

Prus, Maryna; Filová, Lenka

Article — Published Version

Computational aspects of experimental designs in multiple-group mixed models

Statistical Papers

Provided in Cooperation with:

Springer Nature

Suggested Citation: Prus, Maryna; Filová, Lenka (2023) : Computational aspects of experimental designs in multiple-group mixed models, Statistical Papers, ISSN 1613-9798, Springer, Berlin, Heidelberg, Vol. 65, Iss. 2, pp. 865-886,
<https://doi.org/10.1007/s00362-023-01416-1>

This Version is available at:

<https://hdl.handle.net/10419/316988>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/4.0/>



Computational aspects of experimental designs in multiple-group mixed models

Maryna Prus^{1,2} · Lenka Filová³

Received: 12 April 2022 / Revised: 9 December 2022 / Published online: 13 March 2023

© The Author(s) 2023

Abstract

We extend the equivariance and invariance conditions for construction of optimal designs to multiple-group mixed models and, hence, derive the support of optimal designs for first- and second-order models on a symmetric square. Moreover, we provide a tool for computation of D - and L -efficient exact designs in multiple-group mixed models by adapting the algorithm of Harman et al. (Appl Stoch Models Bus Ind, 32:3–17, 2016). We show that this algorithm can be used both for size-constrained problems and also in settings that require multiple resource constraints on the design, such as cost constraints or marginal constraints.

Keywords Optimal design · Exact design · Random coefficient regression · Equivariance · Invariance · Resource constraints

1 Introduction

The aim of this work is computation of highly efficient experimental designs in multiple-group random coefficient regression models. Analytical approach for determining optimal approximate designs for this type of models has been discussed, i.e., in Fedorov and Jones (2005), Schmelter (2007) and Prus (2022). In Fedorov and Jones (2005), optimal designs were obtained for specific regression functions. Schmelter (2007) proposed optimality conditions in the particular case of group-wise identical designs for commonly used linear and determinant criteria. In Prus (2022), equivalence theorems for the general form of multiple-group models have been formulated.

✉ Maryna Prus
maryna.prus@uni-hohenheim.de

¹ Institute of Crop Science, University of Hohenheim, Stuttgart, Germany

² Division of Statistics and Machine Learning, Linköping University, Linköping, Sweden

³ Faculty of Mathematics, Physics and Informatics, Comenius University in Bratislava, Bratislava, Slovakia

For computing optimal designs in mixed-effect models, several solutions are available. However, most of them focus on computing *approximate* designs for several traditional criteria, and we are not aware of any work that considers additional constraints besides the size of the design.

Namely, Dumont et al. (2018) created an R package for computing designs in mixed-effects model that is predominantly focused on nonlinear models that are used in drug development. However, they only consider the D -optimality criterion without any additional constraints that may arise in the experiment (such as budget, material or other types of resources). The software package (Aliev et al. 2012) is aimed at computing *approximate* designs, mainly for mixed-effects models arising in pharmacokinetic applications. Finally, the software solution of Nyberg et al. (2012) seems to be the most versatile, as it admits user-defined criteria, including their Bayesian versions, but the focus is on approximate designs in nonlinear mixed-effect models and no additional constraints can be included.

In our paper, we propose to use the algorithm of Harman et al. (2016), originally developed for computing D -efficient *exact* designs in the linear regression model with possibly multiple *resource constraints*, for the general form of multiple-group models. We also propose analytical solutions based on equi- and invariance properties of optimal designs for several particular models.

The paper has the following structure: In Sect. 2, we shortly introduce the multiple group model, the design problem and the optimality conditions that are subsequently used in Sect. 3 to show the equivariance and invariance properties of D - and L -optimal designs. In Sect. 4, we show that the problem of computing optimal designs in the multiple-group mixed model can be reformulated as a problem of computing optimal designs with respect to a monotonous criterion function with resource constraints on weights, and, hence, a modification of a recent algorithm for computing resource constrained designs can be used to obtain efficient exact designs in our model. In Sects. 5 and 6, we compute the D - and $IMSE$ -optimal designs in bilinear and quadratic models and show that we can easily solve problems with additional constraints that cannot be solved analytically.

2 Multiple-group RCR model

2.1 Model specification

In multiple-group random coefficient regression model the h -th observation of the j -th observational unit in the i -th group is given by

$$\begin{aligned} \mathbf{Y}_{ijh} &= \mathbf{F}_{(i)}(x_{ih})\boldsymbol{\beta}_{ij} + \boldsymbol{\varepsilon}_{ijh}, \quad x_{ih} \in \mathcal{X}_i, \quad i = 1, \dots, s, \quad j = 1, \dots, n_i, \\ h &= 1, \dots, m_i, \end{aligned} \quad (1)$$

where n_i is the number of observational units in group i , m_i is the number of observations per unit in group i , observational settings x_{ih} come from some experimental region \mathcal{X}_i . In this work we allow for multivariate (l -variate) response and $\mathbf{F}_{(i)}$ denotes a group-specific ($l \times p$) matrix of known regression functions in group i . In the particular

case of univariate response we deal with “classical” regression functions: $\mathbf{F}_{(i)} = \mathbf{f}_{(i)}^\top$, and $l = 1$. Unit-specific random parameters $\boldsymbol{\beta}_{ij} = (\beta_{ij1}, \dots, \beta_{ijp})^\top$ have unknown mean $\boldsymbol{\beta}_0$ and given $(p \times p)$ covariance matrix \mathbf{D}_i , $\boldsymbol{\varepsilon}_{ijh}$ denote observational errors with zero mean and non-singular $(l \times l)$ covariance matrix Σ_i . All observational errors and all random parameters are assumed to be uncorrelated.

The covariance matrix of the best linear unbiased estimator for $\boldsymbol{\beta}_0$ is given by

$$\text{Cov}(\hat{\boldsymbol{\beta}}_0) = \left[\sum_{i=1}^s n_i ((\tilde{\mathbf{F}}_i^\top \tilde{\mathbf{F}}_i)^{-1} + \mathbf{D}_i)^{-1} \right]^{-1}, \quad (2)$$

where $\tilde{\mathbf{F}}_i = (\tilde{\mathbf{F}}_{(i)}^\top(x_{i1}), \dots, \tilde{\mathbf{F}}_{(i)}^\top(x_{im_i}))^\top$ for $\tilde{\mathbf{F}}_{(i)}(x_{ih}) = \Sigma_i^{-1/2} \mathbf{F}_{(i)}(x_{ih})$, $h = 1, \dots, m_i$, and the symmetric positive definite matrix $\Sigma_i^{1/2}$ with the property $\Sigma_i = \Sigma_i^{1/2} \Sigma_i^{1/2}$.

2.2 Design criteria

The experimental settings x_{i1}, \dots, x_{im_i} in formula 1 are not necessarily all distinct. We define an exact design in group i as

$$\xi_i = \begin{pmatrix} x_{i1}, \dots, x_{ik_i} \\ m_{i1}, \dots, m_{ik_i} \end{pmatrix}, \quad (3)$$

where x_{i1}, \dots, x_{ik_i} are the distinct support points in \mathcal{X}_i with the corresponding numbers of observations $m_{i1}, \dots, m_{ik_i} \in \mathbb{N}$, $\sum_{k=1}^{k_i} m_{ik} = m_i$.

For analytical purposes we also introduce approximate designs:

$$\xi_i = \begin{pmatrix} x_{i1}, \dots, x_{ik_i} \\ w_{i1}, \dots, w_{ik_i} \end{pmatrix},$$

where $w_{ik} \geq 0$ denotes the weight of observations at x_{ik} , $k = 1, \dots, k_i$, and $\sum_{k=1}^{k_i} w_{ik} = 1$.

We will use the following notation for the moment (or information) matrix in group i :

$$\mathbf{M}_i(\xi_i) = m_i \sum_{k=1}^{k_i} w_{ik} \tilde{\mathbf{F}}_{(i)}(x_{ik})^\top \tilde{\mathbf{F}}_{(i)}(x_{ik}). \quad (4)$$

For exact designs we have $w_{ik} = m_{ik}/m_i$ and

$$\mathbf{M}_i(\xi_i) = \tilde{\mathbf{F}}_i^\top \tilde{\mathbf{F}}_i,$$

which follows from formula (4) and the definition of $\tilde{\mathbf{F}}_i$ below formula (2).

