	0.692
d^*_1	4	5	6	6	7	8	8	9
$w^*_{d^*_1}$	0.333	0.286	0.250	0.333	0.300	0.273	0.333	0.308

It is worthwhile mentioning that generally for the situation when $S = K$ the optimal comparison depths d^* and d^*_1, and the corresponding optimal weights $w^*_{d^*}$ satisfy the condition $w^*_{d^*} = d^*_1/(d^* + d^*_1)$ for $d^* = [(K+1)/3]$ and $d^* + d^*_1 = K + 1$. We further note that the particular case when $S = K = 4$ can be found in the corresponding Theorem 8.3 where the design ξ^* is uniform on all four comparison depth.

We now show that by direct maximization of $\ln(\det(\mathbf{M}(w_d\xi_d + (1-w_d)\xi_{d_1})))$ for the corresponding optimal comparison depth d^* and optimal weights $w^*_{d^*}$ where $1 - w^*_{d^*} = w^*_{d^*_1}$, the corresponding numerical values of the optimal weights $w^*_{d^*}$ in Table 8.2 can be determined analytically as follows.

For the case, when $S = 5$ we consider the intermediate comparison depth $d^* = (S-1)/2$ as a candidate for beign optimal. Now by Lemma 8.1 the entries of the information matrix $\mathbf{M}(\xi)$ are specified as

$$\begin{aligned}
h_1(\xi) &= wh_1(d^*) + (1-w)h_1(d_1^*) \\
&= \frac{2S - 8w + 6}{K}, \\
h_2(\xi) &= wh_2(d^*) + (1-w)h_2(d_1^*) \\
&= \frac{2S^2 + 16w - 18}{K(K-1)}, \\
h_3(\xi) &= wh_3(d^*) + (1-w)h_3(d_1^*) \\
&= \frac{2S^3 - 6S^2 + 24Sw - 14S - 72w + 66}{K(K-1)(K-2)}, \quad \text{and} \\
h_4(\xi) &= wh_4(d^*) + (1-w)h_4(d_1^*) \\
&= \frac{2S^3 - 6S^2 - 2S - 96w + 102}{K(K-1)(K-2)(K-3)}.
\end{aligned}$$

Now as the determinant of the information matrix $\mathbf{M}(\xi)$ is proportional to $h_1(\xi)^{p_1}$ $h_2(\xi)^{p_2} h_3(\xi)^{p_3} h_4(\xi)^{p_4}$, we obtain

$$\begin{aligned}
\ln\det(\mathbf{M}(\xi)) = c &+ p_1 \ln(2S - 8w + 6) + p_2 \ln(2S^2 + 16w - 18) \\
&+ p_3 \ln(2S^3 - 6S^2 + 24Sw - 14S - 72w + 66) \\
&+ p_4 \ln(2S^3 - 6S^2 - 2S - 96w + 102),
\end{aligned}$$

where c is a constant independent of the weight w. Taking derivatives with respect to w we obtain

$$\begin{aligned}
\frac{\partial}{\partial w} \ln\det(\mathbf{M}(\xi)) = &-\frac{8p_1}{2S - 8w + 6} + \frac{16p_2}{2S^2 + 16w - 18} \\
&+ \frac{p_3(24S - 72)}{2S^3 - 6S^2 + 24Sw - 14S - 72w + 66} \\
&- \frac{96p_4}{2S^3 - 6S^2 - 2S - 96w + 102},
\end{aligned}$$

which has root

$$w = w^*_{d^*} = \frac{S+3}{2S+2}.$$

This root $w^*_{d^*}$ gives a maximum for the determinant. The design ξ^* is thus D-optimal when we consider the reduced design region $\mathcal{X}^{(S)}_{d^*} \cup \mathcal{X}^{(S)}_{d_1^*}$.

Further for the case when $S = 6$ and 8 we consider the intermediate comparison depth $d^* = (S-2)/2$ as a candidate for beign optimal. Similarly, the entries of the

information matrix $\mathbf{M}(\xi)$ are specified as

$$\begin{aligned}
h_1(\xi) &= wh_1(d^*) + (1-w)h_1(d_1^*) \\
&= \frac{2S - 12w + 8}{K}, \\
h_2(\xi) &= wh_2(d^*) + (1-w)h_2(d_1^*) \\
&= \frac{2S^2 + 24w - 32}{K(K-1)}, \\
h_3(\xi) &= wh_3(d^*) + (1-w)h_3(d_1^*) \\
&= \frac{2S^3 - 6S^2 + 36Sw - 20S - 168w + 144}{K(K-1)(K-2)}, \quad \text{and} \\
h_4(\xi) &= wh_4(d^*) + (1-w)h_4(d_1^*) \\
&= \frac{2S^4 - 12S^3 + 16S^2 - 144Sw + 192S + 672w - 768}{K(K-1)(K-2)(K-3)}.
\end{aligned}$$

Now as the determinant of the information matrix $\mathbf{M}(\xi)$ is proportional to $h_1(\xi)^{p_1}$ $h_2(\xi)^{p_2} h_3(\xi)^{p_3} h_4(\xi)^{p_4}$, we obtain

$$\begin{aligned}
\ln\det(\mathbf{M}(\xi)) = c &+ p_1 \ln(2S - 12w + 8) + p_2 \ln(2S^2 + 24w - 32) \\
&+ p_3 \ln(2S^3 - 6S^2 + 36Sw - 20S - 168w + 144) \\
&+ p_4 \ln(2S^4 - 12S^3 + 16S^2 - 144Sw + 192S + 672w - 768),
\end{aligned}$$

where c is a constant independent of the weight w. Taking derivatives with respect to w we obtain

$$\begin{aligned}
\frac{\partial}{\partial w} \ln\det(\mathbf{M}(\xi)) = &-\frac{12p_1}{2S - 12w + 8} + \frac{24p_2}{2S^2 + 24w - 32} \\
&+ \frac{p_3(36S - 168)}{2S^3 - 6S^2 + 36Sw - 20S - 168w + 144} \\
&+ \frac{p_4(-144S + 672)}{2S^4 - 12S^3 + 16S^2 - 144Sw + 192S + 672w - 768},
\end{aligned}$$

which has root

$$w = w_{d^*}^* = \frac{S+4}{2S+2}$$

This root $w_{d^*}^*$ gives a maximum for the determinant. The design ξ^* is thus D-optimal when we consider the reduced design region $\mathcal{X}_{d^*}^{(S)} \cup \mathcal{X}_{d_1^*}^{(S)}$.

Also for the case when $S = 7, 9$ and 11 we consider the intermediate comparison depth $d^* = (S-3)/2$ as a candidate for beign optimal. The entries of the corresponding

information matrix $\mathbf{M}(\xi)$ are specified as

$$
\begin{aligned}
h_1(\xi) &= wh_1(d^*) + (1-w)h_1(d_1^*) \\
&= \frac{2S - 16w + 10}{K}, \\
h_2(\xi) &= wh_2(d^*) + (1-w)h_2(d_1^*) \\
&= \frac{2S^2 + 32w - 50}{K(K-1)}, \\
h_3(\xi) &= wh_3(d^*) + (1-w)h_3(d_1^*) \\
&= \frac{2S^3 - 6S^2 + 48Sw - 26S - 336w + 270}{K(K-1)(K-2)}, \quad \text{and} \\
h_4(\xi) &= wh_4(d^*) + (1-w)h_4(d_1^*) \\
&= \frac{2S^4 - 12S^3 + 16S^2 - 192Sw + 300S + 1344w - 1650}{K(K-1)(K-2)(K-3)}.
\end{aligned}
$$

Now as the determinant of the information matrix $\mathbf{M}(\xi)$ is proportional to $h_1(\xi)^{p_1}$ $h_2(\xi)^{p_2} h_3(\xi)^{p_3} h_4(\xi)^{p_4}$, we obtain

$$
\begin{aligned}
\ln\det(\mathbf{M}(\xi)) = c &+ p_1 \ln(2S - 16w + 10) + p_2 \ln(2S^2 + 32w - 50) \\
&+ p_3 \ln(2S^3 - 6S^2 + 48Sw - 26S - 336w + 270) \\
&+ p_4 \ln(2S^4 - 12S^3 + 16S^2 - 192Sw + 300S + 1344w - 1650),
\end{aligned}
$$

where c is a constant independent of the weight w. Taking derivatives with respect to w we obtain

$$
\begin{aligned}
\frac{\partial}{\partial w} \ln\det(\mathbf{M}(\xi)) = &-\frac{16p_1}{2S - 16w + 10} + \frac{32p_2}{2S^2 + 32w - 50} \\
&+ \frac{p_3(48S - 336)}{2S^3 - 6S^2 + 48Sw - 26S - 336w + 270} \\
&+ \frac{p_4(-192S + 1344)}{2S^4 - 12S^3 + 16S^2 - 192Sw + 300S + 1344w - 1650},
\end{aligned}
$$

which has root

$$
w = w_{d^*}^* = \frac{S+5}{2S+2}
$$

This root $w_{d^*}^*$ gives a maximum for the determinant. The design ξ^* is thus D-optimal when we consider the reduced design region $\mathcal{X}_{d^*}^{(S)} \cup \mathcal{X}_{d_1^*}^{(S)}$.

Finally, for the case when $S = 10$ and 12 we consider the intermediate comparison depth $d^* = (S-4)/2$ as a candidate for beign optimal. The entries of the corresponding

information matrix $\mathbf{M}(\xi)$ are specified as

$$\begin{aligned}
h_1(\xi) &= wh_1(d^*) + (1-w)h_1(d_1^*) \\
&= \frac{2S - 20w + 12}{K}, \\
h_2(\xi) &= wh_2(d^*) + (1-w)h_2(d_1^*) \\
&= \frac{2S^2 + 40w - 72}{K(K-1)}, \\
h_3(\xi) &= wh_3(d^*) + (1-w)h_3(d_1^*) \\
&= \frac{2S^3 - 6S^2 + 60Sw - 32S - 600w + 456}{K(K-1)(K-2)}, \quad \text{and} \\
h_4(\xi) &= wh_4(d^*) + (1-w)h_4(d_1^*) \\
&= \frac{2S^4 - 12S^3 + 16S^2 - 240Sw + 432S + 2400w - 3168}{K(K-1)(K-2)(K-3)}.
\end{aligned}$$

Now as the determinant of the information matrix $\mathbf{M}(\xi)$ is proportional to $h_1(\xi)^{p_1}$ $h_2(\xi)^{p_2} h_3(\xi)^{p_3} h_4(\xi)^{p_4}$, we obtain

$$\begin{aligned}
\ln\det(\mathbf{M}(\xi)) = c &+ p_1 \ln(2S - 20w + 12) + p_2 \ln(2S^2 + 40w - 72) \\
&+ p_3 \ln(2S^3 - 6S^2 + 60Sw - 32S - 600w + 456) \\
&+ p_4 \ln(2S^4 - 12S^3 + 16S^2 - 240Sw + 432S + 2400w - 3168),
\end{aligned}$$

where c is a constant independent of the weight w. Taking derivatives with respect to w we obtain

$$\begin{aligned}
\frac{\partial}{\partial w} \ln\det(\mathbf{M}(\xi)) = &-\frac{20p_1}{2S - 20w + 12} + \frac{40p_2}{2S^2 + 40w - 72} \\
&+ \frac{p_3(60S - 600)}{2S^3 - 6S^2 + 60Sw - 32S - 600w + 456} \\
&+ \frac{p_4(-240S + 2400)}{2S^4 - 12S^3 + 16S^2 - 240Sw + 432S + 2400w - 3168},
\end{aligned}$$

which has root

$$w = w_{d^*}^* = \frac{S+6}{2S+2}$$

This root $w_{d^*}^*$ gives a maximum for the determinant. The design ξ^* is thus D-optimal when we consider the reduced design region $\mathcal{X}_{d^*}^{(S)} \cup \mathcal{X}_{d_1^*}^{(S)}$.

