Optimal Designs for Paired Comparisons in Main Effects Analysis of Variance Models

ULRIKE GRASSHOFF
*Institut für Mathematik I, Freie Universität Berlin,
Arnimallee 2–6, D–14195 Berlin, Germany*

HEIKO GROSSMANN, HEINZ HOLLING
*Psychologisches Institut IV, Westfälische Wilhelms–Universität,
Fliednerstr. 21, D–48149 Münster, Germany*

RAINER SCHWABE[1]
*Institut für Medizinische Biometrie, Eberhard–Karls–Universität Tübingen,
Westbahnhofstr. 55, D–72070 Tübingen, Germany
& Institut für Mathematik I, Freie Universität Berlin*

Abstract

In psychological research paired comparisons which demand judges to evaluate the trade-off between two alternatives have been shown to yield valid estimates of the judges' preferences. For this situation we present optimal designs in a discrete setting where the alternatives are modeled by an analysis of variance model with main effects only. While product type designs are theoretically optimal they bear the disadvantage of exponentially increasing sample sizes. We employ orthogonality and combinatorial tools like Hadamard matrices to achieve optimal designs which are saturated or which have sufficiently small sample sizes increasing linearly in the number of factors involved. Moreover, optimal designs are constructed when the number of varying factors is restricted for each pair of alternatives.

Keywords:
paired comparisons, main effects model, optimal design, equivalence theorem, symmetrization, additive model, marginal design, product design, Hadamard matrix, orthogonal array, saturated design, profile strength

1 Introduction

In social sciences and especially in marketing research paired comparisons constitute a prominent mode for data collection (Baier and Gaul, 1999). Applications of the method usually require a large number of comparisons. To keep this number at a manageable size efficient experimental designs are called for.

[1]corresponding author, email: rainer.schwabe@uni-tuebingen.de

In this note, we present optimal designs for paired comparisons in a discrete setting where the alternatives are described by an analysis of variance model with main effects only. We start with the optimal designs for the one–way layout (Section 3) and use them to show that the corresponding product type designs are theoretically optimal (Section 4). However, these product designs bear the disadvantage of tremendously large sample sizes which are increasing exponentially in the number of components. Therefore we employ an orthogonality property together with combinatorial tools like Hadamard matrices and orthogonal arrays to construct optimal designs which have sufficiently small sample sizes (Section 5). These sample sizes increase linearly in the number of components, i.e. they do not grow faster than the number of parameters in the model. In some cases it is even possible to find optimal designs which are saturated (Section 6).

It has to be noted that the optimal designs require the alternatives to differ in each component at every paired comparison. However, in psychological applications it is usually too difficult for a judge to express his preference between two alternatives which differ with respect to too many components. Therefore only a prespecified maximal number of components, the so–called profile strength, is allowed to vary simultaneously within a single paired comparison. Corresponding optimal designs are generated in Section 7.

2 General setting

As in the monograph by van Berkum (1987) we confine ourselves to the solution of the design problem for the approximating linear model. According to Offen and Littell (1987) this can be seen to be the linearization of the binary model by Bradley and Terry (1952) where the utility is given by the logistic response function in the particular case of a zero parameter vector of equal preference. Moreover, there are situations where paired comparisons occur in a natural way for linear models when strong unknown intra–individual effects have to be excluded from the analysis.

To be more specific we consider the general linear model

$$\mu(t) = \mu_0 + f(t)^\top \beta$$

with an explicit constant term μ_0 for modeling the value of an alternative $t \in \mathcal{T}$. Here f is a vector of known regression functions and β is the vector of parameters of interest. In contrast to standard design problems the values for the alternatives are not directly observable. Only observations $Y(s,t)$ are available for comparing pairs of alternatives s and t. Nevertheless, the observations are still properly described by a linear model

$$Y(s,t) = (f(s) - f(t))^\top \beta + Z(s,t)$$

with settings $x = (s,t)$ for the paired comparisons chosen from the design region $\mathcal{X} = \mathcal{T} \times \mathcal{T}$. Note that for paired comparisons the regression function is given by $x = (s,t) \to F(s,t) := f(s) - f(t)$. Different experiments are assumed to result in uncorrelated and homoscedastic errors Z.

