

Article

Optimal Designs for Multi-Group Linear Models with Measurement Errors

Min-Jue Zhang ^{1,†}, Min-Qian Liu ^{2,†}  and Xue-Ping Chen ^{1,*,†}¹ Department of Statistics, Jiangsu University of Technology, Changzhou 213001, China; zmjstat@jsut.edu.cn² NITFID, LPMC & KLMDASR, School of Statistics and Data Science, Nankai University, Tianjin 300071, China; mqliu@nankai.edu.cn

* Correspondence: chenxueping@jsut.edu.cn

† These authors contributed equally to this work.

Abstract

Multi-group linear models with measurement errors are frequently employed in situations where covariates cannot be precisely measured, thereby compromising the validity of between-group comparisons. However, the study of experimental design theory for these models is currently at an underdeveloped stage. This paper is concerned with the problem of constructing locally c -, D_A - and D -optimal designs of multi-group linear models with measurement errors for estimating parameters or contrasts in the model parameters. Equivalence theorems are established to confirm the optimality of the designs for such models under each criterion, and the generalization of Elfving's theorem is proved to describe the geometrical characterization of locally c -optimal designs for such models. Furthermore, the locally D -optimal designs for a class of multi-group linear models with measurement errors can be explicitly determined. It is shown that the locally D -optimal design for such models is given by the product of the locally D -optimal designs for linear measurement error models corresponding to those groups. To illustrate these concepts, several examples are provided.

Keywords: optimal design; equivalence theorem; Elfving's theorem; multi-group linear models with measurement errors

MSC: 62K05; 62K99

1. Introduction

In fields such as clinical medicine, agricultural science, and ecology, comparative analysis of parameters in multi-group linear models is a widely used analytical approach. This method quantifies differences in model parameters across various sub-groups (e.g., patient cohorts, crop varieties, and treatment phases) to investigate systematic variations among them. For instance, ref. [1] examined efficacy differences between two racial groups using the two-group linear dose–response regression model:

$$y_l = \beta_{l1} + \beta_{l2}x_l + e_l, \quad l = 1, 2, \quad (1)$$

where x_l denotes the drug dose for the l th group, y_l represents the corresponding observed efficacy, and β_{l1} and β_{l2} are the intercept and slope parameters of the model for that group, respectively. The intercept indicates the baseline effect, while the slope reflects the treatment effect. The difference in treatment effects between the two racial groups in model (1) can



Academic Editor: Félix Belzunce

Received: 13 February 2026

Revised: 6 March 2026

Accepted: 11 March 2026

Published: 13 March 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

be quantified as $\beta_{12} - \beta_{22}$. For further information on analytical methods for comparing parameters in classical multi-group linear models, please refer to [2–4].

In practice, however, covariates are often subject to measurement errors (MEs) due to human negligence, instrument inaccuracies, and systematic errors. For example, if there is an error in the measurement of drug dosage in model (1), the data can be modeled and estimated using the following two-group linear model with MEs:

$$\begin{aligned} y_l &= \beta_{l1} + \beta_{l2}x_l + e_l, \\ z_l &= x_l + u_l, \end{aligned} \quad l = 1, 2, \quad (2)$$

where z_l denotes the observed value of the true covariate x_l . Such multi-group linear models with MEs (2) have been applied in various fields, such as agriculture, chemistry, and medicine. As an illustration, ref. [5] employed this model to test for a significant difference in proficiency between two operators in measuring human lung capacity. Similar applications can be found in [6,7] and others. The existing literature on multi-group linear models with MEs has proposed several modeling and data analysis methods (see [8–10]). Nonetheless, the accuracy of these methods depends critically on the experimental settings used to collect the data.

Currently, the theory of experimental designs for multi-group linear models is primarily articulated within the framework of mixed models. For example, ref. [11] investigated this issue within the scope of random intercept models, while [12] constructed optimal designs for estimating model parameters and predicting individual treatment effects in multi-group linear random coefficient regression models without covariates. Additionally, ref. [13] derived optimal designs for estimating parameter contrasts in a linear mixed model incorporating group-specific treatment effects, as proposed by [14].

In contrast to mixed models, multi-group linear models with MEs are distinguished because both the response variable and the covariates are subject to errors. This fundamental distinction leads to significant differences in their theoretical foundations, such as the information matrix. Therefore, this paper aims to develop and evaluate optimal experimental designs for multi-group linear models with MEs. Specifically, we focus on local c -, D_A -, and D -optimality criteria, which are essential for estimating linear combinations of model parameters. This paper is organized as follows: Section 2 introduces a mathematical framework for multi-group linear models with MEs. Sections 3 and 4 explore the concepts of locally c -, D_A -, and D -optimal designs within the context of these models. This exploration includes the derivation of equivalence theorems to evaluate the optimality of approximate designs under each criterion and extends Elfving's theorem to provide a geometrical characterization of locally c -optimal designs. Furthermore, for a class of multi-group linear models with MEs where the regression functions share an identical structure across groups, we demonstrate that the locally D -optimal design is the product of the locally D -optimal designs for the corresponding linear ME models. Finally, concluding remarks are presented in Section 5, and some proofs of theorems are deferred to Appendix A.

2. Model Specification and Preliminaries

In this paper, we focus on a general form of multi-group linear models with MEs, formulated as follows:

$$y_{ij} = f^T(x_{ij})K_i\beta + e_{ij}, \quad i = 1, \dots, n, \quad j = 1, \dots, r_i, \quad (3)$$

where y_{ij} is a real-valued response, $x_{ij} \in \mathcal{X}$ indicates the j th true observation of the i th individual, with $\mathcal{X} \subset \mathbb{R}$ denoting the design space, n signifies the number of individuals and r_i

is the number of observations from individual i . The known regression function vector is $f(x_{ij}) = (f_1(x_{ij}), \dots, f_p(x_{ij}))^T$ and β is a q -dimensional unknown model parameter vector supposed to be the same for all individuals. The matrix K_i denotes the group membership of individual i , independent of the experimental setting x_{ij} chosen. Consequently, the individual parameter vector for individual i is expressed as $K_i\beta$, and e_{ij} is the ME in the response y_{ij} . The true observation x_{ij} , which is unknown, is also subject to ME u_{ij} . In particular, it is assumed that we observe

$$z_{ij} = x_{ij} + u_{ij}, \quad i = 1, \dots, n, \quad j = 1, \dots, r_i. \quad (4)$$

The vectors $\varepsilon_{ij} = (e_{ij}, u_{ij})^T$ of response errors e_{ij} and covariate errors u_{ij} for $i = 1, \dots, n$ and $j = 1, \dots, r_i$ are considered to be independent and normally distributed with mean $0_{2 \times 1}$ and a covariance matrix of the form

$$\Sigma_{eu} = \begin{pmatrix} \sigma_e^2 & \sigma_{eu} \\ \sigma_{eu} & \sigma_u^2 \end{pmatrix},$$

where $\sigma_{eu} = \sigma_{eu}$. To clarify models (3) and (4), we take model (2) from the introduction as an illustrative example, defining $f(x_{ij}) = (1, x_{ij})^T$ and $\beta = (\beta_{11}, \beta_{21}, \beta_{12}, \beta_{22})^T$. Thus, model (2) can be expressed as follows:

$$\begin{aligned} y_{ij} &= f^T(x_{ij})K_i\beta + e_{ij}, \\ z_{ij} &= x_{ij} + u_{ij}, \end{aligned} \quad i = 1, 2, \quad j = 1, \dots, r_i, \quad (5)$$

where $K_{(1)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ and $K_{(2)} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ correspond to the first and second groups, respectively.

Since the vector of MEs is normally distributed, the maximum likelihood estimator $\hat{\beta}$ is defined as the value of β that minimizes Formula (6),

$$\sum_{i=1}^n \sum_{j=1}^{r_i} (y_{ij} - f^T(x_{ij})K_i\beta, z_{ij} - x_{ij}) \Sigma_{eu}^{-1} (y_{ij} - f^T(x_{ij})K_i\beta, z_{ij} - x_{ij})^T. \quad (6)$$

This is based on the following assumptions:

- (i) The function $f^T(x)K_i\beta$ is continuous and has continuous first- and second-order derivatives with respect to both β and x for each individual i ;
- (ii) For every $\zeta > 0$, there exists a constant $\delta_\zeta > 0$ such that

$$Q_r(\beta) = \frac{1}{r} \sum_{i=1}^n \sum_{j=1}^{r_i} \inf_{y_{ij}=f^T(x_{ij})K_i\beta} (y_{ij} - f^T(x_{ij})K_i\beta, z_{ij} - x_{ij}) \Sigma_{eu}^{-1} (y_{ij} - f^T(x_{ij})K_i\beta, z_{ij} - x_{ij})^T > \delta_\zeta,$$

for β satisfying $|\beta - \beta_\star| > \zeta$, where $\beta_\star$ denotes the true value of the model parameter vector.

