Optimal Designs for Two-Color Microarray Experiments in Multi-Factorial Models

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

Diplom-Mathematikerin
Katharina Schiffl
aus Viersen

Berichter: Universitätsprofessor Ralf-Dieter Hilgers
Universitätsprofessor Eberhard Triesch

Tag der mündlichen Prüfung: 29. März 2011

Abstract

Two-color microarray experiments form an important tool in gene expression analysis. They are often used to identify candidate genes that can be made accountable for the genesis of a certain disease. Due to the high costs of microarray experiments it is fundamental to design these experiments carefully and specifically give instructions, which samples should be allocated on the same microarray. Thereby, two samples are hybridized together on one array and the assignment of samples to arrays influences the precision of the results. Therefore, design issues for microarray experiments have been investigated intensively in the last years. However, only few authors, e.g., Stanzel [37], focused on more than one factor of interest. We extend Stanzel's work and derive approximate optimal designs for estimating interactions in multi-factorial settings. Thereby, optimality of candidate designs is shown using equivalence theorems (Pukelsheim [33]). Another practical important but less studied topic is the derivation of exact optimal designs. Most research considers approximate designs or exact designs for special contrast sets and selected numbers of arrays. Therefore, we focus on exact designs and present a method to construct A-optimal microarray designs for arbitrary numbers of arrays and arbitrary contrast sets. This method is applied to derive optimal designs for estimating treatment-control comparisons, all-to-next contrasts, Helmert contrasts and all pairwise comparisons. Furthermore, we derive robust designs, which achieve efficient results even if observations are missing. Missing values are a crucial topic in the context of microarray experiments, since they often occur due to scratches on the slide or other damaging. In applications recommendations for the choice of efficient experimental layouts can be derived from our constructed designs.

Acknowledgment

First of all, I would like to thank Prof. Hilgers for giving me very interesting and challenging problems to work on and for introducing me to the field of optimal design. Especially, I would like to thank him for many motivating and worthwhile Monday morning meetings, where I was allowed to ask all questions filling my mind. Secondly, I want to thank Prof. Triesch for refereeing my thesis. I appreciate all helpful advices and important discussions with him which were always very motivating. Furthermore, thank you to Prof. Cheng who gave me the opportunity to work with him at the University of California, Berkeley for six months. I thank him for many motivating meetings and valuable discussions as well as many advices to get settled in the new environment. I thank Prof. Hilgers, who gave me the chance to research at the UCB as well as the opportunity to gain a lot of valuable experiences during different conferences. Thanks to Prof. Bailey, who gave me very useful information at the conference "Advances in Model-Oriented Design and Analysis 9". Furthermore, I would like to thank everyone in our department for talking to me about statistics as well as other non-mathematics related issues. The working atmosphere was always very pleasant, friendly and chocolaty. I never want to miss this time. Especially, I want to thank Sven Stanzel who introduced me to the work in our department and the field of optimal design.

I thank the "Bundesministerium für Bildung und Forschung" (BMBF) who financially supported this work within the joint research project SKAVOE.

Last and surely not least, I warmly thank my parents and especially Felix for being there all the time.

Contents

Chapter 1

Introduction

Many diseases, such as Alzheimer's disease or Huntington's disease [31], can be traced back to particular strongly expressed genes [12]. Therefore, it is desirable to identify these genes that can be made accountable for a certain disease in order to generate novel drugs. In recent years microarray technology has become one of the most prominent tools in gene expression analysis due to the fact that gene expressions of thousands of genes can be measured simultaneously. The microarrays consist of thousands of spots, where each spot contains e.g. the genetic information of one gene of the human genome. After processing the experiment, gene expression measurements for each gene are available and researchers can analyze which genes are higher expressed in diseased cells compared to healthy cells. Several microarray technologies are commonly used, the most prominent ones are oligonucleotide arrays and cDNA microarrays, also called two-color microarrays. Oligonucleotide arrays measure gene expressions of one sample per array, whereas cDNA microarrays hybridize two samples on one array by coloring one sample green and the other sample red. For cDNA microarrays, two important design questions arise in order to achieve precise parameter estimates in the underlying statistical model. Which samples should be allocated together on one microarray? Which samples should be labeled with the green or red dye? For instance, Figure 1.1 illustrates a simplification of a two-color microarray process with one array. Here, mRNA transcripts from a tumor cell and from a healthy cell are extracted and labeled with green (Cy3)

and red (Cy5) dyes, respectively, and are placed on the microarray. The thus labeled mRNA molecules of each gene bind to the complementary DNA strands of the corresponding gene spot on the array. Gene spots are illustrated as points on the microarray in Figure 1.1. Afterwards, a laser scanner measures the amount of hybridized mRNA for each color and each gene and gives dye fluorescence intensities, which correspond to the gene expression levels of the considered genes. Here, higher intensities indicate higher gene expressions, e.g., if the red labeled sample has twice as much of a transcript as the green labeled sample, then the red signal should be twice as much as the green signal [21].

If genes are found, for which the mRNA amount of the tumor cells is extremely high or low regulated in comparison to the healthy cells, they represent candidate genes that could be accountable for the considered disease. A detailed description of microarray experiments can be found in Klug et al. [27], Simon et al. [35], Parmigiani et al. [32], Draghici [10] or Wit and McClure [44]. In this thesis, samples from cells with a known disease or samples prepared with a specific treatment are only referred to as treatment, e.g., a sample from a healthy cell is called treatment zero, whereas a sample from a tumor cell is called treatment one.

However, microarrays are very expensive, thus it is fundamental to use appropriate designs to get most precise parameter estimates in the underlying statistical model. Optimal designs assign treatments and dyes in such a way to the microarrays that unbiased estimates with minimal variances of the effects of interest are ensured. Thereby, microarray experiments correspond to incomplete block designs with block size two, whereas each microarray illustrates a block. Design issues for microarray experiments have been investigated intensively in recent years, see for example Kerr and Churchill [23], Glonek and Solomon [15] or Yang and Speed [46]. However, most authors focus on the contrast set of all pairwise treatment comparisons. Further contrast sets are seldom addressed. In addition, only few authors, e.g., Stanzel [37], consider more than one experimental factor of interest, although in medical applications scientists are often interested in many factors and their interactions. For instance, Churchill [8] investigated several mouse cell lines medicated with different treatments. He was interested

in the cell line effect as well as in the treatment effect and in corresponding two-way interactions. Furthermore, Taylor et al. [39] and Stamatakis et al. [36] were interested in three-way interactions. Therefore, the interesting question of optimal designs for estimating interactions in multi-factorial settings arises and is considered in this work. To this end, we extend the investigations of Stanzel [37], who focused on two factors of interest.

The thesis is structured as follows: Chapter 2 introduces the statistical model, which is used to describe microarray experiments, and gives a short overview of design of experiments in Section 2.3. For instance, we define approximate and exact designs and introduce tools, e.g., information matrices, information functions and the equivalence theorems, to illustrate the principles of optimal design. Exact A-optimal designs for different contrast sets, including the comparisons with a control-treatment, all-to-next contrasts and Helmert contrasts, are derived in Chapter 3. Moreover, in Chapter 4 approximate optimal designs are investigated for several contrast sets in the one-factorial and multi-factorial setting. The dye effect is explicitly studied in Section 4.7. In Chapter 3 and in the first part of Chapter 4 we neglect the impact of the dyes temporarily. Chapter 5 considers robustness issues and provides efficient designs in scenarios with missing values. A conclusion and perspective is given in Chapter 6.

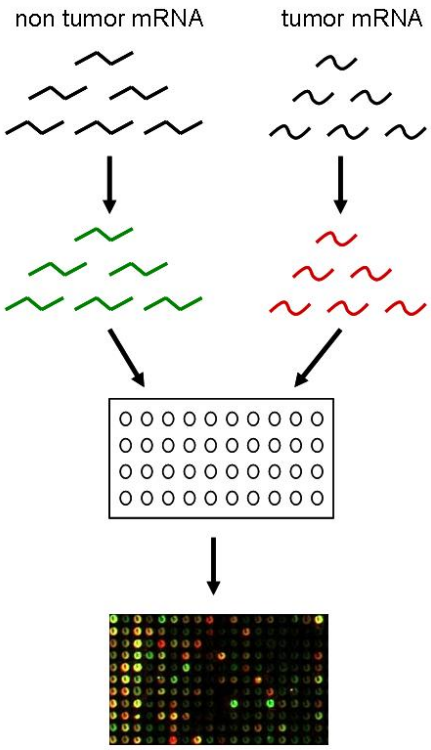


Figure 1.1: Simplified illustration of a microarray experiment.

Chapter 2

Fundamentals

In this chapter we introduce the graphical representation of microarray experiments, as well as the underlying statistical model, which is used throughout this work. Furthermore, we give a short overview of the main definitions and main concepts of optimal design.

2.1 Graphical representation of microarray experiments

Two color microarray experiments can be represented as directed graphs with multiple edges. A directed graph $D = (V, A)$ is a pair of a set V, whose elements are called vertices, and a set A of ordered pairs of vertices, called directed edges. Multiple edges connect the same vertices. Each treatment is illustrated as a vertex and each microarray is illustrated as a directed edge of the graph. The tail of each edge corresponds to the red labeled sample, the head to the green labeled sample. For instance, Figure 2.1 displays a microarray experiment with

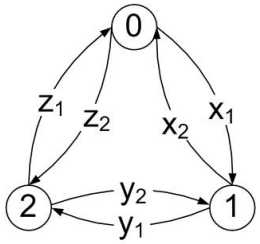


Figure 2.1: Graph representation of a microarray experiment

three different treatments and $x_1 + x_2 + y_1 + y_2 + z_1 + z_2$ arrays. Treatment zero labeled in red and treatment one labeled in green are hybridized together on x_1 arrays and so on. Many designs receive their name due to their graphical representation, for example the loop design is represented as a loop in the graph representation and the star design is represented as a star (Figure 2.2). Star designs allocate each treatment together with the control-treatment on the same array. Ignoring the dye effect, microarray experiments can be displayed as graphs $G = (V, E)$ with undirected edges, i.e. each edge $e \in E$ corresponds to a set of two vertices $e = \{v_1, v_2\}$, $v_1, v_2 \in V$. In this case, two treatment effects can be compared, i.e. their difference can be estimated, if and only if there exists a path between the corresponding two vertices. A path is an alternating sequence of distinct vertices and edges in the graph. The precision of the estimate of this treatment difference depends on the number of paths between the two vertices. For example, in Figure 2.2 all pairwise comparisons can be estimated in the right and left design, since both designs are connected. A design is called connected if every pair of vertices is joined by a path. The definitions become more complex if we include the dye effect. The dye effect can be estimated if there exists at least one loop in the graph representation. A loop is a path whose endvertices coincide. The estimate of the dye effect becomes more precise when the length of the loop increases. For instance, in the case with dye effect, all pairwise comparisons can be estimated in the left design in Figure 2.2, whereas treatment zero is confounded with the red dye in the right design.

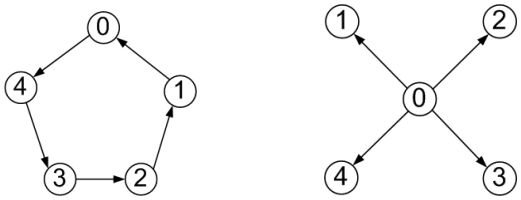


Figure 2.2: Graphical representation of the loop design and the star design.

2.2 Statistical modeling of microarray experiments

Many authors have focused on the statistical analysis and modeling of microarray experiments. Kerr et al. [24] firstly examined two-color microarray data by analysis of variance (ANOVA) and recommended a model describing the logarithms of the measured intensities dependent on the array-, treatment-, dye- and gene-effect, including treatment interactions of interest. Their work has been extended by many authors. For instance, Landgrebe et al. [29], Bailey [2] and Latif et al. [30] ignored the gene effect and considered gene specific models

$$\log_2(y) = T\tau + A\alpha + D\delta + \epsilon \tag{2.1}$$

where $y = (y_1, \ldots, y_{2a})$ is the vector of all observed dye intensities for a particular gene. These logarithmized dye intensities depend on the treatment effect $\tau = (\tau_0, \tau_1, \ldots, \tau_t)$, the array effect $\alpha = (\alpha_1, \ldots, \alpha_a)$ and the utilized dye $\delta = (\delta_{\text{green}}, \delta_{\text{red}})$. $[T \mid A \mid D]$ denotes the $2a \times (t + 1 + a + 2)$ design matrix and $\epsilon = (\epsilon_1, \ldots, \epsilon_{2a})$ denotes the vector of error terms.

Further model modifications are ascribed to Wolfinger et al. [45], who modeled the array effect as random, or Landgrebe et al. [29], who analyzed the logarithmized ratios of dye intensities obtained for each microarray separately for each gene. Instead of the two observations

$$\log_2(y_{ij\text{green}}) = \tau_i + \alpha_j + \delta_{\text{green}} + \epsilon_{ij\text{green}}, \tag{2.2}$$

$$\log_2(y_{kj\text{red}}) = \tau_k + \alpha_j + \delta_{\text{red}} + \epsilon_{kj\text{red}}, \tag{2.3}$$

Landgrebe et al. considered the log ratio

$$\log_2\left(\frac{y_{ij\text{green}}}{y_{kj\text{red}}}\right) = \tau_i - \tau_k + \delta_{\text{green}} - \delta_{\text{red}} + \epsilon_{ij\text{green}} - \epsilon_{kj\text{red}}. \tag{2.4}$$

Therefore, they introduced the model

$$z = X\tau + W\delta + \eta \tag{2.5}$$

where $z = (z_1, \ldots, z_a)$ is the vector of log ratios of the dye intensities measured for a particular gene on all a arrays. This vector is dependent on the treatment effect $\tau = (\tau_0, \tau_1, \ldots, \tau_t)$ and the dye effect $\delta = (\delta_{\text{green}}, \delta_{\text{red}})$. $[X \mid W]$ is the design matrix, where each row of X consists of exactly one 1 and one -1, whereas all other entries are equal to zero. W is equal to $(\mathbb{1}_a, -\mathbb{1}_a)$, where $\mathbb{1}_a$ is the a-dimensional column vector with all entries equal to one. The term η is the random error vector. We assume all η_i, $i \in \{1, \ldots, a\}$ to be independently identically distributed with mean zero and variance σ^2. Throughout this work we assume that $\sigma^2 = 1/2$ without loss of generality. This assumption simplifies the calculations of variances in Chapter 3 and in the following chapters, because a factor of two can be eliminated from all calculations.

2.3 Design of experiments

In this section, we introduce some basics on optimal design of experiments, which are required throughout this work. Firstly, we focus on differences between approximate and exact designs. Secondly, we present the main definitions of optimal design of experiments and finally we describe a tool to confirm optimality of a given candidate design.

2.3.1 Approximate and exact designs

An approximate design $\xi \in \Xi$ over a design region $\mathcal{X}$ can be described as

$$\xi = \begin{Bmatrix} x_1 & \ldots & x_l \\ p_1 & \ldots & p_l \end{Bmatrix}$$

with support points $x_i \in \mathcal{X}$ and weights $0 \leq p_i \leq 1$, $\sum_{i=1}^{l} p_i = 1$ representing the proportion of realizations of point x_i. The design region $\mathcal{X}$ is determined by the

values of the explanatory variables of the given statistical model. Approximate designs are not restricted to specific numbers of observations, they are defined for infinity observations. Synonyms for approximate designs are continuous or asymptotic designs, see e.g. Goos [16]. On the other hand, exact designs $\xi_n \in \Xi_n$ with n observations can be represented as

$$\xi_n = \begin{Bmatrix} x_1 & \dots & x_l \\ n_1 & \dots & n_l \end{Bmatrix}$$

where $\sum_{i=1}^{l} n_i = n$ and n_i is the number of observations at design point x_i. In practice exact designs are used; they are also called discrete designs. Efficient exact designs with n observations can be achieved with the help of optimal approximate theory. All weights p_i of a given approximate design are multiplied with n and rounded. This procedure often yields good exact designs and sometimes even optimal exact designs, see Goos [16].

2.3.2 Optimal designs

We present the main definitions of optimal design of experiments on the basis of the simple model

$$y = X\theta + \epsilon \tag{2.6}$$

where y is the $n \times 1$ response vector dependent on the v experimental conditions $\theta = (\theta_1, \dots, \theta_v)^T$, X is the corresponding $n \times v$ *design matrix*. The random error terms ϵ_i, $i \in \{1, \dots, v\}$ are assumed to be independently identically distributed with mean zero and variance $\tilde{\sigma}^2 = 1$. An unbiased estimator of the fixed parameter vector θ is calculated with the method of ordinary least squares as

$$\hat{\theta} = (X^T X)^- X^T y \tag{2.7}$$

with variance

$$\mathrm{Var}(\hat{\theta}) = (X^T X)^-, \tag{2.8}$$

where $(X^TX)^-$ denotes a generalized inverse of X^TX. The information matrix for estimating the unknown parameter vector θ, given the design ξ_n with design matrix X is

$$M_{\xi_n} = (X^TX). \tag{2.9}$$

Considering approximate designs ξ the information matrix is defined as

$$M_\xi = (\grave{X}^TP\grave{X}) \tag{2.10}$$

with weight matrix $P = \text{diag}(p_1, \ldots, p_l)$ and design matrix $\grave{X}$ containing the l support points of the design. If only the parameters $\theta_1, \ldots, \theta_w$ with $w < v$ are of interest, the model can be restated as

$$y = X\theta + \epsilon = \left(\tilde{X}_1 \ \tilde{X}_2\right)\begin{pmatrix}\tilde{\theta}_1 \\ \tilde{\theta}_2\end{pmatrix} + \epsilon \tag{2.11}$$

with $\tilde{\theta}_1 = (\theta_1, \ldots, \theta_w)^T$ and $\tilde{\theta}_2 = (\theta_{w+1}, \ldots, \theta_v)^T$. $\tilde{X}_1$ and $\tilde{X}_2$ are the $n \times w$ and $n \times (v-w)$ submatrices of X. Therefore, the information matrix for estimating the unknown parameters $\tilde{\theta}_1 = (\theta_1, \ldots, \theta_w)^T$ is

$$M_{\xi_n} = \tilde{X}_1^T\tilde{X}_1 - \tilde{X}_1^T\tilde{X}_2(\tilde{X}_2^T\tilde{X}_2)^-\tilde{X}_2^T\tilde{X}_1 \tag{2.12}$$

due to Harvilles Theorem 9.6.1 [19].

Researchers are often interested in estimating a set of m contrasts $C^T\theta$ of the parameters θ; $C = (c_1, c_2, \ldots, c_m)$ is a $v \times m$ matrix with $c_i^T = (c_{i1}, \ldots, c_{iv})$ and $\sum_{j=1}^{v} c_{ij} = 0$ for $1 \leq i \leq m$. In this case, the ordinary least squares estimator for $C^T\theta$ has variance

$$\text{Var}(C^T\hat{\theta}) = C^T(X^TX)^-C \tag{2.13}$$

and the information matrix changes to

$$M_{\xi_n}^C = \left(C^T(X^TX)^-C\right)^{-1} \tag{2.14}$$

if the $v \times m$ matrix C has full column space. Pukelsheim [33] extended this definition for rank deficient subsystems and defined the generalized information matrix for $C^T\theta$ as

$$\tilde{M}_{\xi_n}^C = \min_{Q \in \mathbb{R}^{v \times v} : Q^T C = C} Q^T (X^T X) Q \tag{2.15}$$

with a $v \times m$ matrix C that may be rank deficient. The minimum is taken relative to the Loewner ordering. This partial ordering is defined as $A \geq B$ if and only if $A - B$ is nonnegative definite for symmetric matrices A and B. Thus, the Gauss-Markov Theorem provides

$$\tilde{M}_{\xi_n}^C = (X^T X) - (X^T X) R^T (R(X^T X) R^T)^- R(X^T X) \tag{2.16}$$

with $R = I_v - CG$ with an arbitrary generalized inverse G of C. Pukelsheim shows the equality $\tilde{M}_{\xi}^C = C M_{\xi}^C C^T$ for all C with a full column rank. It is easy to verify that these contrast information matrices are nonnegative definite with zero row and column sums. Druilhet and Markiewicz [11] give another definition of generalized information matrices, since Pukelsheims definition does not lead to the usual information matrix for full rank subsystems. They proposed

$$M_{\xi_n}^C = \min_{Q \in \mathbb{R}^{v \times m} : Q^T C = C^T (C C^T)^+ C} Q^T (X^T X) Q \tag{2.17}$$

and showed $M_{\xi_n}^C = (C^T (X^T X)^- C)^+$ if $C^T\theta$ is estimable, i.e. if

$$\text{Range}(C) \subset \text{Range}(X^T X).$$

$(C^T (X^T X)^- C)^+$ denotes the Moore-Penrose-Inverse of $(C^T (X^T X)^- C)$. This definition leads to the usual information matrix for full rank subsystems. Since the Loewner ordering is a partial ordering, we define information functions in order to compare arbitrary information matrices. Let NND_v^0 be the set of $v \times v$ nonnegative definite matrices with zero row and column sums. Information functions are defined as in Kiefer [25].