We will also use the notation $\boldsymbol{\xi}$ for the tuple of all group-designs ξ_i : $\boldsymbol{\xi} = (\xi_1, \dots, \xi_s)$.

Further we extend the definition of the variance–covariance matrix (2) with respect to approximate designs:

$$\text{Cov}_{\xi} = \left[\sum_{i=1}^s n_i \left(\mathbf{M}_i(\xi_i)^{-1} + \mathbf{D}_i \right)^{-1} \right]^{-1}. \quad (5)$$

We generally search for the designs which minimize the variance-covariance matrix. Instead of the minimization of the matrix itself (which is in general not possible), we instead minimize suitable functions of this matrix which we call optimality criteria. We focus on the commonly used linear (L -) and determinant (D -) criteria for the estimation of the population parameters β_0 , which are given by

$$\phi_L(\xi) = \text{tr} \left(\left[\sum_{i=1}^s n_i \left(\mathbf{M}_i(\xi_i)^{-1} + \mathbf{D}_i \right)^{-1} \right]^{-1} \mathbf{V} \right), \quad (6)$$

where \mathbf{V} is some non-negative definite ($p \times p$) matrix, and

$$\phi_D(\xi) = -\ln \det \left(\sum_{i=1}^s n_i \left(\mathbf{M}_i(\xi_i)^{-1} + \mathbf{D}_i \right)^{-1} \right), \quad (7)$$

respectively (see Prus 2022). (Note that matrix $\mathbf{M}_i(\xi_i)$ here differs from that in Prus 2022 by constant m_i .)

Frequently used particular cases of the L -criterion are the c - and A -criterion, which are of the form (6) with $\mathbf{V} = \mathbf{c}\mathbf{c}^\top$, $\mathbf{c} \in \mathbb{R}^p$, and $\mathbf{V} = \mathbb{I}_p$, where \mathbb{I}_p is the $p \times p$ identity matrix, respectively. Another frequently used linear criterion is the $IMSE$ -criterion. For the estimation of the mean parameters β_0 in multiple-group model (1) we define this criterion as follows:

$$\begin{aligned} \phi_{IMSE}(\xi) &= \sum_{i=1}^s a_i \text{tr} \left(\int_{\mathcal{X}_i} \mathbb{E} \left[\left(\mathbf{F}_{(i)}(x) \hat{\beta}_0 - \mathbf{F}_{(i)}(x) \beta_0 \right) \left(\mathbf{F}_{(i)}(x) \hat{\beta}_0 - \mathbf{F}_{(i)}(x) \beta_0 \right)^\top \right] v_i(dx) \right), \end{aligned} \quad (8)$$

where v_i is some suitable measure on the experimental region \mathcal{X}_i (typically uniform on \mathcal{X}_i) with $v_i(\mathcal{X}_i) = 1$ and a_i is a coefficient related to group i , $\sum_{i=1}^s a_i = 1$. The coefficients a_1, \dots, a_s may depend on the group sizes or, alternatively, equal weight may be given to each group. $IMSE$ -criterion (8) may be rewritten in form

$$\phi_{IMSE}(\xi) = \text{tr} \left(\text{Cov}(\hat{\beta}_0) \sum_{i=1}^s a_i \int_{\mathcal{X}_i} \mathbf{F}_{(i)}(x)^\top \mathbf{F}_{(i)}(x) v_i(dx) \right). \quad (9)$$

Then we extend it for approximate designs by using the extended variance-covariance matrix (5) and we obtain the particular linear criterion with $\mathbf{V} = \sum_{i=1}^s a_i \int_{\mathcal{X}_i} \mathbf{F}_{(i)}(x)^\top \mathbf{F}_{(i)}(x) v_i(dx)$, which simplifies to $\mathbf{V} = \int_{\mathcal{X}_1} \mathbf{F}_{(1)}(x)^\top \mathbf{F}_{(1)}(x) v_1(dx)$ if the regression matrices $\mathbf{F}_{(i)}$, the experimental regions \mathcal{X}_i and the weighting measures v_i are the same among all groups: $\mathbf{F}_{(i)} = \mathbf{F}_{(1)}$, $\mathcal{X}_i = \mathcal{X}_1$ and $v_i = v_1$ for $i = 1, \dots, s$.

2.3 Optimality conditions

The optimality conditions for the L - and D -criteria are provided by the following theorems (see Prus 2022):

Theorem 1 Approximate designs $\xi^* = (\xi_1^*, \dots, \xi_s^*)$ are L -optimal for estimation of the mean parameters β_0 iff

$$\begin{aligned} & m_i \operatorname{tr} \left\{ \tilde{\mathbf{F}}_{(i)}(x_i) \left[\mathbf{M}_i(\xi_i^*)^{-1} \left(\mathbf{M}_i(\xi_i^*)^{-1} + \mathbf{D}_i \right)^{-1} \left[\sum_{r=1}^s n_r \left(\mathbf{M}_r(\xi_r^*)^{-1} + \mathbf{D}_r \right)^{-1} \right]^{-1} \mathbf{V} \right. \right. \\ & \quad \cdot \left. \left[\sum_{r=1}^s n_r \left(\mathbf{M}_r(\xi_r^*)^{-1} + \mathbf{D}_r \right)^{-1} \right]^{-1} \left(\mathbf{M}_i(\xi_i^*)^{-1} + \mathbf{D}_i \right)^{-1} \mathbf{M}_i(\xi_i^*)^{-1} \right] \tilde{\mathbf{F}}_{(i)}(x_i)^\top \left. \right\} \\ & \leq \operatorname{tr} \left\{ \mathbf{M}_i(\xi_i^*)^{-1} \left(\mathbf{M}_i(\xi_i^*)^{-1} + \mathbf{D}_i \right)^{-1} \left[\sum_{r=1}^s n_r \left(\mathbf{M}_r(\xi_r^*)^{-1} + \mathbf{D}_r \right)^{-1} \right]^{-1} \mathbf{V} \right. \\ & \quad \cdot \left. \left[\sum_{r=1}^s n_r \left(\mathbf{M}_r(\xi_r^*)^{-1} + \mathbf{D}_r \right)^{-1} \right]^{-1} \left(\mathbf{M}_i(\xi_i^*)^{-1} + \mathbf{D}_i \right)^{-1} \right\} \end{aligned} \quad (10)$$

for $x_i \in \mathcal{X}_i$, $i = 1, \dots, s$.

For support points of ξ_i^* equality holds in (10).

Theorem 2 Approximate designs $\xi^* = (\xi_1^*, \dots, \xi_s^*)$ are D -optimal for estimation of the mean parameters β_0 iff

$$\begin{aligned} & m_i \operatorname{tr} \left\{ \tilde{\mathbf{F}}_{(i)}(x_i) \left[\mathbf{M}_i(\xi_i^*)^{-1} \left(\mathbf{M}_i(\xi_i^*)^{-1} + \mathbf{D}_i \right)^{-1} \left[\sum_{r=1}^s n_r \left(\mathbf{M}_r(\xi_r^*)^{-1} + \mathbf{D}_r \right)^{-1} \right]^{-1} \right. \right. \\ & \quad \cdot \left. \left(\mathbf{M}_i(\xi_i^*)^{-1} + \mathbf{D}_i \right)^{-1} \mathbf{M}_i(\xi_i^*)^{-1} \right] \tilde{\mathbf{F}}_{(i)}(x_i)^\top \left. \right\} \\ & \leq \operatorname{tr} \left\{ \mathbf{M}_i(\xi_i^*)^{-1} \left(\mathbf{M}_i(\xi_i^*)^{-1} + \mathbf{D}_i \right)^{-1} \left[\sum_{r=1}^s n_r \left(\mathbf{M}_r(\xi_r^*)^{-1} + \mathbf{D}_r \right)^{-1} \right]^{-1} \right. \\ & \quad \cdot \left. \left(\mathbf{M}_i(\xi_i^*)^{-1} + \mathbf{D}_i \right)^{-1} \right\} \end{aligned} \quad (11)$$

for $x_i \in \mathcal{X}_i$, $i = 1, \dots, s$.

For support points of ξ_i^* equality holds in (11).