Let $F_i = (f^T(x_{i1}), \dots, f^T(x_{ir_i}))^T$ and $Y_i = (y_{i1}, \dots, y_{ir_i})^T$; from Theorem 3.2.1 of [8], we can obtain the asymptotic property of the maximum likelihood estimator $\hat{\beta}$ for the model parameter vector β :

$$\sqrt{r}(\hat{\beta} - \beta_\star) \xrightarrow{L} N(0, M_{all}^{-1}(\zeta, \beta_\star)),$$

where $\beta_\star$ denotes the true value of the model parameter vector, r is the total number of trials, $\xrightarrow{L}$ denotes convergence in distribution and $M_{all}(\zeta, \beta)$ is called the information matrix

of design ξ . Here we assume that there exists a design ξ , such that $M_{all}(\xi, \beta)$ is nonsingular. The information matrix $M_{all}(\xi, \beta)$ is given by

$$M_{all}(\xi, \beta) = \frac{(GK)^T V^{-1}(\beta) GK}{r}, \quad (7)$$

where $K = (K_1^T, \dots, K_n^T)^T$, $Y = (Y_1^T, \dots, Y_n^T)^T$, G is block diagonal with the matrices F_i as blocks and $V(\beta)$ is block diagonal with the matrices $V_i(\beta)$ as diagonal elements,

$$V_i(\beta) = \text{diag} \left\{ \sigma_i(x_{i1}, \beta), \sigma_i(x_{i2}, \beta), \dots, \sigma_i(x_{ir_i}, \beta) \right\},$$

$$\sigma_i(x_{ij}, \beta) = \left(1, \frac{\partial f^T(x_{ij})}{\partial x_{ij}} K_i \beta \right) \Sigma_{eu} \left(1, \frac{\partial f^T(x_{ij})}{\partial x_{ij}} K_i \beta \right)^T,$$

for $i = 1, \dots, n$, $j = 1, \dots, r_i$, and we require that both the matrices K and GK have a full column rank.

For the sake of generality, we assume that there are only L distinct groups, that is, L versions of K_i : $K_{(1)}, \dots, K_{(L)}$ with sizes $n_{(1)}, \dots, n_{(L)}$, summing up to n , and the number of observations is the same for all individuals of one group, hereafter denoted by $r_{(1)}, \dots, r_{(L)}$ for these groups, which is a common practice. In this circumstance, we study the group-wise identical designs, which require that all members of the same group follow the identical design settings, see [14]. In this paper, the group-wise identical design for the l th group is denoted by

$$\xi_{(l)} = (x_{l1}, \dots, x_{lk_l}; w_{l1}, \dots, w_{lk_l}; r_{(l)}), \quad l = 1, 2, \dots, L,$$

where $x_{l1}, \dots, x_{lk_l}$ are different experimental settings, with $r_{l1}, \dots, r_{lk_l}$ denoting the number of observations made at each corresponding setting, such that $\sum_{j=1}^{k_l} r_{lj} = r_{(l)}$. Moreover, $w_{lj} = r_{lj}/r_{(l)}$ signifies the proportion of observations taken at the setting x_{lj} .

Therefore, the overall information matrix in (7) for design $\xi = (\xi_{(1)}, \dots, \xi_{(L)})$ is

$$M_{all}(\xi, \beta) = \sum_{l=1}^L \tilde{\omega}_{(l)} K_{(l)}^T M_{(l)}(\xi_{(l)}, \beta) K_{(l)}, \quad (8)$$

where $\tilde{\omega}_{(l)} = n_{(l)} r_{(l)} / r$ for $l = 1, \dots, L$. This term represents the proportion of individual observations within the l th group relative to the total number of trials, and it is evident that $\sum_{l=1}^L \tilde{\omega}_{(l)} = 1$. The matrix $M_{(l)}(\xi_{(l)}, \beta) = F_{(l)}^T W_{(l)} V_{(l)}^{-1}(\beta) F_{(l)}$ denotes the information matrix for the design $\xi_{(l)}$ in the l th group under maximum likelihood estimation. Additionally, $W_{(l)}$ and $V_{(l)}(\beta)$ are diagonal matrices with w_{ij} and $\sigma_{(l)}(x_{lj}, \beta)$ as their diagonal elements, respectively, for $j = 1, \dots, k_l$. The design matrix $F_{(l)}$ is evaluated at the experimental settings $x_{l1}, \dots, x_{lk_l}$ (see [15]).

When comparing the difference in drug efficacy between groups, such as the treatment difference $\beta_{12} - \beta_{22}$ in model (5), a natural objective is to minimize the variance of its estimate $\text{Var}(c^T \hat{\beta}) = c^T M_{all}^{-1}(\xi, \beta) c / r$ or equivalently minimize the function

$$\psi_c(\xi, \beta) = c^T M_{all}^{-1}(\xi, \beta) c. \quad (9)$$

In general, for fixed $\tilde{\omega}_{(1)}, \dots, \tilde{\omega}_{(L)}$, the design $\xi^* = (\xi_{(1)}^*, \dots, \xi_{(L)}^*)$ is called c -optimal if it minimizes the function $\psi_c(\xi, \beta)$ over the space of all approximate groups of designs $(\xi_{(1)}, \dots, \xi_{(L)})$ on $\mathcal{X}^L$. Note that the information matrix of the multi-group linear model with MEs defined in (3) and (4) typically depends on the values of the parameters. One way to overcome this dependence is to replace these parameters by their ‘best guesses’, which could be based on prior experiments. The resulting optimal designs are known as locally optimal designs, see [16].

3. Locally c -Optimal Design

In this section, we focus on constructing locally c -optimal designs for some given contrast $c^T \beta$. The following equivalence theorem establishes the sufficient and necessary condition to verify whether a design is locally c -optimal.

Theorem 1. For the multi-group linear models with MEs defined in (3) and (4), let

$$\phi_c(x_1, \dots, x_L, \xi) = \sum_{l=1}^L \tilde{\omega}_{(l)} \left(\frac{f^T(x_l)}{\sqrt{\sigma_{(l)}(x_l, \beta)}} N_{(l)}(\xi, \beta) c \right)^2 \quad (10)$$

and

$$N_{(l)}(\xi, \beta) = K_{(l)} M_{all}^{-1}(\xi, \beta),$$

Then a design $\xi^* = (\xi_1^*, \xi_2^*, \dots, \xi_L^*)$ is locally c -optimal if and only if the inequality

$$\phi_c(x_1, \dots, x_L, \xi^*) \leq c^T \left(\sum_{l=1}^L \tilde{\omega}_{(l)} N_{(l)}^T(\xi^*, \beta) M_{(l)}(\xi_{(l)}^*, \beta) N_{(l)}(\xi^*, \beta) \right) c \quad (11)$$

holds for all $x_1, \dots, x_L \in \mathcal{X}$. Moreover, the equality in (11) is achieved for any $(x_1^*, \dots, x_L^*) \in \text{supp}(\xi_{(1)}^*) \times \dots \times \text{supp}(\xi_{(L)}^*)$, where $\text{supp}(\xi_{(l)})$ denotes the support points of design $\xi_{(l)}$ on $\mathcal{X}$, $l = 1, \dots, L$.

