Optimal Design in the Presence of Random Block Effects at Baseline

Jesús Alonso Cabrera
Universidad del Norte, Barranquilla

Rainer Schwabe
Otto-von-Guericke-Universität Magdeburg

Abstract

Optimal experimental designs are developed for linear regression models with both qualitative and quantitative factors of influence. In particular we generate a characterization of optimal designs for random blocked regression experiments where under few assumptions, this characterization allows to find the weights of the optimal design analytically by means of convex optimization. It is worth-while noting that the optimal weights depend on the ratio of the variance components. However, in this context, we show that in practical applications limiting optimal designs show a high efficiency, when the variance ratio approaches zero or infinity.

Keywords: D-optimal designs, random block effects, partial interactions, qualitative and quantitative factors, product type design.

1 Introduction

In the framework of statistical experiments the theory of optimal designs has been developed over the last decades. In general the subject of this theory is that for an appropriate model the experimental settings should be chosen according to certain criteria with statistical meaning, when we want to put special emphasis on the quality of the parameter estimates. More specifically it happens very often that for a more realistic analysis of the data a regression experiment has to be designed which involves both qualitative and quantitative factors of influence. For example such general two-factor linear models includes intra-class regression with identical partial models in each class (see Searle, 1971, p 355 or Kurotschka, 1984), blocked response surface designs (see Gilmour and Trinca, 2000, Goos, 2002, and Waite and Woods, 2012).

The problem of constructing optimal experimental designs for estimating the parameter vector in a two-factor linear model is more complex than for single factor models. However, the question under what conditions we can find optimal designs for two-factor models in terms of optimal designs for their single factor models has been considered in the literature. For example, in multi-factor models with homoscedastic errors Schwabe (1996) presents optimal designs for a large variety of cases. Based on these ideas the aim of the present work is to generate D-optimal designs for multi-factor models in the additional presence of random block effects. Of particular interest is the limiting behavior, when the variance of the random effects gets large or close to zero. Moreover, we apply essentially analytical methods that involve convex optimization in a continuous setup to models which involve a discrete structure, a fact that has not been treated in the literature with few exceptions.

This paper is organized as follows: In section 2 we introduce the random block effect models with baseline considered throughout the paper. Section 3 provides the corresponding information matrices and the essential structure of optimal designs. In section 4 we give a characterization of D-optimal designs in terms of marginal properties of the within block arrangements. These findings are illustrated by an example on polynomial regression with common baseline and random block effects in Section 5, and the paper is concluded by a discussion of the results and an outlook to possible future research.

2 The random block effects model

When an experiment includes both quantitative and qualitative factors, the effects between the quantitative and qualitative factors should be taken into consideration. We consider the particular situation of a two-factor model with v treatments $k = 1, \ldots, v$ and an additional explanatory variable x with common intercept at $x = 0$ for all treatments The jth observation at treatment k is given by

$$Y_{kj} = \beta_0 + \mathbf{f}(x_{kj})^\top \boldsymbol{\beta}_k + \varepsilon_{kj}, \tag{1}$$

for $j = 1, \ldots, n_k$, where n_k is the number of observations for treatment k, $x_{kj} \in \mathcal{X}$ is the corresponding setting of the explanatory variables, $\mathbf{f}(x) = (f_1(x), \ldots, f_q(x))^\top$ denotes the vector of (regression functions, $\beta_0 \in \mathbb{R}$ and $\boldsymbol{\beta}_k = (\beta_{k1}, \ldots, \beta_{kq})^\top \in \mathbb{R}^q$ are the unknown parameters. Here β_0 denotes a common intercept and the other parameters $\boldsymbol{\beta}_k$ may vary with the treatment k. Furthermore, while the regression function $\mathbf{f}(x)$ is the same across all treatment levels, the quantitative variable x may be chosen for each k independently. Finally, it is assumed that the error terms ε_{kj} are uncorrelated with zero means and equal variances σ^2. The model (1) and corresponding optimal designs had been proposed by Buonaccorsi and Iyer (1986) for linear regression with two treatments and had been extended to $v > 2$ treatments by Schwabe (1996, Section 7).

In this work, we focus on the construction of designs when the response depends on a random blocking variable. Hence we suppose that for each treatment k the n_k observations are blocked into b_k blocks of size m. The block effects are assumed to be random. Then the jth observation in block i for treatment k at setting x_{kij} for the explanatory variable has the form

$$Y_{kij} = \beta_0 + \gamma_{ki} + \mathbf{f}(x_{kij})^\top \boldsymbol{\beta}_k + \varepsilon_{kij}\,, \tag{2}$$

where γ_{ki} is the ith random block effect for treatment k and ε_{kij} is a random error as described above. The random effects are assumed to be uncorrelated to each other and to the random errors with zero means and common variance σ_γ^2.