Example 1 We consider the two-groups model of general form (1) with the regression functions $\mathbf{F}_{(i)}(x) = (1, x)$, $x \in \mathcal{X}_i$:

$$Y_{ijh} = \beta_{ij1} + \beta_{ij2}x_{ih} + \varepsilon_{ijh}, \quad j = 1, \dots, n_i, \quad h = 1, \dots, m_i, \quad i = 1, 2, \quad (12)$$

on the design regions $\mathcal{X}_i = [0, 1]$. The covariance structures of random effects and observational errors are given by $\mathbf{D}_i = \text{diag}(d_{i1}, d_{i2})$ and $\Sigma_i = 1$ for both groups. For this model the left hand sides of the optimality conditions (10) and (11) are parabolas with positive leading terms. Therefore, D - and L -optimal approximate group-designs have the form

$$\xi_i = \begin{pmatrix} 0 & 1 \\ 1 - w_{i1} & w_{i1} \end{pmatrix}, \quad (13)$$

where w_{i1} denotes the weight of observations in point 1 for the i -th group and may depend on the choice of the design criterion as well as on model parameters. The moment matrices are given by

$$\mathbf{M}_i(\xi_i) = \begin{pmatrix} m_i & m_{i1} \\ m_{i1} & m_{i1} \end{pmatrix}, \quad (14)$$

where $m_{i1} = w_{i1}m_i$. Optimal designs for random intercept and random slope have been considered in more detail in Prus (2022).

3 Equi- and invariance considerations for construction of optimal designs

Equi- and invariance of design criteria play an important role for determining optimal designs in fixed effects models (see e.g. Heiligers 1992 or Schwabe 1996, ch. 3). Prus and Schwabe (2016) investigated the related properties of designs, which are optimal for prediction of individual random parameters in single-group mixed effects models. Here we extend those results to multiple-group models.

We consider a one-to-one transformation g of the experimental regions \mathcal{X}_i for all $i = 1, \dots, s$ simultaneously with $g(\mathcal{X}_i) = \mathcal{X}_i^g$. We assume the regression matrices $\mathbf{F}_{(i)}$ to be defined on both \mathcal{X}_i and \mathcal{X}_i^g . We also assume the existence of a non-singular $p \times p$ matrix \mathbf{Q}_g such that

$$\tilde{\mathbf{F}}_{(i)}(g(x)) = \mathbf{Q}_g \tilde{\mathbf{F}}_{(i)}(x), \quad \forall x \in \mathcal{X}_i, \quad i = 1, \dots, s, \quad (15)$$

i.e. all $\tilde{\mathbf{F}}_{(i)}$ are linearly equivariant with respect to the transformation g (see e.g. Schwabe 1996, ch. 3). We denote by ξ_i^g the following transformation of an approximate design ξ_i :

$$\xi_i^g = \begin{pmatrix} g(x_{i1}), \dots, g(x_{ik_i}) \\ w_{i1}, \dots, w_{ik_i} \end{pmatrix}, \quad (16)$$

where the weight w_{ik} is the same for both ξ_i and ξ_i^g and only the design points x_{ik} are transformed. Then we obtain the next property of the moment matrices:

$$\mathbf{M}_i(\xi_i^g) = \mathbf{Q}_g \mathbf{M}_i(\xi_i) \mathbf{Q}_g^\top, \quad i = 1, \dots, s. \quad (17)$$

Further we use the notations $\mathcal{D} = (\mathbf{D}_1, \dots, \mathbf{D}_s)$ and $\mathbb{X} = \times_{i=1}^s \mathcal{X}_i$ for the tuple of covariance matrices and the Cartesian product of the experimental regions, respectively, in all groups. For the covariance matrix (5) the following relation can be easily verified:

$$\text{Cov}_{\xi^g}(\mathcal{D}^g) = \mathbf{Q}_g^{-\top} \text{Cov}_{\xi}(\mathcal{D}) \mathbf{Q}_g^{-1}, \quad (18)$$

where $\xi^g = (\xi_1^g, \dots, \xi_s^g)$, $\mathcal{D}^g = (\mathbf{D}_1^g, \dots, \mathbf{D}_s^g)$, $\mathbf{D}_i^g = \mathbf{Q}_g^{-\top} \mathbf{D}_i \mathbf{Q}_g^{-1}$ and $\mathbf{Q}_g^{-\top} = (\mathbf{Q}_g^\top)^{-1}$. We use the notation $\text{Cov}_{\xi}(\mathcal{D})$ [instead of Cov_{ξ} as in formula (5)] to emphasize the dependence on the covariance matrices \mathcal{D} of random effects.

Then the equivariance of the D - and L -criteria with respect to a transformation g can be established.

Theorem 3 *If the approximate designs ξ^* are D -optimal for the estimation of β_0 on the experimental regions \mathbb{X} under the dispersion matrices \mathcal{D} , then the induced approximate designs ξ^g are D -optimal for the estimation of β_0 on the experimental regions $\mathbb{X}^g = \times_{i=1}^s \mathcal{X}_i^g$ under the induced dispersion matrices \mathcal{D}^g .*

Proof From the definition of the D -criterion for the estimation of β_0 and formula (18) we obtain

$$\phi_D(\xi^g, \mathcal{D}^g) = -2 \ln |\det(\mathbf{Q}_g)| + \phi_D(\xi, \mathcal{D}),$$

which proves the optimality of ξ^g on \mathbb{X}^g for ξ optimal on \mathbb{X} . \square

Theorem 4 *If the approximate designs ξ^* are L -optimal for the estimation of β_0 on the experimental regions \mathbb{X} under the dispersion matrices \mathcal{D} with respect to the transformation matrix \mathbf{V} , then the induced approximate designs ξ^{*g} are L -optimal for the estimation of β_0 on the experimental regions \mathbb{X}^g under the induced dispersion matrices \mathcal{D}^g with respect to the induced transformation matrix $\mathbf{V}_g = \mathbf{Q}_g \mathbf{V} \mathbf{Q}_g^\top$.*

Proof Using formulas (6) and (18) it can be easily verified that

$$\phi_L(\xi^g, \mathcal{D}^g, \mathbf{V}_g) = \phi_L(\xi, \mathcal{D}, \mathbf{V}),$$

which proves the optimality of ξ^g on \mathbb{X}^g . \square

Corollary 1 *The A -criterion for the estimation of β_0 is equivariant with respect to a transformation g if \mathbf{Q}_g is orthogonal, i.e.:*

$$\mathbf{Q}_g \mathbf{Q}_g^\top = \mathbf{Q}_g^\top \mathbf{Q}_g = \mathbb{I}_p. \quad (19)$$

To verify the equivariance of the $IMSE$ -criterion we assume, besides the transformed regression matrices $\tilde{\mathbf{F}}_{(i)}$, the original regression matrices $\mathbf{F}_{(i)}$ to be linearly equivariant with respect to the transformation g :

$$\mathbf{F}_{(i)}(g(x)) = \mathbf{Q}_g \mathbf{F}_{(i)}(x), \quad \forall x \in \mathcal{X}_i, \quad i = 1, \dots, s. \quad (20)$$

Then if the measure ν_i is transformed to its image ν_i^g , we obtain

$$\mathbf{V}_g = \sum_{i=1}^s a_i \int_{\mathcal{X}_i^g} \mathbf{F}_{(i)}(x)^\top \mathbf{F}_{(i)}(x) \nu_i^g(dx) = \mathbf{Q}_g \mathbf{V} \mathbf{Q}_g^\top.$$

Corollary 2 *The $IMSE$ -criterion for the estimation of β_0 is equivariant with respect to a transformation g if condition (20) is satisfied.*

Example 1 (continued). We consider again the two-groups linear regression model (12) on $\mathcal{X}_i = [0, 1]$ with diagonal covariance structure of random effects. For the $IMSE$ -criterion we chose the uniform weighting $\nu_i = \lambda_{[0,1]}$, $i = 1, 2$, where $\lambda_{[c_1, c_2]}$ denotes the Lebesgue measure on $[c_1, c_2]$. Let ξ_i^* be D -, A - or $IMSE$ -optimal group-designs of form (13) with the optimal weight of observations w_{i1}^* (which generally depends on the choice of the design criterion).