Again, the comparison depth d^* is an integer solution for the maximum of the variance function which proofs the D-optimality of the design ξ^* by virtue of the Kiefer and Wolfowitz (1960) equivalence theorem. □

In the following Table 8.3 we present values of the normalized variance function

$V(d, \xi^*)/p$, which shows D-optimality of the corresponding design ξ^* in Table 8.2 by virtue of the equivalence theorem by Kiefer and Wolfowitz (1960).

TABLE 8.3: Values of the Variance Function of the D-Optimal Design ξ^* for $S = K$ Binary Attributes (Boldface **1** Corresponds to the Optimal Comparison Depths d^* and d_1^*)

	d											
K	1	2	3	4	5	6	7	8	9	10	11	12
5	0.938	**1**	0.938	**1**	0.938							
6	0.850	**1**	0.950	0.950	**1**	0.850						
7	0.792	**1**	0.982	0.952	0.982	**1**	0.792					
8	0.759	0.998	**1**	0.954	0.954	**1**	0.998	0.759				
9	0.693	0.958	**1**	0.966	0.945	0.966	**1**	0.958	0.693			
10	0.644	0.925	**1**	0.985	0.958	0.958	0.985	**1**	0.925	0.644		
11	0.609	0.901	0.999	**1**	0.973	0.960	0.973	**1**	0.999	0.901	0.609	
12	0.566	0.860	0.979	**1**	0.982	0.963	0.963	0.982	**1**	0.979	0.860	0.566

Remark 8.2. *In Table 8.4 the values of the normalized variance function $V(d, \xi^*)/p$ show that at most two comparison depths d^* and d_1^* may be needed for D-optimality. In particular, for the case when $S = 4$ and $K = 5$, one has to consider the corresponding invariant design $\xi^* = w_{d^*}^* \xi_{d^*} + (1 - w_{d^*}^*)\xi_{d_1^*}$ with $d^* = 1$ and $d_1^* = 3$ and corresponding weights $w_{d^*}^* = 0.936$ and $w_{d_1^*}^* = 0.064$, respectively. Accordingly, for the case when $S = 4$ and $K > 5$ the D-optimal design $\xi^* = \xi_{d^*}$ has to be considered which is uniform on the optimal comparison depth d^* (in boldface).*

TABLE 8.4: Values of the Variance Function of the D-Optimal Design ξ^* for $S < K$ Binary Attributes (Boldface **1** Corresponds to the Optimal Comparison Depths d^* and d_1^*)

		d			
K	S	1	2	3	4
5	4	**1**	0.972	**1**	0.994
6	4	**1**	0.810	0.976	0.905
7	4	**1**	0.667	0.905	0.762
8	4	**1**	0.560	0.868	0.658
9	4	**1**	0.478	0.851	0.580
10	4	**1**	0.416	0.844	0.519
11	4	**1**	0.366	0.843	0.471
12	4	**1**	0.326	0.845	0.430

9 Optimal Designs for Third-Order Interactions General Level Models

This chapter is a generalization of the results in Chapter 8 for the case of two-level attributes to the case of common number of general levels $i_k = 1, \ldots, v$ for each attribute $k = 1, \ldots, K$. For this situation we introduce an appropriate model and derive optimal designs in the presence of third-order interactions.

In the present setting of direct response (utility) $\tilde{Y}_{na}(\mathbf{i})$, model (8.1) can be reformulated as

$$\tilde{Y}_{na}(\mathbf{i}) = \mu_n + \sum_{k=1}^{K} \alpha_{i_k}^{(k)} + \sum_{k<\ell} \alpha_{i_k i_\ell}^{(k\ell)} + \sum_{k<\ell<m} \alpha_{i_k i_\ell i_m}^{(k\ell m)} + \sum_{k<\ell<m<r} \alpha_{i_k i_\ell i_m i_r}^{(k\ell mr)} + \tilde{\varepsilon}_{na}, \tag{9.1}$$

where $\alpha_{i_k}^{(k)}$, $\alpha_{i_k i_\ell}^{(k\ell)}$, $\alpha_{i_k i_\ell i_m}^{(k\ell m)}$ are defined in (8.1), and $\alpha_{i_k i_\ell i_m r}^{(k\ell mr)}$ is the third-order interaction effect of the k-th, ℓ-th, m-th and r-th attribute when the correpsonding levels are $i_k = 1, \ldots, v, i_\ell = 1, \ldots, v$, $i_m = 1, \ldots, v$ and $i_r = 1, \ldots, v$. Then by the common identifiability conditions of effects-coding the following equalities, in particular for the third-order interactions effects $\alpha_{i_k i_\ell i_m i_r}^{(k\ell mr)} = \beta_{i_k i_\ell i_m i_r}^{(k\ell mr)}$ for $i_k, i_\ell, i_m, i_r = 1, \ldots, v-1$ hold:

$$\alpha_{i_k i_\ell i_m v}^{(k\ell mr)} = -\sum_{i_r=1}^{v-1} \beta_{i_k i_\ell i_m i_r}^{(k\ell mr)}, \quad \alpha_{i_k i_\ell v i_r}^{(k\ell mr)} = -\sum_{i_m=1}^{v-1} \beta_{i_k i_\ell i_m i_r}^{(k\ell mr)}, \quad \alpha_{i_k v i_m i_r}^{(k\ell mr)} = -\sum_{i_\ell=1}^{v-1} \beta_{i_k i_\ell i_m i_r}^{(k\ell mr)},$$
$$\alpha_{v i_\ell i_m i_r}^{(k\ell mr)} = -\sum_{i_k=1}^{v-1} \beta_{i_k i_\ell i_m i_r}^{(k\ell mr)}, \quad \alpha_{i_k i_\ell v v}^{(k\ell mr)} = \sum_{i_m=1}^{v-1}\sum_{i_r=1}^{v-1} \beta_{i_k i_\ell i_m i_r}^{(k\ell mr)}, \quad \alpha_{i_k v v i_r}^{(k\ell mr)} = \sum_{i_\ell=1}^{v-1}\sum_{i_m=1}^{v-1} \beta_{i_k i_\ell i_m i_r}^{(k\ell mr)},$$
$$\alpha_{v v i_m i_r}^{(k\ell mr)} = \sum_{i_k=1}^{v-1}\sum_{i_\ell=1}^{v-1} \beta_{i_k i_\ell i_m i_r}^{(k\ell mr)}, \quad \alpha_{v i_\ell v i_r}^{(k\ell mr)} = \sum_{i_k=1}^{v-1}\sum_{i_m=1}^{v-1} \beta_{i_k i_\ell i_m i_r}^{(k\ell mr)}, \quad \alpha_{i_k v i_m v}^{(k\ell mr)} = \sum_{i_\ell=1}^{v-1}\sum_{i_r=1}^{v-1} \beta_{i_k i_\ell i_m i_r}^{(k\ell mr)},$$
$$\alpha_{v i_\ell i_m v}^{(k\ell mr)} = \sum_{i_k=1}^{v-1}\sum_{i_r=1}^{v-1} \beta_{i_k i_\ell i_m i_r}^{(k\ell mr)}, \quad \alpha_{i_k v v v}^{(k\ell mr)} = -\sum_{i_\ell=1}^{v-1}\sum_{i_m=1}^{v-1}\sum_{i_r=1}^{v-1} \beta_{i_k i_\ell i_m i_r}^{(k\ell mr)},$$
$$\alpha_{v i_\ell v v}^{(k\ell mr)} = -\sum_{i_k=1}^{v-1}\sum_{i_m=1}^{v-1}\sum_{i_r=1}^{v-1} \beta_{i_k i_\ell i_m i_r}^{(k\ell mr)}, \quad \alpha_{v v i_m v}^{(k\ell mr)} = -\sum_{i_k=1}^{v-1}\sum_{i_\ell=1}^{v-1}\sum_{i_r=1}^{v-1} \beta_{i_k i_\ell i_m i_r}^{(k\ell mr)},$$
$$\alpha_{v v v i_r}^{(k\ell mr)} = -\sum_{i_k=1}^{v-1}\sum_{i_\ell=1}^{v-1}\sum_{i_m=1}^{v-1} \beta_{i_k i_\ell i_m i_r}^{(k\ell mr)} \quad \text{and} \quad \alpha_{v v v v}^{(k\ell mr)} = \sum_{i_k=1}^{v-1}\sum_{i_\ell=1}^{v-1}\sum_{i_m=1}^{v-1}\sum_{i_r=1}^{v-1} \beta_{i_k i_\ell i_m i_r}^{(k\ell mr)}.$$

The parameters for the main effects, the first-order interactions, the second-order interactions presented in (7.2) and the third-order interactions, respectively, can be summarized as follows

$$\begin{aligned}
\boldsymbol{\beta}_k &= (\beta^{(k)}_{i_k})_{i_k=1,\ldots,v-1},\\
\boldsymbol{\beta}_{k\ell} &= (\beta^{(k\ell)}_{i_k i_\ell})_{i_k=1,\ldots,v-1,\ i_\ell=1,\ldots,v-1},\\
\boldsymbol{\beta}_{k\ell m} &= (\beta^{(k\ell m)}_{i_k i_\ell i_m})_{i_k=1,\ldots,v-1,\ i_\ell=1,\ldots,v-1,\ i_m=1,\ldots,v-1} \text{ and}\\
\boldsymbol{\beta}_{k\ell mr} &= (\beta^{(k\ell mr)}_{i_k i_\ell i_m i_r})_{i_k=1,\ldots,v-1,\ i_\ell=1,\ldots,v-1,\ i_m=1,\ldots,v-1, i_r=1,\ldots,v-1},
\end{aligned}$$

where e. g. $\boldsymbol{\beta}_{k\ell mr}$ describes the effect of the third-order interaction of the k-th, ℓ-th, m-th and r-th attribute. As a consequence, the minimal vector of parameters of dimension $p = K(v-1) + \binom{K}{2}(v-1)^2 + \binom{K}{3}(v-1)^3 + \binom{K}{4}(v-1)^4$ is given by

$$\boldsymbol{\beta} = ((\boldsymbol{\beta}_k)^\top_{k=1,\ldots,K}, (\boldsymbol{\beta}_{k\ell})^\top_{k<\ell}, (\boldsymbol{\beta}_{k\ell m})^\top_{k<\ell<m}(\boldsymbol{\beta}_{k\ell mr})^\top_{k<\ell<m<r})^\top. \tag{9.2}$$