Definition 2.1:

$\phi : \text{NND}_v^0 \to \mathbb{R}$ is an information function, if it satisfies the following conditions:

(a) ϕ is convex,

(b) ϕ is invariant under simultaneous permutations of rows and the same columns,

(c) $\phi(\alpha C)$ is non increasing in the scalar $\alpha \geq 0$ for $C \in NND_v^0$.

The smaller the value of $\phi(M_\xi)$, the smaller the variance of the estimated parameters and the more efficient is the design. The purpose of optimal design theory is to determine the designs with the highest information or equivalent with the minimal variance. This corresponds to the minimization of the function $\phi(M_\xi^C)$ for $\xi \in \Xi$. A design ξ is said to be universal optimal in a set Ξ of designs if it minimizes $\phi(M_\xi)$ for all information functions ϕ (Kiefer [26]). In many cases universal optimal designs do not exist since they depend on the information function. Therefore, we have to restrict to special optimality criteria. The most prominent criteria are the matrix means ϕ_q'.

Example 2.2*:*
A design ξ is ϕ_q'-optimal for $q \in [-1, \infty) \backslash \{0\}$, if it minimizes the expression

$$\phi_q'(M_\xi^C) = \left(\frac{\sum_{i=1}^{v-1} \lambda_i^{-q}}{v-1} \right)^{\frac{1}{q}} \tag{2.18}$$

for the $v-1$ eigenvalues λ_i of the information matrix M_ξ^C. Define $\phi_0'(M_\xi^C) = \prod_i \lambda_i^{-1/(v-1)}$ and $\phi_\infty'(M_\xi^C) = \max \lambda_i^{-1}$.

A special case of the ϕ_q' -criteria are the D-optimality criterion for $q = 0$ and the A-optimality criterion for $q = 1$. D-optimal designs minimize the determinant of the variance covariance matrix of the parameter estimates, their main advantage is that they are invariant to a change of scale in the factors. A-optimal designs minimize the trace of the variance covariance matrix or equivalently minimize the sum of the variances of the parameters of interest $\sum_{i=1}^{v} \mathrm{Var}(\hat{\theta}_i)$. Considering the contrast set $(c_1, \ldots, c_m)$ the A-optimal design minimizes the term $\mathrm{Tr}\left(\mathrm{Var}(C^T\hat{\theta})\right) = \sum_{l=1}^{m} \mathrm{Var}({c_l}^T\hat{\theta})$ for $\hat{\theta} = (\hat{\theta}_1, \ldots, \hat{\theta}_v)^T$.

Pukelsheim [33] generalized the ϕ_q'-criteria for rank deficient subsystems with singular information matrices. Matrix means ϕ_q for rank deficient subsystems

are defined as follows

$$\phi_q(M_\xi^C) = \phi_q'(\lambda_1, \ldots, \lambda_r)$$

whereas $\lambda_1, \ldots, \lambda_r$ are the positive eigenvalues of the singular information matrix M_ξ^C.

2.3.3 Equivalence theorem

We present a central result of optimal design theory for approximate designs, the equivalence theorem. This theorem can be applied to show ϕ_q-optimality of a given design for the estimation of an arbitrary contrast set C.

Theorem 2.3:
A design is ϕ_{-p}-optimal, $p \in (-\infty, 1]$, for the estimation of the contrast set C if and only if there exists a generalized inverse $G = (\grave{X}^T P \grave{X})^-$ of $\grave{X}^T P \grave{X}$ that satisfies the normality inequality

$$x^T GC \left(C^T GC\right)^+ \left(C^T GC\right)^{1-p} \left(C^T GC\right)^+ C^T G^T x \leq Tr\left(\left(C^T GC\right)^+ \left(C^T GC\right)^{1-p}\right) \tag{2.19}$$

for all possible design points $x \in \mathcal{X}$, $\mathcal{X} = \{x \in \{-1, 0, 1\}^v : \exists! i \text{ with } x_i = 1 \wedge \exists! j \text{ with } x_j = -1\}$. The expression $\exists!$ stands for "'there exists exactly one"'. P is the diagonal matrix containing the optimal weights for all design points listed in the design matrix $\grave{X}$. In case of optimality, equality holds in the normality inequality (2.19) for all support points x of all optimal designs.

For the proof, a detailed discussion and the equivalence theorem for $p = \infty$ we refer to Pukelsheim [33].

Another interesting theorem which we use in this work is due to Kiefer [25].

Theorem 2.4:
Let $\phi : \mathrm{NND}_v^0 \to \mathbb{R}$ be an information function. If there is a design with a completely symmetric information matrix, which has maximum trace in a class of designs, then it is universally optimal in the given class.

A matrix is said to be completely symmetric if it is of the form $eI_v + fJ_v$, where I_v is the $v \times v$ identity matrix and J_v is the $v \times v$ matrix with all entries equal

to one, e and f are scalars. It can be easily shown that

$$(eI_v + fJ_v)^+ = \frac{1}{e}I_v - \frac{f}{e(e+vf)}J_v \tag{2.20}$$

if $e + vf \neq 0$ and $e \neq 0$.

$$(eI_v + fJ_v)^+ = \frac{1}{e}I_v - \frac{1}{ve}J_v \tag{2.21}$$

if $e + vf = 0$. The eigenvalues of $(eI_v + fJ_v)$ are e with multiplicity $v - 1$ and $e + vf$ with multiplicity 1.

In the special case of model (2.5) with $\tilde{R} := X^T X$ yields $\tilde{r}_{ii}$ is the number of occurrences of treatment i overall, while $\tilde{r}_{ij}$ denotes the number of blocks which contain both treatments i and j. Therefore, an information matrix maximizes the trace and is completely symmetric if and only if each treatment occurs equally often and every two treatments are contained in the same number of blocks. A block of size two corresponds to an array in the microarray setting.

Define P_k recursively by $P_2 := [1, -1]$ and $P_k := \left[\begin{array}{c|c} \mathbb{1}_{k-1} & -I_{k-1} \\ \hline 0_{\binom{k-1}{2}} & P_{k-1} \end{array}\right]$ for all $k \in \mathbb{N}^{\geq 3}$, whereas $\mathbb{1}_k$ and 0_k are k-dimensional column vectors with all entries equal to 1 and 0, respectively. I_k is the $k \times k$ identity matrix.

Obviously the design with design matrix $X := P_{t+1}$

$$\mathrm{P}_{t+1} := \left(\begin{array}{c|ccc|cc} \mathbf{1}_t & \multicolumn{5}{c}{-\boldsymbol{I}_t} \\ \hline \mathbf{0}_{t-1} & \mathbf{1}_{t-1} & \multicolumn{4}{c}{-\boldsymbol{I}_{t-1}} \\ \hline \mathbf{0}_{t-2} & \mathbf{0}_{t-2} & \mathbf{1}_{t-2} & \multicolumn{3}{c}{-\boldsymbol{I}_{t-2}} \\ \hline \vdots & \ddots & \ddots & \multicolumn{3}{c}{\ddots} \\ \hline \mathbf{0}_2 & \cdots & \mathbf{0}_2 & \mathbf{1}_2 & \multicolumn{2}{c}{-\boldsymbol{I}_2} \\ \hline \mathbf{0}_1 & \multicolumn{2}{c}{\cdots} & \mathbf{0}_1 & \mathbf{1}_1 & -\boldsymbol{I}_1 \end{array}\right)$$

fulfills this condition. Thus, it is particularly ϕ_q-optimal, $q \in [-1, \infty]$, for the estimation of all orthogonal contrast sets, which we will use later on. We refer

to Stanzel [37] for helpful properties of the matrix P_{t+1}. We use the facts

$$(P_{t+1}P_{t+1}^T)^q = (t+1)^{q-1}P_{t+1}P_{t+1}^T \text{ and} \tag{2.22}$$

$$\mathrm{Tr}((P_{t+1}P_{t+1}^T)^+(P_{t+1}P_{t+1}^T)) = t \tag{2.23}$$

for $q \in \mathbb{R}^{\geq 0}$ and $t \in \mathbb{N}$ in Section 4.

Chapter 3

Exact A-Optimal Designs

In this chapter we propose a method to derive exact A-optimal designs for practical situations with a given number of arrays and small numbers of treatments, since in many applications the number of treatments does not exceed a known limit. We apply this approach to several contrast settings, including the comparisons with a control treatment, all-to-next contrasts and Helmert contrasts. A-optimal designs for pairwise treatment comparisons are derived in Tsai et al. [41] and Bailey [2].

Throughout this chapter we consider the A-optimality criterion, since it is still the most popular one for block designs, see Atkinson [1]. Another important criterion, especially in robustness investigations, is the D-optimality criterion. However, Bailey [2] already has investigated D-optimal design for the estimation of all pairwise comparisons and this criterion is independent on the contrast set of interest. All derivations in this chapter are based on model 2.1, whereas the dye effect is ignored in a first step referring to Bailey [2]. Dyes are reintroduced in Chapter 4.7.

3.1 Calculation of A-optimal designs

Let $(c_1, \ldots, c_m)$ with $c_l^T = (c_{l0}, \ldots, c_{lt})$ be the contrast set of interest for the parameters $\tau = (\tau_0, \ldots, \tau_t)^T$. Then the A-optimal design minimizes the term

$$\sum_{l=1}^{m} \mathrm{Var}({c_l}^T \hat{\tau}) \tag{3.1}$$

as outlined in Section 2.3.

This sum is in particular dependent on the number of used arrays a and on the two treatments combined on each array. Since each contrast $c_l \in \mathbb{R}^t$, $c_l^T = (c_{l0}, \ldots, c_{lt})$, $l \in \{1, \ldots, m\}$ fulfills the equation $\sum_{i=0}^{t} c_{li} = 0$ by definition, we get $\mathrm{Var}({c_l}^T \hat{\tau}) = \mathrm{Var}(\sum_{i=0}^{s} c_{li}\hat{\tau}_i + \sum_{i=s+1}^{t} c_{li}\hat{\tau}_i)$ for $c_{l0}, \ldots, c_{ls} \geq 0$ and $c_{l(s+1)}, \ldots, c_{lt} < 0$ with $\sum_{i=0}^{s} c_{li} = -\sum_{i=s+1}^{t} c_{li}$ without loss of generality. Hence, we rephrase $\mathrm{Var}\left({c_l}^T \hat{\tau}\right) = \mathrm{Var}\left(\sum_{i<j} a_{ij}\left(\hat{\tau}_i - \hat{\tau}_j\right)\right)$ with appropriate a_{ij} resulting from the values of c_{li}. Thus, it is sufficient to calculate the expressions $\mathrm{Var}(\hat{\tau}_i - \hat{\tau}_j)$ and $\mathrm{Cov}(\hat{\tau}_i - \hat{\tau}_j, \hat{\tau}_k - \hat{\tau}_l)$ to determine $\sum_{l=1}^{m} \mathrm{Var}({c_l}^T \hat{\tau})$. Therefore we will give a formula to calculate these expressions in the following theorem. This theorem uses the fact that each microarray experiment with $t+1$ treatments and a arrays can be illustrated by a multigraph with $t+1$ vertices and a edges, whereas every two treatments tested on the same array in the experiment are connected by an edge in the graph, as outlined in Section 2.1.

Theorem 3.1:
Let $V = \{0, \ldots, t\}$ be the set of vertices (treatments) and $E = \{x_{01}, x_{02}, \ldots, x_{(t-1)t}\}$ the set of edges of a given graph with $|E| = \binom{t+1}{2}$.
The function $b : x_{ij} \longmapsto b(x_{ij}) : E \longrightarrow \mathbb{N}_0$ specifies the number of arrays $b(x_{ij})$ comparing treatments i and j for each treatment pair (i, j), i.e. the graph with vertex set V and with the a edges in the multi set

$$\tilde{E} = \{\underbrace{x_{01}, \ldots, x_{01}}_{b(x_{01}) \text{ times}}, \underbrace{x_{02}, \ldots, x_{02}}_{b(x_{02}) \text{ times}}, \ldots, \underbrace{x_{(t-1)t}, \ldots, x_{(t-1)t}}_{b(x_{(t-1)t}) \text{ times}}\}$$

describes a specific microarray experiment with a arrays. Then

$$\mathrm{Var}(\hat{\tau}_i - \hat{\tau}_j) = \frac{\sum\limits_{\substack{A\subset E\setminus\{x_{ij}\}:|A|=t-1:\\ (V,A\cup\{x_{ij}\}) \text{ has no loops}}} b(a_1)b(a_2)\cdots b(a_{t-1})}{\sum\limits_{\substack{A\subset E:|A|=t:\\ (V,A) \text{ has no loops}}} b(a_1)b(a_2)\cdots b(a_t)} \tag{3.2}$$

with $i, j \in \{0, \ldots, t\}$, $i \neq j$ and $A = \{a_1, a_2, \ldots\} \subset E$.

Theorem 3.1 can be proven with results from physical networks, especially with the help of resistance matrices. For a detailed description see Bailey and Cameron [3]. Furthermore, it can be shown easily that

$$\mathrm{Cov}(\hat{\tau}_i-\hat{\tau}_j, \hat{\tau}_k-\hat{\tau}_l) = \frac{1}{2}(Var(\hat{\tau}_i-\hat{\tau}_l)+Var(\hat{\tau}_j-\hat{\tau}_k)-Var(\hat{\tau}_j-\hat{\tau}_l)-Var(\hat{\tau}_i-\hat{\tau}_k)).$$

We demonstrate the application of Theorem 3.1 in the following examples.

Example 3.2*:*
If we consider three treatments the experiment with $a = x + y + z$ arrays is illustrated in Figure 3.1. Let $b(x_{01}) = x$, $b(x_{02}) = y$, $b(x_{12}) = z$, i.e. treatment 0 is combined with treatment 1 on x arrays etc. Consequently we get

$$\begin{aligned}
&\mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_1) = \tfrac{y+z}{xy+xz+yz}, && \mathrm{Cov}(\hat{\tau}_0 - \hat{\tau}_1, \hat{\tau}_0 - \hat{\tau}_2) = \tfrac{z}{xy+xz+yz},\\
&\mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_2) = \tfrac{x+z}{xy+xz+yz}, && \mathrm{Cov}(\hat{\tau}_0 - \hat{\tau}_1, \hat{\tau}_1 - \hat{\tau}_2) = \tfrac{y}{xy+xz+yz},\\
&\mathrm{Var}(\hat{\tau}_1 - \hat{\tau}_2) = \tfrac{x+y}{xy+xz+yz}, && \mathrm{Cov}(\hat{\tau}_0 - \hat{\tau}_2, \hat{\tau}_1 - \hat{\tau}_2) = \tfrac{x}{xy+xz+yz}.
\end{aligned}$$

The denominator of these terms corresponds to the number of spanning trees of the underlying graph. A spanning tree is a connected subgraph without any loop, which contains every vertex of the underlying graph. The numerator of $\mathrm{Var}(\hat{\tau}_i - \hat{\tau}_j)$ specifies the number of spanning thickets of the graph with i and j in different components, whereby a spanning thicket is a spanning forest with exactly two components. The variances of these estimators depend on the number of paths joining the two vertices.

Example 3.3*:*
Considering four treatments we denote $b(x_{ij}) = B_{ij}$, i.e. treatment i is combined

with treatment j on B_{ij} arrays. The experiment with $a = B_{01} + B_{02} + B_{03} + B_{12} + B_{13} + B_{23}$ arrays is illustrated in Figure 3.1. Consequently we can calculate the denominator of $\mathrm{Var}(\hat{\tau}_i - \hat{\tau}_j)$ as

$$\begin{aligned}
&B_{01}B_{12}B_{23} + B_{01}B_{12}B_{03} + B_{01}B_{12}B_{13} + B_{01}B_{23}B_{03} + B_{01}B_{23}B_{13} + B_{01}B_{23}B_{02} \\
&+B_{01}B_{03}B_{02} + B_{01}B_{13}B_{02} + B_{12}B_{23}B_{03} + B_{12}B_{23}B_{02} + B_{12}B_{03}B_{13} + B_{12}B_{03}B_{02} \\
&+B_{12}B_{13}B_{02} + B_{23}B_{03}B_{13} + B_{23}B_{13}B_{02} + B_{03}B_{13}B_{02} =: d
\end{aligned}$$

for all $i \neq j$. The numerators of $\mathrm{Var}(\hat{\tau}_i - \hat{\tau}_j)$ can be extracted from the following terms.

$$\begin{aligned}
d \cdot \mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_1) &= B_{12}B_{23} + B_{12}B_{03} + B_{12}B_{13} + B_{23}B_{03} + B_{23}B_{13} \\
&\quad +B_{23}B_{02} + B_{03}B_{02} + B_{13}B_{02}, \\
d \cdot \mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_2) &= B_{01}B_{23} + B_{01}B_{03} + B_{01}B_{13} + B_{12}B_{23} + B_{12}B_{03} \\
&\quad +B_{12}B_{13} + B_{23}B_{13} + B_{03}B_{13}, \\
d \cdot \mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_3) &= B_{01}B_{12} + B_{01}B_{23} + B_{01}B_{02} + B_{12}B_{23} + B_{12}B_{13} \\
&\quad +B_{12}B_{02} + B_{23}B_{13} + B_{13}B_{02}, \\
d \cdot \mathrm{Var}(\hat{\tau}_1 - \hat{\tau}_2) &= B_{01}B_{23} + B_{01}B_{03} + B_{01}B_{13} + B_{23}B_{03} + B_{23}B_{02} \\
&\quad +B_{03}B_{13} + B_{03}B_{02} + B_{13}B_{02}, \\
d \cdot \mathrm{Var}(\hat{\tau}_1 - \hat{\tau}_3) &= B_{01}B_{12} + B_{01}B_{23} + B_{01}B_{02} + B_{12}B_{03} + B_{12}B_{02} \\
&\quad +B_{23}B_{03} + B_{23}B_{02} + B_{03}B_{02}, \\
d \cdot \mathrm{Var}(\hat{\tau}_2 - \hat{\tau}_3) &= B_{01}B_{12} + B_{01}B_{03} + B_{01}B_{13} + B_{01}B_{02} + B_{12}B_{03} \\
&\quad +B_{12}B_{02} + B_{03}B_{13} + B_{13}B_{02},
\end{aligned}$$

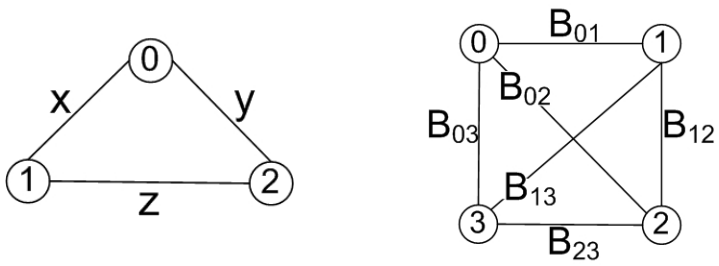


Figure 3.1: Graph representation of microarray experiments with three and four treatments.

$$d \cdot \mathrm{Cov}(\hat{\tau}_0 - \hat{\tau}_1, \hat{\tau}_0 - \hat{\tau}_2) = B_{12}B_{23} + B_{12}B_{03} + B_{12}B_{13} + B_{23}B_{13},$$

$$d \cdot \mathrm{Cov}(\hat{\tau}_0 - \hat{\tau}_1, \hat{\tau}_0 - \hat{\tau}_3) = B_{12}B_{23} + B_{12}B_{13} + B_{23}B_{13} + B_{13}B_{02},$$

$$d \cdot \mathrm{Cov}(\hat{\tau}_0 - \hat{\tau}_2, \hat{\tau}_0 - \hat{\tau}_3) = B_{01}B_{23} + B_{12} + B_{23} + B_{12}B_{13} + B_{23}B_{13},$$

$$d \cdot \mathrm{Cov}(\hat{\tau}_1 - \hat{\tau}_2, \hat{\tau}_1 - \hat{\tau}_3) = B_{01}B_{23} + B_{23}B_{03} + B_{23}B_{02} + B_{03}B_{02},$$

and so on.

Theorem 3.1 can theoretically be applied to all values of treatments t, but the computation time increases immensely. However, in many applications the number of treatments is small and exact optimal designs can be derived with Theorem 3.1 for all numbers of arrays a. For larger values of t approximate optimal designs are proposed in Chapter 4. The corresponding approximate optimality results can be used to construct nearly optimal designs for all values of t and a.

3.2 Optimal designs for treatment-control comparisons

Although in medical applications scientists are often interested in comparing several treatments to a control-treatment, only few authors considered design problems for treatment-control comparisons in microarray experiments. Kunert et al. [28] derived approximate optimal designs in this scenario, exact designs were not computed. Thus, we will construct exact A-optimal designs for estimating treatment-control comparisons and we will show that these designs are

more efficient than the star designs which are often used in practice. Star designs allocate each treatment together with the control-treatment on the same array, e.g. Figure 2.2. For example, Zieker et al.[47] used a star design with two treatments and a control to compare gene expressions of marathon runners, before, immediately after and 24 hours after exercise. The control-treatment in a star design is of practical interest in contrast to the reference in a common reference design. Common reference designs are widely used in practice (e.g. Callow et al. [6]) and compare competing treatments via a reference sample, which is not of interest itself. Star designs always perform better than the common reference designs, since they do not waste resources in order to estimate the effect of the uninteresting reference treatment.