Now we consider the linear transformation $g(x) = ax$, $a > 0$, for which we obtain $\mathbf{Q}_g = \text{diag}(1, a)$. Then the D -, A - or $IMSE$ -optimal group-designs in model (12) on $\mathcal{X}_i^g = [0, a]$ for $\mathbf{D}_i^g = \text{diag}(d_{i1}, d_{i2}/a^2)$ and $\nu_i^g = \frac{1}{a} \lambda_{[0,a]}$ are given by

$$\xi_i^{*g} = \begin{pmatrix} 0 & a \\ 1 - w_{i1}^* & w_{i1}^* \end{pmatrix}. \quad (21)$$

Same behavior of optimal designs has been established for the prediction of random effects in single-group model in Prus and Schwabe (2016).

Further we consider a finite group G of transformations $g : \mathcal{X}_i \rightarrow \mathcal{X}_i$ of the experimental regions \mathcal{X}_i onto themselves for all $i = 1, \dots, s$ simultaneously. We assume the equivariance condition (15) to be satisfied and the dispersion matrices to be invariant: $\mathbf{D}_i^g = \mathbf{D}_i$, for all $g \in G$, $i = 1, \dots, s$. For the linear criteria we additionally assume the invariance of the transformation matrices: $\mathbf{V}_g = \mathbf{V}$. Then the D - and L -criteria are invariant with respect to all $g \in G$ and the following statement can be formulated:

Theorem 5 *If the approximate designs ξ^* are D - or L -optimal for the estimation of β_0 , then the symmetrized designs $\tilde{\xi}^* = (\tilde{\xi}_1^*, \dots, \tilde{\xi}_s^*)$ for $\tilde{\xi}_i^* = \frac{1}{\#G} \sum_{g \in G} \xi_i^{*g}$ are also D - or L -optimal for the estimation of β_0 .*

Proof Let the designs ξ^* be D -optimal for the estimation of β_0 . Then it follows from Theorem 3 and the invariance of the dispersion matrices that the induced designs ξ^{*g} are also D -optimal, i.e.

$$\phi_D(\xi^{*g}, \mathcal{D}) = \phi_D(\xi^*, \mathcal{D}), \quad \forall g \in G.$$

From the convexity of the criterion we obtain

$$\phi_D(\bar{\xi}^*, \mathcal{D}) \leq \phi_D(\xi^*, \mathcal{D}),$$

which implies the D -optimality of the designs $\bar{\xi}^*$.

For the linear criterion the proof is similar. \square

The invariance of the A -criterion is straightforward if condition (19) is satisfied for all $g \in G$.

Corollary 3 *If the approximate designs ξ^* are A -optimal for the estimation of β_0 and condition (19) is satisfied for all $g \in G$, then the symmetrized designs $\bar{\xi}^*$ are also A -optimal for the estimation of β_0 .*

For the $IMSE$ -criterion we require the invariance of the weighting measures: $v_i^g = v_i$, which leads to $\mathbf{V}_g = \mathbf{V}$.

Corollary 4 *If the approximate designs ξ^* are $IMSE$ -optimal for the estimation of β_0 and condition (20) is satisfied for all $g \in G$, then the symmetrized designs $\bar{\xi}^*$ are also $IMSE$ -optimal for the estimation of β_0 .*

Example 2 We consider the multiple-group model of the form (1) with the regression functions $\mathbf{F}_{(i)}(x) = (1, x, x^2)$ on a symmetric design region $\mathcal{X}_i = [-a, a]$, $a > 0$, $i = 1, \dots, s$:

$$Y_{ijh} = \beta_{ij1} + \beta_{ij2}x_{ih} + \beta_{ij3}x_{ih}^2 + \varepsilon_{ijh}, \quad j = 1, \dots, n_i, \quad h = 1, \dots, m_i. \quad (22)$$

For this model the left hand sides of the optimality conditions (10) and (11) for the L - and D -criterion, respectively, are polynomial functions of degree four. Consequently, the corresponding optimal group-designs ξ_i are supported by not more than three design points including the two endpoints of the experimental region:

$$\xi_i^* = \begin{pmatrix} -a & o_i & a \\ w_{i1}^* & 1 - w_{i1}^* - w_{i2}^* & w_{i2}^* \end{pmatrix}, \quad (23)$$

where $o_i \in (-a, a)$ may differ for different design criteria or for different groups. We assume covariance structures of random effects and observational errors to be given by

$$\mathbf{D}_i = \begin{pmatrix} d_{i11} & 0 & d_{i13} \\ 0 & d_{i22} & 0 \\ d_{i13} & 0 & d_{i33} \end{pmatrix} \quad (24)$$

and $\Sigma_i = \mathbf{I}$ for all groups. For the $IMSE$ -criterion we chose the uniform weighting measure $v_i = \frac{1}{2a}\lambda_{[-a,a]}$ for all $i = 1, \dots, s$.

Further we consider the group of transformations $G = \{g_1, g_2\}$ with $g_1(x) = -x$ and $g_2(x) = x$. Then we obtain $\mathbf{Q}_{g_1} = \text{diag}(1, -1, 1)$ and \mathbf{Q}_{g_2} is equal to the identity matrix. Hence, the dispersion matrices \mathbf{D}_i and the measures ν_i are invariant and conditions (19) and (20) are satisfied for both g_1 and g_2 . Then by Theorem 5 and Corollary 3 group-designs of the general form

$$\bar{\xi}_i^* = \begin{pmatrix} -a & 0 & a \\ w_{i1}^* & 1 - 2w_{i1}^* & w_{i1}^* \end{pmatrix} \quad (25)$$

are D -, A - and $IMSE$ -optimal for the estimation of the mean parameters β_0 . The optimal weights of observations w_{i1}^* at points $x = a$ and $x = -a$ generally depend on the design criterion, the variance parameters, the group sizes, the numbers of observations and the length of the interval (see Sect. 6 for examples of the designs).

Further we consider some examples of multiple polynomial regression. For models without random effects optimal designs for multiple polynomial regression have been discussed, e.g., in Galil and Kiefer (1977) and Heiligers (1992).

Example 3 We consider the multiple-group bi-linear model with the regression functions $\mathbf{F}_{(i)}(x) = (1, x_1, x_2)$ on a design region $\mathcal{X}_i = [-a, a]^2$, $a > 0$, $i = 1, \dots, s$:

$$Y_{ijh} = \beta_{ij1} + \beta_{ij2}x_{ih1} + \beta_{ij3}x_{ih2} + \varepsilon_{ijh}, \quad j = 1, \dots, n_i, \quad h = 1, \dots, m_i. \quad (26)$$

For this model the left hand sides of the optimality conditions (10) and (11) for the L - and D -criterion, respectively, are convex paraboloids. Therefore, the only admissible support points for optimal designs are $x_{i1} = (a, a)$, $x_{i2} = (a, -a)$, $x_{i3} = (-a, a)$, $x_{i4} = (-a, -a)$:

$$\bar{\xi}_i^* = \begin{pmatrix} x_{i1} & x_{i2} & x_{i3} & x_{i4} \\ w_{i1}^* & w_{i2}^* & w_{i3}^* & w_{i4}^* \end{pmatrix}, \quad (27)$$

where $\sum_{k=1}^4 w_{ik} = 1$. For the $IMSE$ -criterion we use the product measure $\nu_i = \frac{1}{2a} \lambda_{[-a,a]} \times \frac{1}{2a} \lambda_{[-a,a]}$ for all $i = 1, \dots, s$.

Further we assume the same covariance structures of random effects and observational errors as in Example 2 of quadratic regression and we consider the group of transformations $G = \{g_1, g_2\}$ with $g_1(x) = (-x_1, x_2)^\top$ and $g_2(x) = (x_1, x_2)^\top$. We obtain $\mathbf{Q}_{g_1} = \text{diag}(1, -1, 1)$ and \mathbf{Q}_{g_2} is equal to the 3×3 identity matrix. Then the dispersion matrices \mathbf{D}_i and the weighting measures ν_i are invariant and conditions (19) and (20) are satisfied for both g_1 and g_2 and, consequently, group-designs of the general form

$$\bar{\xi}_i^* = \begin{pmatrix} x_{i1} & x_{i2} & x_{i3} & x_{i4} \\ w_{i1}^* & w_{i2}^* & w_{i1}^* & w_{i2}^* \end{pmatrix} \quad (28)$$

with $w_{i2}^* = \frac{1}{2}(1 - 2w_{i1}^*)$ are D -, A - and $IMSE$ -optimal for the estimation of the mean parameters β_0 . Only the optimal weights of observations w_{i1}^* have to be determined.