With the above notation the model (9.1) can be reformulated as

$$\begin{aligned}
\tilde{Y}_{na}(\mathbf{i}) = \mu_n &+ \sum_{k=1}^{K} \mathbf{f}_1(i_k)^\top \boldsymbol{\beta}_k + \sum_{k<\ell} (\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell))^\top \boldsymbol{\beta}_{k\ell}\\
&+ \sum_{k<\ell<m} (\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) \otimes \mathbf{f}_1(i_m))^\top \boldsymbol{\beta}_{k\ell m}\\
&+ \sum_{k<\ell<m<r} (\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) \otimes \mathbf{f}_1(i_m) \otimes \mathbf{f}_1(i_r))^\top \boldsymbol{\beta}_{k\ell mr} + \tilde{\varepsilon}_{na},
\end{aligned} \tag{9.3}$$

where $\otimes$ denotes the Kronecke product of vectors or matrices, respectively, which results in the p dimensional vector

$$\begin{aligned}
\mathbf{f}(\mathbf{i}) = (&\mathbf{f}_1(i_1)^\top, \ldots, \mathbf{f}_1(i_K)^\top, \mathbf{f}_1(i_1)^\top \otimes \mathbf{f}_1(i_2)^\top, \ldots, \mathbf{f}_1(i_{K-1}) \otimes \mathbf{f}_1(i_K)^\top,\\
&\mathbf{f}_1(i_1)^\top \otimes \mathbf{f}_1(i_2)^\top \otimes \mathbf{f}_1(i_3)^\top, \ldots, \mathbf{f}_1(i_{K-2})^\top \otimes \mathbf{f}_1(i_{K-1})^\top \otimes \mathbf{f}_1(i_K)^\top,\\
&\mathbf{f}_1(i_1) \otimes \mathbf{f}_1(i_2) \otimes \mathbf{f}_1(i_3) \otimes \mathbf{f}_1(i_4), \ldots, \mathbf{f}_1(i_{K-3}) \otimes \mathbf{f}_1(i_{K-2}) \otimes \mathbf{f}_1(i_{K-1}) \otimes \mathbf{f}_1(i_K)^\top),^\top
\end{aligned} \tag{9.4}$$

where the first components $\mathbf{f}_1(i_1), \ldots, \mathbf{f}_1(i_K)$, the second components $\mathbf{f}_1(i_1) \otimes \mathbf{f}_1(i_2), \ldots, \mathbf{f}_1(i_{K-1}) \otimes \mathbf{f}_1(i_K)$, the third components $\mathbf{f}_1(i_1) \otimes \mathbf{f}_1(i_2) \otimes \mathbf{f}_1(i_3), \ldots, \mathbf{f}_1(i_{K-2}) \otimes \mathbf{f}_1(i_{K-1}) \otimes \mathbf{f}_1(i_K)$ of $\mathbf{f}(\mathbf{i})$ are defined in (7.4), and the remaining components $\mathbf{f}_1(i_1) \otimes \mathbf{f}_1(i_2) \otimes \mathbf{f}_1(i_3) \otimes \mathbf{f}_1(i_4), \ldots, \mathbf{f}_1(i_{K-3}) \otimes \mathbf{f}_1(i_{K-2}) \otimes \mathbf{f}_1(i_{K-1}) \otimes \mathbf{f}_1(i_K)$ of $\mathbf{f}(\mathbf{i})$ are associated with the third-order interactions and have $p_4 = (1/24)K(K-1)(K-2)(K-3)(v-1)^4$ parameter vector $(\boldsymbol{\beta}_{k\ell mr})_{1\leq k<\ell<m<r\leq K}$.

The corresponding paired comparison model is given by

$$
\begin{aligned}
Y_n(\mathbf{i},\mathbf{j}) = &\sum_{k=1}^{K}(\mathbf{f}_1(i_k)-\mathbf{f}_1(j_k))^\top \boldsymbol{\beta}_k + \sum_{k<\ell}((\mathbf{f}_1(i_k)\otimes\mathbf{f}_1(i_\ell)) - (\mathbf{f}_1(j_k)\otimes\mathbf{f}_1(j_\ell)))^\top \boldsymbol{\beta}_{k\ell} \\
&+ \sum_{k<\ell<m}((\mathbf{f}_1(i_k)\otimes\mathbf{f}_1(i_\ell)\otimes\mathbf{f}_1(i_m)) - (\mathbf{f}_1(j_k)\otimes\mathbf{f}_1(j_\ell)\otimes\mathbf{f}_1(j_m)))^\top \boldsymbol{\beta}_{k\ell m} \\
&+ \sum_{k<\ell<m<r}((\mathbf{f}_1(i_k)\otimes\mathbf{f}_1(i_\ell)\otimes\mathbf{f}_1(i_m)\otimes\mathbf{f}_1(i_r)) \\
&\qquad - (\mathbf{f}_1(j_k)\otimes\mathbf{f}_1(j_\ell)\otimes\mathbf{f}_1(j_m)\otimes\mathbf{f}_1(j_r)))^\top \boldsymbol{\beta}_{k\ell mr} + \varepsilon_n. \qquad (9.5)
\end{aligned}
$$

In the present setting we similarly point out that for the case of partial profiles the k-th attribute that is not shown has corresponding level $i_k = 0$, and that the corresponding regression functions are given by $\mathbf{f}_k(0) = \mathbf{f}_1(0) = \mathbf{0}$. Hence, the paired comparison model (9.5) is thus restricted to those paired alternatives for which exactly S attributes are presented (as similarly defined in (7.23))

$$
\begin{aligned}
\mathcal{X}^{(S)} = \{(\mathbf{i},\mathbf{j});\ & i_k, j_k \in \{1,\dots,v\} \text{ for } S \text{ components and} \\
& i_k = j_k = 0 \text{ for exactly } K-S \text{ components}\}.
\end{aligned} \qquad (9.6)
$$

We similarly note that the design region $\mathcal{X}^{(S)} = \mathcal{X}^{(K)}$ (in the case of full profiles $S = K$) can be partitioned into disjoint sets such that the pairs in each set differ only in a fixed number d of the attributes. Specifically, for a comparison depth $d = 0,\dots,S$, let

$$
\mathcal{X}_d^{(S)} = \{(\mathbf{i},\mathbf{j}) \in \mathcal{X}^{(S)} : |\{k : i_k \neq j_k\}| = d\}, \qquad (9.7)
$$

be the set of all pairs of alternatives which differ in exactly d attributes. These sets also constitute the orbits with respect to permutations. Here we similarly note that the regression functions $\mathbf{f}$ in (9.4) extended to the design region $\mathcal{X}^{(S)}$ are still linearly equivariant i. e. also here relabeling does not affect D-optimality as well as D-optimality of invariant subvectors $\boldsymbol{\beta}$ presented in (9.2). As a result, it is sufficient to look for optimality in the class of invariant designs which are uniform on the orbits of fixed comparison depth d.

In what follows, let $N_d = \binom{K}{S}\binom{S}{d}v^S(v-1)^d$ be the number of different pairs in $\mathcal{X}_d^{(S)}$ which vary in exactly d attributes and denote by ξ_d the uniform approximate design which assigns equal weight $\xi_d(\mathbf{i},\mathbf{j}) = 1/N_d$ to each pair in $\mathcal{X}_d^{(S)}$ and weight zero to all remaining pairs in $\mathcal{X}^{(S)}$. In the following we derive the information matrices for the aforementioned invariant designs.

Lemma 9.1. *Let d be a fixed comparison depth. The uniform design ξ_d on the set $\mathcal{X}_d^{(S)}$ of comparison depth d has block diagonal information matrix*

$$\mathbf{M}(\xi_d) = diag(h_q(d)\mathbf{Id}_{p_q} \otimes \mathbf{M}^{\otimes q})_{q=1,\ldots,4},$$

where $\mathbf{M}^{\otimes q}$ denotes the q-fold Kronecker product of $\mathbf{M}$ and

$$h_1(d) = \frac{d}{K}, \; h_2(d) = \frac{d((d-1)(v-2) + 2(S-d)(v-1))}{2vK(K-1)},$$
$$h_3(d) = \frac{d\lambda_1(d)}{4v^2K(K-1)(K-2)}, \; h_4(d) = \frac{d\lambda_2(d)}{8v^3K(K-1)(K-2)(K-3)},$$
$$\lambda_1(d) = (d-1)(d-2)(v^2-3v+3) + 3(S-d)(d-1)(v-1)(v-2) + 3(S-d)(S-d-1)(v-1)^2,$$
$$\begin{aligned}\lambda_2(d) = {} & (d-1)(d-2)(d-3)(v^3-4v^2+6v-4) + 4(S-d)(d-1)(d-2)(v^2-3v+3)(v-1) \\ & + 6(S-d)(S-d-1)(d-1)(v-1)^2(v-2) + 4(S-d)(S-d-1)(S-d-2)(v-1)^3.\end{aligned}$$

Proof. Note that the corresponding functions $h_q(d)$, $q = 1, 2, 3$ can be found in Lemma 7.3 for the second-order interaction models. Analogously, for the third-order interactions we consider attributes k, ℓ, m and r, say, and distinguish between pairs in which all four attributes are distinct, pairs in which three of these attributes k, ℓ and m, say, have distinct levels in the alternatives while the same level is presented in both alternatives for the remaining attribute, pairs in which two of these attributes k, ℓ, say, have distinct levels in the alternatives while the same level is presented in both alternatives for the remaining attribute two attributes and, finally, pairs in which only one of the attributes, say, k has distinct levels in the alternatives while the same level is presented in both alternatives for the three remaining attributes:

$$
\begin{aligned}
&\sum_{i_k\neq j_k}\sum_{i_\ell\neq j_\ell}\sum_{i_m\neq j_m}\sum_{i_r\neq j_r}(\mathbf{f}_1(i_k)\otimes\mathbf{f}_1(i_\ell)\otimes\mathbf{f}_1(i_m)\otimes\mathbf{f}_1(i_r)-\mathbf{f}_1(j_k)\otimes\mathbf{f}_1(j_\ell)\otimes\mathbf{f}_1(j_m)\otimes\mathbf{f}_1(j_r))\\
&\qquad\cdot(\mathbf{f}_1(i_k)\otimes\mathbf{f}_1(i_\ell)\otimes\mathbf{f}_1(i_m)\otimes\mathbf{f}_1(i_r)-\mathbf{f}_1(j_k)\otimes\mathbf{f}_1(j_\ell)\otimes\mathbf{f}_1(j_m)\otimes\mathbf{f}_1(j_r))^\top\\
&=\sum_{i_k=1}^{v}\sum_{j_k\neq i_k}\sum_{i_\ell=1}^{v}\sum_{j_\ell\neq i_\ell}\sum_{i_m=1}^{v}\sum_{j_m\neq i_m}\sum_{i_r=1}^{v}\sum_{j_r\neq i_r}(\mathbf{f}_1(i_k)\mathbf{f}_1(i_k)^\top\otimes\mathbf{f}_1(i_\ell)\mathbf{f}_1(i_\ell)^\top\otimes\mathbf{f}_1(i_m)\mathbf{f}_1(i_m)^\top\otimes\mathbf{f}_1(i_r)\mathbf{f}_1(i_r)^\top\\
&\qquad+\mathbf{f}_1(j_k)\mathbf{f}_1(j_k)^\top\otimes\mathbf{f}_1(j_\ell)\mathbf{f}_1(j_\ell)^\top\otimes\mathbf{f}_1(j_m)\mathbf{f}_1(j_m)^\top\otimes\mathbf{f}_1(j_r)\mathbf{f}_1(j_r)^\top\\
&\qquad-\mathbf{f}_1(i_k)\mathbf{f}_1(j_k)^\top\otimes\mathbf{f}_1(i_\ell)\mathbf{f}_1(j_\ell)^\top\otimes\mathbf{f}_1(i_m)\mathbf{f}_1(j_m)^\top\otimes\mathbf{f}_1(i_r)\mathbf{f}_1(j_r)^\top\\
&\qquad-\mathbf{f}_1(j_k)\mathbf{f}_1(i_k)^\top\otimes\mathbf{f}_1(j_\ell)\mathbf{f}_1(i_\ell)^\top\otimes\mathbf{f}_1(j_m)\mathbf{f}_1(i_m)^\top\otimes\mathbf{f}_1(j_r)\mathbf{f}_1(i_r)^\top)\\
&=2(v-1)^4\sum_{i_k=1}^{v}\mathbf{f}_1(i_k)\mathbf{f}_1(i_k)^\top\otimes\sum_{i_\ell=1}^{v}\mathbf{f}_1(i_\ell)\mathbf{f}_1(i_\ell)^\top\otimes\sum_{i_m=1}^{v}\mathbf{f}_1(i_m)\mathbf{f}_1(i_m)^\top\otimes\sum_{i_m=1}^{v}\mathbf{f}_1(i_r)\mathbf{f}_1(i_r)^\top\\
&\qquad-2\sum_{i_k\neq j_k}\mathbf{f}_1(i_k)\mathbf{f}_1(j_k)^\top\otimes\sum_{i_\ell\neq j_\ell}\mathbf{f}_1(i_\ell)\mathbf{f}_1(j_\ell)^\top\otimes\sum_{i_m\neq j_m}\mathbf{f}_1(i_m)\mathbf{f}_1(j_m)^\top\otimes\sum_{i_r\neq j_r}\mathbf{f}_1(i_r)\mathbf{f}_1(j_r)^\top\\
&=\frac{1}{8}v(v-1)^4(v^3-4v^2+6v-4)\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M},
\end{aligned}
\tag{9.8}
$$

also

$$
\begin{aligned}
&\sum_{i_k\neq j_k}\sum_{i_\ell\neq j_\ell}\sum_{i_m\neq j_m}\sum_{i_r=j_r}(\mathbf{f}_1(i_k)\otimes\mathbf{f}_1(i_\ell)\otimes\mathbf{f}_1(i_m)\otimes\mathbf{f}_1(i_r)-\mathbf{f}_1(j_k)\otimes\mathbf{f}_1(j_\ell)\otimes\mathbf{f}_1(j_m)\otimes\mathbf{f}_1(j_r))\\
&\qquad\cdot(\mathbf{f}_1(i_k)\otimes\mathbf{f}_1(i_\ell)\otimes\mathbf{f}_1(i_m)\otimes\mathbf{f}_1(i_r)-\mathbf{f}_1(j_k)\otimes\mathbf{f}_1(j_\ell)\otimes\mathbf{f}_1(j_m)\otimes\mathbf{f}_1(j_r))^\top\\
&=\sum_{i_k=1}^{v}\sum_{j_k\neq i_k}\sum_{i_\ell=1}^{v}\sum_{j_\ell\neq i_\ell}\sum_{i_m=1}^{v}\sum_{j_m\neq i_m}\sum_{i_r=1}^{v}\sum_{j_r=i_r}(\mathbf{f}_1(i_k)\mathbf{f}_1(i_k)^\top\otimes\mathbf{f}_1(i_\ell)\mathbf{f}_1(i_\ell)^\top\otimes\mathbf{f}_1(i_m)\mathbf{f}_1(i_m)^\top\otimes\mathbf{f}_1(i_r)\mathbf{f}_1(i_r)^\top\\
&\qquad+\mathbf{f}_1(j_k)\mathbf{f}_1(j_k)^\top\otimes\mathbf{f}_1(j_\ell)\mathbf{f}_1(j_\ell)^\top\otimes\mathbf{f}_1(j_m)\mathbf{f}_1(j_m)^\top\otimes\mathbf{f}_1(j_r)\mathbf{f}_1(j_r)^\top\\
&\qquad-\mathbf{f}_1(i_k)\mathbf{f}_1(j_k)^\top\otimes\mathbf{f}_1(i_\ell)\mathbf{f}_1(j_\ell)^\top\otimes\mathbf{f}_1(i_m)\mathbf{f}_1(j_m)^\top\otimes\mathbf{f}_1(i_r)\mathbf{f}_1(j_r)^\top\\
&\qquad-\mathbf{f}_1(j_k)\mathbf{f}_1(i_k)^\top\otimes\mathbf{f}_1(j_\ell)\mathbf{f}_1(i_\ell)^\top\otimes\mathbf{f}_1(j_m)\mathbf{f}_1(i_m)^\top\otimes\mathbf{f}_1(j_r)\mathbf{f}_1(i_r)^\top)\\
&=2(v-1)^3\sum_{i_k=1}^{v}\mathbf{f}_1(i_k)\mathbf{f}_1(i_k)^\top\otimes\sum_{i_\ell=1}^{v}\mathbf{f}_1(i_\ell)\mathbf{f}_1(i_\ell)^\top\otimes\sum_{i_m=1}^{v}\mathbf{f}_1(i_m)\mathbf{f}_1(i_m)^\top\otimes\sum_{i_m=1}^{v}\mathbf{f}_1(i_r)\mathbf{f}_1(i_r)^\top\\
&\qquad-2\sum_{i_k\neq j_k}\mathbf{f}_1(i_k)\mathbf{f}_1(j_k)^\top\otimes\sum_{i_\ell\neq j_\ell}\mathbf{f}_1(i_\ell)\mathbf{f}_1(j_\ell)^\top\otimes\sum_{i_m\neq j_m}\mathbf{f}_1(i_m)\mathbf{f}_1(j_m)^\top\otimes\sum_{i_r=j_r}\mathbf{f}_1(i_r)\mathbf{f}_1(j_r)^\top\\
&=\frac{1}{8}v(v-1)^4(v^2-3v+3)\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M},
\end{aligned}
\tag{9.9}
$$

further

$$
\begin{aligned}
&\sum_{i_k \neq j_k} \sum_{i_\ell \neq j_\ell} \sum_{i_m = j_m} \sum_{i_r = j_r} (\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) \otimes \mathbf{f}_1(i_m) \otimes \mathbf{f}_1(i_r) - \mathbf{f}_1(j_k) \otimes \mathbf{f}_1(j_\ell) \otimes \mathbf{f}_1(j_m) \otimes \mathbf{f}_1(j_r)) \\
&\qquad\qquad \cdot (\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) \otimes \mathbf{f}_1(i_m) \otimes \mathbf{f}_1(i_r) - \mathbf{f}_1(j_k) \otimes \mathbf{f}_1(j_\ell) \otimes \mathbf{f}_1(j_m) \otimes \mathbf{f}_1(j_r))^\top \\
&= \sum_{i_k=1}^{v} \sum_{j_k \neq i_k} \sum_{i_\ell=1}^{v} \sum_{j_\ell \neq i_\ell} \sum_{i_m=1}^{v} \sum_{j_m = i_m} \sum_{i_r=1}^{v} \sum_{j_r = i_r} (\mathbf{f}_1(i_k)\mathbf{f}_1(i_k)^\top \otimes \mathbf{f}_1(i_\ell)\mathbf{f}_1(i_\ell)^\top \otimes \mathbf{f}_1(i_m)\mathbf{f}_1(i_m)^\top \otimes \mathbf{f}_1(i_r)\mathbf{f}_1(i_r)^\top \\
&\qquad\qquad\qquad + \mathbf{f}_1(j_k)\mathbf{f}_1(j_k)^\top \otimes \mathbf{f}_1(j_\ell)\mathbf{f}_1(j_\ell)^\top \otimes \mathbf{f}_1(j_m)\mathbf{f}_1(j_m)^\top \otimes \mathbf{f}_1(j_r)\mathbf{f}_1(j_r)^\top \\
&\qquad\qquad\qquad - \mathbf{f}_1(i_k)\mathbf{f}_1(j_k)^\top \otimes \mathbf{f}_1(i_\ell)\mathbf{f}_1(j_\ell)^\top \otimes \mathbf{f}_1(i_m)\mathbf{f}_1(j_m)^\top \otimes \mathbf{f}_1(i_r)\mathbf{f}_1(j_r)^\top \\
&\qquad\qquad\qquad - \mathbf{f}_1(j_k)\mathbf{f}_1(i_k)^\top \otimes \mathbf{f}_1(j_\ell)\mathbf{f}_1(i_\ell)^\top \otimes \mathbf{f}_1(j_m)\mathbf{f}_1(i_m)^\top \otimes \mathbf{f}_1(j_r)\mathbf{f}_1(i_r)^\top) \\
&= 2(v-1)^2 \sum_{i_k=1}^{v} \mathbf{f}_1(i_k)\mathbf{f}_1(i_k)^\top \otimes \sum_{i_\ell=1}^{v} \mathbf{f}_1(i_\ell)\mathbf{f}_1(i_\ell)^\top \otimes \sum_{i_m=1}^{v} \mathbf{f}_1(i_m)\mathbf{f}_1(i_m)^\top \otimes \sum_{i_m=1}^{v} \mathbf{f}_1(i_r)\mathbf{f}_1(i_r)^\top \\
&\qquad - 2 \sum_{i_k \neq j_k} \mathbf{f}_1(i_k)\mathbf{f}_1(j_k)^\top \otimes \sum_{i_\ell \neq j_\ell} \mathbf{f}_1(i_\ell)\mathbf{f}_1(j_\ell)^\top \otimes \sum_{i_m = j_m} \mathbf{f}_1(i_m)\mathbf{f}_1(j_m)^\top \otimes \sum_{i_r = j_r} \mathbf{f}_1(i_r)\mathbf{f}_1(j_r)^\top \\
&= \frac{1}{8} v(v-1)^4(v-2)\mathbf{M} \otimes \mathbf{M} \otimes \mathbf{M} \otimes \mathbf{M}, \qquad (9.10)
\end{aligned}
$$

