Optimal Designs for Hierarchical Interaction Structures

Rainer Schwabe

Technische Hochschule Darmstadt, Department of Mathematics,
Schlossgartenstr. 7, 64 289 Darmstadt, Germany

In linear models with hierarchical interaction structures product type designs are shown to be D–optimal simultaneously for the whole parameter vector and for testing against any hierarchical submodel. Moreover, product type designs are also Φ_q–optimal for the highest order interactions.

1. Introduction. In most experimental situations the outcome Y of an experiment is influenced by a number of different factors which may interact with each other. The aim of the present paper is to establish that, in the setting of generalized designs in the sense of Kiefer's (1974) approximate theory, product type designs prove to be optimal in models with general hierarchical interaction structures, i. e. in models in which higher order interactions appear together will all corresponding lower order interactions. For these optimal product type designs the components can be chosen to be optimal in the corresponding marginal models in which only one factor is active. (For additional readings on generalized designs we refer e. g. to the monographs by Fedorov, 1972, Bandemer, 1977, and Pukelsheim, 1993.)

Example 1. Multilinear regression on $[-1,1]^K$. Typically hierarchical interaction structures occur in multilinear regression. In this setting the most basic model $E(Y(x_1,\ldots,x_K)) = \beta_0 + \sum_{k=1}^K \beta_k x_k$ just contains additive effects β_k of the single factors. If additional interactions are present a first–order interaction model $E(Y(x_1,\ldots,x_K)) = \beta_0 + \sum_{k=1}^K \beta_k x_k + \sum_{k=1}^{K-1}\sum_{\ell=k+1}^K \beta_{k\ell} x_k x_\ell$ may be appropriate. Besides these *symmetric* interaction structures also models like $E(Y(x_1,x_2,x_3)) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2$ possess a hierarchical interaction structure in which some, but not all, first–order interactions are present.

In Section 2 some general notations and assumptions are introduced, while further examples are provided in Section 3 for hierarchical models which are covered by the present approach. In Section 4 we will show in detail that the product type designs are optimal for the highest order interaction terms with respect to the broad class of Kiefer's (1974) eigenvalue criteria which include the commonly used A–, D– and E–criteria of minimizing the expected Euclidean distance, of minimizing the generalized variance and of minimizing the maximal normalized variance, respectively.

AMS 1991 *subject classification.* 62K05.

Key words and phrases. Product type designs, multi–factor linear models, Φ_q–optimality, D–optimality, hierarchical interactions, maximal interaction.

Abbreviated title. Designs for interactions.

For the D–criterion stronger results will be established which show that product type designs are D–optimal for the whole parameter vector and simultaneously optimal for testing hierarchical models against submodels which retain the hierarchical interaction structure.

2. Definitions and notations. Generally, a linear model is described by its response function

$$E(Y(t_1,\ldots,t_K)) = \beta_0 + f(t_1,\ldots,t_K)^\top \beta_f, \tag{1}$$

$t=(t_1,\ldots,t_K)\in T$, where f is a vector of known regression functions on the design region T and $\beta=(\beta_0,\beta_f^\top)^\top$ is the vector of the corresponding unknown parameters. Throughout this paper we will assume that a constant term β_0 is included in each model equation.

We start our considerations with the marginal models

$$E(Y_k(t_k)) = \beta_0 + f_k(t_k)^\top \beta_k, \tag{2}$$

$t_k\in T_k$, in which only one factor of influence is active. These models can be obtained from the more complex hierarchical models by conditioning, i. e. by keeping all but one factor at a fixed setting. The marginal parameters $\beta^{(k)}=(\beta_0,\beta_k^\top)^\top$ consist of a constant term β_0 and the effect β_k of the kth factor. The marginal information matrices are defined by $\mathbf{I}_k(\delta_k)=\int\binom{1}{f_k}(1,f_k^\top)\,d\delta_k$, and for the complete marginal parameter vector $\beta^{(k)}$ and the marginal effects $\beta_k=L_k\beta^{(k)}$ the marginal covariance matrices are denoted by $\mathbf{C}_k(\delta_k)=\mathbf{I}_k(\delta_k)^{-1}$ and $\mathbf{C}_k(\beta_k;\delta_k)=L_k\mathbf{C}_k(\delta_k)L_k^\top$, respectively, where L_k is a suitably chosen selection matrix. Note that β_k is identifiable under δ_k iff $\mathbf{I}_k(\delta_k)$ is regular. Moreover, the matrices $\mathbf{C}_k$ are proportional to the covariances of the best linear unbiased estimators