Note that besides the choice of the design criterion these numbers may also depend on the model parameters (see Sect. 5 for illustrative examples).

For the particular case with diagonal covariance structure of random effects: $d_{i13} = 0$, we consider the group of transformations $G = \{g_1, g_2, g_3, g_4\}$ with $g_3(x) = (x_1, -x_2)^\top$ and $g_4(x) = (-x_1, -x_2)^\top$, for which we obtain $\mathbf{Q}_{g_3} = \text{diag}(1, 1, -1)$ and $\mathbf{Q}_{g_4} = \text{diag}(1, -1, -1)$. The dispersion matrices \mathbf{D}_i and the measures v_i are invariant and conditions (19) and (20) are satisfied for all transformation in G . Then the balanced group-designs

$$\bar{\xi}_i^* = \begin{pmatrix} x_{i1} & x_{i2} & x_{i3} & x_{i4} \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix} \quad (29)$$

are D -, A - and $IMSE$ -optimal.

Example 4 We consider the multiple-group bi-quadratic model with the regression functions $\mathbf{F}_{(i)}(x) = (1, x_1, x_2, x_1x_2, x_1^2, x_2^2)$ on a design region $\mathcal{X}_i = [-a, a]^2, a > 0$:

$$Y_{ijh} = \beta_{ij1} + \beta_{ij2}x_{ih1} + \beta_{ij3}x_{ih2} + \beta_{ij4}x_{ih1}x_{ih2} + \beta_{ij5}x_{ih1}^2 + \beta_{ij6}x_{ih2}^2 + \varepsilon_{ijh} \quad (30)$$

for $j = 1, \dots, n_i, h = 1, \dots, m_i$ and $i = 1, \dots, s$. For this model the left hand sides of the optimality conditions (10) and (11) are quadric surfaces in (x_1, x_2) , for which the projections on both $x_1 = 0$ and $x_2 = 0$ are polynomials of degree four. Then the only admissible support points for L - and D -optimal designs are $x_{i1} = (a, a)$, $x_{i2} = (a, -a)$, $x_{i3} = (-a, a)$, $x_{i4} = (-a, -a)$, $x_{i5} = (o_{i1}, a)$, $x_{i6} = (o_{i2}, -a)$, $x_{i7} = (a, o_{i3})$, $x_{i8} = (-a, o_{i4})$ and $x_{i9} = (o_{i5}, o_{i6})$, where $o_{il} \in (-a, a), l = 1, \dots, 6$. For the $IMSE$ -criterion we use the same weighting measures as in Example 3.

Further we assume the following simple covariance structure of the random effects and the observational errors: $\mathbf{D}_i = \text{diag}(d_{i1}, \dots, d_{i6})$ and $\Sigma_i = \mathbf{I}, i = 1, \dots, s$. Then we consider the same group of transformations G as in the previous example and we obtain $\mathbf{Q}_{g_1} = \text{diag}(1, -1, 1, -1, 1, 1)$, \mathbf{Q}_{g_2} is equal to the 6×6 identity matrix, $\mathbf{Q}_{g_3} = \text{diag}(1, 1, -1, -1, 1, 1)$ and $\mathbf{Q}_{g_4} = \text{diag}(1, -1, -1, 1, 1, 1)$. Then conditions (19) and (20) are satisfied and dispersion matrices \mathbf{D}_i and the measures v_i are invariant for all $g_i \in G$. Therefore, D -, A - and $IMSE$ -optimal designs have the general form

$$\bar{\xi}_i^* = \begin{pmatrix} x_{i1} & x_{i2} & x_{i3} & x_{i4} & x_{i5} & x_{i6} & x_{i7} & x_{i8} & x_{i9} \\ w_{i1}^* & w_{i1}^* & w_{i1}^* & w_{i1}^* & w_{i2}^* & w_{i2}^* & w_{i3}^* & w_{i3}^* & w_{i4}^* \end{pmatrix}, \quad (31)$$

where $w_{i4}^* = 1 - 4w_{i1}^* - 2w_{i2}^* - 2w_{i3}^*$ and all $o_{il} = 0$, i.e. $x_{i5} = (0, a)$, $x_{i6} = (0, -a)$, $x_{i7} = (a, 0)$, $x_{i8} = (-a, 0)$ and $x_{i9} = (0, 0)$. The weights of observations w_{i2}^*, w_{i3}^* and w_{i4}^* depend on the choice of the design criterion and on the model parameters and have to be optimized.

Then we additionally assume the conditions $d_{i2} = d_{i3}$ and $d_{i5} = d_{i6}$ to be satisfied and consider the extended group of transformations $G_1 = G \cup \{g_5\}$ with $g_5(x) = (x_2, x_1)^\top$, for which we obtain $\mathbf{Q}_{g_5} = \text{block-diag}(1, \mathbf{P}, 1, \mathbf{P})$, where \mathbf{P} is the (2×2)

permutation matrix:

$$\mathbf{P} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The dispersion matrices \mathbf{D}_i and the measures v_i are also invariant with respect to g_5 and conditions (19) and (20) are satisfied. Then the general form (31) of optimal designs simplifies to

$$\xi_i^* = \begin{pmatrix} x_{i1} & x_{i2} & x_{i3} & x_{i4} & x_{i5} & x_{i6} & x_{i7} & x_{i8} & x_{i9} \\ w_{i1}^* & w_{i1}^* & w_{i1}^* & w_{i1}^* & w_{i2}^* & w_{i2}^* & w_{i2}^* & w_{i2}^* & 1 - 4(w_{i1}^* + w_{i2}^*) \end{pmatrix}. \quad (32)$$

Note that similar behavior has been established for optimal designs for Kiefer's Φ_p -criteria in fixed effects models (see Galil and Kiefer 1977). However, designs obtained in that work depend on the choice of the design criterion only. In the model under investigation optimal designs may also depend on the variance parameters, the group sizes and the numbers of observations per observational unit.

4 Computing the multiple-group mixed models designs

In this Section, we will show how to compute efficient exact designs for model (1). To this end, let's discretize each (possibly continuous) experimental region \mathcal{X}_i , $i = 1, \dots, s$, into k_i points x_{i1}, \dots, x_{ik_i} and denote the corresponding numbers of measurements in these points by $m_{i1}, \dots, m_{ik_i} \in \mathbb{N}_0$, as is customary in optimal design algorithms. Similarly to the notation adopted in Sect. 2, we define the k_i -dimensional vectors $\mathbf{m}_i = (m_{i1}, \dots, m_{ik_i})$ and the $\sum_{i=1}^s k_i = u$ -dimensional vector $\mathbf{m} = (m_1, \dots, m_s)$.

Now, consider the optimization problem presented in Harman et al. (2016):

$$\begin{aligned} & \min_{\mathbf{m}} \Phi(\mathbf{m}) \\ & \text{subject to } \mathbf{A}\mathbf{m} \leq \mathbf{b}. \end{aligned}$$

Here, we minimize the function Φ on the set of permissible designs determined by the linear inequality $\mathbf{A}\mathbf{m} \leq \mathbf{b}$, where $\mathbf{A} \in \mathbb{R}^{k \times u}$ and $\mathbf{b} \in \mathbb{R}^k$ are such that the elements of \mathbf{A} are nonnegative and the elements of \mathbf{b} are positive. These kinds of constraints are called *resource constraints*, i.e., we can view each measurement as consuming some amount of each of the k resources, limit on which are given by the vector \mathbf{b} .

The method described in Harman et al. (2016) is related to the Detmax procedure, employing a tabu search principle. The algorithm is based on excursions in the set of all feasible designs. More precisely, from a design ξ we can either make a forward step to one of its upper neighbours or a backward step to one of its lower neighbours. These excursions are directed by the attribute of each design (which can be, e.g., its criterion value), a tabu list of the attributes of already visited designs and a local heuristic evaluation of the design that roughly estimates how promising a design is as a part of an excursion leading to an efficient design.

Note that although the algorithm is primarily developed for D -optimality in the standard linear regression model, it can be easily adapted for different criteria that are monotonous on the set of all approximate designs, which enables us to compute D - and L -efficient exact designs in model (1).