Throughout this section τ_0 describes the control-treatment and τ_i, $i \in \{1, \ldots, t\}$ describe the other treatments. Therefore, we have to minimize $\sum_{i=1}^{t} \mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_i)$, if we are interested in estimating all treatment-control contrasts $\tau_0 - \tau_i$ for $i = 1, \ldots, t$. We will derive A-optimal designs using Theorem 3.1 for $t \in \{2, 3, 4\}$, because these values are often used in practical settings.

If we consider two treatments and one control-treatment we have to minimize the function

$$\mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_1) + \mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_2) = \frac{x + y + 2z}{xy + xz + yz} \tag{3.3}$$

under the constraints $x + y + z = a$ and $x, y, z \in \{0, \ldots, a\}$. The results of this minimization obtained for $a \in \{6, 8, 10, 12, 15\}$ arrays are displayed in Table 3.1. Similar results for other values of a can be obtained easily by minimizing expression (3.3). Certainly, all designs remain optimal, if the values of $x = B_{01}$ and $y = B_{02}$ are interchanged. For instance, if we use $a = 10$ arrays we get the optimal design, which investigates both treatment-control comparisons on four microarrays each and the non-interesting treatment by treatment comparison on the remaining two microarrays. In Table 3.2 and Table 3.3 we listed similar results for $t \in \{3, 4\}$ (using Mathematica 7.0.1.0, Wolfram Research). Altogether, we realize that it is efficient to hybridize the control-treatment on more arrays than the other treatments, but it is recommendable to use also some

Table 3.1: A-Optimal designs for comparisons with a control, $t = 2$

	$a = 6$	$a = 8$	$a = 10$	$a = 12$	$a = 15$
B_{01}	3	4	4	5	6
B_{02}	2	3	4	5	6
B_{12}	1	1	2	2	3

Table 3.2: A-Optimal designs for comparisons with a control, $t = 3$

	$a = 9$	$a = 11$	$a = 12$	$a = 15$	$a = 20$	$a = 25$
B_{01}	2	3	3	4	5	7
B_{02}	2	2	3	4	5	6
B_{03}	2	3	3	4	5	6
B_{12}	1	1	1	1	1	2
B_{13}	1	1	1	1	2	2
B_{23}	1	1	1	1	2	2

arrays without the control. As mentioned above, star designs are often used by researchers in medical and biological applications. Using the designs proposed in this paper instead of the star designs, we observe a gain in efficiency of at least 4% for $t = 2$ and at least 10% for $t \in \{3, 4\}$, see Table 3.4. We get similar results for other values of t and a. Therefore, the star design is not advisable, even if we are interested in the treatment-control comparisons. The poor performance of the star design in other contrast settings is considered in Vinciotti [43] for example.

Table 3.3: A-Optimal designs for comparisons with a control, $t = 4$

	$a = 14$	$a = 15$	$a = 16$	$a = 20$	$a = 25$
B_{01}	2	3	2	4	5
B_{02}	2	2	2	3	4
B_{03}	2	3	3	4	5
B_{04}	2	2	3	3	4
B_{12}	1	1	1	1	1
B_{13}	1	0	1	1	1
B_{14}	1	1	1	1	1
B_{23}	1	1	1	1	1
B_{24}	1	1	1	1	2
B_{34}	1	1	1	1	1

Table 3.4: Comparisons of the variances obtained for the constructed optimal designs and the star designs, $t \in \{2, 3, 4\}$ treatments, a arrays.

$t = 2$	$a = 6$	$a = 8$	$a = 10$	$a = 12$	$a = 15$
Var. opt. design	0.64	0.47	0.38	0.31	0.25
Var. star design	0.67	0.5	0.4	0.33	0.31
$\frac{\text{Var. opt. design}}{\text{Var. star design}}$	0.96	0.94	0.95	0.94	0.80

$t = 3$	$a = 9$	$a = 11$	$a = 12$	$a = 15$	$a = 20$	$a = 25$
Var. opt. design	0.9	0.74	0.67	0.54	0.40	0.32
Var. star design	1	0.83	0.75	0.6	0.45	0.36
$\frac{\text{Var. opt. design}}{\text{Var. star design}}$	0.9	0.89	0.89	0.9	0.88	0.89

$t = 4$	$a = 14$	$a = 15$	$a = 16$	$a = 20$	$a = 25$
Var. opt. design	1.00	0.93	0.87	0.69	0.55
Var. star design	1.17	1.08	1	0.8	0.64
$\frac{\text{Var. opt. design}}{\text{Var. star design}}$	0.86	0.86	0.87	0.86	0.86

3.3 Optimal designs for all-to-next contrasts

Another interesting contrast set are the all-to-next contrasts $\tau_{i-1} - \tau_i$ for $i \in \{1, \ldots, t\}$. They are often used in time course experiments to compare consecutive points in time. We will compute A-optimal designs for these contrasts with the same method used in the previous section. Thus, the derivation of A-optimal designs for this scenario leads to the minimization of the target function $\sum_{i=1}^{t} \text{Var}(\hat{\tau}_{i-1} - \hat{\tau}_i)$ for a given number of a arrays. Therefore, for $t = 2$ the target function

$$\frac{B_{01} + 2B_{02} + B_{12}}{B_{01}B_{02} + B_{01}B_{12} + B_{02}B_{12}}$$

has to be minimized under the constraints $B_{01}+B_{02}+B_{12} = a$ and $B_{01}, B_{02}, B_{12} \in \mathbb{N}_0$, which is equivalent to expression (3.3) with interchanged variables. For $t = 3$ the function

$$\begin{aligned}&(B_{01}B_{12} + B_{01}B_{23} + B_{12}B_{23} + 2B_{01}B_{03} + 2B_{12}B_{03} + 2B_{23}B_{03} + 2B_{01}B_{13}\\ &+B_{12}B_{13} + B_{23}B_{13} + 2B_{03}B_{13} + B_{01}B_{02} + B_{12}B_{02} + 2B_{23}B_{02} + 2B_{03}B_{02}\\ &+3B_{13}B_{02})/d\end{aligned}$$

has to be minimized under the constraints $B_{01} + B_{02} + B_{03} + B_{12} + B_{13} + B_{23} = a$ and $B_{01}, B_{02}, B_{03}, B_{12}, B_{13}, B_{23} \in \mathbb{N}_0$. The solutions of these minimizations are

Table 3.5: A-Optimal designs for the estimation of all-to-next contrasts, $t = 2$

	$a = 6$	$a = 8$	$a = 10$	$a = 12$	$a = 15$
B_{01}	2	3	4	5	6
B_{02}	1	1	2	2	3
B_{12}	3	4	4	5	6

Table 3.6: A-Optimal designs for the estimation of all-to-next contrasts, $t = 3$

	$a = 9$	$a = 11$	$a = 12$	$a = 15$	$a = 20$	$a = 25$
B_{01}	2	3	3	4	5	7
B_{02}	1	1	1	1	2	2
B_{03}	1	1	1	1	1	1
B_{12}	2	2	3	4	5	6
B_{13}	1	1	1	1	2	2
B_{23}	2	3	3	4	5	7

listed in Table 3.5 and Table 3.6. For example, considering 15 microarrays and four treatments the samples of consecutive points in time are compared on four slides and the non-consecutive treatment comparisons are hybridized only on one slide. Independent on the number of arrays and treatments, we observe that A-optimal designs comprise more microarrays hybridizing consecutive treatment comparisons than microarrays hybridizing more distant treatments.

3.4 Optimal designs for Helmert contrasts

Helmert contrasts compare each treatment to the mean of the treatments with subsequent treatment indices, i.e. $\tau_i - \frac{1}{t-i}\sum_{l=i+1}^{t}\tau_l$, $i \in \{0, 1, \ldots, t\}$. These contrasts are also very useful for ordered treatment arrangements, for instance time course experiments. For $t = 2$ the Helmert contrasts are given by $\tau_0 - \frac{1}{2}\tau_1 - \frac{1}{2}\tau_2$ and $\tau_1 - \tau_2$. Thus, A-optimal designs for estimating these contrasts can be obtained by the following minimization

$$
\begin{aligned}
&\min \left(\mathrm{Var}(\hat{\tau_0} - \frac{1}{2}\hat{\tau_1} - \frac{1}{2}\hat{\tau_2}) + \mathrm{Var}(\hat{\tau_1} - \hat{\tau_2})\right) \\
&= \min \left(\frac{1}{4}\mathrm{Var}(\hat{\tau_0} - \hat{\tau_1}) + \frac{1}{4}\mathrm{Var}(\hat{\tau_0} - \hat{\tau_2}) + \mathrm{Var}(\hat{\tau_1} - \hat{\tau_2}) + \frac{1}{2}\mathrm{Cov}(\hat{\tau_0} - \hat{\tau_1}, \hat{\tau_0} - \hat{\tau_2})\right) \\
&= \min \frac{5B_{01} + 5B_{02} + 4B_{12}}{4B_{01}B_{02} + 4B_{01}B_{12} + 4B_{02}B_{12}}
\end{aligned}
$$

Table 3.7: A-Optimal designs for the estimation of Helmert contrasts, $t = 2$

	$a = 6$	$a = 8$	$a = 10$	$a = 12$	$a = 15$
B_{01}	2	2	3	4	4
B_{02}	2	3	3	4	5
B_{12}	2	3	4	4	6

Table 3.8: A-Optimal designs for the estimation of Helmert contrasts, $t = 3$

	$a = 9$	$a = 11$	$a = 12$	$a = 15$	$a = 20$	$a = 25$
B_{01}	1	1	2	2	3	4
B_{02}	2	2	2	3	3	4
B_{03}	1	2	2	2	3	4
B_{12}	1	2	2	2	3	4
B_{13}	2	2	2	3	4	4
B_{23}	2	2	2	3	4	5

under the constraints $B_{01} + B_{02} + B_{12} = a$ and $B_{01}, B_{02}, B_{12} \in \mathbb{N}_0$. The results of this minimization for $a \in \{6, 8, 10, 12, 15\}$ arrays are displayed in Table 3.7. The first Helmert contrast vector $(1, -1/2, -1/2)$ has the norm $\sqrt{3/2}$ and the second vector $(0, 1, -1)$ has the norm $\sqrt{2}$. Therefore, the comparison of treatments one and two is more important than the comparison of treatment zero and treatment one. This is reflected in Table 3.7by the fact that for any value of a the value of B_{12} is always the highest. According results for $t = 3$ are listed in Table 3.8 (using Mathematica 7.0.1.0, Wolfram Research). Again, comparisons between treatments with larger treatment indices are hybridized together on more slides due to the norms of the Helmert contrasts. For example, the value B_{23} is the highest for all values of a in Table 3.8. Helmert contrasts belong to the set of orthogonal contrasts, hence it is straightforward to find the optimal designs for normalized Helmert contrasts $\left(\tau_i - \frac{1}{t-i} \sum_{l=i+1}^{t} \tau_l\right) / \sqrt{\frac{t-i+1}{t-i}}$ due to the fact that the corresponding contrast matrix is orthogonal. In this section we considered usual non-normalized Helmert contrasts $\tau_i - \frac{1}{t-i} \sum_{l=i+1}^{t} \tau_l$, which assign higher weights to posterior comparisons.

3.5 Optimal designs for all pairwise treatment comparisons

The approach outlined in Section 3.1 can also be used to derive A-optimal designs for all pairwise treatment comparisons. Since exact optimal designs for these contrast set are already considered in Tsai et al. [41] and in Bailey [2], we will not state the numerically solutions of this problem here. Tsai et al. proposed an algorithm to find exact optimal designs for this contrast set, which is based on an exhaustive search on non-isomorphic graphs. This approach can only be applied for very small numbers of arrays and treatments and is very time consuming. Our approach yields the same designs as stated in Tsai et al. [41]. Another way to explore optimal designs for the pairwise treatment comparisons $C = P_{t+1}^T$ for selected numbers of arrays can be traced back to Kiefers derivation of universal optimal designs [25] stated in Theorem 2.4.

Theorem 3.4:
Considering model (2.5) without dye effect, the design ξ with design matrix $\grave{X} := P_{t+1}$ is ϕ_q-optimal, $q \in [-1, \infty]$ for the estimation of the contrasts $C^T\tau$ with $C = P_{t+1}^T$.

Proof: As mentioned at the end of Section 2.3.3, a design with design matrix $X := P_{t+1}$ is ϕ_q-optimal, $q \in [-1, \infty]$ for all contrast sets with orthogonal contrast matrices C. Since the pairwise treatment comparison contrasts are not orthogonal, we have to show that the eigenvalues of the information matrix are invariant to pre- and post-multiplication with P_{t+1} and P_{t+1}^T, respectively. We will prove that the eigenvalues remain the same except for multiplication with a constant factor. In this case the optimal design remains the same, since the matrix means do only depend on the eigenvalues of the considered matrices. For that purpose we consider the singular value decomposition of $C^T = P_{t+1} = U\Sigma V^T$ with unitary matrices U and V and a diagonal matrix Σ. Obviously it holds that

$$M := P_{t+1}^T P_{t+1} = (t+1)I_{t+1} - J_{t+1}.$$

Thus, $P_{t+1}^T P_{t+1}$ has t positive eigenvalues equal to $t+1$ as mentioned in Section 2.3.3. Hence, the positive singular values of P_{t+1} correspond to $\sqrt{t+1}$ and we obtain

$$\Sigma = \begin{pmatrix} \sqrt{t+1} & & & 0 \\ & \sqrt{t+1} & & 0 \\ & & \ddots & \vdots \\ & & & \sqrt{t+1}\ 0 \\ 0 & \cdots & & 0 \\ \vdots & & & \vdots \\ 0 & \cdots & & 0 \end{pmatrix}.$$

We are interested in the set of eigenvalues, i.e. the spectrum $\tilde{\sigma}$, of the matrix $P_{t+1} M_\xi^- P_{t+1}^T$. This matrix does not depend on the generalized inverse of M_ξ and equals $P_{t+1} M_\xi^+ P_{t+1}^T$. This can be easily shown with Pukelsheim's [33] Theorem I.17. Due to $\text{Range}(P_{t+1}) \subset \text{Range}(I_{\binom{t+1}{2}})$ the equality

$$\text{Range}(P_{t+1}^T P_{t+1}) = \text{Range}(P_{t+1}^T)$$

holds. In particular it holds that $\text{Range}(P_{t+1}^T P_{t+1}) \subset \text{Range}(P_{t+1}^T)$ and hence $P_{t+1} M^- P_{t+1}^T$ does not depend on the generalized inverse of M using Pukelsheim's Theorem I.17. Supposing that

$$I_0 = \begin{pmatrix} & & 0 \\ & I & \\ & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

is the identity matrix with an additional row with zeros and an additional column with zeros, we can show

$$\begin{aligned} &\tilde{\sigma}(P_{t+1} M^- P_{t+1}^T) = \tilde{\sigma}(U \Sigma V^T M^+ V \Sigma^T U^T) = \tilde{\sigma}(\Sigma V^T M^+ V \Sigma^T) \\ &= v\tilde{\sigma}(I_0 V^T M^+ V I_0) \cup \{0, \ldots, 0\} = v\tilde{\sigma}(M^+) \cup \{0, 0, \ldots, 0\}. \end{aligned}$$

The last equality holds because $M^+ = \frac{1}{t+1} I_{t+1} - \frac{1}{(t+1)^2} J_{t+1}$ has row and column sums equal to zero and the last column of V is a multiple of $(1, \ldots, 1)^T$, since the last column of V corresponds to the singular value 0 of P_{t+1}.We get $\phi_q((P_{t+1} M^- P_{t+1}^T)^+) = v\phi_q(M)$ due to the definition of the matrix means and we achieve the same optimization tasks for our contrast set $C = P_{t+1}^T$ as for

orthogonal contrasts or no contrasts. Thus, we get the same optimal design. □

This result is comprehensible, we expect that the optimal design hybridizes all treatment combinations together on a microarray.

Chapter 4

Approximate Optimal Designs

For large values of t the computation time of exact optimal designs increases immensely, hence we will propose approximate optimal designs for these situations in this chapter. The corresponding approximate optimality results can be used to construct nearly optimal exact designs for all values of $t, a \in \mathbb{N}$. We will derive approximate optimal designs for one and multi-factorial settings. Multi-factorial settings contain more than one experimental factor of interest. For example, in addition to the treatment effect researchers are often interested in the cell line effect, in the effect of gender, and in appropriate interactions. For instance, Churchill [8] investigated several mouse cell lines medicated with different treatments. He was interested in the cell line effect as well as in the treatment effect and in treatment by cell line interactions. However, only few authors have considered optimal designs in multi-factorial settings. Some authors have investigated optimal designs for the estimation of main effects and first-order interactions. See, for example, Glonek and Solomon [15] or Banerjee and Mukerjee [4], who have examined factorial designs for microarray experiments under the baseline parametrization. Furthermore, Kerr [22] as well as Grossmann and Schwabe [17] have derived efficient designs for the estimation of main effects and two way interaction effects when all factors have two levels. Another reference is Stanzel and Hilgers [38], who give approximate designs for the estimation of two-factor interactions. However, in medical applications scientists are often interested in many factors. For instance, Taylor et al. [39]

and Stamatakis et al. [36] have investigated scenarios with more than two factors. They are interested in three-way interactions. Therefore, the interesting question of constructing approximate optimal designs for estimating interaction effects in multi-factorial settings arises and will be considered in this chapter. We will extend the investigations of Stanzel and Hilgers [38], who focused on two factors of interest and all pairwise treatment comparisons, by examining further contrast sets including treatment-control comparisons, all-to-next contrasts and Helmert contrasts. In addition, we consider multi-factorial layouts with more than two factors of interest in Section 4.4.

4.1 Generalized statistical model

In this section, we will extend model (2.5) such that it can be applied to experiments with n experimental factors of interest. Accordingly we can describe the vector of all observed log ratios of the dye intensities of each array $z = (z_1, \ldots, z_a)$ as

$$z = X\tau + W\delta + \eta \tag{4.1}$$

where $\tau = (\tau_{11\ldots1}, \ldots, \tau_{11\ldots k_n}; \ldots; \tau_{k_1 k_2 \ldots k_{n-1} 1}, \ldots, \tau_{k_1 k_2 \ldots k_n})$ is the vector of all the effects of factor level combinations of the n factors of interest, k_i denotes the number of factor levels of factor i, and $\delta = (\delta_{\text{green}}, \delta_{\text{red}})$ are the dye effects. $[X \mid W]$ is the design matrix, where each row of X consists of exactly one 1 and one -1, whereas all other entries are equal to zero. W is equal to $(\mathbb{1}_a, -\mathbb{1}_a)$ and η is the random error vector. For $n = 1$ model (4.1) yields model 2.5, for $n = 2$ it yields the model considered in Stanzel [37]. Again, we will ignore the dyes in the the following sections, but all theorems can be shown with dye effect analogously.

4.2 Main effect contrasts in two- and multi-factorial settings

Throughout this section model (4.1) is considered. First, we derive optimal experimental designs for the estimation of the main effect contrasts in the two-

factorial setting, i.e. $n = 2$ experimental factors of interest. If two factors with k_1 and k_2 levels are considered and one likes to estimate arbitrary contrasts $C^T(\tau_1, \ldots, \tau_{k_2})^T$ of the k_2 levels of the second factor in all of the k_1 levels of the first factor, then the contrasts $(\mathbb{1}_{k_1} \otimes C)^T(\tau_{11}, \ldots, \tau_{k_1k_2})^T$ are of interest. The following theorem is valid for all contrast sets. We assume that an optimal design in the one-factorial setting for the estimation of the contrasts set $C^T(\tau_1, \ldots, \tau_{k_2})^T$ is known for a factor with k_2 factor levels in the one-factorial setting.

Theorem 4.1:
Suppose that a design ξ with $l \times k_2$ design matrix $\grave{X}$ and $l \times l$ weight matrix P is ϕ_{-p}-optimal, $p \in (-\infty, 1]$, for the estimation of the contrast set $C^T(\tau_1, \ldots, \tau_{k_2})^T$ in the one-factorial model (4.1) with $n = 1$. Then, the design with design matrix $\tilde{X} = I_{k_1} \otimes \grave{X}$ and weight matrix $\tilde{P} = \frac{1}{k_1}(I_{k_1} \otimes P)$ is ϕ_{-p}-optimal in the two-factorial model (4.1) with $n = 2$ for the estimation of the contrast set $\tilde{C}^T\tau$ with $\tilde{C} = \mathbb{1}_{k_1} \otimes C$ and $\tau = (\tau_{11}, \ldots, \tau_{k_1k_2})^T$.