To cover all of the hierarchical interaction models we introduce the general multi–factor linear model

$$E(Y_{\mathcal{H}}(t_1,\ldots,t_K)) = \sum\nolimits_{H\in\mathcal{H}} \bigotimes\nolimits_{k\in H} f_k(t_k)^\top \beta_H, \tag{3}$$

with interaction structure $\mathcal{H}$ on the design region of a pseudo–rectangle (Cartesian product) $T=\times_{k=1}^K T_k=\{(t_1,\ldots,t_K)^\top;\ t_k\in T_k,\ k=1,\ldots,K\}$. Thus, the settings of the single component factors may vary independently of each other. Note, however, that each component can have an arbitrary shape and can be higher–dimensional itself.

Here, each interaction structure $\mathcal{H}$ is a subset of the power set $\mathcal{P}=\{H;H\subset\{1,\ldots,K\}\}$ of possible factor combinations, and $\bigotimes_{k\in H} f_k(t_k)=f_{k_1}(t_{k_1})\otimes\ldots\otimes f_{k_m}(t_{k_m})$ is the regression function associated with the interactions for each particular factor combination $H=\{k_1,\ldots,k_m\}$, where '$\otimes$' denotes the Kronecker product. In particular, a constant term is included in the model for $H=\emptyset$. The singleton $\{k\}$ corresponds to the main effect of the kth factor, and pairs $\{k_1,k_2\}$ are related to first–order interactions (see e. g. Collombier, 1996, pp 34ff, for factorial experiments).

Particular classes of such models include Kronecker product type models

$$E(Y_{\mathcal{P}}(t_1,\ldots,t_K)) = (1, f_1(t_1)^\top) \otimes \ldots \otimes (1, f_K(t_K)^\top)\,\beta, \tag{4}$$

additive models

$$E(Y_{\{\emptyset,\{1\},\ldots,\{K\}\}}(t_1,\ldots,t_K)) = \beta_0 + \sum\nolimits_{k=1}^{K} f_k(t_k)^\top \beta_k, \tag{5}$$

and complete M–factor interaction models,

$$\begin{aligned} E(Y_{\mathcal{H}}(t_1,\ldots,t_K)) \;=\; \beta_0 \;&+\; \sum\nolimits_{k=1}^{K} f_k(t_k)^\top \beta_k \\ &+\; \ldots \\ &+\; \sum\nolimits_{1\le k_1<\ldots<k_M\le K} (f_{k_1}(t_{k_1}) \otimes \ldots \otimes f_{k_M}(t_{k_M}))^\top \beta_{k_1,\ldots,k_M}, \end{aligned} \tag{6}$$

where $\mathcal{H} = \{H; H \subset \{1,\ldots,K\},\ |H| \le M\}$. The general hierarchical model (3) can be written, for short, as $E(Y_{\mathcal{H}}(t_1,\ldots,t_K)) = f_{\mathcal{H}}(t_1,\ldots,t_K)^\top\beta$, where the regression functions are given by $f_{\mathcal{H}}(t_1,...,t_K) = (\bigotimes_{k\in H} f_k(t_k))_{H\in\mathcal{H}}$ with corresponding information matrix $\mathbf{I}(\delta) = \int f_{\mathcal{H}} f_{\mathcal{H}}^\top\, d\delta$.