In addition, geometric characterization is also a highly effective tool for identifying optimal designs. The most celebrated geometric result in the theory of optimal design is attributed to [17], which offered a characterization of c -optimal designs, known as Elfving's theorem. This theorem asserts that the optimal design is located at the intersection of vector lines and the boundary of a convex set, referred to as the Elfving set. Elfving's theorem is recognized for its importance in various scientific fields, as evidenced by its citation in notable works such as [18–20]. In this paper, we will extend Elfving's theorem to the case of multi-group linear models with MEs. To achieve this goal, let us give the definition of the generalized Elfving set

$$\mathcal{R} = \text{conv} \left\{ \left(K_{(1)}^T, \dots, K_{(L)}^T \right) \begin{pmatrix} \frac{\sqrt{\tilde{\omega}_{(1)}} f(x_1)}{\sqrt{\sigma_{(1)}(x_1, \beta)}} & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \frac{\sqrt{\tilde{\omega}_{(L)}} f(x_L)}{\sqrt{\sigma_{(L)}(x_L, \beta)}} \end{pmatrix} \varepsilon \mid x_1, \dots, x_L \in \mathcal{X}, \|\varepsilon\|^2 = 1 \right\}, \quad (12)$$

where $\mathbf{0}$ represents the $p \times 1$ zero matrix, $\text{conv}(B)$ denotes the convex hull of the set B , and $\|\varepsilon\|$ refers to the Frobenius norm of the vector ε ; i.e., $\|\varepsilon\|^2 = \varepsilon^T \varepsilon$. Hence we obtain the following result.

Theorem 2. A design $\xi^* = (\xi_1^*, \xi_2^*, \dots, \xi_L^*)$ is locally c -optimal for multi-group linear models with MEs defined in (3) and (4) if and only if there exists a positive constant γ and an $L \times 1$ -dimensional vector ε satisfying

$$\|\varepsilon\|^2 = 1, \quad (13)$$

such that the point $\gamma c \in \mathbb{R}^q$ lies on the boundary of the generalized Elfving set $\mathcal{R}$ defined in (12) and has the representation

$$\gamma c = \sum_{u=1}^{k_1} \cdots \sum_{v=1}^{k_L} \omega_{1u}^* \times \cdots \times \omega_{Lv}^* \left(K_{(1)}^T, \dots, K_{(L)}^T \right) \begin{pmatrix} \frac{\sqrt{\tilde{\omega}_{(1)}} f(x_{1u}^*)}{\sqrt{\sigma_{(1)}^2(x_{1u}^*, \beta)}} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\sqrt{\tilde{\omega}_{(L)}} f(x_{Lv}^*)}{\sqrt{\sigma_{(L)}^2(x_{Lv}^*, \beta)}} \end{pmatrix} \varepsilon. \quad (14)$$

Example 1. In the case described in model (5), we wish to find the locally c -optimal designs for comparing baseline and treatment effects in model (5) respectively, where the design space is $\mathcal{X}^2 = [a, b]^2 \subset \mathbb{R}_+^2$.

1. Construction of the locally c -optimal design for comparing baseline effects.
The c -optimal design for comparing baseline effects in this model was obtained by direct calculation in [21]. We now determine the c -optimal design using Theorem 2. The generalized Elfving set in (12) for this model is

$$\mathcal{R} = \text{conv} \left\{ \begin{pmatrix} \frac{\sqrt{\tilde{\omega}_{(1)}}}{\sqrt{\sigma_e^2 + 2\beta_{12}\sigma_{eu} + \beta_{12}^2\sigma_u^2}} \\ 0 \\ \frac{\sqrt{\tilde{\omega}_{(1)}}x_1}{\sqrt{\sigma_e^2 + 2\beta_{12}\sigma_{eu} + \beta_{12}^2\sigma_u^2}} \\ 0 \end{pmatrix} \varepsilon_1 + \begin{pmatrix} \frac{0}{\sqrt{\tilde{\omega}_{(2)}}} \\ \frac{\sqrt{\tilde{\omega}_{(2)}}}{\sqrt{\sigma_e^2 + 2\beta_{22}\sigma_{eu} + \beta_{22}^2\sigma_u^2}} \\ 0 \\ \frac{\sqrt{\tilde{\omega}_{(2)}}x_2}{\sqrt{\sigma_e^2 + 2\beta_{22}\sigma_{eu} + \beta_{22}^2\sigma_u^2}} \end{pmatrix} \varepsilon_2 \mid x_1, x_2 \in \mathcal{X}, \varepsilon = (\varepsilon_1, \varepsilon_2)^T, \sum_{l=1}^2 \varepsilon_l^2 = 1 \right\}$$

and the boundary of $\mathcal{R}$ is obtained from the endpoints of the design region, a and b . Consequently, the optimal design for the l th group takes the form $\xi_{(l)}^* = (a, b; 1 - \omega_l, \omega_l; r_{(l)})$. In what follows, we define

$$q_1 = \frac{\tilde{\omega}_{(1)}}{\sigma_e^2 + 2\beta_{12}\sigma_{eu} + \beta_{12}^2\sigma_u^2} \quad \text{and} \quad q_2 = \frac{\tilde{\omega}_{(2)}}{\sigma_e^2 + 2\beta_{22}\sigma_{eu} + \beta_{22}^2\sigma_u^2} \quad (15)$$

to simplify some expressions.

For $\xi^* = (\xi_{(1)}^*, \xi_{(2)}^*)$ and $c = (1, -1, 0, 0)^T$, we have

$$\psi_c(\xi^*, \beta) = \frac{\omega_1(b^2 - a^2) + a^2}{q_1(a - b)^2\omega_1(1 - \omega_1)} + \frac{\omega_2(b^2 - a^2) + a^2}{q_2(a - b)^2\omega_2(1 - \omega_2)}. \quad (16)$$

A direct calculation shows that the weights ω_1^* and ω_2^* of the locally c -optimal design at $x = b$ are

$$\omega_1^* = \omega_2^* = \frac{a}{a + b},$$

and

$$\phi_c(x_1, x_2, \xi^*) = \frac{(a + b)^2(a + b - 2x_1)^2}{q_1(a - b)^4} + \frac{(a + b)^2(a + b - 2x_2)^2}{q_2(a - b)^4}.$$

Obviously, $\phi_c(x_1, x_2, \xi^*)$ attains its maximum $\frac{(a + b)^2(q_1 + q_2)}{(a - b)^2q_1q_2}$ at the four vertexes of $[a, b]^2$. Hence, ξ^* is locally c -optimal by Theorem 1.

2. Construction of the locally c -optimal design for comparing treatment effects.
Similarly to the above operations in part (i), let $\xi_{(l)}^* = (a, b; 1 - \omega_l, \omega_l; r_{(l)})$. For $\xi^* = (\xi_{(1)}^*, \xi_{(2)}^*)$ and $c = (0, 0, 1, -1)^T$, we have

$$\psi_c(\xi^*, \beta) = \frac{1}{q_1(a - b)^2\omega_1(1 - \omega_1)} + \frac{1}{q_2(a - b)^2\omega_2(1 - \omega_2)}. \quad (17)$$

It is straightforward to show that the weights ω_1^* and ω_2^* of the locally c -optimal design at $x = b$ are

$$\omega_1^* = \omega_2^* = \frac{1}{2}.$$

From Theorem 1, it can be easily verified that ξ^* is the locally c -optimal design for the second part of Example 1.

4. Locally D_A -Optimal Design

In practical applications, such as drug discovery, the functional estimation of model parameters, particularly for specific linear contrasts of these parameters, has garnered significant interest. Suppose we are concerned with s important linear contrasts, which are represented by rows in $A^T \beta$, and A is a $q \times s$ user-specified matrix with rank $s < q$. A simple calculation reveals that the covariance matrix for these s estimated linear combinations $A^T \hat{\beta}$ is $\text{Cov}(A^T \hat{\beta}) = A^T M_{all}^{-1}(\xi, \beta) A / r$.

We now define the local D_A -optimality criterion for estimating the s contrasts $A^T \beta$ as follows

$$\psi_{D_A}(\xi, \beta) = \log \left\{ \left| A^T M_{all}^{-1}(\xi, \beta) A \right| \right\}, \quad (18)$$

and a design that minimizes $\psi_{D_A}(\xi, \beta)$ is termed the locally D_A -optimal design. Such a design will ensure that all these contrasts are estimated as precisely as possible given the total sample size for this study. The following theorem provides a characterization of a locally D_A -optimal design.