For technical ease we mainly consider generalized designs ξ which are defined as finitely supported probability measures on the design region $\mathcal{X}$. For the definition and properties of generalized (approximate) designs we refer to the seminal paper by Kiefer (1959). We call a design ξ associated with a sample size N if all weights $\xi(s,t)$ are multiples of $\frac{1}{N}$. The smallest such number $N = N(\xi)$ is the minimum sample size required for ξ. Obviously, the minimum required sample size $N(\xi)$ is, at least, as large as the number $V(\xi) = |\{(s,t);\ \xi(s,t) > 0\}|$ of support points of the design ξ where $\xi(s,t)$ is the weight associated with the paired comparison of s with t. If ξ is uniform on its support then the minimum required sample size $N(\xi)$ coincides with the number of support points $V(\xi)$ which will be typically the case in the present note.

The quality of a design ξ is reflected by its information matrix $M(\xi)$ which is essentially proportional to the inverse of the covariance matrix for the best linear unbiased estimator of the parameter vector β. As a performance measure we use the most popular criterion of D–optimality which aims at maximizing the determinant of the information matrix $M(\xi)$. In the present setting the information matrix equals $\sum_{(s,t)} M(s,t)\xi(s,t)$ where $M(s,t) = (f(s) - f(t))(f(s) - f(t))^\top$ is the information contained in the paired comparison of s with t.

Let us note the obvious fact that there is zero information in the pairs (t,t), i. e. in the comparison of t with itself, $M(t,t) = 0$. Moreover, it is also clear that the information for the pair (s,t) is the same as for (t,s), the pair with the internal order reversed, $M(s,t) = M(t,s)$. However, it is worthwhile to distinguish between the comparison of s with t and the comparison of t with s in view of generating higher dimensional designs – although these comparisons seem to be indistinguishable at first glance. In particular, we can always replace a design ξ by its symmetrized version $\bar{\xi}$ with respect to exchanging s and t and get the same information matrix, $M(\bar{\xi}) = M(\xi)$, with the additional benefit that the symmetrized design is balanced with respect to constant terms, $\sum_{(s,t)}(f(s) - f(t))\bar{\xi}(s,t) = 0$. We will take advantage of this property later on, when constructing optimal designs for models where the response depends on more than one component.

3 The one–way layout

For notational convenience we make use of the control parameterization, i. e.

$$\mu(i) = \begin{cases} \mu_0 + \beta_i & \text{for } i = 1, \dots, I-1 \\ \mu_0 & \text{for } i = I \end{cases}$$

with the interpretation that $\mu_0 = \mu(I)$ is the direct response of the control level I and $\beta_i = \mu(i) - \mu(I)$ is the effect of level i compared to control. If we formally set $\beta_I \equiv 0$ then for paired comparisons the mean response $\mu(i) - \mu(j) = \beta_i - \beta_j$ is the effect of level i compared to level j. In this setting the regression functions are given by $f(i) = \mathbf{e}_i$, where $\mathbf{e}_i$ is the ith unit vector of length $I-1$ and, hence, $F(i,j) = \mathbf{e}_i - \mathbf{e}_j$. By convention $\mathbf{e}_I$ is the zero vector of length $I-1$. Note that the D–optimality is not affected by the choice of the parameterization in the underlying model.

Next we mention that in general, invariance and equivariance properties are inherited by passing from the model μ for the direct observations to the model for observing the paired comparisons as long as the constant term is preserved: If the augmented regression function $(1, f^\top)^\top$ is linearly equivariant with respect to a transformation g, i. e. there is a matrix Q_g such that $(1, f(g(t))^\top)^\top = Q_g \cdot (1, f(t)^\top)^\top$ for all $t \in \mathcal{T}$, then the regression function $F : (s,t) \to f(s) - f(t)$ of the paired comparison is also linearly equivariant. The corresponding transformation matrix for F is obtained by deleting the first row and column of Q_g.

When we consider the permutation of all levels this means for the present setting that there is an optimal design with equal weights assigned to each paired comparison for which the alternatives presented are different. Note that the non–informative comparisons (i,i) can be excluded as $M(i,i) = 0$. Hence, the uniform design $\xi^*_{I,0}$ on the set $\{(i,j);\ i \neq j\}$ which assigns equal weight $1/(I(I-1))$ to each ordered pair (i,j) of distinct levels $i \neq j$ is D–optimal. The minimum required sample size is equal to $N(\xi^*_{I,0}) = V(\xi^*_{I,0}) = I(I-1)$.