For the parameters $\beta_H = L_H\beta$ associated with the interaction of the factors in H the generalized covariance matrix is $\mathbf{C}(\beta_H;\delta) = L_H\mathbf{I}(\delta)^- L_H^\top$, where $\mathbf{I}(\delta)^-$ is any generalized inverse of $\mathbf{I}(\delta)$, in case β_H is identifiable which will be assumed throughout. For whole subsets $\beta_{\mathcal{H}^*} = (\beta_H)_{H\in\mathcal{H}^*}$ of parameters we have $\beta_{\mathcal{H}*} = L_{\mathcal{H}^*}\beta$ where $L_{\mathcal{H}^*} = \mathrm{diag}(L_H)_{H\in\mathcal{H}^*}$ and $\mathbf{C}(\beta_{\mathcal{H}^*};\delta) = L_{\mathcal{H}^*}\mathbf{I}(\delta)^- L_{\mathcal{H}^*}^\top$. In case $\mathcal{H}^* = \mathcal{H}$ we write $\mathbf{C}(\delta) = \mathbf{C}(\beta;\delta)$, for short. These matrices $\mathbf{C}$ coincide with the usual covariance matrices for an underlying exact design up to a multiplicative factor.

In case different models with interaction structures $\mathcal{H}$ and $\mathcal{H}'$ are compared we distinguish between the corresponding covariance matrices $\mathbf{C}_{\mathcal{H}}$ and $\mathbf{C}_{\mathcal{H}'}$ etc. by adding the appropriate subscript. In particular, for models with only one active factor, the covariance matrix $\mathbf{C}_{\{k\}}(\delta)$ coincides with the marginal covariance $\mathbf{C}_k(\delta_k)$, where the marginal design δ_k is the projection of δ onto its kth component.

Definition. (i) An interaction structure $\mathcal{H}$ is hierarchical if $H \in \mathcal{H}$ implies $H' \in \mathcal{H}$ for every $H' \subset H$.

(ii) An interaction $H \in \mathcal{H}$ is maximal in $\mathcal{H}$ if $H \subset H' \in \mathcal{H}$ implies $H' = H$.

Equivalently, an interaction structure $\mathcal{H}$ is hierarchical if $\mathcal{P}(H) \subset \mathcal{H}$ for every $H \in \mathcal{H}$, where $\mathcal{P}(H)$ denotes the power set of H. An interaction is maximal if it is not a proper subset of any other interaction included in the model. Finally, for a subsystem $\mathcal{H}^*$ of interactions let $H(\mathcal{H}^*) = \bigcup_{H\in\mathcal{H}^*} H = \{k;\ k \in H$ for some $H \in \mathcal{H}^*\}$ be the collection of all active factors in $\mathcal{H}^*$.

In the following we will show that the product type design $\delta^* = \bigotimes_{k=1}^K \delta_k^*$ is Φ_q–optimal for maximal interactions, i. e. δ^* minimizes the q–"norm" $\sum \lambda_i^q$ of the eigenvalues λ_i of the covariance matrix $\mathbf{C}(\beta_H;\delta)$ of the least squares estimator $\hat\beta_H$ for β_H, if its components δ_k^* are Φ_q–optimal for the parameter vectors β_k of the direct effects in the marginal models. Here, $\bigotimes_{k=1}^K \delta_k = \delta_1 \otimes \ldots \otimes \delta_K$ and '$\otimes$' denotes the measure theoretic product. In particular, for discrete designs δ_k concentrated on a finite number of levels t_k with weights $w_k(t_k)$, the product is concentrated on all level combinations $(t_1, \ldots, t_K)^\top$ with weights $w(t_1, \ldots, t_K) = \prod_{k=1}^K w_k(t_k)$, respectively. More generally, for every pair of designs δ_H on the factor region $\times_{k \in H} T_k$ and $\bar{\delta}$ on the complementary region $\times_{k \notin H} T_k$ corresponding to the remaining factors the resulting product design on the whole design region is denoted by $\pi_H(\delta_H, \bar{\delta})$.

3. Some further examples. The scope of models which are covered by the concept of hierarchical interaction structures is illustrated by the following model classes in addition to the multilinear models mentioned in the introduction.