Theorem 3. For the multi-group linear models with MEs defined in (3) and (4), let

$$\phi_{D_A}(x_1, \dots, x_L, \xi) = \sum_{l=1}^L \frac{\tilde{\omega}_{(l)}}{\sigma_{(l)}(x_l, \beta)} f^T(x_l) N_{(l)}(\xi, \beta) A \left(A^T M_{all}^{-1}(\xi, \beta) A \right)^{-1} A^T N_{(l)}^T(\xi, \beta) f(x_l), \quad (19)$$

Then a design $\xi^* = (\xi_1^*, \xi_2^*, \dots, \xi_L^*)$ is locally D_A -optimal if and only if the inequality

$$\phi_{D_A}(x_1, \dots, x_L, \xi^*) \leq s \quad (20)$$

holds for all $x_1, x_2, \dots, x_L \in \mathcal{X}$. Moreover, the equality in (20) is achieved for any $(x_1^*, x_2^*, \dots, x_L^*) \in \text{supp}(\xi_{(1)}^*) \times \dots \times \text{supp}(\xi_{(L)}^*)$, where $\text{supp}(\xi_{(l)})$ denotes the support points of the design $\xi_{(l)}$ on $\mathcal{X}$, $l = 1, \dots, L$.

The proof follows from similar arguments to those in Theorem 1 and is therefore omitted for brevity.

Example 2. We now explore a more complex scenario than the one presented in Example 1, where our interest extends beyond the difference in the mean of the intercept, $\beta_{11} - \beta_{21}$, to also include the difference in the mean of the slope, $\beta_{12} - \beta_{22}$. The matrix A , which characterizes these contrasts, is as follows:

$$A = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}^T.$$

Let $\xi_{(l)}^*$ be defined as $(a, b; 1/2, 1/2; r_{(l)})$ for $l = 1, 2$, and $\xi^* = (\xi_{(1)}^*, \xi_{(2)}^*)$. We aim to demonstrate that ξ^* is locally D_A -optimal for the model in Example 2. For ξ^* , it is evident that

$$\phi_{D_A}(x_1, x_2, \xi^*) = \frac{q_2(2x_1 - (a + b))^2 + q_1(2x_2 - (a + b))^2 + (q_1 + q_2)(a - b)^2}{(q_1 + q_2)(a - b)^2},$$

in which q_1 and q_2 are as defined in (15). Moreover, it follows that $\phi_{D_A}(x_1, x_2, \xi^*)$ reaches its maximum at the four vertexes of $[a, b]^2$ and equals 2. Consequently, ξ^* is locally D_A -optimal, according to Theorem 3.

Now we investigate a special case of Theorem 3 in which the estimations of all model parameters are considered. This is D -optimality, where $A = I_q$. From Theorem 3, a direct calculation reveals the following corollary.

Corollary 1. For the multi-group linear models with MEs defined in (3) and (4), let

$$\phi_D(x_1, \dots, x_L, \xi) = \sum_{l=1}^L \frac{\tilde{\omega}_{(l)}}{\sigma_{(l)}(x_l, \beta)} f^T(x_l) K_{(l)} M_{all}^{-1}(\xi, \beta) K_{(l)}^T f(x_l), \quad (21)$$

Then a design ξ^* is locally D -optimal if and only if the inequality

$$\phi_D(x_1, \dots, x_L, \xi^*) \leq q \quad (22)$$

holds for all $x_1, \dots, x_L \in \mathcal{X}$. Moreover, the equality in (22) is achieved for any $(x_1^*, \dots, x_L^*) \in \text{supp}(\xi_{(1)}^*) \times \dots \times \text{supp}(\xi_{(L)}^*)$, where $\text{supp}(\xi_{(l)})$ denotes the support points of the design $\xi_{(l)}$ on $\mathcal{X}$, $l = 1, \dots, L$.

In the aforementioned cases, we have obtained the closed-form formulas of the locally c - and D_A -optimal designs for the simple multi-group linear model with MEs and applied the equivalence theorems to ascertain the optimalities of the above designs. However, when dealing with more complex settings, numerical approaches must be employed to determine the optimal design under each criterion. To this end, we utilized the CVX package [22] in MATLAB R2023a, which is specifically designed for disciplined convex programming. In the context of nonlinear models, CVX-based algorithms are widely recognized as more practical and flexible for solving optimal design problems under convex criteria compared to conventional statistical approaches. The specific algorithmic implementation is detailed on page 5 of [23]. It is also worth noting that for a class of multi-group linear models with MEs, the locally D -optimal design can be clearly determined.

Theorem 4. If $\bar{K} = (K_{(1)}^T, \dots, K_{(L)}^T)^T$ is an orthogonal matrix and $\xi_{(l)}^*$ is the locally D -optimal design for the linear ME model in the l th group, $l = 1, \dots, L$, then design $\xi^* = (\xi_{(1)}^*, \dots, \xi_{(L)}^*)$ is locally D -optimal for multi-group linear models with MEs defined in (3) and (4), which is not influenced by the proportion of individual observations within each group relative to the total number of trials.

The optimal design theory for multi-group linear models with MEs is used not only in estimating parameters or contrasts in the model parameters from various populations, but also in estimating them from paired data, as Example 3 will show.

Example 3. This example involves a dataset (cf., [24]) from an ecological study of heavy metal contamination, where twenty-five tanks contained twenty trout, which were exposed to various doses of copper in mg/L. The response variables, denoted as y_1 and y_2 , represent the proportions of dead fish after 8 and 24 h, respectively. The predictor variable, x , is the dose of copper, which was transformed into a logarithmic scale. Based on the dataset, two fitted quadratic polynomial regression models as in (23) were proposed by [25] to compare whether the proportions of dead fish after 8 and 24 h are different over the interval $x \in [5.6, 7.3]$,

$$y_{lj} = (1, x_{lj}, x_{lj}^2) K_l \beta + e_{lj}, \quad l = 1, 2, j = 1, \dots, k_l, \quad (23)$$

where K_l equals either $K_{(1)}$ for the survival status of trout after 8 h or $K_{(2)}$ for the survival status of trout after 24 h with

$$K_{(1)} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \text{ and } K_{(2)} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Suppose that prior information regarding the model parameter value is available from preliminary experiments; that is

$$\boldsymbol{\beta} = (5.7779, -16.2031, -1.9381, 4.6715, 0.1622, -0.3163)^T.$$

In what follows, we consider the construction of the locally D -optimal design over the interval $\mathcal{X} = [5.6, 7.3]$ for the model (23) but with MEs. The MEs are assumed to be

$$z_{lj} = x_{lj} + u_{lj}, \quad l = 1, 2, j = 1, \dots, k_l, \quad (24)$$

For the sake of simplicity, the covariance matrix of e_{lj} and u_{lj} is assumed to be a 2×2 diagonal matrix, i.e., $\Sigma_{eu} = \text{diag}(\sigma_e^2, \lambda\sigma_e^2)$, where $\lambda = \sigma_u^2/\sigma_e^2$ is referred to as the error-variances ratio, see [8,26]. It is remarkable that (7) reduces to the information matrix of the corresponding multi-group linear regression model if $\lambda = 0$.

Locally D -optimal designs. According to the existing conditions, we know that

$$\bar{K} = \begin{pmatrix} K_{(1)} \\ K_{(2)} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

is an orthogonal matrix. The locally D -optimal design for two-group polynomial models with MEs as defined in (23) and (24) can be found straightforwardly using Corollary 1, which is of the following form:

$$\boldsymbol{\xi}^* = (\boldsymbol{\xi}_{(1)}^*, \boldsymbol{\xi}_{(2)}^*).$$

Here $\boldsymbol{\xi}_{(l)}^* = (x_{l1}, x_{l2}, x_{l3}; 1/3, 1/3, 1/3)$ must be the locally D -optimal design for the corresponding polynomial ME model of the l th group for $l = 1, 2$, which can be solved by [15]. Table 1 lists the support points of locally D -optimal designs $\boldsymbol{\xi}^* = (\boldsymbol{\xi}_{(1)}^*, \boldsymbol{\xi}_{(2)}^*)$ for models (23) and (24) with the given $\boldsymbol{\beta}$ for various λ .

Figure 1 presents plots of the function $\phi_D(x_1, x_2, \boldsymbol{\xi}^*)$ defined in (21) for locally D -optimal designs $\boldsymbol{\xi}^*$. These plots illustrate the expected properties of an optimal design, specifically, the function $\phi_D(x_1, x_2, \boldsymbol{\xi}^*)$ being less than or equal to the value on the right-hand side of inequality (22) across the design space, with equality at the support points of $\boldsymbol{\xi}^*$. By the equivalence theorem, the plots confirm the local D -optimality of the designs $\boldsymbol{\xi}^*$.