We can write the vector $\mathbf{Y}_{ki} = (Y_{ki1}, \ldots, Y_{kim})^\top$ of observations of block i for treatment k as

$$\mathbf{Y}_{ki} = \mathbf{1}_m \beta_0 + \mathbf{F}_{ki} \boldsymbol{\beta}_k + \mathbf{1}_m \gamma_{ki} + \boldsymbol{\varepsilon}_{ki}\,, \tag{3}$$

where $\mathbf{1}_m$ denotes a vector of length m with all entries equal to 1, $\mathbf{F}_{ki} = (\mathbf{f}(x_{ki1}), ..., \mathbf{f}(x_{kim}))^\top$ is the design matrix of block i in group k, and $\boldsymbol{\varepsilon}_{ki}$ is the corresponding vector of errors. Then the within block covariance matrix $\mathrm{Cov}(\mathbf{Y}_{ki}) = \sigma^2 \mathbf{V}$ is completely symmetric and equal for all blocks, where

$$\mathbf{V} = \mathbf{I}_m + d \mathbf{1}_m \mathbf{1}_m^\top\,, \tag{4}$$

$d = \sigma_\gamma^2/\sigma^2$ is the variance ratio and $\mathbf{I}_m$ denotes the $m \times m$ identity matrix.

If we stack the observations $\mathbf{Y}_k = (\mathbf{Y}_{k1}^\top, \ldots, \mathbf{Y}_{kb_k}^\top)^\top$ for treatment k we obtain

$$\mathbf{Y}_k = \mathbf{1}_{n_k}\beta_0 + \mathbf{F}_k\boldsymbol{\beta}_k + (\mathbf{I}_{b_k} \otimes \mathbf{1}_m)\boldsymbol{\gamma}_k + \boldsymbol{\varepsilon}_k\,, \tag{5}$$

where $\mathbf{F}_k = (\mathbf{F}_{k1}^\top, \ldots, \mathbf{F}_{kb_k}^\top)^\top$ and $\boldsymbol{\gamma}_k = (\gamma_{k1}, \ldots, \gamma_{kb_k})^\top$ are the stacked design matrix and the vector of random block effects for treatment k, respectively, and "$\otimes$" denotes the Kronecker product of matrices.

Finally, the vector $\mathbf{Y} = (\mathbf{Y}_1^\top, \ldots, \mathbf{Y}_v^\top)^\top$ of all observations across the v treatments can be written as

$$\mathbf{Y} = \mathbf{1}_N\beta_0 + \mathrm{diag}(\mathbf{F}_k)\boldsymbol{\beta} + (\mathbf{I}_B \otimes \mathbf{1}_m)\boldsymbol{\gamma} + \boldsymbol{\varepsilon}\,, \tag{6}$$

where $N = \sum_{k=1}^v n_k$ is the total number of observations, $B = \sum_{k=1}^v b_k$ the total number of blocks, $\mathrm{diag}(\mathbf{F}_k)$ is the block diagonal design matrix, $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \ldots, \boldsymbol{\beta}_v^\top)^\top$ the stacked vector of all parameters besides the intercept and $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^\top, \ldots, \boldsymbol{\gamma}_v^\top)^\top$ the stacked random vector of all random block effects across all treatments. The corresponding covariance matrix $\mathrm{Cov}(\mathbf{Y}) = \sigma^2\mathbf{I}_B \otimes \mathbf{V}$ is block diagonal.

3 Information and design

The quality of the estimator for the fixed effects β_0 and $\boldsymbol{\beta}$ is usually measured in terms of its covariance matrix which is proportional to the inverse of the information matrix and depends on the values of the explanatory variables. If these settings are under the disposition of the experimenter, the quality can be improved by an optimal choice of the design.