3.1. ***K*–way layout.** Every hierarchical K–way layout $E(Y_{\mathcal{H}}(i_1, \ldots, i_K)) = \sum_{H \in \mathcal{H}} \alpha^{(H)}_{(i_k)_{k \in H}}$, $i_k = 1, \ldots, I_k$, $k = 1, \ldots, K$, can be reparametrized to (3) if suitable identifiability conditions are imposed. For example, in the additive model $E(Y(i_1, \ldots, i_K)) = \beta_0 + \sum_{k=1}^K \alpha^{(k)}_{i_k}$ one may require $\sum_{i=1}^{I_k} \alpha^{(k)}_i = 0$. In the first–order interaction model $E(Y(i_1, \ldots, i_K)) = \beta_0 + \sum_{k=1}^K \alpha^{(k)}_{i_k} + \sum_{k=1}^{K-1} \sum_{\ell=k+1}^K \alpha^{(k,\ell)}_{i_k, i_\ell}$ additional conditions like $\sum_{i=1}^{I_k} \alpha^{(k,\ell)}_{i,j} = \sum_{j=1}^{I_\ell} \alpha^{(k,\ell)}_{i,j} = 0$ are needed.

3. 2. Qualitative and quantitative factors. While the theory of optimal designs is well developed for models with purely qualitative factors which may vary only over a finite number of levels (see 3.1) and for models with purely quantitative factors which vary over a whole continuum of levels (like multilinear regression) little is known for those models which contain both kinds of factors despite the fact that they occur in many practical applications as has been pointed out e. g. by Kurotschka (1988).

For illustrative purposes we confine to the simple case of a two–way layout with an additional straight line regression on $[-1, 1]$. Specific models may be additive $E(Y_{\{\emptyset,\{1\},\{2\},\{3\}\}}(i, j, x)) = \alpha^{(1)}_i + \alpha^{(2)}_j + \beta x$ or may contain all interactions $E(Y_{\mathcal{P}}(i, j, x)) = \alpha_{ij} + \beta_{ij} x$ or, as intermediate cases, the additive model is supplemented by some or all of the two–factor interactions α_{ij}, $\beta^{(1)}_i x$ and $\beta^{(2)}_j x$, e. g.

$$E(Y_{\{\emptyset,\{1\},\{2\},\{3\},\{1,2\}\}}(i, j, x)) = \alpha_{ij} + \beta x \tag{7}$$

or

$$E(Y_{\{\emptyset,\{1\},\{2\},\{3\},\{1,3\},\{2,3\}\}}(i, j, x)) = \alpha^{(1)}_i + \alpha^{(2)}_j + \beta^{(1)}_i x + \beta^{(2)}_j x. \tag{8}$$

4. Optimality of product type designs. This section is organized as follows. First, we establish the optimality of product type designs for the highest order interaction in a Kronecker product model. Then this result is extended to subsets of maximal interactions in arbitrary hierarchical models. Finally, for the D–criterion, a more general result is obtained which allows for testing a hierarchical model against any hierarchical submodel. In particular, D–optimality is established for the whole parameter vector in a hierarchical model.

Theorem 1. *Let $\mathcal{H} = \mathcal{P}$ and $H = \{1, \ldots, K\}$. If δ_k^* is Φ_q–optimal for β_k in the kth marginal model, $k = 1, \ldots, K$, then the product type design $\delta^* = \bigotimes_{k=1}^K \delta_k^*$ is Φ_q–optimal for the parameters β_H associated with the interaction of all factors.*

Proof. Note that $\beta_H = L\beta$ where $L = \bigotimes_{k=1}^K L_k$ and $\beta_k = L_k \beta^{(k)}$. Hence, the covariance matrix factorizes according to $\mathbf{C}(\beta_H; \delta^*) = \bigotimes_{k=1}^K \mathbf{C}_k(\beta_k; \delta_k^*)$. The Φ_q–optimality can now be obtained by means of the appropriate equivalence theorems. (For $q < \infty$ see e. g. Schwabe, 1996, Theorem 4.5.) For the E–citerion ($q = \infty$) a more implicit equivalence theorem (Pukelsheim, 1993, p 182) has to be used:

The design δ_k^* is E–optimal for β_k in the kth marginal model iff there exists a non–negative definite, symmetric matrix M_k with $\operatorname{tr} M_k = 1$ such that

$$a_k(t_k)^\top \mathbf{C}_k(\delta_k^*) L_k^\top \mathbf{C}_k(\beta_k; \delta_k^*)^{-1} M_k \mathbf{C}_k(\beta_k; \delta_k^*)^{-1} L_k \mathbf{C}_k(\delta_k^*) a_k(t_k) \leq 1/\lambda_{\max}(\mathbf{C}_k(\beta; \delta_k^*))$$

for all $t_k \in T_k$, where, as usual, tr and $\lambda_{\max}$ denote the trace and the largest eigenvalue of a matrix. Let $M = \bigotimes_{k=1}^K M_k$. Then M is non–negative definite, symmetric, $\operatorname{tr} M = 1$, and

$$a(t)^\top \mathbf{C}(\delta^*) L^\top \mathbf{C}(\beta_H; \delta^*)^{-1} M \mathbf{C}(\beta_H; \delta^*)^{-1} L \mathbf{C}(\delta^*) a(t) \leq 1/\lambda_{\max}(\mathbf{C}(\beta_H; \delta^*)).$$

Hence, the same equivalence theorem proves the E–optimality of the product type design δ^* for β_H. □

Note that $H = \{1, \ldots, K\}$ is the only maximal interaction in the full Kronecker product type interaction model $\mathcal{H} = \mathcal{P}$. In models with general hierarchical interaction structure $\mathcal{H}$ a maximal interaction H will be some proper subset of $\{1, \ldots, K\}$. Then only conditions are required for the active factors $k \in H$ contributing to the maximal interaction H.

Corollary 1. *Let $\mathcal{H}$ be hierarchical and H be maximal in $\mathcal{H}$. If δ_k^* is Φ_q–optimal for β_k in the kth marginal model, $k \in H$, then $\delta^* = \pi_H(\bigotimes_{k \in H} \delta_k^*, \bar{\delta})$ is Φ_q–optimal for β_H, where $\bar{\delta}$ is an arbitrary design on the region $\times_{k \notin H} T_k$ for the remaining factors.*

Proof. For the active factors involved in H the Kronecker product model $E(Y_{\mathcal{P}(H)}(t_k)_{k \in H}) = \sum_{H' \subset H} \bigotimes_{k \in H'} f_k(t_k)^\top \beta_{H'}$ can be regarded as a submodel of the original model with interaction structure $\mathcal{H}$. The interaction H is maximal in $\mathcal{P}(H)$ and, hence, the product type design $\bigotimes_{k \in H} \delta_k^*$ is Φ_q–optimal for β_H in the submodel by Theorem 1. As δ^* is a product of $\bigotimes_{k \in H} \delta_k^*$ and $\bar{\delta}$ the covariance matrices $\mathbf{C}_{\mathcal{P}(H)}(\beta_H; \bigotimes_{k \in H} \delta_k^*) = \mathbf{C}_{\mathcal{H}}(\beta_H; \delta^*)$ coincide for both models which can be seen by an orthogonalizytion argument similar to that used in Schwabe (1996), pp 77. Thus the result follows by a usual refinement argument. □

For general q the main result is the following extension of the Φ_q–optimality to more than one maximal interaction;

Theorem 2. *Let $\mathcal{H}$ be hierarchical, $H_1, \ldots, H_M$ be maximal in $\mathcal{H}$ and $\mathcal{H}^* =$*

$\{H_1, \ldots, H_M\}$. If δ_k^* is Φ_q–optimal for β_k in the kth marginal model, $k \in H(\mathcal{H}^*)$, then the design $\delta^* = \pi_{H(\mathcal{H}^*)}(\bigotimes_{k \in H(\mathcal{H}^*)} \delta_k^*, \bar{\delta})$ is Φ_q–optimal for $\beta_{\{H_1,\ldots,H_M\}}$, where $\bar{\delta}$ is an arbitrary design on the region $\times_{k \notin H(\mathcal{H}^*)} T_k$ for the remaining factors.