and

$$
\begin{aligned}
&\sum_{i_k \neq j_k} \sum_{i_\ell = j_\ell} \sum_{i_m = j_m} \sum_{i_r = j_r} (\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) \otimes \mathbf{f}_1(i_m) \otimes \mathbf{f}_1(i_r) - \mathbf{f}_1(j_k) \otimes \mathbf{f}_1(j_\ell) \otimes \mathbf{f}_1(j_m) \otimes \mathbf{f}_1(j_r)) \\
&\qquad\qquad \cdot (\mathbf{f}_1(i_k) \otimes \mathbf{f}_1(i_\ell) \otimes \mathbf{f}_1(i_m) \otimes \mathbf{f}_1(i_r) - \mathbf{f}_1(j_k) \otimes \mathbf{f}_1(j_\ell) \otimes \mathbf{f}_1(j_m) \otimes \mathbf{f}_1(j_r))^\top \\
&= \sum_{i_k=1}^{v} \sum_{j_k \neq i_k} \sum_{i_\ell=1}^{v} \sum_{j_\ell = i_\ell} \sum_{i_m=1}^{v} \sum_{j_m = i_m} \sum_{i_r=1}^{v} \sum_{j_r = i_r} (\mathbf{f}_1(i_k)\mathbf{f}_1(i_k)^\top \otimes \mathbf{f}_1(i_\ell)\mathbf{f}_1(i_\ell)^\top \otimes \mathbf{f}_1(i_m)\mathbf{f}_1(i_m)^\top \otimes \mathbf{f}_1(i_r)\mathbf{f}_1(i_r)^\top \\
&\qquad\qquad\qquad + \mathbf{f}_1(j_k)\mathbf{f}_1(j_k)^\top \otimes \mathbf{f}_1(j_\ell)\mathbf{f}_1(j_\ell)^\top \otimes \mathbf{f}_1(j_m)\mathbf{f}_1(j_m)^\top \otimes \mathbf{f}_1(j_r)\mathbf{f}_1(j_r)^\top \\
&\qquad\qquad\qquad - \mathbf{f}_1(i_k)\mathbf{f}_1(j_k)^\top \otimes \mathbf{f}_1(i_\ell)\mathbf{f}_1(j_\ell)^\top \otimes \mathbf{f}_1(i_m)\mathbf{f}_1(j_m)^\top \otimes \mathbf{f}_1(i_r)\mathbf{f}_1(j_r)^\top \\
&\qquad\qquad\qquad - \mathbf{f}_1(j_k)\mathbf{f}_1(i_k)^\top \otimes \mathbf{f}_1(j_\ell)\mathbf{f}_1(i_\ell)^\top \otimes \mathbf{f}_1(j_m)\mathbf{f}_1(i_m)^\top \otimes \mathbf{f}_1(j_r)\mathbf{f}_1(i_r)^\top) \\
&= 2(v-1) \sum_{i_k=1}^{v} \mathbf{f}_1(i_k)\mathbf{f}_1(i_k)^\top \otimes \sum_{i_\ell=1}^{v} \mathbf{f}_1(i_\ell)\mathbf{f}_1(i_\ell)^\top \otimes \sum_{i_m=1}^{v} \mathbf{g}(i_m)\mathbf{f}_1(i_m)^\top \otimes \sum_{i_m=1}^{v} \mathbf{f}_1(i_r)\mathbf{f}_1(i_r)^\top \\
&\qquad - 2 \sum_{i_k \neq j_k} \mathbf{f}_1(i_k)\mathbf{f}_1(j_k)^\top \otimes \sum_{i_\ell = j_\ell} \mathbf{f}_1(i_\ell)\mathbf{f}_1(j_\ell)^\top \otimes \sum_{i_m = j_m} \mathbf{f}_1(i_m)\mathbf{f}_1(j_m)^\top \otimes \sum_{i_r = j_r} \mathbf{f}_1(i_r)\mathbf{f}_1(j_r)^\top \\
&= \frac{1}{8} v(v-1)^4\mathbf{M} \otimes \mathbf{M} \otimes \mathbf{M} \otimes \mathbf{M}, \qquad (9.11)
\end{aligned}
$$

respectively.

For given attributes k, ℓ, m and r the pairs with distinct levels in the four attributes occur

$$
\binom{K-4}{S-4}\binom{S-4}{d-4} v^{S-4}(v-1)^{d-4}
$$

times in $\mathcal{X}_d^{(S)}$ with corresponding number of paired comparisons

$$N_d = \binom{K-4}{S-4}\binom{S-4}{d-4}\frac{K(K-1)(K-2)(K-3)}{S(S-1)(S-2)(S-3)}\frac{S(S-1)(S-2)(S-3)}{d(d-1)(d-2)(d-3)}v^S(v-1)^d,$$

while those which differ in the three attributes occur

$$\binom{4}{3}\binom{K-4}{S-4}\binom{S-4}{d-3}v^{S-4}(v-1)^{d-3}$$

times in $\mathcal{X}_d^{(S)}$ with corresponding number of paired comparisons

$$N_d = \binom{K-4}{S-4}\binom{S-4}{d-3}\frac{K(K-1)(K-2)(K-3)}{S(S-1)(S-2)(S-3)}\frac{S(S-1)(S-2)(S-3)}{(S-d)d(d-1)(d-2)}v^S(v-1)^d,$$

while those which differ in the two attributes occur

$$\binom{4}{2}\binom{K-4}{S-4}\binom{S-4}{d-2}v^{S-4}(v-1)^{d-2}$$

times in $\mathcal{X}_d^{(S)}$ with corresponding number of paired comparisons

$$N_d = \binom{K-4}{S-4}\binom{S-4}{d-2}\frac{K(K-1)(K-2)(K-3)}{S(S-1)(S-2)(S-3)}\frac{S(S-1)(S-2)(S-3)}{(S-d)(S-d-1)d(d-1)}v^S(v-1)^d,$$

and, finally, those which differ only in the one attribute occur

$$\binom{4}{1}\binom{K-4}{S-4}\binom{S-4}{d-1}v^{S-4}(v-1)^{d-1}$$

times in $\mathcal{X}_d^{(S)}$ with corresponding number of paired comparisons

$$N_d = \binom{K-4}{S-4}\binom{S-4}{d-1}\frac{K(K-1)(K-2)(K-3)}{S(S-1)(S-2)(S-3)}\frac{S(S-1)(S-2)(S-3)}{(S-d)(S-d-1)(S-d-2)d}v^S(v-1)^d.$$

Hence, for the third-order interactions the diagonal elements $h_4(d)$ in the information matrix are given by

$$\begin{aligned}
h_4(d) &= \frac{1}{N_d}\binom{K-4}{S-4}\Bigg(\frac{1}{8}\tbinom{S-4}{d-4}v(v-1)^4(v^3-4v^2+6v-4)v^{S-4}(v-1)^{d-4}\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M} \\
&\qquad + \frac{1}{2}\tbinom{S-4}{d-3}v(v-1)^4(v^2-3v+3)v^{S-4}(v-1)^{d-3}\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M} \\
&\qquad + \frac{3}{4}\tbinom{S-4}{d-2}v(v-1)^4(v-2)v^{S-4}(v-1)^{d-2}\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M} \\
&\qquad + \frac{1}{2}\tbinom{S-4}{d-1}v(v-1)^4v^{S-4}(v-1)^{d-1}\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M}\Bigg) \\
&= \frac{d}{8v^3K(K-1)(K-2)(K-3)}((d-1)(d-2)(d-3)(v^3-4v^2+6v-4) \\
&\qquad + 4(S-d)(d-1)(d-2)(v^2-3v+3)(v-1) \\
&\qquad + 6(S-d)(S-d-1)(d-1)(v-1)^2(v-2) \\
&\qquad + 4(S-d)(S-d-1)(S-d-2)(v-1)^3)\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M}\otimes\mathbf{M}.
\end{aligned} \tag{9.12}$$

Finally, it can be noted that all off-diagonal entries in the information matrix vanish because the terms in the corresponding sums add up to zero due to the effects-type coding. □

It is worthwhile mentioning that the corresponding function $h_4(0) = 0$ for $d = 0$.

General invariant designs ξ which can be written as convex combination $\xi = \sum_{d=1}^{S} w_d\xi_d$ of the corresponding uniform designs on the comparison depth d with weights $w_d \geq 0$, $\sum_{d=1}^{S} w_d = 1$ have information matrix of the form:

Lemma 9.2. *Let $\xi = \sum_{d=1}^{S} w_d\xi_d$ be an invariant design on $\mathcal{X}^{(S)}$, then ξ has block diagonal information matrix*

$$\mathbf{M}(\xi) = diag(h_q(\xi)\mathbf{Id}_{p_q} \otimes \mathbf{M}^{\otimes q})_{q=1,\dots,4},$$

where $\mathbf{M}^{\otimes q}$ denotes the q-fold Kronecker product of $\mathbf{M}$ and $h_q(\xi) = \sum_{d=1}^{S} w_d h_q(d)$, $q = 1, 2, 3, 4$.

In the following Remark 9.1 we consider optimal designs for the third-order interaction term having entry $h_4(d)$ in the corresponding information matrix. As before the resulting designs may optimize every design criterion which is invariant with respect to both permutations of the levels and permutations of the attributes if one considers the full parameter vector, satisfying the aforementioned identifiabity conditions. Hence, the reduced parameter vector $\boldsymbol{\beta} = ((\boldsymbol{\beta}_k)^\top_{k=1,\dots,K}, (\boldsymbol{\beta}_{k\ell})^\top_{k<\ell}, (\boldsymbol{\beta}_{k\ell m})^\top_{k<\ell<m}(\boldsymbol{\beta}_{k\ell mr})^\top_{k<\ell<m<r})^\top$ presented in (9.2) is also invariant in particular, with the D-criterion. We note that

corresponding results for main effects, first and second-order interactions can be found in Section 7.2.

Remark 9.1. *There exists a single comparison depth d^* subject to the profile strength S such that the uniform design ξ_{d^*} is D-optimal for the third-order interaction effects $(\boldsymbol{\beta}_{k\ell mr})_{k<\ell<m<r}^\top$.*

Accordingly, the corresponding values of d^* presented in Table 9.1 were obtained by first calculating the values of $h_4(d)$ and determining the maximum. It is worthwhile mentioning that the optimal comparison depth $d^* = K-3$ or $d^* = K-4$ for sufficiently large values of v. We further note that for the situation of full profiles ($S = K$) the corresponding results presented in Table 8.1 can be recovered.

TABLE 9.1: Values of the Optimal Comparison Depths d^* of the D-Optimal Uniform Designs ξ_{d^*} for the Third-Order Interactions with $S \leq K$ Attributes and v-Levels

	v									
S	2	3	4	5	6	7	8	9	10	20
4	1	1	1	1	1	1	1	1	1	1
5	1	1	1	2	2	2	2	2	2	2
6	1	2	2	2	2	3	3	3	3	3
7	1	2	3	3	3	3	3	4	4	4
8	2	3	3	4	4	4	4	4	4	5
9	2	3	4	4	5	5	5	5	5	6
10	2	4	4	5	5	6	6	6	6	7
11	3	4	5	6	6	6	7	7	7	8
12	3	5	6	6	7	7	7	7	8	8

The corresponding invariant design ξ has a variance function of the form $V((\mathbf{i},\mathbf{j}),\xi) = (\mathbf{f}(\mathbf{i})-\mathbf{f}(\mathbf{j}))^\top \mathbf{M}(\xi)^{-1}(\mathbf{f}(\mathbf{i})-\mathbf{f}(\mathbf{j}))$. As already noted the variance function is invariant with respect to permutations of levels $i = 1,\ldots,v$ as well as to permutations of attributes $k = 1,\ldots,K$. Now the value of the variance function for the invariant design ξ evaluated at comparison depth d is denoted as $V(d,\xi)$ where $V(d,\xi) = V((\mathbf{i},\mathbf{j}),\xi)$ on the orbits $\mathcal{X}_d^{(S)}$ of fixed comparison depth d.