The information matrix $M(\xi^*_{I,0})$ of the optimal design $\xi^*_{I,0}$ is completely symmetric,

$$M(\xi^*_{I,0}) = \frac{2}{I(I-1)} \begin{pmatrix} I-1 & -1 & \cdots & -1 \\ -1 & I-1 & \cdots & -1 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & \cdots & I-1 \end{pmatrix} = \frac{2}{I-1}(\mathbf{Id}_{I-1} - \frac{1}{I}\mathbf{1}_{I-1}\mathbf{1}_{I-1}^\top)\,,$$

where $\mathbf{Id}_m$ and $\mathbf{1}_m$ denote the $m \times m$ identity matrix and the vector of length m with all entries equal to 1, respectively. Note that the inverse of the information matrix given by

$$M(\xi_{I,0}^*)^{-1} = \frac{I-1}{2}\begin{pmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 2 \end{pmatrix} = \frac{I-1}{2}(\mathbf{Id}_{I-1} + \mathbf{1}_{I-1}\mathbf{1}_{I-1}^\top)$$

shares the property of complete symmetry. As an alternative proof it can be shown that $(f(i)-f(j))^\top M(\xi_{I,0}^*)^{-1}(f(i)-f(j)) = I-1$ for $i \neq j$ which establishes the D–optimality of the design $\xi_{I,0}^*$ by virtue of the celebrated Kiefer–Wolfowitz (1960) equivalence theorem.

As the internal order can be exchanged in each paired comparison without changing the information, $M(i,j) = M(j,i)$, we can confine, for a moment, the experiment to those comparisons where $i < j$. For the resulting design $\xi_{I,1}^*$ which is uniform on $\{(i,j);\ i < j\}$ the information matrix coincides with that of the D–optimal design, $M(\xi_{I,1}^*) = M(\xi_{I,0}^*)$. Hence, $\xi_{I,1}^*$ is also D–optimal and has a smaller minimum required sample size $N(\xi_{I,1}^*) = I(I-1)/2$. As can be seen from the entries of the information matrix $M(\xi_{I,1}^*)$ no further reduction is possible.

For later use we denote by A_I the $\frac{1}{2}I(I-1) \times (I-1)$ design matrix corresponding to the D–optimal design $\xi_{I,1}$, i. e. each row of A_I consists of the regression function $(f(i)-f(j))^\top$ evaluated at distinct pairs of alternatives from $\{(i,j);\ i < j\}$. Hence, $\frac{2}{I(I-1)}A_I^\top A_I = M(\xi_{I,1}^*)$.

Note that in the case of an odd number I of levels it is possible to construct a balanced D–optimal design $\xi_{I,2}^*$ with sample size $N(\xi_{I,2}^*) = I(I-1)/2$ by interchanging the internal order for some of the paired comparisons of the design $\xi_{I,1}^*$, such that $\sum_{i,j}(f(i)-f(j))\xi_{I,2}^*(i,j) = 0$. Consequently, $A^\top \mathbf{1}_{I(I-1)/2} = 0$ for the corresponding design matrix A.

We also want to point out that the designs $\xi_{I,0}^*$ and $\xi_{I,1}^*$ (resp. $\xi_{I,2}^*$) prove to be D–optimal also in the setting of correlated observations coming in pairs which has been investigated by Atkins and Cheng (1999) for the setting of polynomial regression.

4 The K–way layout and product type designs

Usually, in paired comparisons the alternatives are specified by a couple of, say K, components which may vary over I_k levels, $k = 1, \ldots, K$, each. For the control parameterization, $\beta_{1,I_1} = \ldots = \beta_{K,I_K} = 0$, the direct response is modeled by

$$\mu(i_1, \ldots, i_K) = \mu_0 + \sum_{k=1}^{K} \beta_{k,i_k}$$

$i_k = 1, \ldots, I_k$, $k = 1, \ldots, K$, where β_{k,i_k} is the effect of level i_k compared to the control level I_k for the kth component. This is a main effects model without any interactions between the components.

For the response in the pairwise comparisons we obtain

$$\mu(i_1, \ldots, i_K) - \mu(j_1, \ldots, j_K) = \sum_{k=1}^{K} (\beta_{k,i_k} - \beta_{k,j_k}) = \sum_{k=1}^{K} (\mu_k(i_k) - \mu_k(j_k))$$

which breaks down to the sum of the effects for the single components as given by the corresponding marginal one–way layouts

$$\mu_k(i_k) - \mu_k(j_k) = \beta_{k,i_k} - \beta_{k,j_k} \ .$$

According to a result by Rafajłowicz and Myszka (1992) the product type design $\bigotimes_{k=1}^{K} \xi_k^* = \xi_1^* \otimes \ldots \otimes \xi_K^*$ is D–optimal if the components ξ_k^* are D–optimal in the corresponding marginal one–way layouts and if, additionally, the orthogonality condition $\sum_{i_k, j_k} (f_k(i_k) - f_k(j_k)) \xi_k^*(i_k, j_k) = 0$ is satisfied, for all $k = 1, \ldots, K$ (see also Schwabe, 1996, Section 5.2). These conditions are met by the symmetric D–optimal marginal designs $\xi_k^* = \xi_{I_k,0}^*$ and, hence, the product type design $\bigotimes_{k=1}^{K} \xi_{I_k,0}^*$ is D–optimal for the K–way layout.