Proof: Choose $p \in (-\infty, 1]$ fix. Due to Theorem 2.3 and the ϕ_{-p}-optimality of X for the estimation of the contrast set $C^T(\tau_1, \ldots, \tau_{k_2})^T$ we know that a generalized inverse matrix $G = (\grave{X}^T P \grave{X})^-$ exists, which fulfills the inequalities

$$x^T GC \left(C^T GC\right)^+ \left(C^T GC\right)^{1-p} \left(C^T GC\right)^+ C^T G^T x \leq \operatorname{Tr}\left(\left(C^T GC\right)^+ \left(C^T GC\right)^{1-p}\right)$$

for all $x \in \mathcal{X} = \{x \in \{-1, 0, 1\}^{k_2} : \exists! i \text{ with } x_i = 1 \wedge \exists! j \text{ with } x_j = -1\}$. Define

$$A := GC \left(C^T GC\right)^+ \left(C^T GC\right)^{1-p} \left(C^T GC\right)^+ C^T G^T, \tag{4.2}$$

$$\text{const} := \operatorname{Tr}\left(\left(C^T GC\right)^+ \left(C^T GC\right)^{1-p}\right). \tag{4.3}$$

Furthermore, we will use the following equations which can be proved easily:

$$\tilde{G} = \left(\tilde{X}^T \tilde{P} \tilde{X}\right)^- = \left(\frac{1}{k_1}\left(I_{k_1} \otimes \left(X^T P X\right)\right)\right)^-, \tag{4.4}$$

$$= k_1 \left(I_{k_1} \otimes \left(X^T P X\right)^-\right) = k_1 \left(I_{k_1} \otimes G\right), \tag{4.5}$$

$$\tilde{G}\tilde{C} = k_1 \left(\mathbb{1}_{k_1} \otimes GC\right), \tag{4.6}$$

$$\tilde{C}^T \tilde{G} \tilde{C} = \left(\mathbb{1}_{k_1}^T \otimes C^T\right)\left(k_1 \left(I_{k_1} \otimes G\right)\right)\left(\mathbb{1}_{k_1} \otimes C\right) = k_1^2 \left(C^T G C\right), \tag{4.7}$$

$$\left(\tilde{C}^T \tilde{G} \tilde{C}\right)^+ = \frac{1}{k_1^2}\left(C^T G C\right)^+, \tag{4.8}$$

$$\left(\tilde{C}^T \tilde{G} \tilde{C}\right)^{1-p} = k_1^{2-2p}\left(C^T G C\right)^{1-p}. \tag{4.9}$$

For all $y \in \mathcal{Y} = \{y \in \{-1,0,1\}^{k_1 k_2} : \exists! i \text{ with } y_i = 1 \wedge \exists! j \text{ with } y_j = -1\}$ it holds that

$$\begin{aligned} & y^T \tilde{G}\tilde{C}\left(\tilde{C}^T\tilde{G}\tilde{C}\right)^+\left(\tilde{C}^T\tilde{G}\tilde{C}\right)^{1-p}\left(\tilde{C}^T\tilde{G}\tilde{C}\right)^+\tilde{C}^T\tilde{G}^T y \\ & \qquad \leq Tr\left(\left(\tilde{C}^T\tilde{G}\tilde{C}\right)^+\left(\tilde{C}^T\tilde{G}\tilde{C}\right)^{1-p}\right) \\ \Leftrightarrow\ & y^T\left(J_{k_1} \otimes \left(GC\left(C^TGC\right)^+\left(C^TGC\right)^{1-p}\left(C^TGC\right)^+ C^T G^T\right)\right) y \\ & \qquad \leq Tr\left(\left(C^TGC\right)^+\left(C^TGC\right)^{1-p}\right). \end{aligned}$$

From the optimality of the design ξ we know that $x^T A x \leq$ const for all $x \in \mathcal{X} = \{x \in \{-1,0,1\}^{k_2} : \exists! i \text{ with } x_i = 1 \wedge \exists! j \text{ with } x_j = -1\}$. Therefore, we have to show $y^T\left(J_{k_1} \otimes A\right) y \leq$ const for all $y \in \mathcal{Y} = \{y \in \{-1,0,1\}^{k_1 k_2} : \exists! i \text{ with } y_i = 1 \wedge \exists! j \text{ with } y_j = -1\}$. Without loss of generality, let $x_i = 1$ and $x_j = -1$ and thus

$$x^T A x = (a_{i1} - a_{j1}, \ldots, a_{ik_2} - a_{jk_2})x = a_{ii} - a_{ji} - (a_{ij} - a_{jj}),$$

where a_{ij} is the element in row i and column j in the matrix A. We partition $y^T = (y_1^T; \ldots; y_{k_1}^T) = (y_{11}, \ldots, y_{1k_2}; \ldots; y_{k_1 1}, \ldots, y_{k_1 k_2})$ with $y_{hi} = 1$ and $y_{lj} = -1$ without loss of generality. We consider three different cases. Firstly, suppose

$h \neq l$ and $i \neq j$, then

$$\begin{aligned} y^T (J_{k_1} \otimes A)\, y &= \left(\mathbb{1}_{k_1}^T \otimes (a_{i1} - a_{j1}, \ldots, a_{ik_2} - a_{jk_2})\right) y \\ &= a_{ii} - a_{ji} - (a_{ij} - a_{jj}) = x^T A x. \end{aligned}$$

The same result can be shown analogously for $h = l$ and $i \neq j$. Assuming $h = l$ and $i = j$ we get the following inequality $y^T (J_{k_1} \otimes A)\, y = 0 \leq x^T A x$. Therefore, the inequalities in Theorem 2.3 hold for the design with design matrix $\tilde{X}$ and weight matrix $\tilde{P}$ for the estimation of the main effect contrasts. □

Theorem 4.1 deals with two-factorial models. However, it can be applied to multi-factorial settings easily, since $\mathbb{1}_{k_1} \otimes \mathbb{1}_{k_2} \otimes \cdots \otimes \mathbb{1}_{k_{n-1}} \otimes C = \mathbb{1}_{k_1 \cdots k_{n-1}} \otimes C$.

4.3 Interaction effect contrasts in the two-factorial setting

In addition to the estimation of main effects in model (4.1) biologists are often interested in interactions between two or more factors. For example, considering k_1 cell lines and an arbitrary contrast set C of the k_2 treatments, we can specify the two-factor interaction effect contrasts $P_{k_1}^T \otimes C$ with the matrix P_{k_1} defined as in Section 2.3.3.

Theorem 4.2:
Suppose that the design ξ with the $l \times k_2$ design matrix $\grave{X}$ and the $l \times l$ weight matrix P is ϕ_{-p}-optimal, $p \in (-\infty, 1]$, for the estimation of the contrast set $C^T(\tau_1, \ldots, \tau_{k_2})^T$ in the one factorial model (4.1). Then, the design with design matrix $\tilde{X} = I_{k_1} \otimes \grave{X}$ and weight matrix $\tilde{P} = \frac{1}{k_1}(I_{k_1} \otimes P)$ is the ϕ_{-p}-optimal design in the two-factorial model (4.1) for the estimation of the contrast set $\tilde{C}^T\tau$ with $\tilde{C} = P_{k_1}^T \otimes C$, if $a_{ij} \leq 0$, $i \neq j$ and $a_{ii} \leq \frac{k_1 - 1}{2k_1}$const with const $=$ $\mathrm{Tr}((C^T G C)^+ (C^T G C)^{1-p})$ and $A = GC(C^T G C)^+ (C^T G C)^{1-p} (C^T G C)^+ C^T G^T$.

Proof: Choose $p \in (-\infty, 1]$ fix. We will use the equivalence theorems 2.3 to show optimality of the design with design matrix $\tilde{X}$ and weight matrix $\tilde{P}$.

Therefore, we will consider a generalized inverse $G = (\grave{X}^T P \grave{X})^-$, which fulfills the inequalities

$$x^T GC \left(C^T GC\right)^+ \left(C^T GC\right)^{1-p} \left(C^T GC\right)^+ C^T G^T x \leq \mathrm{Tr}\left(\left(C^T GC\right)^+ \left(C^T GC\right)^{1-p}\right)$$

for all $x \in \mathcal{X} = \{x \in \{-1, 0, 1\}^{k_2} : \exists! i \text{ with } x_i = 1 \wedge \exists! j \text{ with } x_j = -1\}$. Analogously to the proof of Theorem 4.1 we can show

$$\tilde{G} = \left(\tilde{X}^T \tilde{P} \tilde{X}\right)^- = k_1 \left(I_{k_1} \otimes G\right), \tag{4.10}$$

$$\tilde{G}\tilde{C} = k_1 (I_{k_1} \otimes G)(P_{k_1}^T \otimes C) = k_1 (P_{k_1}^T \otimes GC), \tag{4.11}$$

$$\tilde{C}^T \tilde{G} \tilde{C} = (P_{k_1} \otimes C^T) k_1 (I_{k_1} \otimes G)(P_{k_1}^T \otimes C) = k_1 (P_{k_1} P_{k_1}^T \otimes C^T GC) \tag{4.12}$$

$$(\tilde{C}^T \tilde{G} \tilde{C})^+ = \frac{1}{k_1} ((P_{k_1} P_{k_1}^T)^+ \otimes (C^T GC)^+), \tag{4.13}$$

$$(\tilde{C}^T \tilde{G} \tilde{C})^{1-p} = {k_1}^{1-p} ((P_{k_1} P_{k_1}^T)^{1-p} \otimes (C^T GC)^{1-p}). \tag{4.14}$$

Since we know from Section 2.3.3 that $(P_{k_1}^T P_{k_1}) = k_1 I_{k_1} - J_{k_1}$ and $(P_{k_1} P_{k_1}^T)^q = {k_1}^{q-1} P_{k_1} P_{k_1}^T$ for $q > 0$, we can demonstrate

$$\begin{aligned}
&\tilde{G}\tilde{C}(\tilde{C}^T \tilde{G} \tilde{C})^+ (\tilde{C}^T \tilde{G} \tilde{C})^{1-p} (\tilde{C}^T \tilde{G} \tilde{C})^+ \tilde{C}^T \tilde{G}^T \\
&= ((P_{k_1}^T (P_{k_1} P_{k_1}^T)^+ (P_{k_1} P_{k_1}^T)^{1-p} (P_{k_1} P_{k_1}^T)^+ P_{k_1}) \\
&\qquad \otimes GC(C^T GC)^+ (C^T GC)^{1-p} (C^T GC)^+ C^T G^T) {k_1}^{1-p} \\
&= (P_{k_1}^T (P_{k_1} P_{k_1}^T)^+ {k_1}^{-p} (P_{k_1} P_{k_1}^T)(P_{k_1} P_{k_1}^T)^+ P_{k_1} \otimes A) {k_1}^{1-p} \\
&= (k_1{}^{-p} P_{k_1}^T (P_{k_1} P_{k_1}^T)^+ P_{k_1} \otimes A) {k_1}^{1-p} = (k_1{}^{-p} \frac{1}{{k_1}^2} P_{k_1}^T (P_{k_1} P_{k_1}^T) P_{k_1} \otimes A) {k_1}^{1-p} \\
&= (k_1{}^{-p} \frac{1}{{k_1}^2} (k_1 I_{k_1} - J_{k_1})^2 \otimes A) {k_1}^{1-p} = (k_1{}^{-p-1} (k_1 I_{k_1} - J_{k_1}) \otimes A) {k_1}^{1-p} \\
&= ((k_1 I_{k_1} - J_{k_1}) \otimes A) {k_1}^{-2p}.
\end{aligned}$$

Using $\mathrm{Tr}((P_{k_1}P_{k_1}^T)^+(P_{k_1}P_{k_1}^T)) = k_1 - 1$ from Section 2.3.3 we see

$$\begin{aligned}
&\mathrm{Tr}((\tilde{C}^T\tilde{G}\tilde{C})^+(\tilde{C}^T\tilde{G}\tilde{C})^{1-p}) \\
&= \mathrm{Tr}((P_{k_1}P_{k_1}^T)^+(P_{k_1}P_{k_1}^T)^{1-p} \otimes (C^TGC)^+(C^TGC)^{1-p}){k_1}^{-p} \\
&= \mathrm{Tr}(((P_{k_1}P_{k_1}^T)^+(P_{k_1}P_{k_1}^T)^{1-p})\,\mathrm{Tr}((C^TGC)^+(C^TGC)^{1-p}){k_1}^{-p} \\
&= \mathrm{Tr}((P_{k_1}P_{k_1}^T)^+{k_1}^{-p}(P_{k_1}P_{k_1}^T))\,\mathrm{const}\;\; {k_1}^{-p} \\
&= {k_1}^{-2p}(k_1-1)\,\mathrm{const}.
\end{aligned}$$

Due to the ϕ_{-p}-optimality of the design ξ, we know that $x^TAx \leq \mathrm{const}$ for all $x \in \mathcal{X} = \{x \in \{-1,0,1\}^{k_2} : \exists! i \text{ with } x_i = 1 \wedge \exists! j \text{ with } x_j = -1\}$. Thus, we have to show

$$y^T\left((k_1 I_{k_1} - J_{k_1}) \otimes A\right) y \leq (k_1 - 1)\,\mathrm{const}$$

for all $y \in \mathcal{Y} = \{y \in \{-1,0,1\}^{k_1k_2} : \exists! i \text{ with } y_i = 1 \wedge \exists! j \text{ with } y_j = -1\}$. Without loss of generality, we assume $x_i = 1$ and $x_j = -1$ and thus $x^TAx = a_{ii} - a_{ji} - (a_{ij} - a_{jj})$ as in the proof of Theorem 4.1. We partition $y^T = (y_1^T; \ldots; y_{k_1}^T) = (y_{11}, \ldots, y_{1k_2}; \ldots; y_{k_11}, \ldots, y_{k_1k_2})$ with $y_{hi} = 1$ and $y_{lj} = -1$ and consider three different cases. Firstly, suppose $h \neq l$, $i \neq j$:

$$\begin{aligned}
&y^T\left((k_1 I_{k_1} - J_{k_1}) \otimes A\right) y \\
&= (\underbrace{-a_{i1} + a_{j1}, \ldots, -a_{ik_2} + a_{jk_2}}_{\text{1. cell line}}, \underbrace{-a_{i1} + a_{j1}, \ldots, -a_{ik_2} + a_{jk_2}}_{\text{2. cell line}}, \ldots, \\
&\quad \underbrace{(k_1-1)a_{i1} + a_{j1}, \ldots, (k_1-1)a_{ik_2} + a_{jk_2}}_{h.\text{ cell line}}, \ldots, \\
&\quad \underbrace{-a_{i1} - (k_1-1)a_{j1}, \ldots, -a_{ik_2} - (k_1-1)a_{jk_2}}_{l.\text{ cell line}}, \ldots)y \\
&= (k_1-1)a_{ii} + a_{ji} - (-a_{ij} - (k_1-1)a_{jj}) \\
&= (k_1-1)(a_{ii} + a_{jj}) + 2a_{ij} \\
&= (k_1-1)(a_{ii} + a_{jj} - 2a_{ij}) + 2(k_1-1)a_{ij} + 2a_{ij} \\
&\leq (k_1-1)\ \mathrm{const} + 2k_1a_{ij} \\
&\leq (k_1-1)\ \mathrm{const}\ .
\end{aligned}$$

The last inequality holds due to the assumption $a_{ij} \leq 0$, for all $i \neq j$. Secondly,

we suppose $h \neq l$, $i = j$ and use the fact that $a_{ii} \leq \frac{k_1-1}{2k_1}$const. Therefore,

$$\begin{aligned}
&y^T \left((k_1 I_{k_1} - J_{k_1}) \otimes A\right) y \\
&= (k_1 - 1)a_{ii} + a_{ii} - (-a_{ii} - (k_1 - 1)a_{ii}) = 2k_1 a_{ii} \\
&\leq (k_1 - 1) \text{ const.}
\end{aligned}$$

Assuming $h = l$, $i \neq j$ we finally get

$$\begin{aligned}
&y^T \left((k_1 I_{k_1} - J_{k_1}) \otimes A\right) y \\
&= (\underbrace{-a_{i1} + a_{j1}, \ldots, -a_{ik_2} + a_{jk_2}}_{\text{1. cell line}}, \ldots, \\
&\quad \underbrace{(k_1 - 1)a_{i1} - (k_1 - 1)a_{j1}, \ldots, (k_1 - 1)a_{ik_2} - (k_1 - 1)a_{jk_2}}_{l. \text{ cell line}}, \ldots) y \\
&= (k_1 - 1)a_{ii} - (k_1 - 1)a_{ji} - ((k_1 - 1)a_{ij} - (k_1 - 1)a_{jj}) \\
&= (k_1 - 1)(a_{ii} + a_{jj} - 2a_{ij}) \leq (k_1 - 1)\text{const.}
\end{aligned}$$

This completes the proof. □

Example 4.3*:*

If we are interested in estimating the specific interaction contrasts $\tilde{C} = P_{k_1}^T \otimes P_{k_2}^T$ of k_1 cell lines and k_2 treatments, i.e. $C = P_{k_2}^T$ in Theorem 4.2, where $P_{k_2}^T$ defines all pairwise treatment comparisons, the inequalities in Theorem 4.2 reduce to $k_2 \leq k_1$. We will show this in the following part. Because of Theorem 3.4 we know that the design with design matrix $\grave{X} = P_{k_2}$ and equal weights is ϕ_{-p}-optimal for estimating the contrasts $C^T\tau$ with $C = P_{k_2}^T$. Thus, the Moore-Penrose-Inverse $G = (\grave{X}^T P \grave{X})^+$ fulfills the normality inequalities of the equiva-

lence Theorem 2.3 for $p \in (-\infty, 1]$ with $P = \frac{1}{\binom{k_2}{2}} I_{\binom{k_2}{2}}$. Hence,

$$
\begin{aligned}
G = (\grave{X}^T P \grave{X})^+ &= \left(\frac{1}{\binom{k_2}{2}} (k_2 I_{k_2} - J_{k_2}) \right)^+ = \binom{k_2}{2} \left(\frac{1}{k_2} I_{k_2} - \frac{1}{{k_2}^2} J_{k_2}) \right), \\
C^T G C &= \frac{k_2 - 1}{2} P_{k_2} P_{k_2}^T, \\
(C^T G C)^+ &= \frac{2}{(k_2 - 1){k_2}^2} P_{k_2} P_{k_2}^T, \\
\text{const} &= \mathrm{Tr}((C^T G C)^+ (C^T G C)^{1-p}) = \mathrm{Tr}(\frac{2^p}{(k_2 - 1)^p {k_2}^{p+1}} (P_{k_2} P_{k_2}^T)) \\
&= \frac{2^p}{(k_2 - 1)^p {k_2}^{p+1}} \mathrm{Tr}((P_{k_2}^T P_{k_2})) = \frac{2^p}{(k_2 - 1)^p {k_2}^{p+1}} \mathrm{Tr}(k_2 I_{k_2} - J_{k_2}) \\
&= \frac{2^p}{(k_2 - 1)^{p-1} {k_2}^p},
\end{aligned}
$$

$$
\begin{aligned}
A &= GC(C^T G C)^+ (C^T G C)^{1-p} (C^T G C)^+ C^T G^T \\
&= \frac{(k_2 - 1)^{1-p}}{2^{1-p} {k_2}^{p+4}} P_{k_2}^T (P_{k_2} P_{k_2}^T)^3 P_{k_2} \\
&= \frac{(k_2 - 1)^{1-p}}{2^{1-p} {k_2}^{p+1}} (k_2 I_{k_2} - J_{k_2}).
\end{aligned}
$$

Therefore, $a_{ij} = -\frac{(k_2-1)^{1-p}}{2^{1-p} {k_2}^{p+1}} \leq 0$ for all $i \neq j$ and $a_{ii} \leq \frac{k_1 - 1}{2k_1} \text{const} \Leftrightarrow k_2 \leq k_1$, i.e. the design with design matrix $I_{k_1} \otimes P_{k_2}$ and equal weights is ϕ_{-p}-optimal for estimating $\tilde{C} = P_{k_1}^T \otimes P_{k_2}^T$, if there are less treatments k_2 than cell lines k_1.

4.4 Interaction effect contrasts in the multi-factorial settings for the estimation of all pairwise comparisons

Up to know we have considered contrasts of the main effects in the two- and the multi-factorial model as well as contrasts of the interaction effects in the two-factorial model. In this section we will derive approximate optimal designs for the estimation of interaction effects in the multi-factorial setting, if we are interested in all pairwise comparisons of the factor levels of all n experimental factors under investigation. Assuming a three-factorial model, Hinkel-

mann et al. [20] have investigated interaction contrast $C^T(\tau_{111},\ldots,\tau_{k_1k_2k_3})^T$ with $C^T = P_{k_1} \otimes P_{k_2} \otimes P_{k_3}$, where k_1, k_2 and k_3 are the levels of the three factors of interest. In general, the interaction contrasts in a multi-factorial model with n experimental factors of interest are defined as $C^T\tau = (P_{k_1} \otimes P_{k_2} \otimes \ldots \otimes P_{k_n})(\tau_{11\ldots1},\ldots,\tau_{k_1k_2\ldots k_n})^T$. We can show the following theorem.