For a single block the contribution to the information for β_0 and $\boldsymbol{\beta}_k$ amounts to

$$\mathcal{M}_{ki} = \begin{pmatrix} \mathbf{1}_m^\top\mathbf{V}^{-1}\mathbf{1}_m & \mathbf{1}_m^\top\mathbf{V}^{-1}\mathbf{F}_{ki} \\ \mathbf{F}_{ki}^\top\mathbf{V}^{-1}\mathbf{1}_m & \mathbf{F}_{ki}^\top\mathbf{V}^{-1}\mathbf{F}_{ki} \end{pmatrix}. \tag{7}$$

Note that $\mathbf{V}^{-1} = \mathbf{I}_m - d/(1+md)\mathbf{1}_m\mathbf{1}_m^\top$. Then $\mathbf{V}^{-1}\mathbf{1}_m = 1/(1+md)\mathbf{1}_m$ and $\mathbf{1}_m^\top\mathbf{V}^{-1}\mathbf{1}_m = m/(1+md)$ such that the marginal entries in the information matrix simplify.

The information depends only on the s_{ki} distinct settings $x_{ki1}, ..., x_{kis_{ki}}$, say, and their corresponding numbers $n_{ki1}, ..., n_{kis_{ki}}$ of replications in the ith block for treatment k. Let $\mathbf{M}_{ki} = \frac{1}{m}\mathbf{F}_{ki}^\top\mathbf{1}_m = \frac{1}{m}\sum_{j=1}^{s_{ki}} n_{kij}\mathbf{f}(x_{kij})\mathbf{f}(x_{kij})^\top$

denote the information per observation for $\boldsymbol{\beta}_k$ in the fixed effects model (1) and by $\mathbf{m}_{ki} = \frac{1}{m}\mathbf{F}_{ki}^\top \mathbf{F}_{ki} = \frac{1}{m}\sum_{j=1}^{s_{ki}} n_{kij}\mathbf{f}(x_{kij})$ the corresponding weighted mean of the regression function. Then the contribution of a single block to the information (7) associated with β_0 and $\boldsymbol{\beta}_k$ can be rewritten as

$$\mathcal{M}_{ki} = \frac{m}{1+md}\begin{pmatrix} 1 & \mathbf{m}_{ki}^\top \\ \mathbf{m}_{ki} & (1+md)\mathbf{M}_{ki} - md\mathbf{m}_{ki}\mathbf{m}_{ki}^\top \end{pmatrix}. \tag{8}$$

The joint information based on all observations is then

$$\mathcal{M} = \frac{m}{1+md}\begin{pmatrix} B & \sum_{i=1}^{c_1} b_{1i}\mathbf{m}_{1i}^\top \quad \cdots & \sum_{i=1}^{c_v} b_{vi}\mathbf{m}_{vi}^\top \\ \sum_{i=1}^{c_1} b_{1i}\mathbf{m}_{1i} & & \\ \vdots & \multicolumn{2}{c}{\mathrm{diag}(\sum_{i=1}^{c_k} b_{ki}((1+md)\mathbf{M}_{ki} - md\mathbf{m}_{ki}\mathbf{m}_{ki}^\top))} \\ \sum_{i=1}^{c_v} b_{vi}\mathbf{m}_{vi} & & \end{pmatrix} \tag{9}$$

for the whole set of parameters β_0 and $\boldsymbol{\beta}$, where b_{ki} is the number of blocks with design matrix $\mathbf{F}_{ki}$ and $\mathbf{F}_{k1}, ..., \mathbf{F}_{kc_k}$ are c_k distinct design matrices within treatment k.

For measuring the quality of a design, i. e. the choice of settings x_{kij}, one has to apply a real-valued functional on the information matrix. We will use here the most common criterion of D-optimality which aims at maximization the determinant $\det(\mathcal{M})$ of the information matrix which is equivalent to minimization of the volume of the confidence ellipsoid under normality assumption.

Because this discrete optimization problem is to hard, in general, we adopt the concept of approximate designs in the spirit of Kiefer (1974). This means that the numbers n_{kij} of replications of a setting x_{kij} within a block ki, the numbers b_{ki} of equal blocks within a treatment k and the numbers b_k of blocks are relaxed to be non-negative real numbers satisfying $\sum_{j=1}^{s_{ki}} n_{kij} = m$, $\sum_{i=1}^{c_k} b_{ki} = b_k$ and $\sum_{k=1}^{v} b_k = B$, respectively. Optimization in this continuous setup is much easier. If the obtained optimal solutions are not integer, efficient exact (integer) designs can often be obtained from the optimal (continuous) solutions by appropriate rounding.

By convexity arguments (cf. Schmelter, 2007) symmetry considerations lead to optimal designs in the continuous setup which are invariant with respect to permutations of the treatments as well as permutations of the blocks within treatments. This mean, in particular that for such an optimal solution all numbers of blocks within each treatment are equal ($b_1 = ... = b_v = B/v$) and that all within block designs coincide ($c_1 = ... = c_v = 1$,

$s_{ki} = s$, $x_{kij} = x_j$ and $n_{kij} = n_j$ for all k and i). We, thus, only have to find an optimal approximate within block design.