Moreover, it follows that the information matrix

$$M(\xi_1^* \otimes \ldots \otimes \xi_K^*) = \begin{pmatrix} M(\xi_1^*) & 0 & \cdots & 0 \\ 0 & M(\xi_2^*) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & M(\xi_K^*) \end{pmatrix}$$

of a D–optimal design is block diagonal where the blocks on the diagonal are equal to the information matrices $M(\xi_k^*) = M(\xi_{I_k,0}^*)$ for the marginal D–optimal designs ξ_k^* in the corresponding marginal one–way layouts.

In general, a product type design $\xi_1 \otimes \ldots \otimes \xi_K$ is defined by its individual weights $\xi_1 \otimes \ldots \otimes \xi_K(x_1, \ldots, x_K) = \prod_{k=1}^{K} \xi_k(x_k)$ for each combination $(x_1, \ldots, x_K)$ of support points x_k of the marginal designs ξ_k. In particular, the product type design is supported on the Cartesian product of the supports of the marginal designs. From this it follows that $V(\xi_1 \otimes \ldots \otimes \xi_K) = \prod_{k=1}^{K} V(\xi_k)$ and a similar result holds for the minimum required sample size $N(\xi_1 \otimes \ldots \otimes \xi_K)$. Note also that the product design is uniform on the Cartesian product if the marginal designs are uniform on their respective marginal supports. In the present setting the design points x are pairs of K–dimensional alternatives $(i_1, \ldots, i_K)$ and $(j_1, \ldots, j_K)$, respectively, whereas the marginal design points x_k are pairs of alternatives (i_k, j_k) for the kth component, $k = 1, \ldots, K$.

The D–optimal product type design $\xi^*_{I_1,0} \otimes \ldots \otimes \xi^*_{I_K,0}$ is uniform on the set $\{((i_1, \ldots, i_K), (j_1, \ldots, j_K));\ i_k \neq j_k,\ k = 1, \ldots, K\}$ of all pairs of K–dimensional alternatives which differ in each component. Thus its minimum required sample size equals $N(\xi^*_{I_1,0} \otimes \ldots \otimes \xi^*_{I_K,0}) = \prod_{k=1}^{K} I_k(I_k - 1)$.

Typically, in many applications the number of levels is constant across the components, $I_1 = \ldots = I_K = I$. For this case the minimum required sample size of the above product type design equals

$$N(\xi^*_{I,0} \otimes \ldots \otimes \xi^*_{I,0}) = I^K (I-1)^K$$

which is seen to increase exponentially while the number of parameters increases linearly, $p_K = K(I-1)$, as the number K of components is growing.

A slight reduction of the sample size is possible as for a product type design to be D–optimal the orthogonality condition is required only for all but one component. Hence, in the product type design, say, the first component may be replaced by $\xi^*_{I,1}$ still resulting in a D–optimal design which has minimum required sample size $N(\xi^*_{I,1} \otimes \xi^*_{I,0} \otimes \ldots \otimes \xi^*_{I,0}) = \frac{1}{2} I^K (I-1)^K$. In the case of an odd number of levels the components of the D–optimal product type design may be chosen equal to the balanced design $\xi^*_{I,2}$ with a minimum required sample size

$$N(\xi^*_{I,2} \otimes \ldots \otimes \xi^*_{I,2}) = \left(\frac{1}{2} I(I-1)\right)^K$$

which is still increasing exponentially in the dimension K.

5 Reduction by Hadamard structure

In this section we investigate the case of equal numbers of levels in more detail, $I_1 = \ldots = I_K = I$. As the number of support points and, hence, the minimum required sample size is increasing exponentially in the number of components, when a product type design is used, there is a need to search for techniques to reduce the sample size.