Theorem 4.4:

Consider the multi-factorial model (4.1) with $n \in \mathbb{N}$ experimental factors of interest. Let k_i, $i \in \{1,\ldots,n\}$ denote the number of levels of factor i and assume without loss of generality $2 \leq k_1 \leq \ldots \leq k_n$. The design with design matrix $X = P_{k_1} \otimes I_{k_2} \otimes \ldots \otimes I_{k_n}$ and weight matrix $P = \frac{1}{k_2\cdots k_n\binom{k_1}{2}} I_{k_2\cdots k_n\binom{k_1}{2}}$ is a ϕ_{-p}-optimal design ($p \in (-\infty, 1]$) for the estimation of the interaction contrasts $C^T\tau = (P_{k_1} \otimes P_{k_2} \otimes \ldots \otimes P_{k_n})\tau$ with $\tau = (\tau_{11\ldots1},\ldots,\tau_{k_1k_2\ldots k_n})^T$.

Proof: We will proof this statement per induction for fixed $p \in (-\infty, 1]$. Example 4.3 provides the basis of the induction for $n = 2$. We assume that the statement is true for arbitrary $l \leq n-1$, i.e. for the estimation of $C_l^T(\tau_{11\ldots1},\ldots,\tau_{k_1k_2\ldots k_l})^T = (P_{k_1} \otimes \ldots \otimes P_{k_l})(\tau_{11\ldots1},\ldots,\tau_{k_1k_2\ldots k_l})^T$ with $\min\{k_1,\ldots,k_l\} = k_1 \geq 2$ is the design with design matrix $X_l = P_{k_1} \otimes I_{k_2} \otimes \ldots \otimes I_{k_l}$ and with weight matrix $P^{(l)} = \frac{1}{k_2\cdots k_l\binom{k_1}{2}} I_{k_2\cdots k_l\binom{k_1}{2}}$ ϕ_{-p}-optimal for $l \leq n-1$. We will derive that the statement is also true for n factors and $k_n \geq k_1$.

$$
\begin{aligned}
G_n &= (X_n^T P^{(n)} X_n)^+ = \left(((X_{n-1}) \otimes I_{k_n})^T \frac{1}{k_n}(P^{n-1)} \otimes I_{k_n})((X_{n-1}) \otimes I_{k_n})\right)^+ \\
&= k_n(G_{n-1} \otimes I_{k_n}), \\
G_nC_n &= k_n(G_{n-1}C_{n-1} \otimes P_{k_n}^T), \\
C_n^TG_nC_n &= k_n(C_{n-1}^TG_{n-1}C_{n-1} \otimes P_{k_n}P_{k_n}^T), \\
E_n &:= G_nC_n(C_n^TG_nC_n)^+(C_n^TG_nC_n)^{1-p}(C_n^TG_nC_n)^+C_n^TG_n^T \\
&= k_n^{1-p}(E_{n-1} \otimes P_{k_n}^T(P_{k_n}P_{k_n}^T)^+(P_{k_n}P_{k_n}^T)^{1-p}(P_{k_n}P_{k_n}^T)^+P_{k_n}) \\
&= k_n^{1-p}(E_{n-1} \otimes k_n^{-p}(I_{k_n} - \frac{1}{k_n}J_{k_n})) = k_n^{1-2p}(E_{n-1} \otimes (I_{k_n} - \frac{1}{k_n}J_{k_n})) \\
&= 2^{p-1}k_1^{-p}(k_1-1)^{1-p}\prod_{i=2}^{n} k_i^{1-2p}((I_{k_1} - \frac{1}{k_1}J_{k_1}) \otimes \ldots \otimes (I_{k_n} - \frac{1}{k_n}J_{k_n})).
\end{aligned}
$$

Thereby we use the fact from Example 4.3 and the proof of Theorem 4.2 that

$$E_2 = 2^{p-1}k_2^{1-2p}k_1^{-p}(k_1-1)^{1-p}((I_{k_1} - \frac{1}{k_1}J_{k_1}) \otimes (I_{k_2} - \frac{1}{k_2}J_{k_2})).$$

Furthermore, we get

$$\begin{aligned}
\text{const}_n &:= \text{Tr}((C_n^T G_n C_n)^+ (C_n^T G_n C_n)^{1-p}) \\
&= \text{Tr}((C_{n-1}^T G_{n-1} C_{n-1})^+ (C_{n-1}^T G_{n-1} C_{n-1})^{1-p} \otimes (P_{k_n} P_{k_n}^T)^+ (P_{k_n} P_{k_n}^T)^{1-p}) k_n^{-p} \\
&= \text{const}_{n-1} \text{Tr}(k_n^{-p-1}(P_{k_n} P_{k_n}^T)) k_n^{-p} \\
&= k_n^{-2p-1} \text{const}_{n-1} 2\frac{k_n(k_n-1)}{2} = k_n^{-2p}(k_n-1)\text{const}_{n-1} \\
&= 2^p k_1^{-p}(k_1-1)^{1-p} \prod_{i=2}^{n} k_i^{-2_p}(k_i-1) \\
&= \left(\frac{k_1-1}{2k_1}\right)^{-p} \prod_{i=1}^{n} k_i^{-2p}(k_i-1),
\end{aligned}$$

since we know from the proof of Theorem 4.2, Example 4.3, and $k_1 \leq k_2$ that

$$\text{const}_2 = k_2^{-2p}(k_2-1)\text{const}_1 = 2^p k_1^{-p}(k_1-1)^{1-p}(k_2-1)k_2^{-2_p}.$$

Knowing $x^T E_{n-1} x \leq \text{const}_{n-1}$ for all $x \in \mathcal{X}_{n-1} = \{x \in \{-1,0,1\}^{k_1 \ldots k_{n-1}} | \exists! i : x_i = -1 \wedge \exists! j : x_j = 1\}$ we have to show $y^T E_n y \leq \text{const}_n$ for all $y \in \mathcal{X}_n = \{y \in \{-1,0,1\}^{k_1 \ldots k_n} | \exists! i : y_i = -1 \wedge \exists! j : y_j = 1\}$.

$$\begin{aligned}
& y^T E_n y \leq \text{const}_n \\
\Leftrightarrow\ & y^T k_n^{1-2p}(E_{n-1} \otimes (I_{k_n} - \frac{1}{k_n}J_{k_n}))y \leq k_n^{-2p}(k_n-1)\text{const}_{n-1} \\
\Leftrightarrow\ & y^T (E_{n-1} \otimes (I_{k_n} - \frac{1}{k_n}J_{k_n}))y k_n \leq (k_n-1)\text{const}_{n-1}
\end{aligned}$$

Assume $1 \leq h, l \leq k_1 k_2 \cdots k_{n-1}$ and $1 \leq i, j \leq k_n$ and partition

$$y^T = (y_1^T; \ldots; y_{k_1 k_2 \ldots k_{n-1}}^T) = (y_{11}, \ldots, y_{1k_n}; \ldots; y_{k_1 k_2 \ldots k_{n-1} 1}, \ldots, y_{k_1 k_2 \ldots k_{n-1} k_n}).$$

The inequality $x^T E_{n-1} x \leq \text{const}_{n-1}$ for $x \in \mathcal{X}_{n-1}$ with $x_h = 1$ and $x_l = -1$ is

equivalent to

$$e_{hh}^{(n-1)} + e_{ll}^{(n-1)} - e_{hl}^{(n-1)} - e_{lh}^{(n-1)} \leq \text{const}_{n-1},$$

where $e_{ij}^{(n-1)}$ is the ij-th element of the matrix E_{n-1}.

We will distinguish three cases. Firstly, suppose $h = l$ and $i \neq j$, without loss of generality $i < j$ and $y_{hi} = 1$ as well as $y_{lj} = -1$. Thus,

$$\begin{aligned}
&y^T k_n (E_{n-1} \otimes (I_{k_n} - \frac{1}{k_n} J_{k_n}))y \\
&= (e_{h1}^{(n-1)}(-1) - e_{h1}^{(n-1)}(-1), \ldots, \underbrace{e_{h1}^{(n-1)}(k_n - 1) - e_{h1}^{(n-1)}(-1)}_{\text{i.th entry}}, \ldots, \\
&\qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots, \underbrace{e_{h1}^{(n-1)}(-1) - e_{h1}^{(n-1)}(k_n - 1)}_{\text{j.th entry}}, \ldots, \\
&\quad e_{h2}^{(n-1)}(-1) - e_{h2}^{(n-1)}(-1), \ldots, \underbrace{e_{h2}^{(n-1)}(k_n - 1) - e_{h2}^{(n-1)}(-1)}_{k_n + i.\text{th entry}}, \ldots, \\
&\qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots, \underbrace{e_{h2}^{(n-1)}(-1) - e_{h2}^{(n-1)}(k_n - 1)}_{k_n + j.\text{th entry}}, \ldots)y \\
&= 2k_n e_{hh}^{(n-1)}
\end{aligned}$$

Due to the structure of E_n we know

$$\begin{aligned}
e_{hh}^{(n)} &= 2^{p-1} k_1^{-p} (k_1 - 1)^{1-p} \frac{(k_1 - 1)}{k_1} \prod_{i=2}^{n} k_i^{-2p} (k_i - 1) \\
&= \left(\frac{k_1 - 1}{2k_1} \right)^{1-p} \prod_{i=1}^{n} k_i^{-2p} (k_i - 1).
\end{aligned}$$

Using all results, we can show that $y^T E_n y \leq \text{const}_n$ is equivalent to

$$\begin{aligned}
&2k_n \left(\frac{k_1 - 1}{2k_1} \right)^{1-p} \prod_{i=1}^{n-1} k_i^{-2p} (k_i - 1) \\
&\leq (\left(\frac{k_1 - 1}{2k_1} \right)^{-p} \prod_{i=1}^{n-1} k_i^{-2p} (k_i - 1))(k_n - 1) \\
&\Leftrightarrow k_n (k_1 - 1) \leq k_1 (k_n - 1) \\
&\Leftrightarrow k_1 \leq k_n.
\end{aligned}$$

This inequality $y^T E_n y \leq \text{const}_n$ is fulfilled due to our assumption $k_1 \leq k_n$.

Secondly, we assume $h \neq l$ and $i = j$.

$$\begin{aligned}
&y^T k_n (E_{n-1} \otimes (I_{k_n} - \frac{1}{k_n} J_{k_n}))y \\
&= (e_{h1}^{(n-1)}(-1) - e_{l1}^{(n-1)}(-1), \ldots, \underbrace{e_{h1}^{(n-1)}(k_n - 1) - e_{l1}^{(n-1)}(k_n - 1)}_{\text{i.th entry}}, \ldots, \\
&\qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots, e_{h1}^{(n-1)}(-1) - e_{l1}^{(n-1)}(-1), \ldots, \\
&\quad e_{h2}^{(n-1)}(-1) - e_{l2}^{(n-1)}(-1), \ldots, \underbrace{e_{h2}^{(n-1)}(k_n - 1) - e_{l2}^{(n-1)}(k_n - 1)}_{k_n + i.\text{th entry}}, \ldots, \\
&\qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots, e_{h2}^{(n-1)}(-1) - e_{l2}^{(n-1)}(-1), \ldots)y \\
&= (k_n - 1)(e_{hh}^{(n-1)} + e_{ll}^{(n-1)} - 2e_{hl}^{(n-1)}) \leq (k_n - 1)\text{const}_{n-1}.
\end{aligned}$$

And last we suppose $h \neq l$ and $i \neq j$.

$$\begin{aligned}
&y^T k_n (E_{n-1} \otimes (I_{k_n} - \frac{1}{k_n} J_{k_n}))y \\
&= (e_{h1}^{(n-1)}(-1) - e_{l1}^{(n-1)}(-1), \ldots, \underbrace{e_{h1}^{(n-1)}(k_n - 1) - e_{l1}^{(n-1)}(-1)}_{\text{i.th entry}}, \ldots, \\
&\qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots, \underbrace{e_{h1}^{(n-1)}(-1) - e_{l1}^{(n-1)}(k_n - 1)}_{\text{j.th entry}}, \ldots, \\
&\quad e_{h2}^{(n-1)}(-1) - e_{l2}^{(n-1)}(-1), \ldots, \underbrace{e_{h2}^{(n-1)}(k_n - 1) - e_{l2}^{(n-1)}(-1)}_{k_n + i.\text{th entry}}, \ldots, \\
&\qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots, \underbrace{e_{h2}^{(n-1)}(-1) - e_{l2}^{(n-1)}(k_n - 1)}_{k_n + j.\text{th entry}}, \ldots)y \\
&= (k_n - 1)(e_{hh}^{(n-1)} + e_{ll}^{(n-1)}) + 2e_{lh}^{(n-1)} \\
&= (k_n - 1)(e_{hh}^{(n-1)} + e_{ll}^{(n-1)} - 2e_{hl}^{(n-1)}) + 2k_n e_{hl}^{(n-1)} \\
&\leq (k_n - 1)\text{const}_{n-1} + 2k_n e_{hl}^{(n-1)}.
\end{aligned}$$

Due to the structure of E_{n-1} we know for $h \neq l$

$$e_{lh}^{(n-1)} \in \left\{ \frac{(\prod_{i=2}^{n} k_i^{1-2_p}) \cdot (k_1-1)^{q_1}(k_2-1)^{q_2}\dots(k_{n-1}-1)^{q_{n-1}}(-1)^{q_0}}{2^{1-p}k_1^p(k_1-1)^{p-1}\cdot k_1k_2\dots k_{n-1}} \right.$$
$$\left. \text{for } i \in \{1,\dots,n-1\} \text{ is } q_i \in \{0,1\} \text{ and } q_0 = (n-1) - \sum_{i=1}^{n-1} q_i \right\},$$
$$e_{hh}^{(n-1)} = \left(\frac{k_1-1}{2k_1}\right)^{1-p} \prod_{i=1}^{n-1} k_i^{-2p}(k_i-1).$$

Obviously, $e_{hl}^{(n-1)} \leq 0$ for q_0 odd. Hence let q_0 be even. Define

$$h_n(k_1,\dots,k_n) := 2^{p-1}k_1^{-p}(k_1-1)^{1-p}(\prod_{i=2}^{n} k_i^{1-2_p})$$

and therefore $\text{const}_{n-1} = 2h_{n-1}(k_1,\dots,k_{n-1})(\prod_{i=2}^{n-1}\frac{k_i-1}{k_i})$. We have to prove the following inequality.

$$\frac{2(k_n-1)h_{n-1}(k_1,\dots,k_{n-1})(k_1-1)(k_2-1)\cdots(k_{n-1}-1)}{k_1k_2\cdots k_{n-1}} +$$
$$\frac{2h_{n-1}(k_1,\dots,k_{n-1})(k_1-1)^{q_1}\cdots(k_{n-1}-1)^{q_{n-1}}(-1)^{q_0}}{k_1k_2\cdots k_{n-1}}$$
$$\leq 2(k_n-1)h_{n-1}(k_1,\dots,k_{n-1})(\prod_{i=2}^{n-1}\frac{k_i-1}{k_i})\frac{k_1}{k_1}$$
$$\Leftrightarrow (k_1-1)(k_2-1)\dots(k_{n-1}-1)(k_n-1) +$$
$$(k_1-1)^{q_1}(k_2-1)^{q_2}\cdots(k_{n-1}-1)^{q_{n-1}}(-1)^{q_0}$$
$$\leq k_1(k_2-1)\cdots(k_{n-1}-1)(k_n-1)$$
$$\Leftrightarrow (k_1-1)^{q_1}(k_2-1)^{q_2}\cdots(k_{n-1}-1)^{q_{n-1}}(-1)^{q_0} \leq (k_2-1)(k_3-1)\cdots(k_n-1)$$
$$\Leftrightarrow \frac{(k_2-1)^{q_2}}{k_2-1}\frac{(k_3-1)^{q_3}}{k_3-1}\cdots\frac{(k_{n-1}-1)^{q_{n-1}}}{(k_{n-1}-1)}\cdot 1\cdot(k_1-1)^{q_1} \leq (k_n-1).$$

This inequality is fulfilled for $k_n \geq k_1 \geq 2$ and for all $q_i \in \{0,1\}$ for $i \in \{1,\dots,n-1\}$ and $q_0 = n - \sum_{i=1}^{n} q_i$, since $\frac{(k_i-1)^{q_i}}{k_i-1} \leq 1$. □

At this point we know the ϕ_{-p}-optimal designs for the estimation of interaction effects, that are associated with studying all pairwise comparisons of the factor levels of all the experimental factors examined in a multi-factorial setting. In the next sections we will turn to further contrast sets.

4.5 Comparisons with a control-treatment

A very important contrast set for biological applications is the set of all treatment-control comparisons introduced in Section 3.2, where exact A-optimal designs were derived. In the following subsections, we will use the method explained in Section 3.1 to derive approximate A-optimal designs for the treatment-control comparisons. We will consider a setting with one factor of interest as well as a multi-factorial layout. To be able to apply Theorem 4.1 and Theorem 4.2, an approximate optimal design associated with the one-factorial model should be known. Therefore, we will first construct A-optimal designs for estimating all treatment-control comparisons in the one-factorial setting in the next subsection.

4.5.1 A-optimal designs in the one-factorial setting

The following construction of A-optimal approximate designs is based on the gene-specific model (2.5), where the dye effect is removed. The dye effect is reintroduced in Section 4.7. Again we denote the control-treatment by τ_0 and the other treatments by τ_i, $i \in \{1, \ldots, t\}$.

The derivation of the approximate A-optimal designs is also based on the results presented in Section 3.1, i.e. we will use Theorem 3.1 to minimize $\sum_{i=1}^{t} \mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_i)$, but, in contrast to Chapter 3 we will consider continuous values $\frac{b(x_{ij})}{a} \in [0, 1]$. The value $\frac{b(x_{ij})}{a}$ is denoted by $\tilde{B}_{ij}^{(t)}$, if we consider t treatments and one control-treatment. $\tilde{B}_{ij}^{(t)}$ describes the rate of microarrays hybridizing treatments i and j. Therefore, we will minimize

$$\sum_{i=1}^{t} \mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_i)$$

under the constraints

$$\tilde{B}_{01}^{(t)} + \tilde{B}_{02}^{(t)} + \cdots + \tilde{B}_{(t-1)t}^{(t)} = 1,$$

$$\tilde{B}_{ij}^{(t)} \in [0, 1] \text{ for all } 0 \leq i < j \leq t.$$

For $t = 3$ and $1 \leq i, j \leq t$, $j \neq i$ we obtain the weights $\tilde{B}_{0i}^{(t)} = \frac{1}{4}$ and $\tilde{B}_{ij}^{(t)} = \frac{1}{12}$.

For $t \in \mathbb{N}^{\geq 4} \cup \{2\}$ the weights can be expressed as

$$\tilde{B}_{0i}^{(t)} = \frac{2((t-1)\sqrt{t+1} - (t+1))}{t(t+1)(t-3)}, \tag{4.15}$$

$$\tilde{B}_{ij}^{(t)} = \frac{2(t - 2\sqrt{t+1} + 1)}{t(t+1)(t-3)}. \tag{4.16}$$

We define $r_1^{(t)} := \tilde{B}_{0i}^{(t)}$ and $r_2^{(t)} := \tilde{B}_{ij}^{(t)}$ since these weights are independent on i and j. Using the equivalence theorem 2.3 we show the A-optimality of the designs that are constructed according to these weights.

Theorem 4.5:

Approximate A-optimal designs for the estimation of all treatment-control comparisons $\tau_0 - \tau_i$, $i \in \{1, \ldots, t\}$ in the one-factorial setting allocate the control with each of the other treatment on the rate $r_1^{(t)}$ of all a microarrays, all other treatment comparisons are allocated on the rate $r_2^{(t)}$ of all a microarrays with $tr_1^{(t)} + \binom{t}{2} r_2^{(t)} = 1$. In other words, designs with design matrix $\grave{X} = P_{t+1}$ and weight matrix $P := \text{diag}(\underbrace{r_1^{(t)}, \ldots, r_1^{(t)}}_{t}, \underbrace{r_2^{(t)}, \ldots, r_2^{(t)}}_{\binom{t}{2}})$ are A-optimal.