In agreement with standard notation we define a standardized (within block) design by

$$\xi = \begin{pmatrix} x_1 & x_2 & \dots & x_s \\ w_1 & w_2 & \dots & w_s \end{pmatrix}, \tag{10}$$

where $x_1, ..., x_s \in \mathcal{X}$ are mutually distinct experimental settings for the explanatory variable x and $w_1, ..., w_s$ are the corresponding weights satisfying $\sum_{j=1}^{s} w_j = 1$. While for exact designs these weights are the relative frequencies of replications $w_j = n_j/m$ within the block, they need not be multiple of $1/m$ for an approximate design.

Similar to exact within block designs we define for the approximate design ξ by $\mathbf{M}(\xi) = \sum_{j=1}^{s} w_j \mathbf{f}(x_j)\mathbf{f}(x_j)^\top$ the information per observation for $\boldsymbol{\beta}_k$ in the fixed effects model (1) and by $\mathbf{m}(\xi) = \sum_{j=1}^{s} w_j \mathbf{f}(x_j)$ the corresponding weighted mean of the regression function. With this notation we can define the standardized information (per observation) $\mathcal{M}(\xi)$ for the design with equal numbers of blocks per treatment and equal within block designs ξ in each block by

$$\mathcal{M}(\xi) = \frac{1}{1+md} \begin{pmatrix} 1 & \frac{1}{v}\mathbf{1}_v^\top \otimes \mathbf{m}(\xi)^\top \\ \frac{1}{v}\mathbf{1}_v \otimes \mathbf{m}(\xi) & \frac{1}{v}\mathbf{I}_v \otimes \mathcal{M}_{22}(\xi) \end{pmatrix}, \tag{11}$$

where

$$\mathcal{M}_{22}(\xi) = (1+md)\mathbf{M}(\xi) - md\mathbf{m}(\xi)\mathbf{m}(\xi)^\top \tag{12}$$

is the diagonal block associated with each treatment effect $\boldsymbol{\beta}_k$ in the information matrix. For exact designs the definition (11) is in accordance with (9) in the sense that then due to the standardization per observation $\mathcal{M}(\xi) = \frac{1}{N}\mathcal{M}$.

4 Characterization of optimal designs

To give a condition for the D-optimality of an invariant design generated by a marginal within block design ξ we note that

$$\mathcal{M}_1(\xi) = \frac{1}{1+md} \begin{pmatrix} 1 & \mathbf{m}(\xi)^\top \\ \mathbf{m}(\xi) & \mathcal{M}_{22}(\xi) \end{pmatrix} \tag{13}$$

is the information matrix associated with the mixed model (2) in the case of $v = 1$ treatment.

By standard inversion rules for partitioned matrices the determinant of the information matrix $\mathcal{M}(\xi)$ can be derived as

$$\begin{aligned} \det(\mathcal{M}(\xi)) &= c_1 \left(1 - \mathbf{m}(\xi)^\top \mathcal{M}_{22}(\xi)^{-1}\mathbf{m}(\xi)\right) \left(\det(\mathcal{M}_{22}(\xi))\right)^v \\ &= c_2 \det(\mathcal{M}_1(\xi)) \left(\det(\mathcal{M}_{22}(\xi))\right)^{v-1} \end{aligned} \tag{14}$$

provided that $\mathcal{M}_{22}(\xi)$ is non-singular, where c_1 and c_2 are generic constants not depending on ξ. Otherwise, when $\mathcal{M}_{22}(\xi)$ is singular, the determinant of the information matrix $\mathcal{M}(\xi)$ is equal to 0. This establishes the following result.

Theorem 1 *In the mixed model (2) with random block effects and common baseline the invariant design with equal numbers of blocks per treatment and equal within block designs ξ^* is D-optimal, if ξ^* maximizes*

$$\det(\mathcal{M}_1(\xi)) \left(\det(\mathcal{M}_{22}(\xi))\right)^{v-1} . \tag{15}$$

In view of Theorem 1 the D-optimality criterion in the mixed model can be identified as a compound criterion for the within block design ξ by taking logarithms.