In the setting of direct observations this is done by employing fractional factorials and, more generally, orthogonal arrays (see Raghavarao, 1971, and Hedayat et al., 1999, for details). In this context Hadamard matrices which consist of orthogonal columns (and rows) with all entries equal to $+1$ or -1 play an important role. From an $n \times n$ Hadamard matrix a fractional factorial can be constructed with sample size n when there are less than n components which may vary on two levels each. Note that the order n of a Hadamard matrix is either 1, 2 or a multiple of 4. For multiples of 4 Hadamard matrices are known to exist for all moderate $n = 4m$ up to 424 (see Hedayat et al., 1999, p.147).

While for direct observations the use of Hadamard matrices is restricted to the number of levels I equal to 2 and, eventually, higher powers of 2, I equal to 2^q, we are able to combine the Hadamard structure with the immanent balance of pairwise comparisons for an arbitrary number of levels I.

To this end let g_K be the smallest integer for which a Hadamard matrix exists with order not less than K. As already noted $g_K = K$ for $K = 1$, 2 or moderate multiples of 4 and g_K will usually not be substantially larger than K for moderate numbers of components. In fact, g_K is always smaller than $2K$. Let H_K be a Hadamard matrix of order g_K and denote by $h_1, \ldots, h_K$ the first K columns of H_K. Using the $\frac{1}{2}I(I-1) \times (I-1)$ design matrix A_I of the D–optimal marginal design $\xi^*_{I,1}$ we can construct a $\frac{1}{2}g_K I(I-1) \times K(I-1)$ design matrix $A_{I,K}$ for the K–way layout

$$A_{I,K} = \left(\; h_1 \otimes A_I, \quad h_2 \otimes A_I, \quad \ldots, \quad h_K \otimes A_I \; \right),$$

where "$\otimes$" denotes the usual Kronecker product of vectors and matrices. Note that by the immanent balance of paired comparisons the matrix $-A_I$ corresponds to the design which is obtained by interchanging the internal order of alternatives for each comparison in A_I. Hence, the rows of the matrix $A_{I,K}$ can, indeed, be identified as experimental settings, i. e. as the regression function evaluated at some pair of alternatives. More specifically, if an entry of the vector h_k equals -1 then the internal order of the corresponding alternatives is interchanged for the kth component.

Denote by $\xi^{(K)}$ the design which is uniform on those $\frac{1}{2}g_K I(I-1)$ pairwise comparisons represented by the rows of the design matrix $A_{I,K}$. Then we obtain for the corresponding information matrix

$$\begin{aligned} M(\xi^{(K)}) &= \frac{1}{\frac{1}{2}g_K I(I-1)} (A_{I,K})^\top A_{I,K} \\ &= \mathbf{Id}_K \otimes \left(\frac{1}{\frac{1}{2}I(I-1)} (A_I)^\top A_I \right) \; = \; \mathbf{Id}_K \otimes M(\xi^*_{I,0}) \end{aligned}$$

by means of the common multiplication rules for Kronecker products and the orthogonality property of the columns of the Hadamard matrix, $h_k^\top h_k = g_K$ and $h_k^\top h_\ell = 0$, for $k \neq \ell$. As the information matrix coincides with that of

Table 1: Required sample sizes for D–optimal designs with Hadamard structure compared to product type designs

I	2			3			7		
K	2	4	8	2	4	8	2	4	8
product design	2	8	128	9	81	6 561	441	194 481	$3.8 \cdot 10^{10}$
Hadamard design	2	4	8	6	12	24	42	84	168
# parameters	2	4	8	4	8	16	12	24	48

the D–optimal product type design, $M(\xi^{(K)}) = M(\xi^*_{I,0} \otimes \ldots \otimes \xi^*_{I,0})$ the design $\xi^{(K)}$ is also D–optimal with a sample size

$$N(\xi^{(K)}) = \frac{1}{2} g_K I(I-1)$$

which increases linearly in K. The reduction in the required sample size is illustrated in Table 1 for various numbers of levels and components.

In particular, for $I = 2$, the Hadamard design is saturated, i. e. the sample size is as small as the number of parameters in the K–way layout, if $K = 2$ or K is a moderate multiple of 4. It is worth noting that for the case $I = 2$ the model of paired comparisons is equivalent to the situation of weighing K objects on a chemical balance. Hence, for $I = 2$ the results can be directly derived from the theory of weighing designs (see e. g. Shah and Sinha, 1989).

6 Orthogonal arrays and partially balanced optimal designs

As a generalization of Hadamard matrices orthogonal designs were developed for generating fractional factorials in the setting of directly observable responses when the number of levels is larger than 2.