Theorem 9.1. *For every invariant design ξ on the orbits $\mathcal{X}_d^{(S)}$ the variance function $V(d,\xi)$ is given by*

$$\begin{aligned}
V(d,\xi) = d(v-1)\Bigg(\frac{1}{h_1(\xi)} &+ \frac{v-1}{4vh_2(\xi)}\Big((d-1)(v-2)+2(S-d)(v-1)\Big)\\
&+ \frac{(v-1)^2}{24v^2h_3(\xi)}\big((d-1)(d-2)(v^2-3v+3)\\
&\qquad + 3(S-d)(d-1)(v-1)(v-2)\\
&\qquad + 3(S-d)(S-d-1)(v-1)^2\big)\\
&+ \frac{(v-1)^4}{192v^3h_4(\xi)}\Big((d-1)(d-2)(d-3)(v^3-4v^2+6v-4)\\
&\qquad + 4(S-d)(d-1)(d-2)(v^2-3v+3)(v-1)\\
&\qquad + 6(S-d)(S-d-1)(d-1)(v-1)^2(v-2)\\
&\qquad + 4(S-d)(S-d-1)(S-d-2)(v-1)^3\Big)\Bigg).
\end{aligned}$$

Proof. In view of Theorem 7.3, it follows that for the regression function associated with the interaction of the attributes k, ℓ, m and r, say, we obtain

$$\begin{aligned}
&(\mathbf{f}_1(i_k)\otimes\mathbf{f}_1(i_\ell)\otimes\mathbf{f}_1(i_m)\otimes\mathbf{f}_1(i_r)-\mathbf{f}_1(j_k)\otimes\mathbf{f}_1(j_\ell)\otimes\mathbf{f}_1(j_m)\otimes\mathbf{f}_1(j_r))^\top\mathbf{M}^{-1}\otimes\mathbf{M}^{-1}\otimes\mathbf{M}^{-1}\otimes\mathbf{M}^{-1}\\
&\qquad\cdot\mathbf{f}_1(i_k)\otimes\mathbf{f}_1(i_\ell)\otimes\mathbf{f}_1(i_m)\otimes\mathbf{f}_1(i_r)-\mathbf{f}_1(j_k)\otimes\mathbf{f}_1(j_\ell)\otimes\mathbf{f}_1(j_m)\otimes\mathbf{f}_1(j_r))\\
&=\mathbf{f}_1(i_k)^\top\mathbf{M}^{-1}\mathbf{f}_1(i_k)\cdot\mathbf{f}_1(i_\ell)^\top\mathbf{M}^{-1}\mathbf{f}_1(i_\ell)\cdot\mathbf{f}_1(i_m)^\top\mathbf{M}^{-1}\mathbf{f}_1(i_m)\cdot\mathbf{f}_1(i_r)^\top\mathbf{M}^{-1}\mathbf{f}_1(i_r)\\
&\qquad+\mathbf{f}_1(j_k)^\top\mathbf{M}^{-1}\mathbf{f}_1(j_k)\cdot\mathbf{f}_1(j_\ell)^\top\mathbf{M}^{-1}\mathbf{f}_1(j_\ell)\cdot\mathbf{f}_1(j_m)^\top\mathbf{M}^{-1}\mathbf{f}_1(j_m)\cdot\mathbf{f}_1(j_r)^\top\mathbf{M}^{-1}\mathbf{f}_1(j_r)\\
&\qquad-\mathbf{f}_1(i_k)^\top\mathbf{M}^{-1}\mathbf{f}_1(j_k)\cdot\mathbf{f}_1(i_\ell)^\top\mathbf{M}^{-1}\mathbf{f}_1(j_\ell)\cdot\mathbf{f}_1(i_m)^\top\mathbf{M}^{-1}\mathbf{f}_1(j_m)\cdot\mathbf{f}_1(i_r)^\top\mathbf{M}^{-1}\mathbf{f}_1(j_r)\\
&\qquad-\mathbf{f}_1(j_k)^\top\mathbf{M}^{-1}\mathbf{f}_1(i_k)\cdot\mathbf{f}_1(j_\ell)^\top\mathbf{M}^{-1}\mathbf{f}_1(i_\ell)\cdot\mathbf{f}_1(j_m)^\top\mathbf{M}^{-1}\mathbf{f}_1(i_m)\cdot\mathbf{f}_1(j_r)^\top\mathbf{M}^{-1}\mathbf{f}_1(i_r)\\
&=\begin{cases}
\frac{1}{8v^3}(v-1)^4(v^3-4v^2+6v-4) & \text{for}\quad i_k\neq j_k, i_\ell\neq j_\ell, i_m\neq j_m, i_r\neq j_r\\
\frac{1}{8v^3}(v-1)^5(v^2-3v+3) & \text{for}\quad i_k\neq j_k, i_\ell\neq j_\ell, i_m\neq j_m, i_r= j_r\\
\frac{1}{8v^3}(v-1)^6(v-2) & \text{for}\quad i_k\neq j_k, i_\ell\neq j_\ell, i_m= j_m, i_r= j_r\\
\frac{1}{8v^3}(v-1)^7 & \text{for}\quad i_k\neq j_k, i_\ell= j_\ell, i_m= j_m, i_r= j_r.
\end{cases}
\end{aligned}\tag{9.13}$$

Now for a pair of alternatives $(\mathbf{i},\mathbf{j})\in\mathcal{X}_d^{(S)}$ of comparison depth d: there are $d(d-1)(d-2)(d-3)$ third-order interaction terms for which $(i_ki_\ell i_mi_r)$ and $(j_kj_\ell j_mj_r)$ differ in all four attributes k, ℓ, m and r, there are $(1/6)(S-d)d(d-1)(d-2)$ third-order interaction terms for which $(i_ki_\ell i_mi_r)$ and $(j_kj_\ell j_mj_r)$ differ in exactly three of the associated four attributes, there are $(1/4)(S-d)(S-d-1)d(d-1)$ third-order interaction terms for which $(i_k i_\ell i_m i_r)$ and $(j_k j_\ell j_m j_r)$ differ in exactly two of the associated four attributes and finally there are $(1/6)(S-d)(S-d-1)(S-d-2)d$ third-order interaction terms for which $(i_k i_\ell i_m i_r)$ and $(j_k j_\ell j_m j_r)$ differ in exactly one of the associated four attributes. Hence from (7.16), (7.17) and (7.18) and by taking the inverse of the corresponding information matrix $\mathbf{M}(\xi)$ in Lemma 9.2 we obtain

$$\begin{aligned}
V(d,\xi) &= (\mathbf{f}(\mathbf{i}) - \mathbf{f}(\mathbf{j}))^\top \mathbf{M}(\xi)^{-1}(\mathbf{f}(\mathbf{i}) - \mathbf{f}(\mathbf{j})) \\
&= \frac{d(v-1)}{h_1(\xi)} + \frac{d(v-1)^2}{4vh_2(\xi)}\Big((d-1)(v-2) + 2(S-d)(v-1)\Big) \\
&\quad + \frac{d(v-1)^3}{24v^2h_3(\xi)}\big((d-1)(d-2)(v^2-3v+3) + 3(S-d)(d-1)(v-1)(v-2) \\
&\qquad\qquad + 3(S-d)(S-d-1)(v-1)^2\big) \\
&\quad + d(d-1)(d-2)(d-3)\frac{(v-1)^4(v^3-4v^2+6v-4)}{8v^3h_4(\xi)} \\
&\quad + \frac{4(S-d)d(d-1)(d-2)}{24}\,\frac{(v-1)^5(v^2-3v+3)}{8v^3h_4(\xi)} \\
&\quad + \frac{6(S-d)(S-d-1)d(d-1)}{24}\,\frac{(v-1)^6(v-2)}{8v^3h_4(\xi)} \\
&\quad + \frac{4(S-d)(S-d-1)(S-d-2)d}{24}\,\frac{(v-1)^7}{8v^3h_4(\xi)} \\
&= \frac{d(v-1)}{h_1(\xi)} + \frac{d(v-1)^2}{4vh_2(\xi)}\Big((d-1)(v-2) + 2(S-d)(v-1)\Big) \\
&\quad + \frac{d(v-1)^3}{24v^2h_3(\xi)}\big((d-1)(d-2)(v^2-3v+3) + 3(S-d)(d-1)(v-1)(v-2) \\
&\qquad\qquad + 3(S-d)(S-d-1)(v-1)^2\big) \\
&\quad + \frac{d(v-1)^4}{192v^3h_4(\xi)}\Big((d-1)(d-2)(d-3)(v^3-4v^2+6v-4) \\
&\qquad\qquad + 4(S-d)(d-1)(d-2)(v^2-3v+3)(v-1) \\
&\qquad\qquad + 6(S-d)(S-d-1)(d-1)(v-1)^2(v-2) \\
&\qquad\qquad + 4(S-d)(S-d-1)(S-d-2)(v-1)^3\Big),
\end{aligned}$$

for $(\mathbf{i},\mathbf{j}) \in \mathcal{X}_d^{(S)}$ which proofs the proposed formula. □

Note that the corresponding variance function $V(0,\xi) = 0$ for comparison depth $d = 0$.

The representation of the corresponding variance function $V(d,\xi)$ simplifies if the general invariant design ξ is concentrated on a single comparison depth d:

Corollary 9.1. *For a uniform design $\xi_{d'}$ on a single comparison depth d' the variance function is given by*

$$V(d,\xi_{d'})=\frac{d}{d'}\left(p_1+p_2\tfrac{(d-1)(v-2)+2(S-d)(v-1)}{(d'-1)(v-2)+2(S-d')(v-1)}+p_3\tfrac{\lambda_1(d)}{\lambda_1(d')}+p_4\tfrac{\lambda_2(d)}{\lambda_2(d')}\right),$$

where

$$\lambda_1(d)=(d-1)(d-2)(v^2-3v+3)+3(S-d)(d-1)(v-1)(v-2)+3(S-d)(S-d-1)(v-1)^2,$$

$$\begin{aligned}\lambda_2(d)=&(d-1)(d-2)(d-3)(v^3-4v^2+6v-4)+4(S-d)(d-1)(d-2)(v^2-3v+3)(v-1)\\&+6(S-d)(S-d-1)(d-1)(v-1)^2(v-2)+4(S-d)(S-d-1)(S-d-2)(v-1)^3.\end{aligned}$$

Proof. This representation of the variance function follows immediately by inserting the values of $h_q(\xi_d)$ from Lemma 9.1 and $p_q=\binom{K}{q}(v-1)^q$, $q=1,2,3,4$ into the formula of Theorem 9.1. $\square$

Note that for $d=d'$, we obtain $V(d,\xi_d)=p_1+p_2+p_3+p_4=p$ for $p_1=K(v-1)$, $p_2=\binom{K}{2}(v-1)^2$, $p_3=\binom{K}{3}(v-1)^3$ and $p_4=\binom{K}{4}(v-1)^4$ which shows the D-optimality of ξ_d on $\mathcal{X}_d^{(S)}$.

The following Theorem 9.2 gives an upper bound on the number of comparison depths required for a D-optimal design. In other words for the identifiability of all model parameters for the main effects, the first-order interactions, the second-order interactions and the third-order interactions simultaneously.