The first term in (15) can be rewritten as

$$\det(\mathcal{M}_1(\xi)) = \frac{1}{1+md} \det(\mathbf{M}(\xi) - \mathbf{m}(\xi)\mathbf{m}(\xi)^\top) = \frac{1}{1+md} \det(\mathcal{M}_0(\xi)) \tag{16}$$

by replicate application of the standard formula for the determinant of a partitioned matrix, where

$$\mathcal{M}_0(\xi)) = \begin{pmatrix} 1 & \mathbf{m}(\xi)^\top \\ \mathbf{m}(\xi) & \mathbf{M}(\xi) \end{pmatrix} \tag{17}$$

is the information matrix in the fixed effects model (1) which does not depend on the variability of the blocks. Hence. we can recover the well-known fact that for $v = 1$ treatment the D-optimal design for the fixed effects model retains its optimality also in the presence of random block effects.

5 Example: Polynomial regression

To illustrate the usefulness of the above approach we restrict to polynomial regression for the explanatory variable with baseline at $x = 0$. To be more specific we consider the mixed effects model

$$Y_{kij} = \beta_0 + \gamma_{ki} + \beta_{k1} x_{kij} + \beta_{k2} x_{kij}^2 + ... + \beta_{kq} x_{kij}^q + \varepsilon_{kij} \tag{18}$$

with the regression function $\mathbf{f}(x) = (x, x^2, ..., x^q)^\top$ of monomials up to degree q on a non-negative design region $\mathcal{X} = [0, x_{\max}]$.

Using standard arguments of the theory of Equivalence Theorems for approximate designs (cf. Pukelsheim, 1993, Chapter 9) we can deduce that the D-optimal within block design ξ^* has minimal support, i. e. $s = q + 1$, including the two end-points $x_1^* = 0$ and $x_{q+1}^* = x_{\max}$. Hence, we have only to consider designs of the form

$$\xi = \begin{pmatrix} 0 & x_2 & \dots & x_q & x_{q+1} \\ w_1 & w_2 & \dots & w_q & w_{q+1} \end{pmatrix}. \tag{19}$$

We will make use of the fact that for designs with minimal support the information matrix can be decomposed into square matrices which depend either on the design points $x_1, ..., x_{q+1}$ or on the weights $w_1, ..., w_{q+1}$. First we note that $\mathbf{f}(0) = \mathbf{0}$, where $\mathbf{0}$ is a vector of appropriate length will all entries equal to 0. Denote by $\mathbf{F}_{-0} = (\mathbf{f}(x_2), ..., \mathbf{f}(x_{q+1}))^\top$ and $\mathbf{W}_{-0} = \mathrm{diag}(w_2, ..., w_{q+1})$ the design and weight matrix, respectively, associated with the design ξ reduced by the first support point $x_1 = 0$. Then the information matrix of ξ in the fixed effects model can be written as

$$\mathcal{M}_0(\xi) = \begin{pmatrix} 1 & \mathbf{1}_q^\top \\ \mathbf{0} & \mathbf{F}_{-0}^\top \end{pmatrix} \begin{pmatrix} w_1 & \mathbf{0}^\top \\ \mathbf{0} & \mathbf{W}_{-0} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{0}^\top \\ \mathbf{1}_q & \mathbf{F}_{-0} \end{pmatrix}. \tag{20}$$

All occurring matrices are square and, hence, the determinant of the information matrix factorizes

$$\det(\mathcal{M}_0) = w_1 \det(\mathbf{W}_{-0}) \det(\mathbf{F}_{-0})^2. \tag{21}$$

Similarly, the determinant of the second component $\mathcal{M}_{22}(\xi)$ in condition (15) of Theorem 1 can be calculated to

$$\det(\mathcal{M}_{22}(\xi)) = (1 - \frac{md}{1+md}\mathbf{m}(\xi)^\top \mathbf{M}(\xi)^{-1}\mathbf{m}(\xi)) \det((1+md)\mathbf{M}(\xi)) \tag{22}$$

by the standard formula for the determinant of a sum (difference) of matrices. As $\mathbf{M}(\xi) = \mathbf{F}_{-0}^\top \mathbf{W}_{-0} \mathbf{F}_{-0}$ and $\mathbf{m}(\xi) = \mathbf{F}_{-0}^\top \mathbf{W}_{-0} \mathbf{1}_q$ we obtain

$$\begin{aligned} \mathbf{m}(\xi)^\top \mathbf{M}(\xi)^{-1}\mathbf{m}(\xi) &= \mathbf{1}_q^\top \mathbf{W}_{-0}\mathbf{F}_{-0} \left(\mathbf{F}_{-0}^{-1}\mathbf{W}_{-0}^{-1}(\mathbf{F}_{-0}^\top)^{-1}\right) \mathbf{F}_{-0}^\top \mathbf{W}_{-0}\mathbf{1}_q \\ &= \mathbf{1}_q^\top \mathbf{W}_{-0}\mathbf{1}_q \;=\; 1 - w_1 \end{aligned} \tag{23}$$

and hence

$$\det(\mathcal{M}_{22}(\xi)) = (1+md)^{q-1}(1+mdw_1)\det(\mathbf{W}_{-0})\det(\mathbf{F}_{-0})^2. \tag{24}$$