Proof. The key observation is that the covariance matrix is block diagonal: $\mathbf{C}(\beta_{\mathcal{H}^*}; \delta^*) = \mathrm{diag}(\mathbf{C}(\beta_H; \delta^*))_{H \in \mathcal{H}^*}$ because each $H \in \mathcal{H}^*$ is maximal. This follows again by an orthogonalization argument. By Fan's (1954) result on the eigenvalues of block matrices the Φ_q–optimality of δ^* can be deduced for $\beta_{\mathcal{H}^*}$ from the Φ_q–optimality of δ^* for every single β_H, $H \in \mathcal{H}^*$, established in Corollary 1. □

For additive models, $\mathcal{H} = \{\emptyset, \{1\}, \ldots, \{K\}\}$ all direct effects $\{k\}$ are maximal, and we recover Theorem 5.13 in Schwabe (1996). As indicated there no general results can be expected for the Φ_q–optimality of product type designs in case $q > 0$, if larger parts of the parameter vector are of interest, including lower order interactions or the constant term β_0. In contrast to that general statement some stronger results can be obtained for the D–criterion ($q = 0$) (see Schwabe, 1996, Section 6.1, and the literature quoted there for the models (4) to (6)).

Theorem 3. *Let $\mathcal{H}$ be hierarchical, $\mathcal{H}' \subset \mathcal{H}$ be a hierarchical subsystem of interactions, and $\mathcal{H}^* = \mathcal{H} \setminus \mathcal{H}'$. If δ_k^* is D–optimal in the kth marginal model, $k \in H(\mathcal{H}^*)$, then the design $\delta^* = \pi_{H(\mathcal{H}^*)}(\bigotimes_{k \in H(\mathcal{H}^*)} \delta_k^*, \bar{\delta})$ is D–optimal for $\beta_{\mathcal{H}^*}$, where $\bar{\delta}$ is an arbitrary design on the region $\times_{k \notin H(\mathcal{H}^*)} T_k$ for the remaining factors.*

Proof. First note that a design δ_k^* is D–optimal iff δ_k^* is D–optimal for β_k (see e. g. Schwabe, 1996, Theorem 3.3). Hence, the design δ^* is D–optimal for any set of maximal interactions in view of Theorem 3. The proof will now be completed by a finite induction argument as follows:

Assume that the theorem is valid for a hierarchical subsystem $\mathcal{H}'' \subset \mathcal{H}$. Let H be maximal in $\mathcal{H}''$. Then it suffices to show that the theorem is also valid for $\mathcal{H}'' \setminus \{H\}$, because $\mathcal{H}'$ can be generated by successively deleting interactions H from $\mathcal{H}''$ which are maximal in $\mathcal{H}''$. In fact, by Lemma 1 in the Appendix, Corollary 1 and the assumption we obtain

$$\begin{aligned} \det \mathbf{C}_{\mathcal{H}}(\beta_{\mathcal{H}^{**} \cup \{H\}}; \delta) &= \det \mathbf{C}_{\mathcal{H}}(\beta_{\mathcal{H}^{**}}; \delta) \ \det \mathbf{C}_{\mathcal{H}''}(\beta_H; \delta) \\ &\leq \ \det \mathbf{C}_{\mathcal{H}}(\beta_{\mathcal{H}^{**}}; \delta^*) \ \det \mathbf{C}_{\mathcal{H}''}(\beta_H; \delta^*) \ = \ \det \mathbf{C}_{\mathcal{H}}(\beta_{\mathcal{H}^{**} \cup \{H\}}; \delta^*), \end{aligned}$$

where $\mathcal{H}^{**} = \mathcal{H} \setminus \mathcal{H}''$, for every design δ which proves the assertion. □