For paired comparisons with I levels orthogonal arrays can be applied directly when their number of symbols (levels) equals $S = I(I-1)$. This can be done by identifying the $I(I-1)$ distinct ordered pairs (i_k, j_k), $i_k \neq j_k$, with the S symbols in the orthogonal array, for each component k separately.

In the case of comparisons with an odd number I of levels for each component orthogonal arrays can be used which have $\frac{1}{2}I(I-1)$ levels if those levels are identified with the pairs (i_k, j_k) taken from the balanced marginal design $\xi^*_{I,2}$. Here the balance guarantees the block diagonal structure of the resulting information matrix and, hence, the D–optimality of the design.

The last observation is particularly helpful when there are $I = 3$ levels for each component. For example, a 3×3 Graeco–Latin Square can be used to derive a D–optimal design for paired comparisons when there are $K = 4$ components. The Graeco–Latin Square requires a sample size of 9 whereas 12 comparisons are needed for the Hadamard structure design of the previous section. Similarly, Taguchi's L_{18} can cope with seven components in 18 comparisons whereas the Hadamard structure design requires a sample size of 24. However, we can take further advantage of combining orthogonal arrays with Hadamard structures. So, the design described by the Kronecker product of the Hadamard matrix H_2 with the 3×3 Graeco–Latin Square outperforms the L_{18} by dealing with $K = 8$ components in 18 comparisons.

The balance of some substructures can be used to construct D–optimal designs with even smaller sample sizes. To illustrate this fact we give an example of a saturated D–optimal design for 2 components with 4 levels each which is uniform on the following comparisons:

(i_1, j_1)	(i_2, j_2)
$(1, 2)$	$(3, 1)$
$(1, 3)$	$(1, 2)$
$(1, 4)$	$(2, 3)$
$(2, 3)$	$(4, 1)$
$(2, 4)$	$(3, 4)$
$(3, 4)$	$(4, 2)$

Note that the sample size required is the same as for a single component according to $\xi^*_{I,1}$. Again, this design can be combined with the Hadamard structure which results in saturated D–optimal designs for $K = 4, 8, 16, 24, \ldots$ components.

Similarly, a D-optimal design of sample size 10 can be constructed for 2 components with 5 levels each:

(i_1, j_1)	(i_2, j_2)
$(1, 2)$	$(5, 1)$
$(1, 5)$	$(1, 2)$
$(1, 3)$	$(2, 4)$
$(2, 4)$	$(3, 1)$
$(1, 4)$	$(4, 5)$
$(4, 5)$	$(4, 1)$
$(2, 3)$	$(5, 2)$
$(2, 5)$	$(2, 3)$
$(3, 5)$	$(3, 4)$
$(3, 4)$	$(5, 3)$

We conclude this section by a few examples of small sample size D–optimal designs when there are 2 components with different numbers of levels.

The first example gives a saturated D–optimal design for $I_1 = 3$ and $I_2 = 2$ levels:

(i_1, j_1)	(i_2, j_2)
$(1,2)$	
$(2,3)$	$(1,2)$
$(3,1)$	

More generally, for I_1 odd and $I_2 = 2$ there is a D–optimal design with $\frac{1}{2} I_1(I_1 - 1)$ comparisons generated from the balanced marginal design $\xi^*_{I_1,2}$. Similarly, we can generate a 21 run design for $I_1 = 7$ and $I_2 = 3$ levels:

(i_1, j_1)	(i_2, j_2)
$(1,2)$	
$(2,3)$	
$(3,4)$	
$(4,5)$	$(1,2)$
$(5,6)$	
$(6,7)$	
$(7,1)$	

(i_1, j_1)	(i_2, j_2)
$(1,3)$	
$(3,5)$	
$(5,7)$	
$(7,2)$	$(2,3)$
$(2,4)$	
$(4,6)$	
$(6,1)$	

(i_1, j_1)	(i_2, j_2)
$(1,4)$	
$(4,7)$	
$(7,3)$	
$(3,6)$	$(3,1)$
$(6,2)$	
$(2,5)$	
$(5,1)$	

7 Profile strength

As can be seen from the previous sections the D–optimal designs necessarily require that in each individual comparison the alternatives differ in every component. But due to the limited information processing capacity (Miller, 1956) in psychological applications it is usually too difficult for a judge to express his preference for two alternatives which differ with respect to a large number of components. Therefore, in this setting, only a prespecified maximal number of components, the profile strength, is allowed to vary simultaneously within a single paired comparison. In our applications the profile strength J usually equals 2 or 3.