Combining all these results we get

$$\det(\mathcal{M}(\xi)) = cw_1(1+mdw_1)^{v-1}\det(\mathbf{W}_{-0})^v\det(\mathbf{F}_{-0})^{2v}, \tag{25}$$

where c is some constant not depending on ξ. Because this determinant factorizes into a factor $w_1(1+mdw_1)^{v-1}\det(\mathbf{W}_{-0})^v$ which only depends on the weights and a factor $\det(\mathbf{F}_{-0})^{2v}$ purely related to the design points, these components can be optimized separately.

Obviously the latter part does not depend on the further model parameters m and d and is optimized by the same design points as in the corresponding fixed effects model. These D-optimal design points can be determined as the roots of certain Legendre polynomials (cf. Pukelsheim, 1993, Chapter 9).

For the first part we notice that the optimal weights will depend on the model parameters through the quantity md. If we fix the weight w_1 at $x_1 = 0$, $\det(\mathbf{W}_{-0})$ is maximized, when all other weights are equal, $w_2, ..., w_{q+1} = (1-w_1)/q$. Taking this into account we obtain for the first factor $w_1(1+mdw_1)^{v-1}(1-w_1)^{vq}$ which is maximized by

$$w_1^* = \frac{mdv - vq - 1 + \sqrt{(mdv - vq - 1)^2 + 4mdv(q+1)}}{2mdv(q+1)}. \tag{26}$$

This establishes the following result

Theorem 2 *Let $\tilde{\xi}^*$ be the D-optimal design in the fixed effects polynomial regression model on the design region $\mathcal{X} = [0, x_{\max}]$ with support points $x_1* = 0$ and $x_2^*, ..., x_{q+1}^*$ and equal weights $1/(q+1)$. Then the design with equal numbers of blocks per treatment and equal within block designs*

$$\xi^* = \begin{pmatrix} 0 & x_2^* & \dots & x_q^* & x_{q+1}^* \\ w_1^* & w_2^* & \dots & w_q^* & w_{q+1}^* \end{pmatrix} \tag{27}$$

with w_1^* given by (26) and $w_2^* = ... = w_{q+1}^* = (1-w_1^*)/q$ is D-optimal for the polynomial mixed model (18) with baseline.

In the particular case of straight line regression ($q = 1$, $f(x) = x$) we recover from Theorem 2 the D-optimal designs proposed by Schwabe and Schmelter

(2008). For $v = 1$ treatment and higher degree q Theorem 2 is related to Atkins and Cheng (1999).

It is easy to check that in general the optimal weight w_1^* at $x_1^* = 0$ increases in the variance ratio d from $1/(vq+1)$ for $d = 0$ to $1/(q+1)$ for $d \to \infty$. In particular, we may derive for $d = 0$ the D-optimal design for the corresponding fixed effects model without block effects.

Corollary 3 *Consider the polynomial model*

$$Y_{kj} = \beta_0 + \beta_{k1}x_{kj} + \beta_{k2}x_{kj}^2 + ... + \beta_{kq}x_{kj}^q + \varepsilon_{kj} \tag{28}$$

with common baseline. Let $\tilde{\xi}^*$ *be the* D*-optimal design in the fixed effects polynomial regression model on the design region* $\mathcal{X} = [0, x_{\max}]$ *with support points* $x_1* = 0$ *and* $x_2^*, ..., x_{q+1}^*$*, and let* ζ^* *be the design which assigns equal weights* $1/(vq+1)$ *to each of the* vq *combination of the* v *treatments with the* q *non-zero design points* $x_2^*, ..., x_{q+1}^*$ *as well as to the common baseline* $x_1^* = 0$ *(no treatment effect). Then* ζ^* *is* D*-optimal.*

On the other hand in the limiting case $d \to \infty$ the determinant of the information matrix, suitably standardized, tends to that in the model with fixed rather than random block effects, where each treatment group has to be considered separately.