By formally letting $\mathcal{H}' = \emptyset$ in Theorem 3, i. e. $\mathcal{H}^* = \mathcal{H}$, we derive the D–optimality of the product type design for the whole parameter vector;

Corollary 2. *Let $\mathcal{H}$ be hierarchical. If δ_k^* is D–optimal in the kth marginal model, then the product type design $\delta^* = \bigotimes_{k=1}^{K} \delta_k^*$ is D–optimal.*

Remark 1. The product type design δ^* is simultaneously D–optimal for $\beta_{\mathcal{H}^*}$ for every hierarchical interaction structure $\mathcal{H}$ and every subset of interactions $\mathcal{H}^*$ as long as the remaining interaction structure $\mathcal{H}' = \mathcal{H} \setminus \mathcal{H}^*$ is hierarchical.

Remark 2. Because of the simultaneous D–optimality established in Corollary 2 the product type design δ^* is also D–optimal for more complex experimental situations with multivariate observations where for each component a different interaction structure have to be assumed as considered by Kurotschka and Schwabe (1996).

Remark 3. The restriction to hierarchical interaction structures can be relaxed when additional assumptions are imposed on the interplay between the marginal regression functions f_k and the optimal marginal designs δ_k^* (see Schwabe, 1996, p 81).

Remark 4. The main criticism of product type designs is related to their rapidly increasing number of support points for moderate to large numbers of factors. However, still in those situations their optimality is of practical interest for the construction of good designs, because the knowledge of the hypothetically optimal information matrix allows for judging the performance of competing designs. In particular, this knowledge can help to construct fractional designs concentrated on a substantially smaller subset of design points which yield the same information matrix and are, hence, optimal. Well–known examples include orthogonal designs in factorial experiments (see e. g. Collombier, 1996). More recently Riccomagno, Schwabe and Wynn (1997) achieved a substantial reduction of the number of support points by using space–filling lattices in multi–factor Fourier models with underlying trigonometric regression.

5. Applications. For all the models exhibited in Example 1 and Section 3 optimal designs can be generated as products of optimal marginals.

5.1. Multilinear regression on $[-1,1]^K$. The full 2^K factorial design is simultaneously Φ_q–optimal for every set of maximal interactions in every hierarchical interactions model

$$E(Y_{\mathcal{H}}(x_1,\ldots,x_K)) = \sum\nolimits_{H\in\mathcal{H}} \beta_H \prod\nolimits_{k\in H} x_k \tag{9}$$

of multilinear regression. Moreover, the 2^K factorial is D–optimal for the whole parameter vector as well as for any set $\mathcal{H}^*$ of higher interactions, such that the remaining interaction structure $\mathcal{H}' = \mathcal{H}\backslash\mathcal{H}^*$ is hierarchical. In particular, in the model of complete M–factor interactions the 2^K factorial is the unique D–optimal design if $2M \geq K$. (It is also the unique D–optimal design for all interactions of m up to M factors, i. e. $\mathcal{H}^* = \{H;\, H \subset \{1,\ldots,K\},\, m \leq |H| \leq M\}$ and the unique Φ_q–optimal design for all M–factor interactions.) In case $2M < K$ also suitable 2^d fractions can be optimal, $2M \leq d < K$. For example, it is well–known that in the additive model ($M = 1$) there exist optimal 2^d fractional factorials for $d \geq \log_2(K+1)$ and moderate $K \geq 3$ constructed from Hadamard matrices.

5.2. K–way layout. In every hierarchical K–way layout $E(Y_{\mathcal{H}}(i_1,\ldots,i_K)) = \sum_{H\in\mathcal{H}} \alpha^{(H)}_{(i_k)_{k\in H}}$ the equireplicated design is D–optimal for testing each model with hierarchical interaction structure against each submodel which itself has a hierarchical interaction structure. (For the construction of optimal fractional factorials see e. g. Collombier, 1996).

5. 3. Qualitative and quantitative factors. For all those different models introduced in Subsection 3.2 the equireplicated design on $\{1,\ldots,I\}\times\{1,\ldots J\}\times\{-1,1\}$ is D–optimal for the whole parameter vector (under suitable identifiability conditions) as well as for estimating (testing) the higher interaction terms.