The D -efficiency. In the process of numerical calculation, we note that the locally D -optimal designs are closely related to the values of λ for fixed $\boldsymbol{\beta}$. To compare the performance of the different designs for two-group polynomial models with MEs as defined in (23) and (24), we employ the D -efficiency defined for a design $\boldsymbol{\xi}$ as

$$\text{eff}_D(\boldsymbol{\xi}; \boldsymbol{\xi}_\lambda) = \left(\frac{|M_{all}(\boldsymbol{\xi}, \boldsymbol{\beta})|}{|M_{all}(\boldsymbol{\xi}_\lambda^*, \boldsymbol{\beta})|} \right)^{1/\nu}, \quad (25)$$

where v is the number of the model parameters and ξ_{λ}^* is the locally D -optimal design for two-group polynomial models with MEs using the given error-variance ratio λ . We have conducted a systematic calculation of the efficiencies of the locally D -optimal designs for two-group polynomial models with MEs as defined in (23) and (24) when the error-variance ratio deviates from the true value. Table 2 shows that the optimal design ξ_{λ}^* is sensitive to the error-variance ratio λ , where λ_{mis} is the misspecified error-variance ratio and λ_* is the true error-variance ratio. It is observed that $\text{eff}_D(\xi_{\lambda_{\text{mis}}}^*; \xi_{\lambda_*}^*)$ decreases moderately as the error-variance ratio deviates from the true value λ_* . Ulteriorly, this characterization becomes more pronounced if the dose range of copper is larger.

Table 1. The support points of locally D -optimal designs $\xi^* = (\xi_{(1)}^*, \xi_{(2)}^*)$ for two-group polynomial models with MEs as defined in (23) and (24) with $\lambda = 0, 0.5, 1, 5, 10, 50$ and the given β .

λ	Support Points with Equal Weights
0	$\{(5.60, 5.60), (5.60, 6.45), (5.60, 7.30), (6.45, 5.60), (6.45, 6.45), (6.45, 7.30), (7.30, 5.60), (7.30, 6.45), (7.30, 7.30)\}$
0.5	$\{(5.60, 5.60), (5.60, 6.50), (5.60, 7.30), (6.44, 5.60), (6.44, 6.50), (6.44, 7.30), (7.30, 5.60), (7.30, 6.50), (7.30, 7.30)\}$
1	$\{(5.60, 5.60), (5.60, 6.54), (5.60, 7.30), (6.43, 5.60), (6.43, 6.54), (6.43, 7.30), (7.30, 5.60), (7.30, 6.54), (7.30, 7.30)\}$
5	$\{(5.60, 5.60), (5.60, 6.69), (5.60, 7.30), (6.38, 5.60), (6.38, 6.69), (6.38, 7.30), (7.30, 5.60), (7.30, 6.69), (7.30, 7.30)\}$
10	$\{(5.60, 5.60), (5.60, 6.76), (5.60, 7.30), (6.33, 5.60), (6.33, 6.76), (6.33, 7.30), (7.30, 5.60), (7.30, 6.76), (7.30, 7.30)\}$
50	$\{(5.60, 5.60), (5.60, 6.89), (5.60, 7.30), (6.17, 5.60), (6.17, 6.89), (6.17, 7.30), (7.30, 5.60), (7.30, 6.89), (7.30, 7.30)\}$

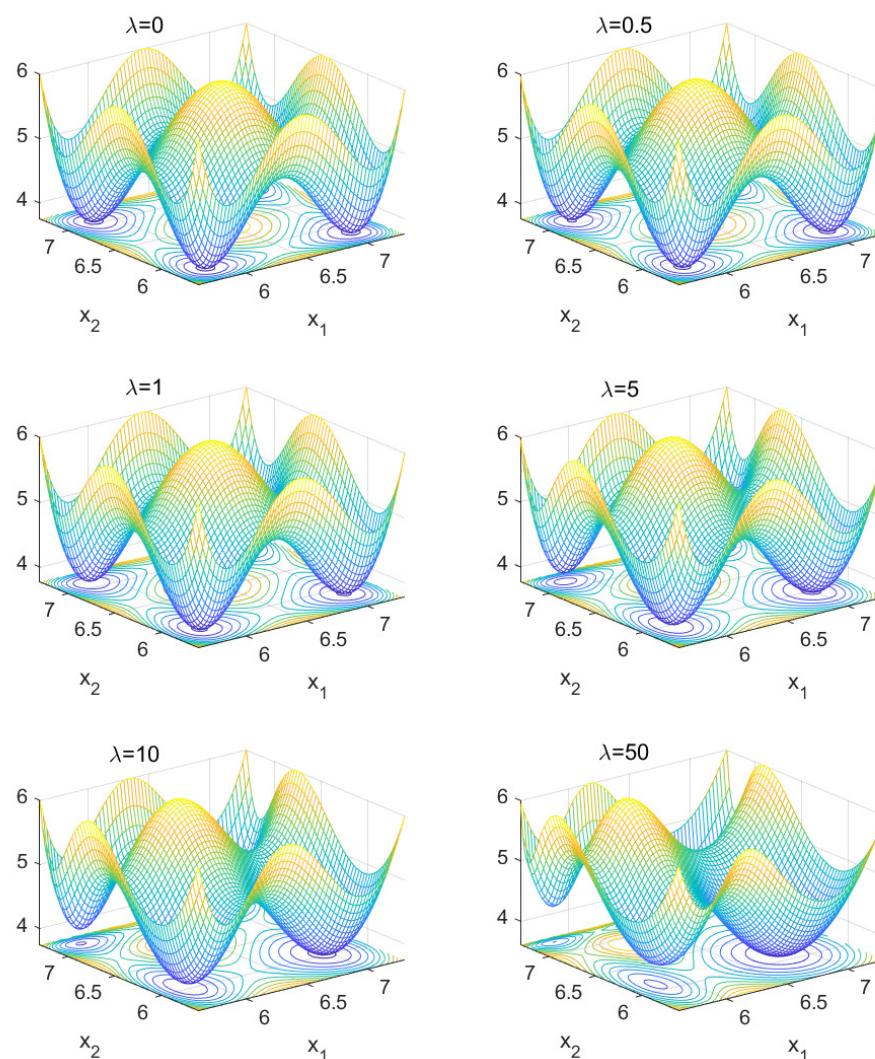


Figure 1. The plots of the function $\phi_D(x_1, x_2, \xi^*)$ defined by (21) for locally D -optimal designs ξ^* in Example 3 with $\lambda = 0, 0.5, 1, 5, 10, 50$ and the given β .

Table 2. Efficiencies of the locally D -optimal designs for two-group polynomial models with MEs as defined in (23) and (24) with the misspecified error-variance ratio λ_{mis} compared to the true error-variance ratio λ_* .

λ_*	λ_{mis}					
	0	0.5	1	5	10	50
0	1.0000	0.9988	0.9961	0.9705	0.9472	0.8674
0.5	0.9985	1.0000	0.9993	0.9808	0.9602	0.8841
1	0.9955	0.9990	1.0000	0.9872	0.9692	0.8966
5	0.9697	0.9804	0.9875	1.0000	0.9952	0.9459
10	0.9472	0.9608	0.9705	0.9961	1.0000	0.9697
50	0.8622	0.8795	0.8933	0.9420	0.9672	1.0000

5. Concluding Remarks

This paper delves into the optimalities for multi-group linear models with MEs under maximum likelihood estimation, including local c -, D_A - and D -optimalities. We present the corresponding equivalence theorems and demonstrate their application in evaluating the optimalities of designs through illustrative examples. It is worth noting that our primary focus is on the assumption of normally distributed errors in multi-group linear models with MEs. However, similar results can also be obtained using the methodology introduced in this paper under alternative distributional assumptions. As previously noted by [13], optimal design problems for multi-group linear models are prevalent in practical applications. Our findings can be beneficial for design problems when multi-group linear models with MEs are employed.

The construction method of the locally D -optimal designs for a class of multi-group linear models with MEs has been provided. This approach is helpful when the model structure is identical across groups, as shown in Example 3. In addition, the numerical results show that the locally D -optimal designs for the multi-group linear models with MEs are sensitive to the error-variance ratio. Hence, it is of particular importance to take MEs into account for planning the experiments in the presence of MEs in variables.