Theorem 9.2. *In the third-order interactions model* (9.5) *the D-optimal design is supported on at most four orbits d^*, d_1^*, d^*+1 and d_1^*+1.*

Proof. The proof follows from Theorem 9.1 and by using analogous arguments in Theorem 8.2. $\square$

Now for fixed number of profile strengths S each at general levels v, intermediate comparison depths d, d_1 and weights w_d, w_{d_1} the numerical results presented in Table 9.2 were obtained by direct maximization of $\ln(\det(\mathbf{M}(w_d\xi_d+(1-w_d)\xi_{d_1})))$ for the corresponding optimal comparison depth d^*, d_1^* and optimal weights $w_{d^*}^*$ where $1-w_{d^*}^*=w_{d_1^*}^*$. In particular, for various choices of the profile strengths S between 5 and 12 and levels $v=2,\ldots,8$ the entries of the form $(d^*, d_1^*, w_{d^*}^*)$ in Table 9.2 indicate that invariant designs $\xi^*=w_{d^*}^*\xi_{d^*}+(1-w_{d^*}^*)\xi_{d_1^*}$ have to be considered, while for single entries d^* the D-optimal design $\xi^*=\xi_{d^*}$ has to be considered which is uniform on the optimal comparison depth d^* (in boldface). Accordingly, the corresponding values of the normalized variance function $V(d,\xi^*)/p$ which shows D-optimality of the design ξ^* in view of the Kiefer and Wolfowitz (1960) equivalence theorem is presented in Table 9.3. It is worthwhile mentioning that the results for the particular case $S=4$ and $v=2$ presented in Table

9.2 correspond to the results in Theorem 8.3 for complete interactions. We further note that in Table 8.2 the corresponding results for $v = 2$ (binary attributes) presented in Table 9.2 can also be found.

TABLE 9.2: Optimal Designs with Intermediate Comparison Depths d^* in Boldface and Optimal Weights $w^*_{d^*}$ of the form $(d^*, d^*_1, w^*_{d^*})$ for $S = K$ Attributes and v-Levels

	v						
K	2	3	4	5	6	7	8
5	$(\mathbf{2}, \mathbf{4},\ 0.667)$	**2**	**2**	**2**	**2**	**2**	**2**
6	$(\mathbf{2}, \mathbf{5},\ 0.714)$	$(\mathbf{2}, \mathbf{5},\ 0.878)$	**3**	**3**	**3**	**3**	**3**
7	$(\mathbf{2}, \mathbf{6},\ 0.750)$	**3**	**3**	**3**	**3**	**4**	**4**
8	$(\mathbf{3}, \mathbf{6},\ 0.667)$	**3**	**4**	**4**	**4**	**4**	**4**
9	$(\mathbf{3}, \mathbf{7},\ 0.700)$	**4**	**4**	**5**	**5**	**5**	**5**
10	$(\mathbf{3}, \mathbf{8},\ 0.727)$	**4**	**5**	**5**	**6**	**6**	**6**
11	$(\mathbf{4}, \mathbf{8},\ 0.667)$	**5**	**5**	**6**	**6**	**7**	**7**
12	$(\mathbf{4}, \mathbf{9},\ 0.692)$	**5**	**6**	**7**	**7**	**7**	**8**

Exhibited in Table 9.3 are the values of the normalized variance function $V(d, \xi^*)/p$ which shows D-optimality of the corresponding design ξ^* in view of the equivalence theorem by Kiefer and Wolfowitz (1960). It should be noted that for the case $S = K = 4$ and $v = 2, \ldots, 8$ all possible comparison depth d proves to be optimal and the corresponding Theorem 4 of Graßhoff et al. (2003) applies.

TABLE 9.3: Values of the Variance Function of the D-Optimal Design ξ^* for $S = K$ Attributes and v-Levels (Boldface **1** Corresponds to the Optimal Comparison Depths d^* and d_1^*)

		d									
K	v	1	2	3	4	5	6	7	8	9	10
5	2	0.938	**1**	0.938	**1**	0.938					
	3	0.881	**1**	0.961	1	0.987					
	4	0.858	**1**	0.965	0.985	0.981					
	5	0.845	**1**	0.970	0.982	0.980					
	6	0.837	**1**	0.974	0.982	0.981					
	7	0.832	**1**	0.977	0.983	0.982					
	8	0.828	**1**	0.980	0.984	0.983					
6	2	0.850	**1**	0.950	0.950	**1**	0.850				
	3	0.793	**1**	0.988	0.970	**1**	0.977				
	4	0.777	0.999	**1**	0.977	0.995	0.987				
	5	0.570	0.984	**1**	0.979	0.990	0.986				
	6	0.734	0.975	**1**	0.982	0.989	0.987				
	7	0.723	0.969	**1**	0.984	0.989	0.988				
	8	0.715	0.964	**1**	0.985	0.989	0.988				
7	2	0.792	**1**	0.982	0.952	0.982	**1**	0.792			
	3	0.723	0.973	**1**	0.972	0.971	0.997	0.965			
	4	0.679	0.945	**1**	0.984	0.976	0.990	0.980			
	5	0.657	0.930	**1**	0.993	0.983	0.992	0.987			
	6	0.643	0.921	**1**	0.999	0.989	0.995	0.993			
	7	0.634	0.914	0.998	**1**	0.991	0.995	0.994			
	8	0.625	0.906	0.994	**1**	0.991	0.995	0.994			
8	2	0.759	0.998	**1**	0.954	0.954	**1**	0.998	0.759		
	3	0.650	0.928	**1**	0.990	0.973	0.981	0.998	0.964		
	4	0.612	0.898	0.993	**1**	0.986	0.984	0.995	0.984		
	5	0.585	0.873	0.982	**1**	0.990	0.986	0.993	0.988		
	6	0.567	0.858	0.974	**1**	0.994	0.989	0.994	0.991		
	7	0.559	0.848	0.969	**1**	0.997	0.992	0.995	0.994		
	8	0.552	0.841	0.965	**1**	0.999	0.994	0.996	0.996		

(To be continued)

Table 9.3 (continued)

		d									
K	v	1	2	3	4	5	6	7	8	9	
9	2	0.693	0.958	**1**	0.966	0.945	0.966	**1**	0.958	0.693	
	3	0.596	0.885	0.989	**1**	0.983	0.976	0.987	0.997	0.960	
	4	0.550	0.841	0.967	**1**	0.995	0.986	0.987	0.995	0.984	
	5	0.528	0.819	0.954	0.998	**1**	0.992	0.991	0.996	0.992	
	6	0.512	0.801	0.941	0.993	**1**	0.994	0.992	0.996	0.993	
	7	0.501	0.789	0.932	0.989	**1**	0.996	0.993	0.996	0.995	
	8	0.493	0.780	0.927	0.986	**1**	0.997	0.994	0.996	0.996	
10	2	0.644	0.925	**1**	0.985	0.958	0.958	0.985	**1**	0.925	0.644
	3	0.544	0.8360	0.965	**1**	0.994	0.982	0.981	0.991	0.996	0.960
	4	0.501	0.791	0.936	0.991	**1**	0.993	0.987	0.990	0.996	0.985
	5	0.478	0.764	0.916	0.982	**1**	0.997	0.992	0.993	0.996	0.992
	6	0.464	0.748	0.904	0.976	0.999	**1**	0.996	0.995	0.998	0.995
	7	0.453	0.735	0.893	0.969	0.997	**1**	0.996	0.995	0.997	0.996
	8	0.446	0.726	0.885	0.965	0.995	**1**	0.997	0.996	0.996	0.997

Remark 9.2. *We first note that for the case when $v = 2$, $S = 4$ and $K = 5$ corresponding results can be found in Remark 8.2. Moreover, for the case when $v > 2$, $S = 4$ and $K \geq 5$ the optimal design $\xi^* = \xi_{d^*}$ has to be considered which is D-optimal on the single comparison depth d^*.*

10 Discussion and Future Research

The main purpose of this thesis was to develop D-optimal designs for paired comparison second-order and third-order interaction models.

By considering the part-worth model, and in particular for the situation when the alternatives to be evaluated are described by only a single attribute as in the one-way analysis of variance model, it has been shown that D-optimal designs can be obtained by assigning equal weights to the set of all distinct ordered pairs. This result was further used as a building block to obtain optimal designs for the situation when the alternatives to be evaluated by respondents are specified by many attributes in the presence of both the second-order interaction model and the third-order interaction model.

In particular, for the second-order interaction model, optimal designs in the case of binary attributes require that both types of pairs (alternatives in the choice sets) should be used in which either all attributes have distinct levels or approximately one half of the attributes are distinct and one half of the attributes coincide to obtain a D-optimal design for the whole parameter vector. For larger number of levels only one type of pairs may be required besides some exceptional cases. In the case that the number of levels get large the optimal pairs seem to be those which are distinct in all but two attributes shown. The resulting optimal designs for the particular situation of two level attributes as considered in Nyarko and Schwabe (2019) and larger common number of levels for each attribute depend both on the profile strength (for example in so-called partial profiles when a subset out of a large number of attributes are presented simultaneously) and on the total number of attributes (for example in so-called full profiles when all out of a large number of attributes are presented simultaneously) available.

On the other hand, for the third-order interaction model, optimal designs in the case of binary attributes require that two types of pairs of alternatives should be used in which certain numbers of attributes are distinct, and these numbers are close to and symmetric with respect to approximately half of the profile strength to obtain a D-optimal design for the whole parameter vector. Similarly to the second-order interaction models also here only one type of pairs may be required for larger number of levels besides one exceptional case. When the number of levels gets large the optimal pairs seem to be those which are distinct in all but three attributes shown.

For future work, we mention that the invariance considerations presented in this thesis have been formulated for continuous or approximate designs which are very useful in proving theorems concerning the optimality of designs, may serve as a benchmark to judge the efficiency of competing designs as well as a starting point to construct (exact) designs or fractions which share the property of optimality and can be realized with a reasonable number of comparisons. This approach has been adopted by Graßhoff et al. (2004) for main effects model in the two-level situation with both full and partial profiles and by van Berkum (1987b) for main effects and first-order interactions model in the two-level situation with full profiles. Corresponding results for methods to construct exact designs for the main effects and first-order interactions model in the two-level situation with both full and partial profiles can also be found (see e.g. Großmann and Schwabe, 2015; Großmann, 2017). Corresponding results of exact designs for the second- and the third-order interaction models remain an open problem.

In addition, optimal approximate as well as exact designs having paired comparisons partitioned into moderate blocks of sizes not varying much and not necessarily equal for the corresponding main effects, first-oder, second-order and third-order models incorporating fixed block effects for the case of both full and partial profiles remain an open problem, which is worth consideration. We mention that several different methods of constructing exact (block) designs for the corresponding main effects model, in particular for the case of direct observations in the two-level situation with full profiles are known (Jacroux, Wong, and Masaro, 1983; Cheng, 1978; Mukerjee, Dey, and Chatterjee, 2002; Jacroux and Kealy-Dichone, 2017; Jacroux and Jacroux, 2016; SahaRay and Dutta, 2018, amongst others). This problem will be further investigated for the case of paired comparisons, because when choosing the design as well as modeling data from paired comparisons, practitioners or reserachers often fail to take into account that the respondents are asked multiple questions and that the resulting answers may therefore be correlated.

Also, higher order interactions may be of interest for further investigation.

References

Atkinson, A., Donev, A., and Tobias, R. (2007). *Optimum Experimental Designs, with SAS*. Oxford University Press, Oxford.

Atkinson, A. C. (2011). Optimum experimental design. In: Lovric, M. (Ed.), *International Encyclopedia of Statistical Science*. Springer, Heidelberg, 1037–1039.

Atkinson, A. C. and Donev, A. N. (1992). *Optimum Experimental Designs*. Oxford University Press, Oxford.