Proof: Choose t fix and denote $r_1^{(t)} = r_1$ and $r_2^{(t)} = r_2$ throughout this proof. For $p = -1$ the ϕ_{-p}-optimality criterion is the same as the A-optimality-criterion. Therefore, we will show that the equivalence theorem 2.3 holds for $p = -1$. We show that the Moore-Penrose-Inverse $G = (\grave{X}^T P \grave{X})^+$ of $\grave{X}^T P \grave{X}$ satisfies the normality inequality

$$x^T GC \left(C^T GC\right)^+ \left(C^T GC\right)^2 \left(C^T GC\right)^+ C^T G^T x \leq Tr\left(\left(C^T GC\right)^+ \left(C^T GC\right)^2\right) \tag{4.17}$$

for all possible design points $x \in \mathcal{X}$, $\mathcal{X} = \{x \in \{-1, 0, 1\}^{t+1} : \exists! i \text{ with } x_i = 1 \wedge \exists! j \text{ with } x_j = -1\}$. $P := \text{diag}(\underbrace{r_1, \ldots, r_1}_{t}, \underbrace{r_2, \ldots, r_2}_{\binom{t}{2}})$ with $tr_1 + \binom{t}{2} r_2 = 1$ is the diagonal matrix containing the optimal weights for all design points listed in $\grave{X} = P_{t+1}$. Since we are interested in the comparisons with a control-treatment,

we consider the contrast matrix

$$C = \begin{pmatrix} 1 & 1 & \dots & 1 \\ -1 & 0 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \dots & -1 \end{pmatrix}.$$

It can be shown easily, that

$$\mathrm{X}^T P X = \left(\begin{array}{c|c} r_1 t & -r_1 \mathbb{1}_t^T \\ \hline -r_1 \mathbb{1}_t & (r_1 + t r_2) I_t - r_2 J_t \end{array}\right),$$

where J_t is the $t \times t$ matrix with all entries equal to 1 and I_t is the $t \times t$ identity matrix. Thus, we get

$$\begin{aligned} G = (X^T P X)^- &= \frac{1}{r_1(t+1)^2} \left(\begin{array}{c|c} t & -\mathbb{1}_t^T \\ \hline -\mathbb{1}_t & \frac{(t+1)^2 r_1}{r_1 + t r_2} I_t + \frac{-(t+2) r_1 + r_2}{r_1 + t r_2} J_t \end{array}\right), \\ \mathrm{C}^T G C &= \frac{1}{r_1 + t r_2} I_t + \frac{r_2}{r_1(r_1 + t r_2)} J_t, \\ (C^T G C)^+ &= (r_1 + t r_2) I_t - r_2 J_t, \\ G C &= \begin{pmatrix} \frac{1}{(t+1) r_1} \mathbb{1}_t^T \\ \\ \frac{1}{r_1 + t r_2} I_t + \frac{r_1 - r_2}{(t+1) r_1 (r_1 + t r_2)} J_t \end{pmatrix}. \end{aligned}$$

Furthermore, we can show

$$\mathrm{Tr}((C^T G C)^+ (C^T G C)^2) = \frac{(r_1 + r_2) t}{r_1(r_1 + t r_2)},$$

and

$$\begin{aligned} & G C \left(C^T G C\right)^+ \left(C^T G C\right)^2 \left(C^T G C\right)^+ C^T G^T \\ &= \left(\begin{array}{c|c} \frac{t}{r_1^2 (t+1)^2} & \frac{-1}{r_1^2 (t+1)^2} \mathbb{1}_t^T \\ \hline \frac{-1}{r_1^2 (t+1)^2} \mathbb{1}_t & \frac{1}{(r_1 + t r_2)^2} I_t + \frac{(r_2 - r_1)((t+2) r_1 + t r_2)}{r_1^2 (t+1)^2 (r_1 + t r_2)^2} J_t \end{array}\right). \end{aligned}$$

If we compare the control-treatment to the i-th treatment, we denote the design

point $x_{(0i)} = (1, 0, \ldots, 0, \underbrace{-1}_{i}, 0, \ldots, 0)$ and therewith

$$\begin{aligned}
&x_{(0i)}^T GC \left(C^T GC\right)^+ \left(C^T GC\right)^2 \left(C^T GC\right)^+ C^T G^T x_{(0i)} \\
&= \left(\frac{t}{(t+1)^2 r_1^2} + \frac{1}{(t+1)^2 r_1^2}\right) - \left(\frac{-1}{(t+1)^2 r_1^2} - \frac{r_1^2(t^2+t-1) + 2r_1 r_2 + t r_2^2}{(t+1)^2 r_1^2 (r_1 + t r_2)^2}\right) \\
&= \frac{1}{2}\left(\sqrt{t+1} - 1)^2 (\sqrt{t+1} + 2\right)^2 \\
&= \frac{(r_1 + r_2)t}{r_1(r_1 + t r_2)} \\
&= \mathrm{Tr}((C^T GC)^+ (C^T GC)^2).
\end{aligned}$$

Analogously we prove equality in (4.17) for $x_{(ij)} = (0, \ldots, 0, \underbrace{1}_{i}, 0, \ldots, 0, \underbrace{-1}_{j}, 0, \ldots, 0)$ for the comparisons of the i-th and j-th treatment.

$$\begin{aligned}
&x_{(ij)}^T GC \left(C^T GC\right)^+ \left(C^T GC\right)^2 \left(C^T GC\right)^+ C^T G^T x_{(ij)} \\
&= 2\left(\frac{r_1^2(t^2+t-1) + 2r_1 r_2 + t r_2^2}{(t+1)^2 r_1^2 (r_1 + t r_2)^2} - \frac{(r_2 - r_1)((t+2)r_1 + t r_2)}{r_1^2 (t+1)^2 (r_1 + t r_2)^2}\right) \\
&= \frac{1}{2}(\sqrt{t+1} - 1)^2 (\sqrt{t+1} + 2)^2 \\
&= \mathrm{Tr}((C^T GC)^+ (C^T GC)^2).
\end{aligned}$$

Therefore, the normality equations (4.17) hold for all possible design points x and the given designs are approximate A-optimal for the estimation of all treatment-control comparisons. □

The weights $r_1^{(t)}$ and $r_2^{(t)}$ of the A-optimal designs for the estimation of all treatment-control contrasts are listed in Table 4.1 for various numbers of treatments. Note, that the A-optimal designs in Theorem 4.5 are not universal optimal. The weights change for other optimality criteria. For example, if we consider the setting of one reference and three treatments, Table 4.1 gives the optimal weights for the A-optimality criteria. These weights are not D-optimal. The value of the D-optimality criteria defined in Example 2.2 for the estimation of the treatment-control contrasts is 2.52. In contrast, if we set all weights to 1/6, the value of the D-optimality criteria changes to 2.38, which represents a better design.

Table 4.1: Weights of the approximate A-optimal designs for various numbers of treatments.

	$r_1^{(t)}$	$r_2^{(t)}$
$t=2$	$\frac{1}{3}(3-\sqrt{3})$	$\frac{1}{3}(-3+2\sqrt{3})$
$t=3$	$\frac{1}{4}$	$\frac{1}{12}$
$t=4$	$\frac{1}{10}(-5+3\sqrt{5})$	$\frac{1}{10}(5-2\sqrt{5})$
$t=5$	$\frac{1}{15}(-3+2\sqrt{6})$	$\frac{1}{15}(3-\sqrt{6})$
$t=6$	$\frac{1}{63}(-7+5\sqrt{7})$	$\frac{1}{63}(7-2\sqrt{7})$

4.5.2 A-optimal designs in the multi-factorial setting

Using Theorem 4.1 and Theorem 4.2, we can expand the results of Section 4.5.1 to the two-factorial setting in model (4.1). Let C denote the contrast matrix comparing t treatments with a control-treatment, i.e. the contrast matrix outlined in the proof of Theorem 4.5. If we are interested in estimating the main effect contrasts $\bar{C}\tau$ with $\bar{C} = 1_{k_1} \otimes C$, it is obvious that the design with design matrix $\bar{X} = I_{k_1} \otimes P_{t+1}$ and weight matrix $\bar{P} = \frac{1}{k_1}(I_{k_1} \otimes \operatorname{diag}(\underbrace{r_1^{(t)}, \ldots, r_1^{(t)}}_{t}, \underbrace{r_2^{(t)}, \ldots, r_2^{(t)}}_{\binom{t}{2}}))$ is A-optimal due to Theorem 4.1 and Theorem 4.5. If we are also interested in the estimation of the interaction effect contrasts $\tilde{C}\tau$ with $\tilde{C} = P_{k_1}^T \otimes C$, we can use Theorem 4.2 and Theorem 4.5 to show that the design with design matrix $\tilde{X} = I_{k_1} \otimes P_{t+1}$ and weight matrix $\tilde{P} = \frac{1}{k_1}(I_{k_1} \otimes \operatorname{diag}(\underbrace{r_1^{(t)}, \ldots, r_1^{(t)}}_{t}, \underbrace{r_2^{(t)}, \ldots, r_2^{(t)}}_{\binom{t}{2}}))$ is also A-optimal for estimating $\tilde{C}\tau$ if $t+1 \leq k_1$. To demonstrate this result, we have to verify the conditions $a_{ij} \leq 0$, $i \neq j$ and $a_{ii} \leq \frac{k_1-1}{2k_1}\text{const}$, with $\text{const} = \operatorname{Tr}((C^TGC)^+(C^TGC)^{1-p})$ and $A = GC(C^TGC)^+(C^TGC)^{1-p}(C^TGC)^+C^TG^T$ due to Theorem 4.2. We know from the proof of Theorem 4.5 that

$$A = \left(\begin{array}{c|c} \frac{t}{r_1^{(t)2}(t+1)^2} & \frac{-1}{r_1^{(t)2}(t+1)^2}\mathbb{1}_t^T \\ \hline \frac{-1}{r_1^{(t)2}(t+1)^2}\mathbb{1}_t & \frac{1}{(r_1^{(t)}+tr_2^{(t)})^2}I_t + \frac{(r_2^{(t)}-r_1^{(t)})((t+2)r_1^{(t)}+tr_2^{(t)})}{r_1^{(t)2}(t+1)^2(r_1^{(t)}+tr_2^{(t)})^2}J_t \end{array}\right),$$

$$\text{const} = \frac{\left(r_1^{(t)}+r_2^{(t)}\right)t}{r_1^{(t)}\left(r_1^{(t)}+tr_2^{(t)}\right)}.$$

Obviously $a_{ij} \leq 0$, $i \neq j$ holds, because of $r_1^{(t)} > r_2^{(t)}$ for all $t \in \mathbb{N}$. Hence, it remains to show $a_{ii} \leq \frac{k_1-1}{2k_1}$const. First, we consider $a_{11} \leq \frac{k_1-1}{2k_1}$const, which is equivalent to

$$\begin{aligned}
& \frac{t}{r_1^{(t)2}(t+1)^2} \leq \frac{k_1-1}{2k_1}\frac{(r_1^{(t)}+r_2^{(t)})t}{r_1^{(t)}(r_1^{(t)}+tr_2^{(t)})} \\
\Leftrightarrow & \frac{tt(t-3)}{2(t+1)(\sqrt{1+t}(-\sqrt{1+t}+t-1))} \leq \frac{-(k_1-1)t(t-3)}{\sqrt{1+t}(\sqrt{1+t}-t+1)2k_1} \\
\Leftrightarrow & \frac{-t}{t+1} \leq \frac{-(k_1-1)}{k_1}.
\end{aligned}$$

This equality holds for $t+1 \leq k_1$. For $i \neq 1$ we can show that the inequality $a_{ii} \leq \frac{k_1-1}{2k_1}$const is equivalent to the following inequality.

$$\begin{aligned}
& \frac{r_1^{(t)2}(t+1)^2 + (r_2^{(t)} - r_1^{(t)})((t+2)r_1^{(t)} + tr_2^{(t)})}{r_1^{(t)2}(t+1)^2(r_1^{(t)}+tr_2^{(t)})^2} \leq \frac{(k_1-1)(r_1^{(t)}+r_2^{(t)})t}{2k_1 r_1^{(t)}(r_1^{(t)}+tr_2^{(t)})} \\
\Leftrightarrow & \frac{-r_1^{(t)2} + 2r_1^{(t)}r_2^{(t)} + r_1^{(t)2}t + r_2^{(t)2}t + r_1^{(t)2}t^2}{r_1^{(t)}(t+1)^2(r_1^{(t)}+tr_2^{(t)})} \leq \frac{(k_1-1)(r_1^{(t)}+r_2^{(t)})t}{2k_1} \\
\Leftrightarrow & \frac{-t(2-t+t^2+2\sqrt{1+t}-2t\sqrt{1+t})}{(1+t)(1-t+\sqrt{1+t})(-1-t-\sqrt{1+t}+t\sqrt{1+t})} \leq \frac{(k_1-1)}{k_1\sqrt{1+t}} \\
\Leftrightarrow & \frac{-t(\sqrt{1+t}+1-t)^2}{(1+t)(\sqrt{1+t}+1-t)(-\sqrt{1+t}-1+t)} \leq \frac{(k_1-1)}{k_1} \\
\Leftrightarrow & \frac{t}{(t+1)} \leq \frac{(k_1-1)}{k_1} \\
\Leftrightarrow & t+1 \leq k_1.
\end{aligned}$$

All transformations hold for $t \neq 3$, but the conditions can also be easily shown for $t = 3$ and $t+1 \leq k_1$. We get

$$\begin{aligned}
\frac{r_1^{(t)2}(t+1)^2 + (r_2^{(t)} - r_1^{(t)})((t+2)r_1^{(t)} + tr_2^{(t)})}{r_1^{(t)2}(t+1)^2(r_1^{(t)} + tr_2^{(t)})^2} &\leq \frac{(k_1-1)(r_1^{(t)} + r_2^{(t)})t}{2k_1 r_1^{(t)}(r_1^{(t)} + tr_2^{(t)})} \\
\Leftrightarrow \qquad 3 &\leq \frac{4(k_1-1)}{k_1} \\
\Leftrightarrow \qquad k_1 &\geq 4,
\end{aligned}$$

$$\begin{aligned}
\frac{t}{r_1^{(t)2}(t+1)^2} &\leq \frac{k_1-1}{2k_1} \frac{(r_1^{(t)} + r_2^{(t)})t}{r_1^{(t)}(r_1^{(t)} + tr_2^{(t)})} \\
\Leftrightarrow \qquad 3 &\leq \frac{4(k_1-1)}{k_1} \\
\Leftrightarrow \qquad k_1 &\geq 4.
\end{aligned}$$

These inequalities are valid, since $k_1 \geq t+1$. Therefore, we know the optimal design for the estimation of the interaction effect contrasts $\tilde{C}\tau$ with $\tilde{C} = P_{k_1}^T \otimes \begin{pmatrix} 1 & 1 & \dots & 1 \\ -1 & 0 & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \dots & -1 \end{pmatrix}$ for practical applications.

4.6 Helmert contrasts and all-to-next contrasts

Analogously to our considerations of treatment-control comparisons, approximate A-optimal designs for Helmert contrasts and all-to-next contrasts can be derived. However, the properties of the resulting optimal designs do not have as nice properties as the designs presented in the previous section. Therefore, it is impossible to state an explicit formula for the optimal weights dependent on t. First, the optimal weights for the Helmert contrasts, which have been introduced in Section 3.4 are derived for $t = 2$ and $t = 3$. As in the previous section, we use the results of Section 3.1, especially Theorem 3.1, to minimize $\sum_{l=1}^{m} \mathrm{Var}(c_l^T \hat{\tau})$ considering continuous values for $\tilde{B}_{ij}^{(t)} \in [0, 1]$. As mentioned in Section 3.4, A-optimal designs for the estimation of Helmert contrasts can be

achieved for $t = 2$ by minimizing the target function:

$$\begin{aligned}
&\min \left(\mathrm{Var}(\hat{\tau}_0 - \frac{1}{2}\hat{\tau}_1 - \frac{1}{2}\hat{\tau}_2) + \mathrm{Var}(\hat{\tau}_1 - \hat{\tau}_2) \right) \\
&= \min \left(\frac{1}{4}\mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_1) + \frac{1}{4}\mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_2) + \frac{1}{2}\mathrm{Cov}(\hat{\tau}_0 - \hat{\tau}_1, \hat{\tau}_0 - \hat{\tau}_2) + \mathrm{Var}(\hat{\tau}_1 - \hat{\tau}_2) \right) \\
&= \min \frac{5\tilde{B}_{01}^{(2)} + 5\tilde{B}_{02}^{(2)} + 4\tilde{B}_{12}^{(2)}}{4\tilde{B}_{01}^{(2)}\tilde{B}_{02}^{(2)} + 4\tilde{B}_{01}^{(2)}\tilde{B}_{12}^{(2)} + 4\tilde{B}_{02}^{(2)}\tilde{B}_{12}^{(2)}}
\end{aligned}$$

under the constraints $\tilde{B}_{01}^{(2)} + \tilde{B}_{02}^{(2)} + \tilde{B}_{12}^{(2)} = 1$ and $\tilde{B}_{01}^{(2)}, \tilde{B}_{02}^{(2)}, \tilde{B}_{12}^{(2)} \in \mathbb{R}^{\geq 0}$. This implies

$$\min \frac{4 + \tilde{B}_{01}^{(2)} + \tilde{B}_{02}^{(2)}}{4\left(-\left(\tilde{B}_{01}^{(2)} + \tilde{B}_{02}^{(2)}\right)^2 + \tilde{B}_{01}^{(2)} + \tilde{B}_{02}^{(2)} + \tilde{B}_{01}^{(2)}\tilde{B}_{02}^{(2)} \right)}$$

under the constraints $\tilde{B}_{01}^{(2)} + \tilde{B}_{02}^{(2)} \leq 1$ and $\tilde{B}_{01}^{(2)}, \tilde{B}_{02}^{(2)} \in \mathbb{R}^{\geq 0}$. Since $\tilde{B}_{01}^{(2)}\tilde{B}_{02}^{(2)}$ reaches its maximum for $\tilde{B}_{01}^{(2)} = \tilde{B}_{02}^{(2)}$ under these constrains, we get

$$\min_{0 \leq \tilde{B}_{01}^{(2)} \leq \frac{1}{2}} \frac{2 + \tilde{B}_{01}^{(2)}}{2\tilde{B}_{01}^{(2)}(2 - 3\tilde{B}_{01}^{(2)})}$$

and therefore $\tilde{B}_{01}^{(2)} = \tilde{B}_{02}^{(2)} = \frac{2}{3}(2\sqrt{3} - 3)$ and $\tilde{B}_{12}^{(2)} = 1 - \frac{4}{3}(2\sqrt{3} - 3)$.

Analogously, we minimize the following function, if we are interested in estimating the Helmert contrasts for $t = 3$.

$$\begin{aligned}
&\min \left(\mathrm{Var}(\hat{\tau}_0 - \frac{1}{3}\hat{\tau}_1 - \frac{1}{3}\hat{\tau}_2 - \frac{1}{3}\hat{\tau}_3) + \mathrm{Var}(\hat{\tau}_1 - \frac{1}{2}\hat{\tau}_2 - \frac{1}{2}\hat{\tau}_3) + \mathrm{Var}(\hat{\tau}_2 - \hat{\tau}_3) \right) \\
&= \min \frac{49ab + 18b^2 + 26ac + 36bc + 49ad + 110bd + 40cd + 20d^2}{18(b + 2c + d)(ab + ad + 2bd)}
\end{aligned}$$

under the constraints $a + 2b + c + 2d = 1$ and $0 \leq a, b, c, d \leq 1$. We set $a = \tilde{B}_{01}^{(3)}$, $b = \tilde{B}_{12}^{(3)} = \tilde{B}_{13}^{(3)}$, $c = \tilde{B}_{23}^{(3)}$, $d = \tilde{B}_{02}^{(3)} = \tilde{B}_{03}^{(3)}$ due to the symmetry properties of the Helmert contrast matrix. The solutions of this minimization, obtained using Mathematica, are illustrated in Table 4.2, whereas g_1 is the second root of the

equation

$$-4 + 32x - 12x^2 - 208x^3 + 263x^4 = 0$$

and g_2 is the second root of the equation

$$-1156 + 8976x - 9828x^2 - 15768x^3 + 21303x^4 = 0.$$

Therefore, $g_3 = 1 - 3g_1 - 2g_2$.

All-to-next contrasts can be deliberated in the same way. For instance, for $t = 2$ the following function is minimized

$$\begin{aligned} &\min \ (\mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_1) + \mathrm{Var}(\hat{\tau}_1 - \hat{\tau}_2)) \\ = &\min \ \frac{\tilde{B}_{01}^{(2)} + 2\tilde{B}_{02}^{(2)} + \tilde{B}_{12}^{(2)}}{\tilde{B}_{01}^{(2)}\tilde{B}_{02}^{(2)} + \tilde{B}_{01}^{(2)}\tilde{B}_{12}^{(2)} + \tilde{B}_{02}^{(2)}\tilde{B}_{12}^{(2)}} \end{aligned}$$

under the constraints $\tilde{B}_{01}^{(2)} + \tilde{B}_{02}^{(2)} + \tilde{B}_{12}^{(2)} = 1$ and $\tilde{B}_{01}^{(2)}, \tilde{B}_{02}^{(2)}, \tilde{B}_{12}^{(2)} \in \mathbb{R}^{\geq 0}$. As a result, we get $\tilde{B}_{01}^{(2)} = \tilde{B}_{12}^{(2)} = \frac{1}{3}(3 - \sqrt{3})$ and $\tilde{B}_{02}^{(2)} = 1 - \frac{2}{3}(3 - \sqrt{3})$.

The resulting optimal weights for $t = 2$ and $t = 3$ are summarized in Table 4.3. Here, f_1 is the third root of

$$0.0625 - 1.00x^2 + 2x^4 = 0.$$

f_2 is the second root of

$$-1.9375 + 15.000x - 31.00x^2 + 8.0x^3 + 2x^4 = 0$$

and f_3 is the third root of

$$0.0625 - 9.00x^2 - 8.0x^3 + 2x^4 = 0.$$

f_4 is calculated as $f_4 = 1 - 2f_1 - f_2 - 2f_3$.