The variety of all comparisons which are allowed in this restricted setting is given by those pairs which differ by not more than J components, $\mathcal{X}(J) = \{((i_1, \ldots, i_K), (j_1, \ldots, j_K));\ i_k \neq j_k$ for, at most, J components$\}$. In this section we are looking for D–optimal designs when the design region is restricted to $\mathcal{X}(J)$.

As D–optimal designs tend to be located on comparisons with a maximum number of alternate components we concentrate on the settings which exhaust the profile strength J, $\mathcal{X}^*(J) = \{((i_1, \ldots, i_K), (j_1, \ldots, j_K));\ i_k \neq j_k$ for exactly J components$\}$. For the uniform design $\xi^{(K)}_{I,\mathcal{X}^*(J)}$ on $\mathcal{X}^*(J)$ the minimum required sample size is $N(\xi^{(K)}_{I,\mathcal{X}^*(J)}) = \binom{K}{J} I^J (I-1)^J$ and the corresponding information matrix becomes

$$M(\xi^{(K)}_{I,\mathcal{X}^*(J)}) = \frac{J}{K}\,\mathbf{Id}_K \otimes M(\xi^*_{I,0}).$$

For every pair $(i,j) = ((i_1,\dots,i_K),(j_1,\dots,j_K)) \in \mathcal{X}(J)$ the variance function breaks down to the sum of its marginal counter–parts,

$$\begin{aligned}
&(f(i)-f(j))^\top M(\xi^{(K)}_{I,\mathcal{X}^*(J)})^{-1}(f(i)-f(j)) \\
=\;& \frac{K}{J}\sum_{k=1}^{K}(f_k(i_k)-f_k(j_k))^\top M(\xi^*_{I,0})^{-1}(f_k(i_k)-f_k(j_k)).
\end{aligned}$$

By the definition of $\mathcal{X}(J)$ at most J terms in this sum are non–zero and the other terms are bounded by $I-1$ in view of the Kiefer–Wolfowitz equivalence theorem because of the D–optimality of the marginal design $\xi^*_{I,0}$. Hence, the variance function is bounded by the number of parameters, $K(I-1)$, which proves the D–optimality of the design $\xi^{(K)}_{I,\mathcal{X}^*(J)}$ by a further application of the Kiefer–Wolfowitz equivalence theorem.

The number of design points can be drastically reduced by making use of a cyclic structure in combination with Hadamard matrices. To achieve this we introduce the cyclic band matrix $C_{K,J}$ of size $K\times K$ and bandwidth J. $C_{K,J}$ is specified by its entries $c_{m,n}$ in row m and column n which are equal to 1 if $m \le n \le \min(m+J-1,K)$ or if $m+J-1>K$ and additionally $n \le (m+J-1)-K$, and 0 otherwise. For $J=3$ and $K=8$ we present $C_{K,J}$ for illustrative purposes

$$C_{8,3} = \begin{pmatrix}
1&1&1&0&0&0&0&0\\
0&1&1&1&0&0&0&0\\
0&0&1&1&1&0&0&0\\
0&0&0&1&1&1&0&0\\
0&0&0&0&1&1&1&0\\
0&0&0&0&0&1&1&1\\
1&0&0&0&0&0&1&1\\
1&1&0&0&0&0&0&1
\end{pmatrix}.$$

Now, $C_{K,J}$ has profile strength J in the sense that an entry $c_{m,n}=1$ shows that the component n is different for the alternatives in run m, whereas the components are equal if $c_{m,n}=0$.

Similarly, we construct a cyclic Hadamard type $Kg_J \times K$ matrix $B_{K,J}$ of profile strength J by defining its $K\times K$ blocks $b_{m,n}$ of size $g_J \times 1$, $B_{K,J} = (b_{m,n})_{m,n=1,\dots,K}$. To achieve this let $h_1,\dots,h_J$ be J distinct columns of a

Table 2: Required sample sizes for reduced D–optimal designs compared to uniform designs; profile strength 2

I	2		3		7	
K	4	8	4	8	4	8
uniform	24	112	216	1 008	10 584	49 392
reduced	4	8	12	24	84	168

Hadamard matrix of order g_J. We set $b_{m,n} = h_{n-m+1}$ for $n \geq m$ and $b_{m,n} = h_{n-m+K+1}$ for $n < m$ respectively, when the corresponding entry $c_{m,n}$ of the cyclic band matrix $C_{K,J}$ equals 1, and $b_{m,n} = 0$ otherwise. Note that each row of $B_{K,J}$ contains exactly J non–zero entries and each column contains exactly J non–zero blocks of length g_J. Moreover, the columns of $B_{K,J}$ are orthogonal and, hence, $B_{K,J}^\top B_{K,J} = J g_J \, \mathbf{Id}_K$.