To illustrate the general dependence of the optimal weight w_1^* on the variance ratio d we look at the particular situation of a quadratic model ($q = 2$, $\mathbf{f}(x) = (x, x^2)^\top$) at baseline for $x \in [0, 1]$ and $v = 2$ treatments. By Theorem 2 the D-optimal within block design

$$\xi_{md}^* = \begin{pmatrix} 0 & 1/2 & 1 \\ w_{md}^*(0) & w_{md}^*(1) & w_{md}^*(2) \end{pmatrix} \tag{29}$$

with

$$w_{md}^*(0) = \frac{1}{12md}\left(2md - 5 + \sqrt{(2md-5)^2 + 24md}\right) \tag{30}$$

and $w_{md}^*(1) = w_{md}^*(2) = (1 - w_{md}^*(0))/2$, where we have slightly changed the notation and the subscript md indicates the dependence of the optimal weights on the product of the block size and the variance ratio.

For this situation the optimal weight $w_{md}^*(0)$ at the baseline $x = 0$ are exhibited in Figure 1.) There the horizontal axis is rescaled to the intra-class correlation $(md)/(1+md)$ in order to cover the whole range of variance ratios

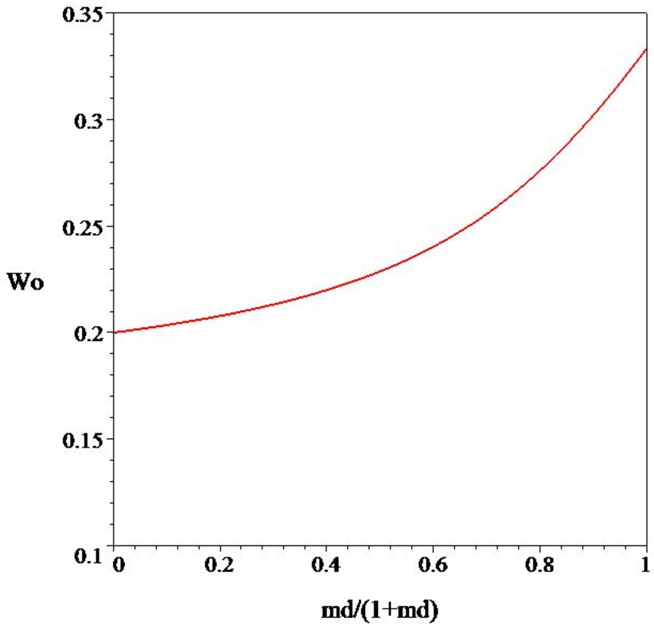

Figure 1: D-optimal weights $w^*_{md}(0)$ in the case of quadratic regression and $v = 2$ treatments

d in a finite interval. This rescaling also suppresses an explicit dependence of the optimal weights on the block size.

The D-optimal designs obtained are locally optimal in the sense that they rely on the value of the magnitude of the variance ratio which is typically not known beforehand. Therefore it is of interest how sensitive the optimal designs are with respect to their performance when the variance ratio is misspecified. This performance will be measured in terms of the D-efficiency

$$\mathrm{eff}_D(\xi) = \left(\frac{\det(\mathcal{M}(\xi))}{\det(\mathcal{M}(\xi^*))} \right)^{1/(vq+1)} \tag{31}$$

which amounts to the number of observations needed when the optimal design ξ^* is used compared to the number of observations for the competing design ξ in order to obtain the same value for the determinant of the information matrix. In Figure 2 we have plotted the D-efficiencies of the limiting optimal designs

$$\xi_0^* = \begin{pmatrix} 0 & 1/2 & 1 \\ 1/5 & 2/5 & 2/5 \end{pmatrix} \tag{32}$$