Referring to Ferrari quartic equations can be solved exactly, but Galois showed that the roots of an fifth power equation cannot be solved exactly. Thus, the

Table 4.2: Approximate A-optimal designs for estimating Helmert contrasts, $t = 2$ and $t = 3$.

	$t = 2$ optimal weight		$t = 3$ optimal weight	
$\tilde{B}_{01}$	$\frac{2}{3}(2\sqrt{3}-3)$	≈ 0.3094	g_1	≈ 0.1522
$\tilde{B}_{02}$	$\frac{2}{3}(2\sqrt{3}-3)$	≈ 0.3094	g_1	≈ 0.1522
$\tilde{B}_{03}$	-	-	g_1	≈ 0.1522
$\tilde{B}_{12}$	$1-\frac{4}{3}(2\sqrt{3}-3)$	≈ 0.3812	g_2	≈ 0.1645
$\tilde{B}_{13}$	-	-	g_2	≈ 0.1645
$\tilde{B}_{23}$	-	-	g_3	≈ 0.2144

Table 4.3: Approximate A-optimal designs for estimating all-to-next contrasts, $t = 2$ and $t = 3$.

	$t = 2$ optimal weight		$t = 3$ optimal weight	
$\tilde{B}_{01}$	$\frac{1}{3}(3-\sqrt{3})$	≈ 0.4227	f_1	≈ 0.2706
$\tilde{B}_{02}$	$1-\frac{2}{3}(3-\sqrt{3})$	≈ 0.1547	f_3	≈ 0.0806
$\tilde{B}_{03}$	-	-	f_4	≈ 0.0538
$\tilde{B}_{12}$	$\frac{1}{3}(3-\sqrt{3})$	≈ 0.4227	f_2	≈ 0.2439
$\tilde{B}_{13}$	-	-	f_3	≈ 0.0806
$\tilde{B}_{23}$	-	-	f_1	≈ 0.2706

complexity of this problem increases immensely for higher values of t and we will terminate the investigations. However, we see in Table 4.2 that posterior treatment comparisons with higher indices i, j in $\tilde{B}_{ij}$ get higher weights, most important is the comparison of treatments two and three and least important are the comparisons with treatment zero. This phenomenon has been explained in Section 3.4. For the all-to-next contrasts we realize again, that the comparisons with the highest weights are the comparisons between treatments with consecutive treatment indices, $\tilde{B}_{01}$, $\tilde{B}_{12}$ and $\tilde{B}_{23}$.

4.7 Dye effect

Up to now the difference in dye intensities caused by different characteristics of the two dyes was not taken into account throughout our optimization considerations. We will reintroduce the dye effect in this chapter. As mentioned in Section 2.1 microarray experiments can be illustrated as directed graphs. Another way to describe a two-color microarray experiment is to consider it as a $2 \times a$ row-column design, where the two dyes are arranged in the rows and the a arrays are displayed in the columns (see Figure 4.1(a)).

It is easy to see that no information is lost if each treatment is colored green and red at the same frequency, i.e. each treatment occurs in the first and second row of the row-column design at the same frequency [23]. Therefore, if all treatments occur with an even quantity, the optimal design remains the same. In the approximate setting, dye swaps can be added to each design point in order to achieve optimal designs, if the primary design was not even. Dye swaps hybridize each treatment comparison on two arrays with the dye assignments reversed in the second comparison. This method always leads to optimal designs including the dye effect. Therefore, many authors recommend dye swaps, e.g. Yang and Speed [46], Stanzel [38] and Kerr [21]. Stanzel [37] proved the optimality of dye swap designs for special contrast sets. He showed the following theorem.

Theorem 4.6:
The design with design matrix $(X \mid W)$ with $X = \left(P_{k_2}^T, -P_{k_2}^T\right)^T$, W defined as $\left(\mathbb{1}_{2\binom{k_2}{2}}, -\mathbb{1}_{2\binom{k_2}{2}}\right)$ and equal weights for each support point listed in the design matrix, is ϕ_{-p} optimal, $p \in (-\infty, 1]$, for the estimation of the contrasts $\left(\mathbb{1}_{k_1}^T \otimes P_{k_2} \middle| 0_{\binom{k_2}{2}}, 0_{\binom{k_2}{2}}\right) \begin{pmatrix} \tau \\ \delta \end{pmatrix}$ and $\left(P_{k_1} \otimes P_{k_2} \middle| 0_{\binom{k_1}{2}\binom{k_2}{2}}, 0_{\binom{k_1}{2}\binom{k_2}{2}}\right) \begin{pmatrix} \tau \\ \delta \end{pmatrix}$. $\mathbb{1}_k$ and 0_k denote the $k \times 1$ column vectors with all entries equal to 1 and 0, respectively.

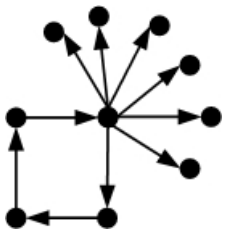
On the other hand Dobbin et al. [9] found out that dye swap assignments are often unnecessary and wasteful of resources. It is more efficient to balance the treatments with respect to the dyes and to avoid repeating comparisons. The designs recommended by Stanzel [37] remain optimal without dye swaps for

odd number of treatments. Thereby, the matrix P_k is substituted by $\tilde{P}_k$. P_k is defined as in Section 2.3.3. We define $\tilde{P}_k$ recursively by $\tilde{P}_2 := [1, -1]$ and $\tilde{P}_k := \left[\begin{array}{c|c} \pm\mathbb{1}_{k-1} & \mp I_{k-1} \\ \hline 0_{\binom{k-1}{2}} & \tilde{P}_{k-1} \end{array}\right]$ for all $k \in \mathbb{N}^{\geq 3}$, where $\pm\mathbb{1}_k$ is the $k \times 1$ vector with alternating values 1 and -1 starting with 1, $(\pm\mathbb{1}_k)_i = (-1)^{i+1}$. $\mp I_k$ is the diagonal matrix with alternating diagonal elements 1 and -1, starting with -1, i.e. $(\mp I_k)_{ii} = (-1)^i$. The proofs can be carried out similarly as in Stanzel [37]. Since all columns of the design matrix $\tilde{P}_k$ are orthogonal to the dyes $(1, \ldots, 1)^T$, the treatment effects can be estimated independently of the dyes [16].

However, the exact setting is more complex. In general, exact results including dyes are not feasible without time-consuming searches. Therefore, we will use heuristics in order to assign the dyes in a good way to our optimal designs presented in the previous sections. Assuming that each treatment i is colored red and green d_{ri} and d_{gi} times, respectively, with $|d_{ri} - d_{gi}| \leq 1$. This approach will lead to optimal and near-optimal designs in most cases. Bailey [2] showed that balanced designs do not perform well in some cases. For instance, she proved that the design displayed in Figure 4.1 (b) is A-optimal for the estimation of all pairwise treatment comparisons. However, in most cases the given approach leads to good designs. Ture [42] showed that row-column designs for treatment-control comparisons perform well if the treatments occur in each row at the same frequency. Therefore, we assign the dyes as balanced as possible to the optimal designs constructed in the previous sections and calculate the variances of the corresponding designs. These variances can be compared to the variances of the optimal designs ignoring the dye effect, which provides a lower bound for the smallest possible variance. If, by including the dye effect, the variance increases slightly, we have found a good microarray design including the dye effect. We use the same approach for the other contrast settings, e.g. Table 4.4, Table 4.5 and Table 4.6. Figure 4.2 and Figure 4.3 illustrate the experiments, which were considered in Table 4.4 and Table 4.5 graphically. The rows "Var. opt. incl. dye" give the variances of the graphically illustrated designs, which take the dye effect into account. We observe that these variances only increase slightly in comparison to the same designs without dye effect, which are displayed in the

	Array 1	Array 2	...	Array a
Green	1	2	...	1
Red	2	1	...	t

(a)



(b)

Figure 4.1: (a)Row-column design. (b) Optimal design for all pairwise treatment comparisons.

Table 4.4: Exact A-optimal and near A-optimal designs for treatment-control comparisons with and without dye effect, $t = 2$ treatments, a arrays. The same variances are obtained for the all-to-next contrasts.

	$a = 6$	$a = 7$	$a = 8$	$a = 9$	$a = 10$	$a = 15$
Var. opt. design	0.64	0.53	0.47	0.42	0.38	0.25
Var. opt. incl. dye	0.66	0.53	0.48	0.42	0.38	0.25

rows "Var. opt. design". Therefore, we found efficient designs for microarray experiments also if the dye effect is taken into account. For a detailed discussion of the dye effect see Bailey [2] or Dobbin [9].

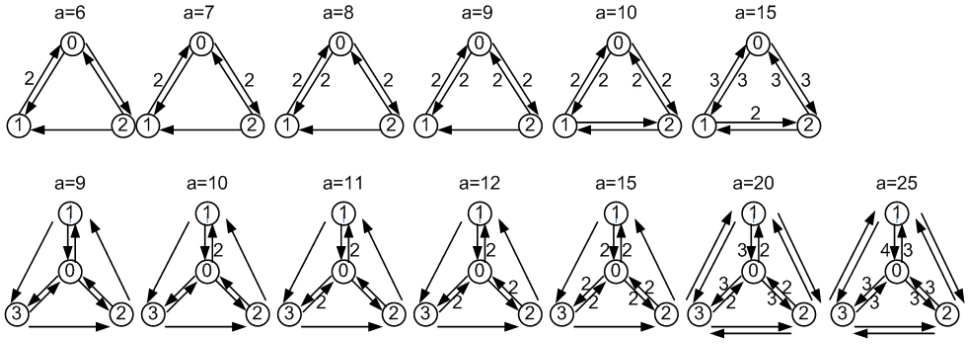


Figure 4.2: Efficient designs for estimating treatment-control comparisons accounting for the dye effect, $t = 3$ treatments, a arrays. Numbers displayed on arrows indicate the number of arrays, on which this treatment comparison is analyzed with the according dye arrangement.

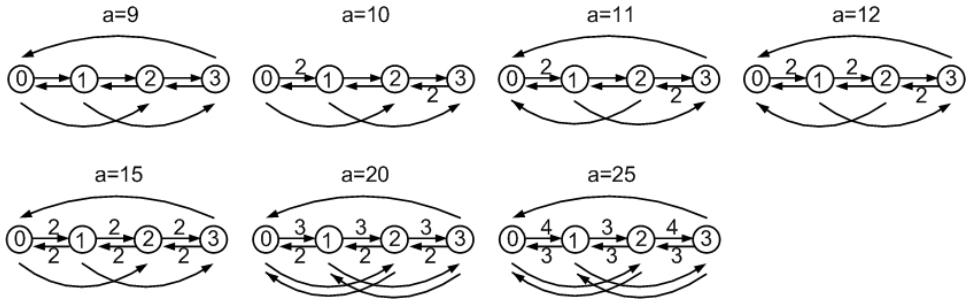


Figure 4.3: Efficient designs for estimating all-to-next contrasts accounting for the dye effect, $t \in \{2, 3\}$ treatments, a arrays. Numbers displayed on arrows indicate the number of arrays, on which this treatment comparison is analyzed with the according dye arrangement.

Table 4.5: Exact A-optimal and near A-optimal designs for treatment-control and all-to-next contrasts with and without dye effect, $t = 3$ treatments, a arrays.

Treatment Control	$a = 9$	$a = 10$	$a = 11$	$a = 12$	$a = 15$	$a = 20$	$a = 25$
Var. opt. design	0.9	0.82	0.74	0.67	0.54	0.40	0.32
Var. opt. incl. dye	0.9	0.82	0.74	0.68	0.54	0.40	0.32

All-to-next	$a = 9$	$a = 10$	$a = 11$	$a = 12$	$a = 15$	$a = 20$	$a = 25$
Var. opt. design	0.92	0.82	0.75	0.68	0.54	0.41	0.32
Var. opt. incl. dye	0.93	0.82	0.75	0.69	0.54	0.41	0.33

Table 4.6: Exact A-optimal and near A-optimal designs for treatment-control and all-to-next contrasts with and without dye effect, $t = 4$ treatments, a arrays.

Treatment-control	$a = 14$	$a = 15$	$a = 16$	$a = 20$	$a = 25$
Var. opt. design	1.00	0.93	0.87	0.69	0.55
Var. opt. incl. dye	1.01	0.94	0.88	0.69	0.55

All-to-next	$a = 14$	$a = 15$	$a = 16$	$a = 20$	$a = 25$
Var. opt. design	1.03	0.96	0.89	0.71	0.57
Var. opt. incl. dye	1.03	0.97	0.90	0.71	0.57

Chapter 5

Robustness Considerations

In many microarray experiments observations are missing and cannot be involved in the statistical analysis of the experiment. Missing values occur for different reasons, such as scratches on the slide, insufficient resolution, image corruption or other damaging (Troyanskaya et al. [40]). It is thus important to use robust experiments, which ensure precise results even if observations are missing. Latif et al. [30] have investigated specific robustness properties of commonly used microarray designs. They proposed two robustness criteria and calculated these criteria for the commonly used designs. But to date no attempts have been made to examine these robustness criteria analytically. We will derive designs with optimal robustness properties and study connections between the robustness criteria introduced by Latif et al. [30] and popular optimality criteria.

5.1 Definition of robustness criteria

This section contains the necessary definitions to describe robustness properties of microarray experiments. We illustrate the robustness criteria proposed in Latif et al. [30] as well as further definitions of robustness (e.g. proposed in Bailey [2]).

We consider model (2.5) without dye effect,

$$z = X\tau + \eta,$$

where $z = (z_1, \ldots, z_a)$ is the vector of log ratios of the dye intensities measured for a particular gene on all a arrays. the log ratios are dependent on the treatment effect $\tau = (\tau_0, \tau_1, \ldots, \tau_t)$. The term η is the random error vector. We assume η_i, $1 \leq i \leq a$ to be independently identically distributed with mean zero and variance σ^2. A design with $a \times (t+1)$ design matrix X is called connected if all contrasts $C^T\tau$ under investigation are estimable, i.e. $C^T(X^TX)^-(X^TX) = C^T$. The breakdown number (BDN) of a design is defined as the minimum number of arrays, whose removal leads to at least one disconnected design. In other words, for a design with breakdown number b the effect of interest is still estimable for all the subdesigns with $b-1$ missing observations, but there exists at least one subdesign with b missing observations with at least one inestimable effect of interest [30]. The breakdown number can be defined with and without dye effect in the same way. In the following, we restrict our investigations to the situation without dye effect. For instance, ignoring the dye effect, the loop design illustrated in Figure 2.2 has breakdown number two, whereas the star design presented in Figure 2.2 has breakdown number equal to one.

The breakdown number is also a well-known number in graph theory. In graph theory edge connectivity of a graph is defined as the minimal number of edges whose removal results in a disconnected graph. Thus, the graph theoretical expression edge connectivity is the same as the breakdown number introduced by Latif et al. [30].

Furthermore, Latif et al. [30] proposed a second robustness criterion to select good designs among designs with the same breakdown number. They defined the residual efficiency measure $\phi^{(b)}$ as the average efficiency of all subdesigns with a given number b of missing observations.

$$\phi^{(b)}(C, X) = \begin{cases} \binom{a}{b}^{-1} \sum\limits_{X_b \in \mathcal{X}_b} \phi\left(\left(C^T \left(X_b^T X_b\right)^- C\right)^{-1}\right) & \text{for } b < BDN \\ \infty & \text{otherwise,} \end{cases}$$

where $\mathcal{X}_b$ is the set of all $\binom{a}{b}$ subdesign matrices resulting from the $a \times (t+1)$ design matrix X by deleting arbitrary b rows. The matrix C is the usual contrast matrix and ϕ is an information function defined in Definition 2.1.

Another robustness criterion is the usual D-optimality criterion defined in Example 2.2, commonly known as the design which minimizes the volume of the confidence ellipsoid for the vector of treatment effects. The D-optimality criterion maximizes the determinant of the information matrix. Cheng [7] emphasized that the value of the D-criterion for block designs is proportional to the number of spanning trees of the corresponding graph. Therefore, D-optimal designs have a strong connection to robustness considerations.

5.2 Optimal breakdown number

Latif et al. [30] introduced the breakdown number for microarray experiments, but they made no attempts to derive the designs with the optimal breakdown number for given values of treatments t and arrays a. The following theorem gives the highest possible breakdown number for given numbers a and t.

Theorem 5.1:
For given values a and t with $0 \leq t-1 \leq a$, the highest possible breakdown number is equal to Opt-BDN $(t,a) = \left\lfloor \frac{2a}{t} \right\rfloor$, where $\lfloor x \rfloor$ denotes the highest integer less or equal to x.

It can be easily seen that Opt-BDN $(t,a) \leq \left\lfloor \frac{2a}{t} \right\rfloor$ since $\sum_{i=1}^{t} \mathrm{Deg}(v_i) = 2a$, where $\mathrm{Deg}(v_i)$ denotes the degree of vertex v_i, i.e. the number of edges incident to the vertex v_i. Thus, there has to be at least one vertex with degree $\mathrm{Deg}(v_i) \leq \left\lfloor \frac{2a}{t} \right\rfloor$. On the other hand, graphs fulfilling equality can be easily derived in many ways. For example, the designs consisting of repeated loops of length t and additional edges joining the vertices with the highest distance in these loops have the optimal breakdown number. Therefore, we know several designs with the best breakdown number. A detailed proof of Theorem 5.1 and further properties regarding edge connectivity can be found in Bollobas [5], Theorem 1.6.

A detailed investigation of the breakdown numbers of D-optimal designs leads

us to the following conjecture.

Conjecture 5.2*:*
D-optimal designs always achieve the optimal breakdown number given in Theorem 5.1.

Although, the conjecture seems reasonable, the proof turns out to be difficult. We show the conjecture for special cases in the following section. Thereby, we restrict ourselves to the consideration of block designs ignoring the dye effect. The general case dealing with row-column designs is a similar task for future research.

5.3 D-optimal designs

Many authors draw their attention to the derivation of D-optimal block designs in different scenarios. We will introduce known D-optimal designs and show that all of them achieve the optimal breakdown number. For example, Kiefer showed the universal optimality of balanced incomplete block designs [26]. Blocks of size k are incomplete in the sense that the block size is smaller than the number of treatments ($k < t$) and that no treatment occurs more than once in any block. An incomplete block design is balanced if the within-block concurrences r_{ij} of any two distinct treatments i and j are equal for all pairs (i, j), $i \neq j$, i.e. the number of blocks containing any two distinct treatments is a constant. Ghosh [14] emphasized that balanced incomplete block designs are robust against the unavailability of all observations in any $\frac{2a}{t} - 1$ blocks. With other words, he stated that the breakdown number for balanced incomplete block designs is the integer $\frac{2a}{t}$.

Another major reference in this regard is Gaffke [13], who derived D-optimal block designs for up to six treatments, when no balanced block design exists. For instance, he showed the D-optimality of the designs illustrated in Figure 5.1. The thick lines in Figure 5.1 illustrate λ edges, $\lambda \in \mathbb{N}$, and the thin lines illustrate one additional edge. Therefore, the first design consists of $a = 6\lambda$ edges, the second design of $a = 6\lambda + 1$ edges and so on.

We calculated the break down numbers of these D-optimal designs by means of Mengers Theorem and represented the results in Figure 5.1. Mengers Theorem states that the maximum number of paths connecting vertices x and y, which have no edge in common, equals the minimal number of edges, whose removal disconnects x and y in a graph (see, e.g., Bollobas [5], Theorem 2.4). We call two paths with no common edge independent paths. The first graph in Fig. 5.1 corresponds to a balanced incomplete block designs. Using Ghoshs [14] statement cited above, the break down number of this design equals 3λ. The break down number of the second graph has to be greater or equal to 3λ. However, deleting the 3λ edges incident to the lower right vertex disconnects the graph. This proves that the break down number equals 3λ. In general, all break down numbers have to be smaller than the minimal degree of a vertex in a graph. In the third graph in Fig. 5.1, the minimal degree of a vertex is $3\lambda+1$. On the other hand, $3\lambda + 1$ independent paths connect every two vertices in this third graph. Therefore, according to Mengers theorem, the break down number equals $3\lambda+1$. This argument can also be used for the other graphs in Fig. 5.1. Altogether, all of these D-optimal designs achieve the optimal break down number given in Theorem 5.1.

In addition, Gaffke derived D-optimal designs for three, five and six treatments. In all cases, it can be shown that these designs achieve the highest breakdown numbers. Thus, our conjecture is true for all D-optimal designs presented in Gaffke [13], which are designs with up to six treatments.

Further work on D-optimal designs includes Cheng [7], who derived the following theorems.