We are now in the situation to define the $\frac{1}{2} K g_J I(I-1) \times K(I-1)$ design matrix $A_{I,K,J} = B_{K,J} \otimes A_I$ which has profile strength J for each of its comparisons. The corresponding uniform design $\xi_{I,K,J}$ is D–optimal because its information matrix coincides with $M(\xi^{(K)}_{I,\mathcal{X}^*(J)})$. However, the design $\xi_{I,K,J}$ requires a substantially smaller sample size, $N(\xi_{I,K,J}) = \frac{1}{2} K g_J I(I-1)$. In particular, for the minimal profile strength $J = 1$ we obtain a D–optimal design with sample size $\frac{1}{2} K I(I-1)$.

Finally, if L is a common divisor of K and J then the cyclic columns in $B_{K,J}$ can be replaced by $\frac{K}{L}$ cyclic blocks covering L columns each. By this method the sample size can be reduced to $\frac{1}{2}\frac{K}{L} g_J I(I-1)$.

The total reduction in sample sizes achievable by the above mentioned methods is illustrated in Tables 2 and 3.

Table 3: Required sample sizes for reduced D–optimal designs compared to uniform designs; profile strength 3

I	2		3		7	
K	4	8	4	8	4	8
uniform	32	448	864	12 096	296 352	$4.1 \cdot 10^6$
reduced	16	32	48	96	336	672

8 Concluding Remarks

In the present note we have investigated D–optimal designs for paired comparisons for situations where the alternatives are modeled by a K–way layout without interactions. In regression models this problem was treated by van Berkum (1987) and Großmann et al. (2000), before. Our main emphasis was on the reduction of the sample size required in the spirit of fractional factorials. This can be regarded as a complexity problem of minimizing the required sample size relatively to the number of parameters in dependence on the number of components involved (see Riccomagno et al., 1997). Here we could construct D–optimal designs for which the support increases linearly in the number of parameters when the number of levels is fixed and the number of components is growing.

The reduction methods also allow for generating D–optimal designs with a sufficiently small sample size when a constraint on the profile strength is employed.

Acknowledgement. This work was partly supported by the research grant Ho 1286/2-1 of the Deutsche Forschungsgemeinschaft.

References

Atkins, J. E. and Cheng, C.-S. (1999), Optimal regression designs in the presence of random block effects. *J. Statist. Plann. Inference* **77**, 321–335

Baier, D. and Gaul, W. (1999), Optimal product positioning based on paired comparison data. *Journal of Econometrics* **89**, 365–392

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

Großmann, H., Holling, H., Graßhoff, U. and Schwabe, R. (2000), Efficient designs for paired comparisons with a polynomial factor. In *Optimum Design 2000* (Atkinson, A. C., Bogacka, B. and Zhigljavsky, A. A., eds.) Kluwer (to appear)

Hedayat, A. S., Sloane, N. J. A. and Stufken, J. (1999), *Orthogonal Arrays. Theory and Applications.* New York, Springer

Kiefer, J. C. (1959), Optimal experimental designs. *J. R. Statist. Soc., Ser. B* **21**, 272–304

Kiefer, J. and Wolfowitz, J. (1960), The equivalence of two extremum problems. *Can. J. Math.* **12**, 363–366

Miller, G. A. (1956), The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* **63**, 81–97

Offen, W. W. and Littell, R. C. (1987), Design of paired comparison experiments when treatments are levels of a single quantitative variable. *J. Statist. Plann. Inference* **15**, 331–346.

Rafajłowicz, E. and Myszka, W. (1992), When product type experimental design is optimal? – Brief survey and new results. *Metrika* **39**, 321–333

Raghavarao, D. (1971), *Constructions and Combinatorial Problems in Design of Experiments.* New York, Wiley

Riccomagno, E., Schwabe, R. and Wynn, H. P. (1997), Lattice–based D–optimum design for Fourier regression. *Ann. Statist.* **25**, 2313–2327

Schwabe, R. (1996), *Optimum Designs for Multi–Factor Models. Lecture Notes in Statistics* **113**. New York, Springer

Shah, K. R. and Sinha, B. K. (1989), *Theory of Optimal Designs. Lecture Notes in Statistics* **54**. New York, Springer

van Berkum, E. E. M. (1987), *Optimal Paired Comparison Designs for Factorial Experiments*, *CWI Tract* **31**