To show this, we rewrite the covariance matrix (5) in the following form:

$$\text{Cov}_\xi = \left[\left(\mathbb{1}_s^\top \otimes \mathbb{I}_p \right) \left(\mathbf{M}_\xi^{-1} + \mathbf{D} \right)^{-1} \left(\mathbb{1}_s \otimes \mathbb{I}_p \right) \right]^{-1}, \quad (33)$$

where $\mathbf{M}_\xi = \text{diag} \left(\tilde{\mathbf{M}}_1(\xi_1), \dots, \tilde{\mathbf{M}}_s(\xi_s) \right)$ is the block-diagonal matrix with the blocks $\tilde{\mathbf{M}}_i(\xi_i) = n_i \mathbf{M}_i(\xi_i)$ and $\mathbf{D} = \text{diag}(\tilde{\mathbf{D}}_1, \dots, \tilde{\mathbf{D}}_s)$ is the block-diagonal matrix with the blocks $\tilde{\mathbf{D}}_i = \frac{1}{n_i} \mathbf{D}_i$, $\mathbb{1}_s$ is the vector of length s with all entries equal to 1 and " \otimes " denotes the Kronecker product.

Then the L - and D -criteria defined by (6) and (7) can be written as the function of the design vector w in the following way:

$$\phi_L(\xi) = \text{tr} \left(\left[\left(\mathbb{1}_s^\top \otimes \mathbb{I}_p \right) \left(\mathbf{M}_\xi^{-1} + \mathbf{D} \right)^{-1} \left(\mathbb{1}_s \otimes \mathbb{I}_p \right) \right]^{-1} \mathbf{V} \right) \quad (34)$$

and

$$\phi_D(\xi) = -\ln \det \left[\left(\mathbb{1}_s^\top \otimes \mathbb{I}_p \right) \left(\mathbf{M}_\xi^{-1} + \mathbf{D} \right)^{-1} \left(\mathbb{1}_s \otimes \mathbb{I}_p \right) \right]. \quad (35)$$

Note that both criteria (34) and (35) are monotonically decreasing with respect to \mathbf{M}_ξ .

Further, model (1) can be viewed as a one-group model on $\mathcal{X} = \times_{i=1}^s \mathcal{X}_i$ with marginal constraints (see, e.g., Cook and Thibodeau 1980) that constrict the number of observations in each group to $\sum_{h=1}^{k_i} m_{ih} = m_i$. This can be formulated in the form of resource constraints by putting $\mathbf{A} = \text{diag}(\mathbb{1}_{k_1}^\top, \dots, \mathbb{1}_{k_s}^\top)$ and $\mathbf{b} = (m_1, \dots, m_s)^\top$.

Hence, the optimization problem to solve is

$$\begin{aligned} & \min_{\mathbf{m}} \phi(\mathbf{m}) \\ & \text{subject to } \text{diag}(\mathbb{1}_{k_1}^\top, \dots, \mathbb{1}_{k_s}^\top) \mathbf{m} \leq \mathbf{b}, \end{aligned} \quad (36)$$

where by ϕ we denote either of the optimality criteria in (34) or (35).

Note that the algorithm used here is heuristic, i.e., it does not guarantee that the resulting design is optimal, although it is demonstrated in Harman et al. (2016) that it is usually highly efficient. Therefore, in the following sections, we will call the designs obtained by the algorithm as *efficient exact designs*.

Further, we will demonstrate that it is of great practical use that the matrix \mathbf{A} and the vector \mathbf{b} can be modified so that they incorporate additional linear resource constraints on the weights, such as the limit on the number of measurements in particular points or cost constraints (see Sect. 6 for an example of such constraints), simply by adding suitable rows to the matrix \mathbf{A} and elements to the vector \mathbf{b} .

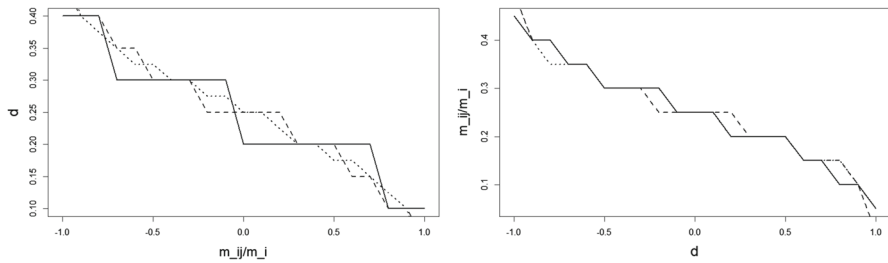


Fig. 1 The dependence of the number of observations in the point $(1, 1)$ on the parameter d in the exact D -efficient designs in bilinear model (26) on $[-1, 1]^2$ with total numbers of observations in the groups given by $m = (10, 20, 40)$ (left) and $m = (20, 20, 20)$ (right). The three lines denote the number of observations in the point $(1, 1)$ normalized by m_i for the first (full line), second (dashed line) and third (dotted line) group

5 Bi-linear regression

Let's consider the model of bi-linear regression (26) with three groups, $\mathcal{X}_i = [-1, 1]^2$, $\Sigma_i = 1$, $n = (1, 1, 1)$ and

$$D_i = \begin{pmatrix} 1 & 0 & d \\ 0 & 1 & 0 \\ d & 0 & 1 \end{pmatrix}, \quad i = 1, 2, 3.$$

As the analytical results in Example 3 show, the approximate optimal designs are supported on the four vertices of the square $[-1, 1]^2$ and the number of observations is identical in the points $(1, 1)$, $(-1, 1)$ and in the points $(-1, -1)$, $(1, -1)$. This phenomenon was confirmed also for the exact D -efficient designs by our algorithm.

In this example, we will numerically illustrate the dependence of efficient designs on the parameter d in the matrices D_i , $i = 1, 2, 3$. To this end, let's consider that in all three groups, the parameter d is the same. Figure 1 shows how the numbers of observation in the point $(1, 1)$ change with d varying from -1 to 1 for two different settings: $m = (10, 20, 40)$ (left) and $m = (20, 20, 20)$ (right). We can see that in both cases, the number of observations in $(1, 1)$ decreases with increasing d .

Now, suppose that

$$D_i = \begin{pmatrix} 1 & 0 & d_i \\ 0 & 1 & 0 \\ d_i & 0 & 1 \end{pmatrix}, \quad i = 1, 2, 3, \quad (37)$$

where $d_i \in \{-0.5, 0, 0.5\}$ are not necessarily the same between groups. In Table 1 we show the behavior of the numbers of observation for several selected d_1, d_2, d_3 in the case $m = (20, 20, 20)$.

Table 1 Exact D -efficient designs in bilinear model (26) on $[-1, 1]^2$ with total numbers of observations in the groups given by $m = (20, 20, 20)$ with $d_i \in \{-0.5, 0, 0.5\}$

(d_1, d_2, d_3)	$(-1, -1)$	$(1, 1)$	$(-1, -1)$	$(1, 1)$	$(-1, -1)$	$(1, 1)$
$(0, 0, 0)$	5	5	5	5	5	5
$(0, 0, 0.5)$	4	6	4	6	7	3
$(0, 0.5, 0.5)$	4	6	7	3	7	3
$(0.5, 0.5, 0.5)$	6	4	6	4	6	4
$(0, 0, -0.5)$	6	4	6	4	3	7
$(0, 0.5, -0.5)$	5	5	7	3	3	7
$(-0.5, 0.5, -0.5)$	3	7	8	2	3	7

6 Quadratic regression on a symmetric interval

Consider the two-groups model of the form (1) with the regression functions $\mathbf{F}_{(i)}(x) = (1, x, x^2)^\top$, $x \in \mathcal{X}_i$, and the design region $\mathcal{X}_i = [-1, 1]$, $i = 1, 2$:

$$Y_{ijh} = \beta_{ij1} + \beta_{ij2}x_{ih} + \beta_{ij3}x_{ih}^2 + \varepsilon_{ijh}, \quad j = 1, \dots, n_i, \quad h = 1, \dots, m_i. \quad (38)$$

The covariance structures of random effects and observational errors are given by $\mathbf{D}_i = \text{diag}(d_{i1}, d_{i2}, d_{i3})$ and $\Sigma_i = \mathbf{I}$ for both groups.