The newly obtained designs are usually locally optimal; that is, they often depend on the model parameter vector and the error-variance ratio. To avoid these, Bayesian and maximin approaches could be alternative ways to overcome the dependence of the information matrix on the model parameter vector and the error-variance ratio; these can be considered in the future.

Author Contributions: Methodology, M.-J.Z., M.-Q.L. and X.-P.C.; formal analysis, M.-J.Z.; writing—original draft, M.-J.Z.; writing—review and editing, M.-J.Z. and M.-Q.L.; supervision, M.-Q.L.; project administration, M.-Q.L. and X.-P.C.; funding acquisition, M.-J.Z. and M.-Q.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (Nos. 12301319, 12131001 and 12371260).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Proofs

Proof of Theorem 1. To illustrate the idea, we present a proof for the case of two distinct groups, i.e., $L = 2$. The argument for the general case will follow the same line. Let Ξ_l encompass all possible approximate designs, where each individual from group l contributes a predetermined number of $r_{(l)}$ observations for $l = 1, 2$, and define for $\xi = (\xi_{(1)}, \xi_{(2)}) \in \Xi_1 \times \Xi_2$

$$\begin{aligned}\overline{M}(\xi, \beta) &= \overline{M}(\xi_{(1)}, \xi_{(2)}, \beta) \\ &= \begin{pmatrix} \tilde{\omega}_{(1)} M_{(1)}(\xi_{(1)}, \beta) & 0_{p \times p} \\ 0_{p \times p} & \tilde{\omega}_{(2)} M_{(2)}(\xi_{(2)}, \beta) \end{pmatrix}.\end{aligned}\quad (\text{A1})$$

Obviously, the set $\mathfrak{M}^{(2)} = \left\{ \overline{M}(\xi_{(1)}, \xi_{(2)}, \beta) : (\xi_{(1)}, \xi_{(2)}) \in \Xi_1 \times \Xi_2 \right\}$ is a convex subset of the set NND ($2p$) of all non-negative definite $2p \times 2p$ matrices. Moreover, if δ_x denotes the Dirac measure at the point $x \in \mathcal{X}$, it is easy to see that $\mathfrak{M}^{(2)}$ is the convex hull of the set $\mathfrak{D}^{(2)} = \left\{ \overline{M}(\delta_{x_1}, \delta_{x_2}, \beta) : x_1, x_2 \in \mathcal{X} \right\}$. Note that for a matrix $\overline{M}(\xi, \beta) \in \mathfrak{M}^{(2)}$, the function $\psi_c(\xi, \beta)$ in (9) can be written as

$$\tilde{\psi}_c(\overline{M}(\xi, \beta)) = c^T \left(\overline{K}^T \overline{M}(\xi, \beta) \overline{K} \right)^{-1} c,$$

where

$$\overline{K} = \begin{pmatrix} K_{(1)} \\ K_{(2)} \end{pmatrix}.\quad (\text{A2})$$

To demonstrate the convexity of the function $\psi_c(\xi, \beta)$, it is equivalently necessary to show that the modified criterion $\tilde{\psi}_c(\overline{M}(\xi, \beta))$ exhibits convexity. Hence we first let $\xi_\alpha = \alpha \xi + (1 - \alpha) \bar{\xi}$, where ξ and $\bar{\xi}$ are any two approximate designs in $\Xi_1 \times \Xi_2$ and $\alpha \in [0, 1]$. Then

$$\overline{M}(\xi_\alpha, \beta) = \alpha \overline{M}(\xi, \beta) + (1 - \alpha) \overline{M}(\bar{\xi}, \beta)$$

and

$$(\overline{K}^T \overline{M}(\xi_\alpha, \beta) \overline{K})^{-1} \leq \alpha (\overline{K}^T \overline{M}(\xi, \beta) \overline{K})^{-1} + (1 - \alpha) (\overline{K}^T \overline{M}(\bar{\xi}, \beta) \overline{K})^{-1}.$$

Consequently,

$$\begin{aligned}\tilde{\psi}_c(\overline{M}(\xi_\alpha, \beta)) &= c^T \left(\overline{K}^T \overline{M}(\xi_\alpha, \beta) \overline{K} \right)^{-1} c \\ &\leq \alpha c^T \left(\overline{K}^T \overline{M}(\xi, \beta) \overline{K} \right)^{-1} c + (1 - \alpha) c^T \left(\overline{K}^T \overline{M}(\bar{\xi}, \beta) \overline{K} \right)^{-1} c \\ &= \alpha \tilde{\psi}_c(\overline{M}(\xi, \beta)) + (1 - \alpha) \tilde{\psi}_c(\overline{M}(\bar{\xi}, \beta)).\end{aligned}$$

Due to the convexity of ψ_c , the design $\xi^* = (\xi_{(1)}^*, \xi_{(2)}^*)$ is considered locally c -optimal if and only if the derivative of $\tilde{\psi}_c(\overline{M})$, evaluated at $\overline{M}^* = \overline{M}(\xi^*, \beta) = \overline{M}(\xi_{(1)}^*, \xi_{(2)}^*, \beta)$, is non-negative for all directions $E_0^* = E - \overline{M}^*$, where $E \in \mathfrak{M}^{(2)}$; i.e., the Fréchet derivative

$$\partial \tilde{\psi}_c(\overline{M}^*, E_0^*) = \lim_{\alpha \rightarrow 0^+} \frac{\tilde{\psi}_c(\overline{M}^* + \alpha E_0^*) - \tilde{\psi}_c(\overline{M}^*)}{\alpha} \geq 0.$$

Since $\mathfrak{M}^{(2)} = \text{conv}(\mathfrak{D}^{(2)})$, it suffices to verify the inequality for all $E \in \mathfrak{D}^{(2)}$. For the locally c -optimal design $\xi^* = (\xi_{(1)}^*, \xi_{(2)}^*)$, let $\xi_\alpha = (1 - \alpha)(\xi_{(1)}^*, \xi_{(2)}^*) + \alpha(\delta_{x_1}, \delta_{x_2})$, where $\alpha \in (0, 1)$. Note that $\overline{M}_\alpha = (1 - \alpha) \overline{M}^* + \alpha \overline{M}(\delta_{x_1}, \delta_{x_2}, \beta)$ and

$$M_{all}(\xi_\alpha, \beta) = \overline{K}^T \overline{M}_\alpha \overline{K}.$$

Assuming that integration and differentiation are interchangeable, the derivative at $\bar{M}^* = \bar{M}(\xi_{(1)}^*, \xi_{(2)}^*, \beta)$ in the direction of $E_0^* = E - \bar{M}^* = \bar{M}(\delta_{x_1}, \delta_{x_2}, \beta) - \bar{M}^*$ is given by

$$\begin{aligned}
 & \partial \tilde{\psi}_c(\bar{M}^*, E_0^*) \\
 &= \lim_{\alpha \rightarrow 0^+} \frac{\tilde{\psi}_c(\bar{M}^* + \alpha E_0^*) - \tilde{\psi}_c(\bar{M}^*)}{\alpha} \\
 &= \lim_{\alpha \rightarrow 0^+} \frac{\tilde{\psi}_c(\bar{M}^* + \alpha(E - \bar{M}^*)) - \tilde{\psi}_c(\bar{M}^*)}{\alpha} \\
 &= \lim_{\alpha \rightarrow 0^+} \frac{\tilde{\psi}_c((1 - \alpha)\bar{M}^* + \alpha E) - \tilde{\psi}_c(\bar{M}^*)}{\alpha} \\
 &= \lim_{\alpha \rightarrow 0^+} \frac{\tilde{\psi}_c((1 - \alpha)\bar{M}^* + \alpha \bar{M}(\delta_{x_1}, \delta_{x_2}, \beta)) - \tilde{\psi}_c(\bar{M}^*)}{\alpha} \\
 &= \frac{d}{d\alpha} \tilde{\psi}_c(\bar{M}_\alpha) \Big|_{\alpha=0^+} \\
 &= \frac{d}{d\alpha} c^T M_{all}^{-1}(\xi_\alpha, \beta) c \Big|_{\alpha=0^+} \\
 &= -c^T M_{all}^{-1}(\xi_\alpha, \beta) \left(\frac{d}{d\alpha} M_{all}(\xi_\alpha, \beta) \right) M_{all}^{-1}(\xi_\alpha, \beta) c \Big|_{\alpha=0^+} \\
 &= c^T M_{all}^{-1}(\xi^*, \beta) \left(\bar{K}^T \bar{M}^* \bar{K} - \bar{K}^T \bar{M}(\delta_{x_1}, \delta_{x_2}, \beta) \bar{K} \right) M_{all}^{-1}(\xi^*, \beta) c \\
 &= c^T \left(\sum_{l=1}^2 n_l r_{(l)} N_{(l)}^T(\xi^*, \beta) M_{(l)}(\xi_{(l)}^*, \beta) N_{(l)}(\xi^*, \beta) \right) c \\
 &\quad - \sum_{l=1}^2 n_l r_{(l)} \left(\frac{f^T(x_l)}{\sqrt{\sigma_{(l)}(x_l, \beta)}} N_{(l)}(\xi^*, \beta) c \right)^2.
 \end{aligned}$$