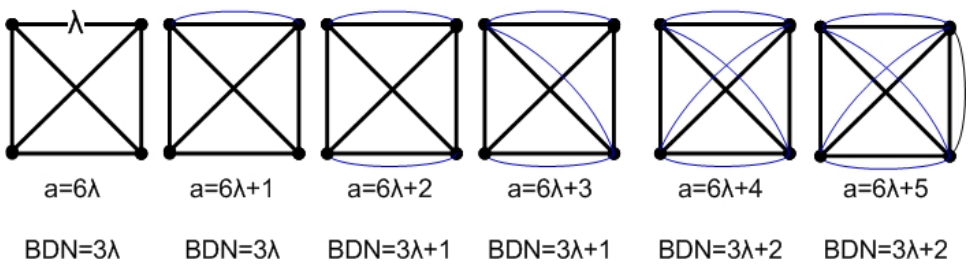


Figure 5.1: D-optimal block designs for four treatments.

Theorem 5.3:

Let G* be a graph with t vertices and a edges which is a regular complete bipartite graph or the graph obtained by adding a constant number λ of edges to each pair of vertices in a regular complete bipartite graph (i.e., the vertices of G* can be divided into two groups of equal size such that there are λ edges between any two vertices in the same group and $\lambda + 1$ edges between any two vertices in different groups, where $\lambda \geq 0$ is an integer). Then G* is the unique (up to isomorphism) graph which has the maximum number of spanning trees among all the graphs with t vertices and a edges.

A regular graph is a graph where each vertex has the same degree. It is called bipartite, if its vertices can be divided into two disjoint parts such that no vertices in the same part are adjacent. Each edge is incident to a vertex from both parts. A bipartite graph is complete, if all vertices from different parts are adjacent. Since these graphs maximize the number of spanning trees, they represent D-optimal microarray designs.

These designs have $a = \left(\frac{t}{2}\right)^2 + \binom{t}{2}\lambda$ arrays and their breakdown number equals $\frac{t}{2} + \lambda\,(t-1)$ due to Mengers Theorem. In a graph obtained by adding a constant number λ of edges to each pair of vertices in a regular complete bipartite graph, each vertex has degree $\frac{t}{2} + \lambda(t-1)$. Therefore, the break down number is less or equal to $\frac{t}{2} + \lambda(t-1)$. On the other hand $\frac{t}{2} + \lambda(t-1)$ paths connecting each pair of vertices can be found. Thus, the break down number equals $\frac{t}{2} + \lambda(t-1)$ and is optimal.

Theorem 5.4:

Let G* be a graph with t vertices and a edges which is a regular complete m-partite graph. Then G* is the unique (up to isomorphism) simple graph that has the maximum number of spanning trees with t vertices and a edges.

A graph is m-partite if its vertices can be divided into m disjoint parts such that no vertices in the same part are adjacent. Each part of the regular complete m-partite graph consists of $\frac{t}{m}$ vertices and each vertex is adjacent to all vertices of the other $m-1$ parts of the graph, but it is not adjacent to the $\frac{t}{m}-1$ vertices of the same part. Therefore, each vertex has degree $\frac{t(m-1)}{m}$ and break down

number $\frac{t(m-1)}{m}$ due Mengers Theorem. This value is the highest possible break down number for $a = \binom{m}{2}\left(\frac{t}{m}\right)^2$.

Furthermore, Cheng proved the following theorem.

Theorem 5.5:

Let G* be a graph with t vertices which is obtained by deleting q mutually nonadjacent edges from a complete graph with $q \leq t/2$. Then G* maximizes the number of spanning trees over all the simple graphs with t vertices and $\binom{t}{2} - q$ edges.

Due to Ghosh [14] the break down number of a complete graph equals $t-1$. By deleting one edge, the graph contains a vertex with degree $t-2$ and the break down number decreases by one. By deleting additional mutually nonadjacent edges up to $q \leq t/2$, the break down number stays the same, because $t-2$ independent paths connecting each pair of vertices can be constructed. Again, the breakdown number $t-2$ of these D-optimal designs is the best possible breakdown number.

Furthermore, the conjecture can be shown in general if the breakdown number is less or equal to two or equivalently $a < 3t/2$. The following theorem is a result of a discussion with R. Bailey at the conference "'Advances in Model-Oriented Design and Analysis 9"' in Bertinoro, Italy in June 2010.

Theorem 5.6:

Every D-optimal connected graph G with t vertices and a edges, $t \leq a$, does not contain a bridge. A bridge is an edge in a connected graph whose removal disconnects the graph.

Proof: Suppose that the edge $\{y, z\}$ is a bridge of G. Removing this bridge splits G into two components Y and Z. Every spanning tree of G consists of a spanning tree for Y, the edge $\{y, z\}$, and a spanning tree for Z. Hence the number of spanning trees for G is sq, where s and q are the numbers of spanning trees in Y and Z respectively. This number is positive, because G is connected. Since $t \leq a$, there is at least one edge e in G which is in a cycle. Without loss of generality, e is in component Y. Let x be the number of spanning trees of Y

which do not contain e. Since e is in a cycle, $x > 0$. Create a new graph Y' by inserting a vertex c into e. Every spanning tree of Y which contains e gives a spanning tree of Y' containing both edges at c; every spanning tree of Y which does not contain e gives two spanning trees of Y', one containing each edge at c. Hence the number of spanning trees of Y' is $s - x + 2x = s + x > s$. Create a new graph G' by replacing Y by Y', removing the bridge, and identifying the vertices y and z. Then G' has t vertices and a edges. Every spanning tree of G' consists of a spanning tree of Y' with a spanning tree of Z. Hence G' has $(s+x)q$ spanning trees. This number is greater than sq, so G cannot be D-optimal. □

Corollary 5.7:
If G is a D-optimal graph with t vertices and a edges, where $t \leq a$, then G has no vertices of degree one.

This corollary was also proved using another method by Bailey and Cameron [3].

Corollary 5.8:
If G is a D-optimal graph with t vertices and a edges, where $t \leq a < 3t/2$, then G has breakdown number two.

Proof: If $a < 3t/2$ then the average degree is strictly less than 3, so G has a vertex v of degree one or two. Since G is D-optimal, it has no vertex of degree one, so v has degree two. Therefore the two edges incident with v form a cutset of size two. A cutset is a set of edges whose removal increases the number of components of the graph. Any cutset of size one is a bridge, but G has no bridge because it is D-optimal. Hence the edge-connectivity of G is two. □

Therefore, Conjecture 5.2 is shown for $a < 3t/2$ and additionally for the given D-optimal designs. It remains an open problem to show the conjecture in general.

Chapter 6

Summary and outlook

In this work, we investigated optimal designs for two-color microarray experiments for many practical relevant scenarios. These optimal designs ensure unbiased parameter estimates with minimal variances. In future, our results can be used to achieve precise estimates in our underlying statistical model with few arrays. We distinguished between two approaches, which complement one another in practical applications. Firstly, we derived exact optimal designs in Chapter 3, and secondly, we constructed approximate optimal designs in Chapter 4. Exact designs provide optimal solutions for a given number of arrays. They can be calculated with the approach stated in Section 3.1 and they are mostly used if the number of treatments does not exceed a given limit, which usually holds in practice. The number of treatments is typically smaller than five in applications, see e.g., Callow et al. [6]. For higher numbers of treatments approximate optimal designs produce relief, since they provide nearly optimal design layouts for all given numbers of treatments and arrays. They may be used, if the computing of exact optimal designs is difficult due to high numbers of treatments. Approximate theory considers scenarios with infinite arrays and assigns weights corresponding to the proportions of arrays investigating a given treatment comparison. For applications with finite arrays, the weights are rounded and multiplied with the number of available arrays. This procedure often yields efficient exact designs and sometimes even optimal exact designs for a fixed number of arrays. Therefore, we derived approximate optimal designs in

one- and multi-factorial settings in Chapter 4. In particular, in multi-factorial settings the construction of approximate designs is very crucial, since, here, many factor level combinations need to be taken into account, which increases the computing time of exact designs. The approximate optimal designs that we have constructed in Chapter 4 can be used to estimate main and interaction effects.

Furthermore, in this thesis we focused on different contrast sets, e.g. the Helmert contrasts, all-to-next contrasts, all pairwise comparisons and treatment-control comparisons. All of these contrast sets are important in different applications and for each contrast set different designs are optimal. Therefore, we derived optimal exact designs for all of these contrast sets in Section 3.2, Section 3.3, Section 3.4 and Section 3.5. In addition, we gave optimal approximate designs for these contrasts in Section 4.5 and Section 4.6. Scientists often use the star design in various situations, although it does not provide precise parameter estimates. We proved that the star design is inefficient, even if we are interested in all treatment-control comparisons. Using the designs proposed in this work instead of the star designs, researchers can save resources while getting the same results. A detailed comparison of star designs and our optimal designs is given in Section 3.2. In addition, we investigated robustness properties of microarray designs against missing values in Chapter 5. Missing values often occur due to different reasons, such as scratches on the array. We found out that D-optimal designs often have good robustness properties.

In this thesis we answered a couple of very interesting questions concerning the design of microarray experiments. Due to the complexity of microarray experiments and due to the quick growth of biotechnology during the last decades, the number of questions in this field increases immensely and there are still several unanswered questions. First of all, we restricted our investigations in Section 3.1 to the A-optimality criterion, since this criterion is the most popular and relevant one for block designs, see Atkinson [1]. The extension of our optimality considerations to other optimality criteria is an issue for future research. Another interesting topic not covered in this thesis addresses technical replicates. Technical replicates are mRNA samples which use a common bio-

logical source, e.g. they are extracted from the same individual. Measurements taken from technical replicates are higher correlated than measurements from biological replicates using different biological sources. To account for technical replicates in our statistical model, we can add a random block effect and consider optimal design theory for mixed models. Under special restrictions optimal designs can be easily detected in this scenario. For instance, optimal designs taking only technical replicates into account are identical to optimal designs with only biological replicates, since the variances of all estimates differ by a constant term. Tsai et al. [41] showed this numerically and presented optimal designs in scenarios with just one biological source. Nevertheless, the calculation of optimal designs is more complex, if a mixture of biological and technical replicates is taken into account. Just in a few cases, if the random block effects can be chosen orthogonal to the other effects, optimal designs can be easily derived from the designs proposed in this work. However, in many settings orthogonality cannot be achieved and the derivation of optimal designs is thus a topic for future research.

Bibliography

[1] A.C. Atkinson, A.N. Donev, and R.D. Tobias. *Optimum Experimental Designs, with SAS.* Oxford Statistical Science Series, New York, 2007.

[2] R.A. Bailey. Designs for two-colour microarray experiments. *Journal of the Royal Statistical Society: Series C*, pages 365–394, 2007.

[3] R.A. Bailey and P. J. Cameron. Combinatorics of optimal designs. *preprint*, page www.maths.qmw.ac.uk/ pjc/preprints/optimal.pdf, 2008.

[4] T. Banerjee and R. Mukerjee. Optimal factorial designs for cdna microarray experiments. *The Annals of Applied Statistics.*, pages 366–385, 2007.

[5] B. Bollobas. *Extremal Graph Theory.* Academic Press Inc., 1978.

[6] M. J. Callow, S. Dudoit, E. L. Gong, T. P. Speed, and E. M. Rubin. Microarray expression profiling identifies genes with altered expression in hdl-deficient mice. *Genome Research*, pages 2022–2029, 2000.

[7] C.-S. Cheng. Maximizing the total number of spanning trees in a graph: two related problems in graph theory and optimum design theory. *Journal of Combinatorial Theory B.*

[8] G.A. Churchill. Fundamentals of experimental design for cdna microarrays. *Nature Genetics*, pages 490–495, 2002.

[9] K. Dobbin, J.H. Shih, and R. Simon. Statistical design of reverse dye microarrays. *Bioinformatics*, 19:803–810, 2003.

[10] S. Draghici. *Data Analysis Tools for DNA Microarrays.* Chapman Hall/CRC, 2003.

[11] P. Druilhet and A. Markiewicz. Information matrices for non full rank subsystems. *Metrika*, pages 171–182, 2007.

[12] V. Emilsson, G. Thorleifsson, B. Zhang, A.S. Leonardson, F. Zink, J. Zhu, S. Carlson, A. Helgason, G. B. Walters, S. Gunnarsdottir, M. Mouy, V. Steinthorsdottir, G. H. Eiriksdottir, G. Bjornsdottir, I. Reynisdottir, D. Gudbjartsson, A. Helgadottir, A. Jonasdottir, A. Jonasdottir, U. Styrkarsdottir, S. Gretarsdottir, K. P. Magnusson, H. Stefansson, R. Fossdal, K. Kristjansson, H. G. Gislason, T. Stefansson, B.G. Leifsson, U. Thorsteinsdottir, J. R. Lamb, J. R. Gulcher, M. L. Reitman, A. Kong, E. E. Schadt, and K. Stefansson. Genetics of gene expression and its effect on disease. *Nature*, 452:423–430, 2008.

[13] N. Gaffke. D-optimal block designs with at most six varieties. *Journal of Statistical Planning and Inference*, pages 183–200, 1982.

[14] S. Ghosh. Robustness of bibd against the unavailability of data. *Journal of Statistical Planning and Inference*, pages 29–32, 1982.

[15] G.F.V. Glonek and P.J. Solomon. Factorial and time course designs for cdna microarray experiments. *Biostatistics*, pages 89–111, 2004.

[16] P. Goos. *The Optimal Design of Blocked and Split-Plot Experiments.* Springer Verlag, New York, Berlin, Heidelberg, 2002.

[17] H. Grossmann and R. Schwabe. The relationship between optimal designs for microarray and paired comparison experiments. *preprint*, pages 1–24, 2007.

[18] S. Gupta. Balanced factorial designs for cdna microarray experiments. *Communications in Statistics - Theory and Methods. 35*, pages 1469–1476, 2006.

[19] D.A. Harville. *Matrix algebra from a statisticians perspective.* Springer, New York, 1997.

[20] K. Hinkelmann and O. Kempthorne. *Designs and Analysis of Experiments. Vol. 2 Advanced Experimental Design.* Wiley, New York, 2005.

[21] M. K. Kerr. Design considerations for efficient and effective microarray studies. *Biometrics*, pages 822–828, 2003.

[22] M. K. Kerr. Efficient 2k factorial designs for blocks of size 2 with microarray applications. *Journal of Quality Technology*, pages 309–317, 2006.

[23] M. K. Kerr and G.A. Churchill. Experimental design for gene expression microarrays. *Biostatistics*, pages 183–201, 2001.

[24] M. K. Kerr, M. Martin, and G. A. Churchill. Analysis of variance for gene expression microarray data. *Journal of Computational Biology*, 7:819–837, 2000.

[25] J. Kiefer. Construction and optimality of generalized youden designs. In J.N.Srivastava, editor, *A Survey of Statistical Design and Linear Models.*, pages 333–353. North-Holland, Amsterdam, 1975.

[26] J. Kiefer. Construction and optimality of generalized youden designs. 1975.

[27] W. S. Klug, M. R. Cummings, and C. A. Spencer. *Genetik.* Pearson Studium, 2007.

[28] J. Kunert, R. J. Martin, and S. Rothe. Optimal designs for treatment-control comparisons in microarray experiments. In B. Schipp and W. Kraemer, editors, *Statistical Inference, Econometric Analysis and Matrix Algebra.* Springer-Verlag, Dortmund, 2009.

[29] J. Landgrebe, F. Bretz, and E. Brunner. Efficient design and analysis of two colour factorial microarray experiments. *Computational Statistics and Data Analysis*, pages 499–517, 2006.

[30] A.H.M.M. Latif, F. Bretz, and E. Brunner. Robustness considerations in selecting efficient two-color microarray designs. *Bioinformatics*, 25:2355–2361, 2009.

[31] R. Luthi-Carter, A. Strand, N.L. Peters, S.M. Solano, Z.R. Hollingsworth, A.S. Menon, A.S. Frey, B.S. Spektor, E.B. Penney, G. Schilling, C.A. Ross,

D.R. Borchelt, S.J. Tapscott, A.B. Young, J.H. Cha, and J.M. Olson. Decreased expression of striatal signaling genes in a mouse model of huntington's disease. *Hum Mol Genet.*, pages 1259–1271, 2000.

[32] G. Parmigiani, E. S. Garrett, R. Irizarry, and S. L. Zeger (eds). *The Analysis of Gene Expression Data: Methods and Software.* Springer Verlag, New York, 2003.

[33] F. Pukelsheim. *Optimal Design of Experiments.* Wiley, New York, 1993.

[34] F. Pukelsheim and S. Rieder. Efficient rounding of approximate designs. *Biometrika*, 79:763–770, 1992.

[35] R. M. Simon, E. L. Korn andL. M. McShane, M.D. Radmacher, G.W. Wright, and Y. Zhao. *Design and Analysis of DNA Microarray Investigations.* Springer Verlag, New York, 2003.

[36] E.A. Stamatakis, W.D. Marslen-Wilson, L.K. Tyler, and P.C. Fletcherc. Cingulate control of fronto-temporal integration reflects linguistic demands: A three-way interaction in functional connectivity. *NeuroImage*, 28:115–121, 2005.

[37] S. Stanzel. *Optimale statistische Versuchsplanung dreifaktorieller Zwei-Farben-cDNA-Microarray-Experimente.* PhD thesis, Technische Universitaet Dortmund, 2007.

[38] S. Stanzel and R.D. Hilgers. The within-b-swap (bs) design is a-and d-optimal for estimating the linear contrast for the treatment effect in 3-factorial cdna microarray experiments. In J. López-Fidalgo, J.M. Rodríguez-Díaz, and B. Torsney, editors, *mODa 8 - Advances in model-oriented design and analysis*, pages 197–204. Physica, Heidelberg, 2007.

[39] J.A. Taylor, D.M. Umbach, E. Stephens, T. Castranio, D. Paulson, G. Robertson, J.L. Mohler, and D.A. Bell. The role of w-acetylation polymorphisms in smoking-associated bladder cancer: Evidence of a gene-gene-exposure three-way interaction. *Cancer Research*, pages 3603–3610, 1998.

[40] O. Troyanskaya, M. Cantor, G. Sherlock, P. Brown, T. Hastie, R. Tibshirani, D. Botstein, and R. B. Altman. Missing value estimation methods for dna microarrays. *Bioinformatics*, pages 520–525, 2001.

[41] S.-F. Tsai, C.-T-Liao, and F.-S. Chai. Statistical designs for two-color microarray experiments involving technical replication. *Computational Statistics and Data Analysis*, pages 2078–2090, 2006.

[42] T.E. Ture. Optimal row-column designs for multiple comparisons with a control: A complete catalog. *Technometrics*, pages 292–299, 1994.

[43] V. Vinciotti, R. Khanin, D. D'Alimonte, X. Liu, N. Cattini, G. Hotchkiss, G. Bucca, O. de Jesus, J. Rasaiyaah, C. P. Smith, P. Kellam, and E. Wit. An experimental evaluation of a loop versus a reference design for two-channel microarrays. *Bioinformatics*, 21:492–501, 2005.

[44] E. Wit and J. McClure. *Statistics for microarrays.* Wiley, Chichester, West Sussex, UK, 2004.

[45] R. D. Wolfinger, G. Gibson, E. D. Wolfinger, L. Bennett, H. Hamadeh, P. Bushel, C. Afshari, and R. S. Paules. Assessing gene significance from cdna microarray expression data via mixed models. *J. Computnl Biol.*, 8:625–637, 2001.

[46] H. Y. Yang and T. Speed. Design issues for cdna microarray experiments. *Nat. Rev. Genet.*, 3:579–588, 2002.

[47] D. Zieker, E. Fehrenbach, J. Dietzsch, J. Fliegner, M. Waidmann, K. Nieselt, P.Gebicke-Haerter, R. Spanagel, P. Simon, A.M. Niess, and H. Northoff. cdna microarray analysis reveals novel candidate genes expressed in human peripheral blood following exhaustive exercise. *Physiol Genomics*, 23:287–294, 2005.

Publications

1. K. Schiffl, R.-D. Hilgers,
   A-Optimal Designs for Two-Color Microarray Experiments for Treatment-Control and All-to-Next contrasts,
   To be published in the special issue of JSPI devoted to the 6th St. Petersburg Workshop of Simulation (part II), (2011)

2. K. Schiffl, R.-D. Hilgers,
   A-Optimal Designs for Two-Color Microarray Experiments for interesting contrasts,
   In: S.M. Ermakov, V.B. Melas, A.N. Pepelyshev (eds), Proceedings of the 6th St. Petersburg Workshop of Simulation, 1103-1108, (2009)

3. K. Schiffl, R.-D. Hilgers,
   Optimal Designs for Two-Color Microarray Experiments for Estimating Interactions,
   In: A. Giovagnoli, A.C. Atkinson, B. Torsney (eds), mODa 9 - Advances in model-oriented design and analysis, 197-204 (2010)

4. K. Schiffl, R.-D. Hilgers,
   Optimal Experimental Designs for Microarray Experiments,
   Talk at the 5th International Conference of EMR, EMR-IBS, Istanbul, Turkey, lecturer: K. Schiffl (2009)

5. K. Schiffl,, S. Stanzel, R.-D. Hilgers,
   Optimale Versuchspläne für Zwei-Farben-Microarray-Experimente,
   Talk at the 55. Biometrischen Kolloquiums, Hannover, Germany, lecturer: K. Schiffl (2009)