The corresponding approximate optimal group-designs ξ_i are supported by three design points $-1, 0, 1$ (see Sect. 3):

$$\xi_i^* = \begin{pmatrix} -1 & 0 & 1 \\ w_{i1}^* & 1 - 2w_{i1}^* & w_{i1}^* \end{pmatrix}. \quad (39)$$

This result was heuristically confirmed to hold also for the exact designs: we discretized the design region into q points $-1 = x_1 < x_2 < \dots < x_q = 1$ and confirmed that for all cases considered below, the support points are indeed $-1, 0$ and 1 .

The exact D - and $IMSE$ -efficient designs for this case and several particular $m = (m_1, m_2)$ are given in Table 2 in the Appendix.

We can see that for the criterion of D -optimality and the values of the diagonal of the matrix D_i equal to either $(1, 1, 1)$ or $(1, 1, 0)$, half of the measurements is in the point 0 and the remaining half is distributed equally among the points -1 and 1 . For (d_1, d_2, d_3) equal either to $(0, 1, 1)$ or $(1, 0, 0)$, the measurements are heavily concentrated in the support point 0 , but the designs are still nonsingular. The case $(1, 0, 1)$ shows opposite phenomenon with the measurements being concentrated in the points -1 and 1 . For the remaining cases, the pattern is not so clear and the weights depend more on m , sometimes even resulting in singular designs.

For the $IMSE$ criterion, we get results identical to D -optimality if the diagonal of D_i is $(1, 1, 1)$. For the rest of the cases, the situation is more varied and we refer reader to Table 2 for details.

From practical point of view, it may not be desirable to only have three support points for each group. Therefore, additional constraints on the design were suggested,

where it is prescribed that, for each group, maximum one half of the measurements can be taken at $-1, 0$ or 1 . Formally, these constraints can be written in the form $\mathbf{A}^{(1)}w \leq b^{(1)}$ (see Sect. 4 for details), where

$$\mathbf{A}^{(1)} = \begin{pmatrix} c_q & 0_q^\top \\ 0_q^\top & c_q \end{pmatrix} \in \mathbb{R}^{2 \times 2q}, \quad b^{(1)} = \begin{pmatrix} m_1/2 \\ m_2/2 \end{pmatrix}, \quad (40)$$

where $c_q = (1, 0, \dots, 0, 1, 0, \dots, 0, 1) \in \mathbb{R}^q$ with 1 on the positions corresponding to the points $-1, 0, 1$. The D - and $IMSE$ -efficient designs for the discretization $\mathcal{X}_i = \{-1, -0.8, \dots, 0.8, 1\}$ of the interval $[-1, 1]$ with the step 0.2 (i.e. $q = 11$) are given in Tables 3 and 4. Note that for both criteria, the tendency is to distribute the measurements as close as possible to the original support points $-1, 0$ and 1 .

Another type of constraint that is often used in practical situations, is the cost constraint: this is natural, for example, in clinical trials, where taking a measurement at a point x consumes a certain number of time, personal or material resources and the total cost of the experiment is limited. In our case, let the measurement at the point x cost $|x| + 0.1$ units, and, for group j , let the maximum admissible cost be $m_j/4$. This leads to adding the constraints $\mathbf{A}^{(2)}w \leq b^{(2)}$ with the following $\mathbf{A}^{(2)}, b^{(2)}$ to the problem (36):

$$\mathbf{A}^{(2)} = \begin{pmatrix} u_q + 0.1\mathbb{1}_q^\top & 0 \\ 0 & u_q + 0.1\mathbb{1}_q^\top \end{pmatrix} \in \mathbb{R}^{2 \times 2q}, \quad b^{(2)} = \frac{1}{4} \begin{pmatrix} m_1 \\ m_2 \end{pmatrix}, \quad (41)$$

where $u_q = (1, 0.8, \dots, 0.8, 1)$.

Again, we computed D - and $IMSE$ -efficient designs with respect to this constraint for the discretization $\mathcal{X}_i = \{-1, -0.8, \dots, 0.8, 1\}$. Now, the designs are supported on $-1, 0$ and 1 , but, compared to the unconstrained designs, much more measurements are made at the point 0 , which is 'cheap': the results are summarized in Table 5.

Finally, it is also feasible and possible to consider both types of constraints together, resulting in $\mathbf{A}^{(3)}w \leq b^{(3)}$ with

$$\mathbf{A}^{(3)} = \begin{pmatrix} \mathbf{A}^{(1)} \\ \mathbf{A}^{(2)} \end{pmatrix} \in \mathbb{R}^{4 \times 4q}, \quad b^{(3)} = \begin{pmatrix} b^{(1)} \\ b^{(2)} \end{pmatrix}. \quad (42)$$

The resulting D - and $IMSE$ -efficient designs for this constraint are given in Tables 6 and 7.

Note that in some cases, the additional constraints on the designs were saturated for a number of measurements that is lower than the maximum attainable number of measurements given by $(m_1, m_2)^\top$. This is demonstrated in a more detailed way in Fig. 2, where we again consider the D -efficient design with $(m_1, m_2) = (20, 40)$ and the cost constraints (41), but now the cost $b^{(2)}$ can vary between 0 and m_i for the i -th group. All the designs are supported in the points $-1, 0$ and 1 and the figure shows that when the maximum allowed cost is too low, the total maximum number of measurements is (sometimes significantly) lower than the corresponding m_i .

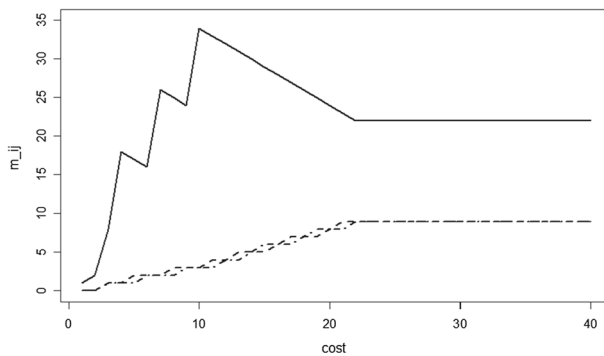


Fig. 2 The numbers of measurements in the point 0 (full line), -1 (dashed line) and 1 (dot-dashed line) for the second group in the D -efficient design in model (38) with $(m_1, m_2) = (20, 40)$ and constraints of the type (41) with the maximum cost $b^{(2)}$ in the second group varying between 0 and 40

7 Discussion

In the paper, we have considered equi- and invariance properties of approximate optimal designs in multiple-group mixed models. We have used these properties to fix the support points and, consequently, to reduce the number of unknown variables in first- and second-order models on a symmetric square. As we currently have no universal computational tool for approximate designs, these results can be used to determine optimal designs analytically in a few isolated and easy cases, as shown in the examples in Sect. 3.

However, from practical point of view, it is more important to be able to compute efficient exact designs, possibly even with some additional constraints given by the experimental conditions. We have shown a modified version of the algorithm of Harman et al. (2016) is a useful tool for such computations, even in the cases where there are several nontrivial constraints on the design.

In the models considered here, covariance matrix of random effects is assumed to be known. A natural question that arises while reading this work is how to perform in the situation where no prior knowledge about variances and covariances is available. In this case an estimation can be used. However, the quality of obtained designs depends on the accuracy of the estimation. For some particular structures of the covariance matrix it may happen that optimal designs turn out to be independent on the variance parameters (consider, for example, compound symmetry structure in Prus and Piepho 2021).

Acknowledgements This research was partially supported by Grant SCHW 531/16 of the German Research Foundation (DFG) and by Grant Number 1/0341/19 and 1/0362/22 from the Slovak Scientific Grant Agency (VEGA). A part of the research has been performed as the second author was visiting Comenius University of Bratislava in framework of the postdoc program (PKZ 91672803) of German Academic Exchange Service (DAAD). The authors are grateful to Radoslav Harman for fruitful discussions. Discussions with Rainer Schwabe on design criteria in multiple group-models were also very helpful.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix

See Tables 2, 3, 4, 5, 6 and 7.