6. Appendix: An auxiliary result. In this section we present a helpful relation between the covariance matrices in a larger model $E(Y(t)) = f_0(t)^\top\beta_0 + f_1(t)^\top\beta_1 + f_2(t)^\top\beta_2$ and those in a proper submodel $E(Y_0(t)) = f_0(t)^\top\beta_0 + f_1(t)^\top\beta_1$. Here, the covariance matrices in the submodel are denoted by $\mathbf{C}_0$. The following statement links the determinants of the covariance matrices in both models.

Lemma 1. *If β_1 and β_2 are identifiable under δ in the larger model $E(Y(t)) = f_0(t)^\top\beta_0 + f_1(t)^\top\beta_1 + f_2(t)^\top\beta_2$, then* $\det\mathbf{C}(\beta_1,\beta_2;\delta) = \det\mathbf{C}_0(\beta_1;\delta)\ \det\mathbf{C}(\beta_2;\delta)$.

Proof. We partition the covariance matrix

$$\mathbf{C}(\beta_1,\beta_2;\delta) = \begin{pmatrix} C_1 & C_{12} \\ C_{12}^\top & C_2 \end{pmatrix}$$

according to the components β_1 and β_2, $C_j = \mathbf{C}(\beta_j;\delta)$. By the formula for generalized inverses of partitioned matrices we obtain $\mathbf{C}_0(\beta_1;\delta) = C_1 - C_{12}C_2^{-1}C_{12}^\top$ and the result follows from $\det\mathbf{C}(\beta_1,\beta_2;\delta) = \det(C_1 - C_{12}C_2^{-1}C_{12}^\top)\ \det C_2$. □

Remark 5. If f_0 is omitted then the result of Lemma 1 specializes to the well–known relation $\det\mathbf{C}(\delta) = \det\mathbf{C}_0(\delta)\ \det\mathbf{C}(\beta_2;\delta)$.

Acknowledgement

The present work was finally completed while the author was visiting the University of Technology at Darmstadt. He wishes to express his sincere thanks for the hospitality he met during his stay. The research was partly supported by grants Ku719/2-1 and Ku719/2-2 of the Deutsche Forschungsgemeinschaft at the Free University, Berlin.

References

Bandemer, H. (1977). *Theorie und Anwendung der optimalen Versuchsplanung I. Handbuch zur Theorie.* Akademie–Verlag, Berlin.

Collombier, D. (1996). *Plans d'expérience factoriels.* Springer, Paris.

Fan, K. (1954). Inequalities for eigenvalues of Hermitian matrices. *Nat. Bur. Standards, Appl. Math. Ser.* **39**, 131–139.

Fedorov, V. V. (1972). *Theory of Optimal Experiments.* Academic Press, New York.

Kiefer, J. (1974). General equivalence theory for optimum designs (approximate theory). *Ann. Statist.* **2**, 849–879.

KUROTSCHKA, V. G. (1988). Characterizations and examples of optimal experiments with qualitative and quantitative factors. In *Model-Oriented Data Analysis, Proceedings Eisenach 1987* (V. Fedorov and H. Läuter, eds.). Springer, Berlin, 53–71.

KUROTSCHKA, V. G. and R. SCHWABE (1996). The reduction of design problems for multivariate experiments to univariate possibilities and their limitations. In *Research Developments in Probability and Statistics. Festschrift in Honor of Madan L. Puri on the Occasion of his 65th Birthday* (E. Brunner and M. Denker, eds.). VSP, Utrecht, 193–204.

PUKELSHEIM, F. (1993). *Optimal Design of Experiments.* Wiley, New York.

RICCOMAGNO, E., R. SCHWABE and H. P. WYNN (1997). Lattice-based optimum design for Fourier regression. *Ann. Statist.* (to appear).

SCHWABE, R. (1996). *Optimum Designs for Multi-factor Models. Lecture Notes in Statistics, Vol. 113.* Springer, New York.