Bradley, R. A. and El-Helbawy, A. T. (1976). Treatment contrasts in paired comparisons: Basic procedures with application to factorials. *Biometrika*, **63**, 255–262.

Bradley, R. A. and Terry, M. E. (1952). Rank analysis of incomplete block designs: I. The method of paired comparisons. *Biometrika*, **39**, 324–345.

Bunch, D. S. and Batsell, R. R. (1989). A Monte Carlo comparison of estimators for the multinomial logit model. *Journal of Marketing Research*, **26**, 56–68.

Burgess, L., Street, D. J., and Wasi, N. (2011). Comparing designs for choice experiments: A case study. *Journal of Statistical Theory and Practice*, **5**, 25–46.

Bush, S. (2014). Optimal designs for stated choice experiments generated from fractional factorial designs. *Journal of Statistical Theory and Practice*, **8**, 367–381.

Carson, R. T., Louviere, J. J., Anderson, D. A., Arabie, P., Bunch, D. S., Hensher, D. A., Johnson, R. M., Kuhfeld, W. F., Steinberg, D., Swait, J., et al. (1994). Experimental analysis of choice. *Marketing Letters*, **5**, 351–367.

Cheng, C.-S. (1978). Optimality of certain asymmetrical experimental designs. *The Annals of Statistics*, **6**, 1239–1261.

Chrzan, K. (2010). Using partial profile choice experiments to handle large numbers of attributes. *International Journal of Market Research*, **52**, 827–840.

Cox, D. R. and Reid, N. (2000). *The Theory of the Design of Experiments*. Chapman and Hall/CRC, Boca Raton.

Davidson, R. R. and Farquhar, P. H. (1976). A bibliography on the method of paired comparisons. *Biometrics*, **32**, 241–252.

El-Helbawy, A. T. and Ahmed, E. A. (1984). Optimal design results for 2^n factorial paired comparison experiments. *Communications in Statistics-Theory and Methods*, **13**, 2827–2845.

El-Helbawy, A. T., Ahmed, E. A., and Alharbey, A. H. (1994). Optimal designs for asymmetrical factorial paired comparison experiments. *Communications in Statistics-Simulation and Computation*, **23**, 663–681.

El-Helbawy, A. T. and Bradley, R. A. (1978). Treatment contrasts in paired comparisons: Large-sample results, applications, and some optimal designs. *Journal of the American Statistical Association*, **73**, 831–839.

Elrod, T., Louviere, J. J., and Davey, K. S. (1992). An empirical comparison of ratings-based and choice-based conjoint models. *Journal of Marketing Research*, **29**, 368–377.

Fechner, G. (1860). *Elemente der Psychophysik. Breitkopf und Härtel.* Leipzig.

Fedorov, V. V. and Hackl, P. (1997). *Model-Oriented Design of Experiments.* Springer, New York.

Graßhoff, U., Großmann, H., Holling, H., and Schwabe, R. (2003). Optimal paired comparison designs for first-order interactions. *Statistics*, **37**, 373–386.

Graßhoff, U., Großmann, H., Holling, H., and Schwabe, R. (2004). Optimal designs for main effects in linear paired comparison models. *Journal of Statistical Planning and Inference*, **126**, 361–376.

Green, P. E. (1974). On the design of choice experiments involving multifactor alternatives. *Journal of Consumer Research*, **1**, 61–68.

Green, P. E. and Rao, V. R. (1971). Conjoint measurement for quantifying judgmental data. *Journal of Marketing Research*, **8**, 355–363.

Green, P. E. and Srinivasan, V. (1990). Conjoint analysis in marketing: New developments with implications for research and practice. *Journal of Marketing*, **54**, 3–19.

Green, P. E. and Srinivasan, V. (1978). Conjoint analysis in consumer research: Issues and outlook. *Journal of Consumer Research*, **5**, 103–123.

Großmann, H. (2003). *Desings für Paarvergleiche in der Metrischen Conjoint-Analyse.* Deutscher Universitätsverlag, Wiesbaden.

Großmann, H. (2017). Partial-profile choice designs for estimating main effects and interactions of two-level attributes from paired comparison data. *Journal of Statistical Theory and Practice*, **11**, 236–253.

Großmann, H. (2018). A practical approach to designing partial-profile choice experiments with two alternatives for estimating main effects and interactions of many two-level attributes. *Journal of Choice Modelling.* https://doi.org/10.1016/j.jocm.2018.04.001.

Großmann, H., Holling, H., and Schwabe, R. (2002). Advances in optimum experimental design for conjoint analysis and discrete choice models. In: Franses, P. H. and Montgomery, A. L. (Eds.), *Advances in Econometrics, Vol. 16. Econometric Models in Marketing.* JAI Press, Amsterdam, 93–117.

Großmann, H. and Schwabe, R. (2015). Design for discrete choice experiments. In: Dean, A., Morris, M., Stufken, J. and Bingham, D. (Eds.), *Handbook of Design and Analysis of Experiments.* Chapman and Hall/CRC, Boca Raton, 787–831.

Huber, J. and Zwerina, K. (1996). The importance of utility balance in efficient choice designs. *Journal of Marketing Research*, **33**, 307–317.

Jacroux, M. and Jacroux, T. (2016). On the E-optimality of blocked main effects plans when $n \equiv 1 \pmod 4$. *Communications in Statistics-Theory and Methods*, **45**, 5584–5589.

Jacroux, M. and Kealy-Dichone, B. (2017). On the E-optimality of blocked main effects plans in blocks of different sizes. *Communications in Statistics-Theory and Methods*, **46**, 2132–2138.

Jacroux, M., Wong, C., and Masaro, J. (1983). On the optimality of chemical balance weighing designs. *Journal of Statistical Planning and Inference*, **8**, 231–240.

Kiefer, J. (1959). Optimum experimental designs. *Journal of the Royal Statistical Society: Series B*, **21**, 272–304.

Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems. *Canadian Journal of Mathematics*, **12**, 363–366.

Kuhfeld, W. F., Tobias, R. D., and Garratt, M. (1994). Efficient experimental design with marketing research applications. *Journal of Marketing Research*, **31**, 545–557.

Louviere, J., Street, D. J., and Burgess, L. (2004). A 20+ years' retrospective on choice experiments. In: Wind, Y. and Green, P. E. (Eds.), *Marketing Research and Modeling: Progress and Prospects*. Kluwer Academic Publishers, Boston, 201–214.

Louviere, J. J., Hensher, D. A., and Swait, J. D. (2000). *Stated Choice Methods: Analysis and Application*. Cambridge University Press, Cambridge.

Louviere, J. J. and Woodworth, G. (1983). Design and analysis of simulated consumer choice or allocation experiments: An approach based on aggregate data. *Journal of Marketing Research*, **20**, 350–367.

Luce, D (1959). *Individual Choice Behavior: A Theoritical Analysis*. Wiley, New York.

Luce, R. D. and Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology*, **1**, 1–27.

Luyten, J., Kessels, R., Goos, P., and Beutels, P. (2015). Public preferences for prioritizing preventive and curative health care interventions: A discrete choice experiment. *Value in Health*, **18**, 224–233.

Marschak, J. (1960). Binary choice constraints on random utility indicators. In: Arrow, K. (Ed.), *Stanford Symposium on Mathematical Methods in the Social Sciences*. Stanford University Press, Stanford, 312–329.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (Ed.), *Frontiers in Econometrics*. Academic Press, New York, 105–142.

Mukerjee, R., Dey, A., and Chatterjee, K. (2002). Optimal main effect plans with non-orthogonal blocking. *Biometrika*, **89**, 225–229.

Nyarko, E. and Schwabe, R. (2019). Optimal Designs for Second-Order Interactions in Paired Comparison Experiments with Binary Attributes. *Journal of Statistical Theory and Practice*, **13**, 51.

Offen, W. and Littell, R. (1987). Design of paired comparison experiments when treatments are levels of a single quantitative variable. *Journal of Statistical Planning and Inference*, **15**, 331–346.

Pázman, A. (1986). *Foundations of Optimum Experimental Design*. Reidel, Dordrecht.

Pukelsheim, F. (1993). *Optimal Design of Experiments*. Wiley, New York.

Quenouille, M. and John, J. (1971). Paired comparison designs for 2^n- factorials. *Journal of the Royal Statistical Society: Series C*, **20**, 16–24.

SahaRay, R. and Dutta, G. (2018). On the optimality of blocked main effects plans with even number of runs. *Journal of Statistical Theory and Practice*, **12**, 136–150.

Sandor, Z. and Wedel, M. (2001). Designing conjoint choice experiments using managers' prior beliefs. *Journal of Marketing Research*, **38**, 430–444.

Scheffé, H. (1952). An analysis of variance for paired comparisons. *Journal of the American Statistical Association*, **47**, 381–400.

Schwabe, R. (1996). *Optimum Designs for Multi-Factor Models*. Springer, New York.

Schwabe, R., Graßhoff, U., Großmann, H., and Holling, H. (2003). Optimal 2^K paired comparison designs for partial profiles. In: *PROBASTAT 2002, Proceedings of the Fourth International Conference on Mathematical Statistics, Smolenice 2002, Tatra Mountains Mathematical Publications,* **26**, 79–86.

Shiao, S.-Y.P. K., Ahn, C. W., and Akazawa, K. (2007). Second-order interactions with the treatment groups in controlled clinical trials. *Computer Methods and Programs in Biomedicine*, **86**, 10–16.

Silvey, S. D (1980). *Optimal Designs*. Chapman and Hall, London.

Smith, K. (1918). On the standard deviations of adjusted and interpolated values of an observed polynomial function and its constants and the guidance towards a proper choice of the distribution of observations. *Biometrika*, **12**, 1–85.

Street, D. J., Bunch, D. S., and Moore, B. J. (2001). Optimal designs for 2^k paired comparison experiments. *Communications in Statistics-Theory and Methods*, **30**, 2149–2171.

Street, D. J. and Burgess, L. (2004). Optimal and near-optimal pairs for the estimation of effects in 2-level choice experiments. *Journal of Statistical Planning and Inference*, **118**, 185–199.

Street, D. J. and Burgess, L. (2007). *The Construction of Optimal Stated Choice Experiments: Theory and Methods*. Wiley, Hoboken.

Street, D. J., Burgess, L., and Louviere, J. J. (2005). Quick and easy choice sets: Constructing optimal and nearly optimal stated choice experiments. *International Journal of Research in Marketing*, **22**, 459–470.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, **34**, 368–389.

Train, K. E. (2003). *Discrete Choice Methods with Simulation.* Cambridge University Press, Cambridge.

van Berkum, E. E. M. (1987a). Optimal paired comparison designs for factorial and quadratic models. *Journal of Statistical Planning and Inference*, **15**, 265–278.

van Berkum, E. E. M. (1987b). *Optimal Paired Comparison Designs for Factorial Experiments*. CWI Tract 31, Amsterdam.

van Berkum, E. E. M. (1989). Reduction of the number of pairs in paired comparison designs and exact designs for quadratic models. *Computational Statistics and Data Analysis*, **8**, 93–107.

Verelst, F., Willem, L., Kessels, R., and Beutels, P. (2018). Individual decisions to vaccinate one's child or oneself: A discrete choice experiment rejecting free-riding motives. *Social Science and Medicine*, **207**, 106–116.