Table 2 Exact D - and $IMSE$ -efficient designs in quadratic model on the interval $[-1, 1]$ with respect to the numbers of observations m_i for $D_i = \text{diag}(d_1, d_2, d_3)$, $i = 1, 2$

criterion (d_1, d_2, d_3)	D								IMSE					
	m_1	m_2	-1	0	1	-1	0	1	-1	0	1	-1	0	1
(1, 1, 1)	20	80	5	10	5	20	40	20	5	10	5	20	40	20
	50	50	12	25	13	13	25	12	13	25	12	12	25	13
	80	20	20	40	20	5	10	5	20	10	20	5	10	5
	40	160	10	20	10	40	80	40	10	20	10	40	80	40
	100	100	25	50	25	25	50	25	25	50	25	25	50	25
	160	40	40	80	40	10	20	10	40	80	40	10	20	10
(1, 1, 0)	20	80	5	10	5	20	40	20	9	2	9	15	49	16
	50	50	12	25	13	13	25	12	13	25	12	12	25	13
	80	20	20	40	20	5	10	5	15	49	16	9	2	9
	40	160	10	20	10	40	80	40	19	2	19	31	98	31
	100	100	25	50	25	25	50	25	25	50	25	25	50	25
	160	40	40	80	40	10	20	10	31	98	31	19	2	19
(0, 1, 1)	20	80	2	16	2	1	78	1	10	0	10	13	54	13
	50	50	1	48	1	1	48	1	12	26	12	13	24	13
	80	20	1	78	1	2	16	2	13	54	13	10	0	10
	40	160	2	36	2	1	158	1	20	0	20	27	106	27
	100	100	1	98	1	1	98	1	25	50	25	25	50	25
	160	40	1	158	1	2	36	2	27	106	27	20	0	20
(1, 0, 1)	20	80	8	4	8	40	1	39	4	12	4	25	30	25
	50	50	25	1	24	24	1	25	13	25	12	12	25	13
	80	20	39	1	40	8	4	8	25	30	25	4	12	4
	40	160	18	4	18	80	1	79	8	25	7	49	61	50
	100	100	50	1	49	49	1	50	25	50	25	25	50	25
	160	40	80	1	79	18	4	18	49	61	50	8	25	7
(1, 0, 0)	20	80	2	16	2	1	78	1	1	19	0	26	28	26
	50	50	1	48	1	1	48	1	13	25	12	12	25	13
	80	20	1	78	1	2	16	2	26	28	26	1	19	0
	40	160	2	36	2	1	158	1	1	38	1	53	54	53
	100	100	1	98	1	1	98	1	25	50	25	25	50	25
	160	40	1	158	1	2	36	2	53	54	53	1	38	1

Table 2 continued

criterion (d_1, d_2, d_3)	D								IMSE					
	m_1	m_2	-1	0	1	-1	0	1	-1	0	1	-1	0	1
(0, 1, 0)	20	80	10	0	10	19	42	19	10	0	10	16	48	16
	50	50	12	26	12	13	24	13	25	12	13	13	25	12
	80	20	19	42	19	10	0	10	16	48	16	10	0	10
	40	160	18	4	18	38	84	38	20	0	20	32	96	32
	100	100	25	50	25	25	50	25	25	50	25	25	50	25
	160	40	38	84	38	18	4	18	32	96	32	20	0	20
(0, 0, 1)	20	80	9	2	9	19	42	19	10	0	10	17	46	17
	50	50	12	26	12	13	24	13	12	26	12	13	24	13
	80	20	19	42	19	9	2	9	17	46	17	10	0	10
	40	160	18	4	18	38	84	38	20	0	20	35	91	34
	100	100	25	50	25	25	50	25	25	50	25	25	50	25
	160	40	38	84	38	18	4	18	34	91	35	20	0	20

Table 3 Exact D -efficient designs in quadratic model on the interval $[-1, 1]$ with constraints given by 40 and numbers of observations m_i for D_i given by 37 with $d_i = 0.5, i = 1, 2$

m_1	m_2	-1	-0.2	0	0.2	1	-1	-0.2	0	0.2	1
20	40	4	5	2	5	4	8	10	4	10	8
40	20	8	10	4	10	8	4	5	2	5	4
25	100	5	7	2	6	5	21	26	9	24	20
100	25	21	27	9	23	20	5	6	2	7	5

Table 4 Exact $IMSE$ -efficient designs in quadratic model on the interval $[-1, 1]$ with constraints given by (40) and numbers of observations m_i for D_i given by (37) with $d_i = 0.5, i = 1, 2$

m_1	m_2	-1	-0.2	0	0.2	1	-1	-0.2	0	0.2	1
20	40	5	5	0	5	5	9	10	1	10	10
40	20	9	10	1	10	10	5	5	0	5	5
25	100	6	6	0	7	6	23	26	3	24	24
100	25	24	25	3	25	23	6	6	0	7	6

Table 5 Exact D - (left) and $IMSE$ -efficient (right) designs in quadratic model on the interval $[-1, 1]$ with constraints given by (41) and numbers of observations m_i for D_i given by (37) with $d_i = 0.5, i = 1, 2$

(m_1, m_2)	-1	0	1	-1	0	1	-1	0	1	-1	0	1
(20,40)	2	17	1	3	34	3	2	6	2	4	23	3
(40,20)	3	34	3	2	17	1	4	23	3	2	6	2
(25,100)	3	14	3	8	74	8	3	14	3	9	52	9
(100,25)	8	74	8	3	14	3	8	63	9	3	14	3

Table 6 Exact D -efficient designs in quadratic model on the interval $[-1, 1]$ with constraints given by (42) and numbers of observations m_i for D_i given by (37) with $d_i = 0.5, i = 1, 2$

(m_1, m_2)	-1	-0.2	0	0.2	1	-1	-0.2	0	0.2	1
(20,40)	2	3	7	0	1	3	3	14	3	3
(40,20)	3	0	13	3	4	2	0	6	0	2
(25,100)	3	5	7	1	2	8	7	34	6	8
(100,25)	8	7	34	6	8	2	4	8	5	2

Table 7 Exact $IMSE$ -efficient designs in quadratic model on the interval $[-1, 1]$ with constraints given by (42) and numbers of observations m_i for D_i given by (37) with $d_i = 0.5, i = 1, 2$

(m_1, m_2)	-1	-0.2	0	0.2	1	-1	-0.2	0	0.2	1
(20,40)	2	0	6	0	2	4	0	12	3	3
(40,20)	4	0	12	0	4	2	0	6	0	2
(25,100)	3	1	7	2	3	8	2	38	0	12
(100,25)	9	1	35	0	11	3	2	7	1	3

References

- Aliev A, Fedorov V, Leonov S, McHugh B, Magee M (2012) Pkstamp library for constructing optimal population designs for pk/pd studies. *Commun Stat Simul Comput* 41(6):717–729
- Cook RD, Thibodeau L (1980) Marginally restricted D-optimal designs. *J Am Stat Assoc* 75(370):366–371
- Dumont C, Lestini G, Nagard HL, Mentré F, Comets E, Nguyen TT (2018) PFIM 4.0, an extended R program for design evaluation and optimization in nonlinear mixed-effect models. *Comput Methods Programs Biomed* 156:217–229
- Fedorov V, Jones B (2005) The design of multicentre trials. *Stat Methods Med Res* 14:205–248
- Galil Z, Kiefer J (1977) Comparison of designs for quadratic regression on cubes. *J Stat Plan Inference* 1:121–132
- Harman R, Bachratá A, Filová L (2016) Construction of efficient experimental designs under multiple resource constraints. *Appl Stoch Models Bus Ind* 32:3–17
- Heiligers B (1992) Admissible experimental designs in multiple polynomial regression. *J Stat Plan Inference* 31:219–233
- Nyberg J, Ueckert S, Strömberg EA, Hennig S, Karlsson MO, Hooker AC (2012) Poped: an extended, parallelized, nonlinear mixed effects models optimal design tool. *Comput Methods Programs Biomed* 108(2):789–805
- Prus M (2022) Equivalence theorems for multiple-design problems with application in mixed models. *J Stat Plan Inference* 217:153–164
- Prus M, Piepho H-P (2021) Optimizing the allocation of trials to sub-regions in multi-environment crop variety testing. *J Agric Biol Environ Stat* 26:267–288

- Prus M, Schwabe R (2016) Optimal designs for the prediction of individual parameters in hierarchical models. *J R Stat Soc B* 78:175–191
- Schmelter T (2007) Considerations on group-wise identical designs for linear mixed models. *J Stat Plan Inference* 137:4003–4010
- Schwabe R (1996) Optimum designs for multi-factor models. Springer, New York

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.