Consequently, the design $\xi^* = (\xi_{(1)}^*, \xi_{(2)}^*)$ is locally c -optimal if and only if the inequality

$$\sum_{l=1}^2 n_l r_{(l)} \left(\frac{f^T(x_l)}{\sqrt{\sigma_{(l)}(x_l, \beta)}} N_{(l)}(\xi^*, \beta) c \right)^2 \leq c^T \left(\sum_{l=1}^2 n_l r_{(l)} N_{(l)}^T(\xi^*, \beta) M_{(l)}(\xi_{(l)}^*, \beta) N_{(l)}(\xi^*, \beta) \right) c$$

is satisfied for all $x_1, x_2 \in \mathcal{X}$.

For the proof of the equality part of Theorem 1, let us assume that

$$\phi_c(x_1, x_2, \xi^*) < c^T \left(\sum_{l=1}^2 n_l r_{(l)} N_{(l)}^T(\xi^*, \beta) M_{(l)}(\xi_{(l)}^*, \beta) N_{(l)}(\xi^*, \beta) \right) c,$$

Then

$$\int_{\mathcal{X}} \int_{\mathcal{X}} \phi_c(x_1, x_2, \xi^*) d\xi_{(1)}^*(x_1) d\xi_{(2)}^*(x_2) < c^T \left(\sum_{l=1}^2 n_l r_{(l)} N_{(l)}^T(\xi^*, \beta) M_{(l)}(\xi_{(l)}^*, \beta) N_{(l)}(\xi^*, \beta) \right) c.$$

Now for any design $\xi = (\xi_{(1)}, \xi_{(2)})$, from expression (10), we have

$$\begin{aligned}
& \int_{\mathcal{X}} \int_{\mathcal{X}} \phi_c(x_1, x_2, \xi) d\xi_{(1)}(x_1) d\xi_{(2)}(x_2) \\
&= \int_{\mathcal{X}} \int_{\mathcal{X}} \sum_{l=1}^2 n_l r_{(l)} \left(\frac{f^T(x_l)}{\sqrt{\sigma_{(l)}(x_l, \beta)}} N_{(l)}(\xi, \beta) c \right)^2 d\xi_{(1)}(x_1) d\xi_{(2)}(x_2) \\
&= \sum_{l=1}^2 \int_{\mathcal{X}} \int_{\mathcal{X}} n_l r_{(l)} \left(\frac{f^T(x_l)}{\sqrt{\sigma_{(l)}(x_l, \beta)}} N_{(l)}(\xi, \beta) c \right)^2 d\xi_{(1)}(x_1) d\xi_{(2)}(x_2) \\
&= \sum_{l=1}^2 \int_{\mathcal{X}} n_l r_{(l)} \left(\frac{f^T(x_l)}{\sqrt{\sigma_{(l)}(x_l, \beta)}} N_{(l)}(\xi, \beta) c \right)^2 d\xi_{(l)}(x_l) \\
&= \sum_{l=1}^2 n_l r_{(l)} c^T N_{(l)}^T(\xi, \beta) \left(\int_{\mathcal{X}} \frac{f(x_l) f^T(x_l)}{\sigma_{(l)}(x_l, \beta)} d\xi_{(l)}(x_l) \right) N_{(l)}(\xi, \beta) c \\
&= \sum_{l=1}^2 n_l r_{(l)} c^T N_{(l)}^T(\xi, \beta) M_{(l)}(\xi_{(l)}, \beta) N_{(l)}(\xi, \beta) c \\
&= c^T \left(\sum_{l=1}^2 n_l r_{(l)} N_{(l)}^T(\xi, \beta) M_{(l)}(\xi_{(l)}, \beta) N_{(l)}(\xi, \beta) \right) c.
\end{aligned}$$

This contradiction shows that equality in (11) must hold whenever $(x_1, x_2) \in \text{supp}(\xi_{(1)}^*) \times \text{supp}(\xi_{(2)}^*)$, which proves the assertion. $\square$

Proof of Theorem 2. Similarly to Theorem 1, we provide a proof for the case of two different groups, as the proof for scenarios with $L \geq 3$ is analogous. Suppose design $\xi^* = (\xi_{(1)}^*, \xi_{(2)}^*)$ minimizes $c^T M_{all}^{-1}(\xi, \beta) c$ or equivalently satisfies Theorem 1, where

$$\xi_{(1)}^* = \begin{pmatrix} x_{11}^* & \cdots & x_{1k_1}^* \\ \omega_{11}^* & \cdots & \omega_{1k_1}^* \end{pmatrix} \quad \text{and} \quad \xi_{(2)}^* = \begin{pmatrix} x_{21}^* & \cdots & x_{2k_2}^* \\ \omega_{21}^* & \cdots & \omega_{2k_2}^* \end{pmatrix}.$$

Based on Equations (A1) and (A2), the function $M_{all}(\xi^*, \beta)$ defined in Equation (8) can be written as

$$M_{all}(\xi^*, \beta) = \bar{K}^T \bar{M}(\xi^*, \beta) \bar{K},$$

and the support points of ξ^* can be denoted by $\text{supp}(\xi^*) = \{(x_{1u}^*, x_{2v}^*) \mid u = 1, \dots, k_1, v = 1, \dots, k_2\}$. We define $\gamma^2 = (c^T M_{all}^{-1}(\xi^*, \beta) c)^{-1}$ and $d = \gamma M_{all}^{-1}(\xi^*, \beta) c \in \mathbb{R}^q$; then, according to $c \in \text{Range}(M_{all}(\xi^*, \beta))$, it follows that

$$\begin{aligned}
\gamma c &= M_{all}(\xi^*, \beta) d \\
&= \bar{K}^T \bar{M}(\xi^*, \beta) \bar{K} d \\
&= \left(\bar{K}^T \begin{pmatrix} \tilde{\omega}_{(1)} \sum_{u=1}^{k_1} \omega_{1u}^* \frac{f(x_{1u}^*) f^T(x_{1u}^*)}{\sigma_{(1)}(x_{1u}^*, \beta)} & 0_{p \times p} \\ 0_{p \times p} & \tilde{\omega}_{(2)} \sum_{v=1}^{k_2} \omega_{2v}^* \frac{f(x_{2v}^*) f^T(x_{2v}^*)}{\sigma_{(2)}(x_{2v}^*, \beta)} \end{pmatrix} \bar{K} \right) d \\
&= \left(\sum_{u=1}^{k_1} \sum_{v=1}^{k_2} \omega_{uv}^* \bar{K}^T \begin{pmatrix} \tilde{\omega}_{(1)} \frac{f(x_{1u}^*) f^T(x_{1u}^*)}{\sigma_{(1)}(x_{1u}^*, \beta)} & 0_{p \times p} \\ 0_{p \times p} & \tilde{\omega}_{(2)} \frac{f(x_{2v}^*) f^T(x_{2v}^*)}{\sigma_{(2)}(x_{2v}^*, \beta)} \end{pmatrix} \bar{K} \right) d \\
&= \sum_{u=1}^{k_1} \sum_{v=1}^{k_2} \omega_{uv}^* \bar{K}^T \begin{pmatrix} \frac{\sqrt{\tilde{\omega}_{(1)}} f(x_{1u}^*)}{\sqrt{\sigma_{(1)}(x_{1u}^*, \beta)}} & 0_{p \times 1} \\ 0_{p \times 1} & \frac{\sqrt{\tilde{\omega}_{(2)}} f(x_{2v}^*)}{\sqrt{\sigma_{(2)}(x_{2v}^*, \beta)}} \end{pmatrix} \varepsilon_{uv},
\end{aligned}$$