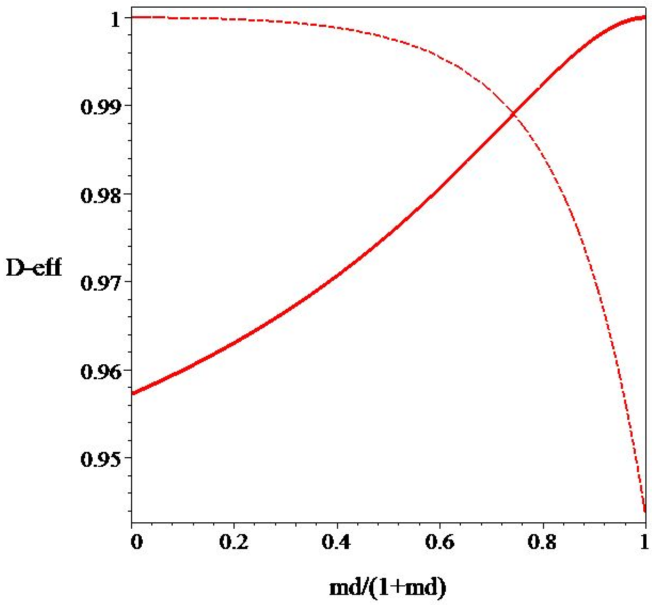

Figure 2: D-efficiencies for ξ_0^* (dashed line) and ξ_∞^* (solid line) in the case of quadratic regression and $v = 2$ treatments

and

$$\xi_\infty^* = \begin{pmatrix} 0 & 1/2 & 1 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}, \tag{33}$$

respectively, in dependence on $\frac{md}{1+md}$.

As can be seen from Figure 2, the D-optimal within block design ξ_0^* based on the linear model without blocks effects is highly efficient over the whole range of variance ratios d with a minimal D-efficiency of 0.9432 when the variance ratio d becomes large. On the other hand the D-optimal fixed block effects design ξ_∞ has also a good performance for all values of the variance ratio d with an even higher minimal D-efficiency of 0.9572 for $d = 0$.

6 Discussion

In this paper we could characterize and develop D-optimal designs for a linear mixed model which incorporates interacting qualitative and quantitative factors of influence in the presence of random block effects and a common baseline. For the illustrating example of polynomial regression within treatments the characterization is based on the equivalence theory for approximate designs besides symmetry considerations. Exact designs have then to be obtained either by suitable rounding the weights of the overall used within block design or by slightly varying the weights between the blocks. It has to be studied which of these strategies is more appropriate and leads to sufficiently efficient designs. Furthermore the dependence of the optimal designs on the variance ratio has to be taken into account and alternatives like maximin efficient or (pseudo) Bayesian optimal designs should be developed which are highly efficient over the whole range of reasonable variance ratios.

Topic of future research should be to investigate how the results elaborated in the example of Section 5 can be extended to situations where either the marginal optimal design is not minimally supported or the baseline is not included in the design region. Also more complicated interaction structures can be considered, when only parts of the regression functions of the quantitative explanatory variable interact with the treatment (see e. g. Schwabe, 1996, Section 6.2 for an extensive listing of corresponding fixed effect models), and the quantitative explanatory variable may be replaced by a categorical (qualitative) one or even by multiple regressors.

References

[1] Atkins J. E. and Cheng, C.-S. (1999). Optimal regression design in the presence of random block effects. *Journal of Statistical Planning and Inference*, **77**, 321–335.

[2] Buonaccorsi, J. P. and Iyer, H. K. (1986). Optimal designs for ratios of linear combinations in the general linear model. *Journal of Statistical Planning and Inference*, **13**, 345–356.

[3] Gilmour, S. G. and Trinca, L. A. (2000). Some practical advice on polynomial regression analysis from blocked response surface designs, *Communications in Statistics: Theory and Methods*, **29**, 2157–2180.

[4] Goos, P. (2002). *The Optimal Design of Blocked and Split-Plot Experiments*. Springer, New York.

[5] Kiefer, J. (1974). General equivalence theory for optimum designs (approximate theory). *The Annals of Statistics*, **2**, 849–879.

[6] Kurotschka, V. G. (1984). A general approach to optimal design of experiments with qualitative and quantitative factors. In: J. K. Ghosh and J. Roy (eds): *Statistics: Applications and New Directions*, 353–368.

[7] Pukelsheim, F. (1993). *Optimal Design of Experiments*. Wiley, New York.

[8] Schmelter, T. (2007). *Experimental Design for Mixed Models with Application to Population Pharmacokinetic Studies*. Dissertation, Faculty of Mathematics, Otto-von-Guericke University, Magdeburg.

[9] Schmelter, T. and Schwabe, R. (2008). On optimal designs in random intercept models. *Tatra Mountains Mathematical Publications*, **39**, 145–153.

[10] Schwabe, R. (1996). *optimal Design for Multi-Factor Models*. Springer, New York.

[11] Searle, S. R. (1971). *Linear Models*. Wiley, New York.

[12] Waite, T. W. and Woods, D. C. (2015). Designs for generalized linear models with random block effects via information matrix approximations. *Biometrika*, **102**, 677–693.