where we have used the notation

$$\omega_{uv}^* = \omega_{1u}^* \times \omega_{2v}^* \quad \text{and} \quad \varepsilon_{uv} = \begin{pmatrix} \frac{\sqrt{\tilde{\omega}_{(1)}} f^T(x_{1u}^*)}{\sqrt{\sigma_{(1)}(x_{1u}^*, \beta)}} & 0_{1 \times p} \\ 0_{1 \times p} & \frac{\sqrt{\tilde{\omega}_{(2)}} f^T(x_{2v}^*)}{\sqrt{\sigma_{(2)}(x_{2v}^*, \beta)}} \end{pmatrix} \bar{K}d,$$

and by Theorem 1

$$\begin{aligned} \|\varepsilon_{uv}\|^2 &= \varepsilon_{uv}^T \varepsilon_{uv} \\ &= d^T \left(\tilde{\omega}_{(1)} K_{(1)}^T \frac{f(x_{1u}^*) f^T(x_{1u}^*)}{\sigma_{(1)}(x_{1u}^*, \beta)} K_{(1)} + \tilde{\omega}_{(2)} K_{(2)}^T \frac{f(x_{2v}^*) f^T(x_{2v}^*)}{\sigma_{(2)}(x_{2v}^*, \beta)} K_{(2)} \right) d \\ &= \gamma^2 c^T \left(\tilde{\omega}_{(1)} N_{(1)}^T(\xi^*, \beta) \frac{f(x_{1u}^*) f^T(x_{1u}^*)}{\sigma_{(1)}(x_{1u}^*, \beta)} N_{(1)}(\xi^*, \beta) + \right. \\ &\quad \left. \tilde{\omega}_{(2)} N_{(2)}^T(\xi^*, \beta) \frac{f(x_{2v}^*) f^T(x_{2v}^*)}{\sigma_{(2)}(x_{2v}^*, \beta)} N_{(2)}(\xi^*, \beta) \right) c \\ &= \gamma^2 c^T M_{all}^{-1}(\xi^*, \beta) c \\ &= 1, \end{aligned}$$

for $u = 1, \dots, k_1, v = 1, \dots, k_2$. Consequently, the conditions (13) and (14) are satisfied and it remains to show that γc lies on the boundary of the generalized Elfving set $\mathcal{R}$ defined in (12). For this purpose, by Cauchy's inequality and Theorem 1, we have

$$\begin{aligned} &\left(d^T \bar{K}^T \begin{pmatrix} \frac{\sqrt{\tilde{\omega}_{(1)}} f(x_1)}{\sqrt{\sigma_{(1)}(x_1, \beta)}} & 0_{p \times 1} \\ 0_{p \times 1} & \frac{\sqrt{\tilde{\omega}_{(2)}} f(x_2)}{\sqrt{\sigma_{(2)}(x_2, \beta)}} \end{pmatrix} \varepsilon \right)^2 \\ &\leq (\varepsilon^T \varepsilon) d^T \bar{K}^T \begin{pmatrix} \frac{\sqrt{\tilde{\omega}_{(1)}} f(x_1)}{\sqrt{\sigma_{(1)}(x_1, \beta)}} & 0_{p \times 1} \\ 0_{p \times 1} & \frac{\sqrt{\tilde{\omega}_{(2)}} f(x_2)}{\sqrt{\sigma_{(2)}(x_2, \beta)}} \end{pmatrix} \begin{pmatrix} \frac{\sqrt{\tilde{\omega}_{(1)}} f^T(x_1)}{\sqrt{\sigma_{(1)}(x_1, \beta)}} & 0_{1 \times p} \\ 0_{1 \times p} & \frac{\sqrt{\tilde{\omega}_{(2)}} f^T(x_2)}{\sqrt{\sigma_{(2)}(x_2, \beta)}} \end{pmatrix} \bar{K}d \\ &= \|\varepsilon\|^2 \gamma^2 c^T \left(\tilde{\omega}_{(1)} N_{(1)}^T(\xi^*, \beta) \frac{f(x_1) f^T(x_1)}{\sigma_{(1)}(x_1, \beta)} N_{(1)}(\xi^*, \beta) + \tilde{\omega}_{(2)} N_{(2)}^T(\xi^*, \beta) \frac{f(x_2) f^T(x_2)}{\sigma_{(2)}(x_2, \beta)} N_{(2)}(\xi^*, \beta) \right) c \\ &\leq \|\varepsilon\|^2 \\ &= 1 \end{aligned}$$

for all $x_1, x_2 \in \mathcal{X}$ and ε satisfying (13). Moreover, $d^T \gamma c = 1$ by the definition of the constant γ and the vector d . Consequently, the vector d defines a supporting hyperplane to the generalized Elfving set $\mathcal{R}$ at the point γc , which implies that γc lies on the boundary of the generalized Elfving set $\mathcal{R}$ defined in (12).

In order to prove the converse, suppose γc lies on the boundary of the generalized Elfving set $\mathcal{R}$ defined in (12), and (13) and (14) are satisfied. In this case, there exists a supporting hyperplane, say $d \in \mathbb{R}^q$, to the generalized Elfving set $\mathcal{R}$ at the boundary point γc ; that is,

$$d^T \gamma c = 1 \tag{A3}$$

and the inequality

$$\left| d^T \bar{K}^T \begin{pmatrix} \sqrt{\tilde{\omega}_{(1)}} \frac{f(x_1)}{\sqrt{\sigma_{(1)}(x_1, \beta)}} & 0_{p \times 1} \\ 0_{p \times 1} & \sqrt{\tilde{\omega}_{(2)}} \frac{f(x_2)}{\sqrt{\sigma_{(2)}(x_2, \beta)}} \end{pmatrix} \varepsilon \right| \leq 1 \quad (\text{A4})$$

holds for all $x_1, x_2 \in \mathcal{X}$ and ε satisfying (13). Defining

$$\varepsilon(x) = (\varepsilon_1(x_1), \varepsilon_2(x_2))^T,$$

where

$$\varepsilon_l(x_l) = \frac{\sqrt{\tilde{\omega}_{(l)}} f^T(x_l) K_{(l)} d}{\sqrt{d^T \left(\tilde{\omega}_{(1)} K_{(1)}^T \frac{f(x_1) f^T(x_1)}{\sigma_{(1)}(x_1, \beta)} K_{(1)} + \tilde{\omega}_{(2)} K_{(2)}^T \frac{f(x_2) f^T(x_2)}{\sigma_{(2)}(x_2, \beta)} K_{(2)} \right) d}}, \quad l = 1, 2, \quad (\text{A5})$$

we have $\|\varepsilon(x)\|^2 = 1$ and obtain from the inequality (A4)

$$d^T \left(\sum_{l=1}^2 \tilde{\omega}_{(l)} K_{(l)}^T \frac{f(x_l) f^T(x_l)}{\sigma_{(l)}(x_l, \beta)} K_{(l)} \right) d = \left| d^T \bar{K}^T \begin{pmatrix} \sqrt{\tilde{\omega}_{(1)}} \frac{f(x_1)}{\sqrt{\sigma_{(1)}(x_1, \beta)}} & 0_{p \times 1} \\ 0_{p \times 1} & \sqrt{\tilde{\omega}_{(2)}} \frac{f(x_2)}{\sqrt{\sigma_{(2)}(x_2, \beta)}} \end{pmatrix} \varepsilon(x) \right|^2 \leq 1 \quad (\text{A6})$$

for all $x_1, x_2 \in \mathcal{X}$. On the other hand, using the Cauchy–Schwarz inequality from (14) and (A3), we obtain

$$\begin{aligned} 1 &= (d^T \gamma c)^2 \\ &= \left(d^T \sum_{u=1}^{k_1} \sum_{v=1}^{k_2} \omega_{uv}^* \bar{K}^T \begin{pmatrix} \sqrt{\tilde{\omega}_{(1)}} \frac{f(x_{1u}^*)}{\sqrt{\sigma_{(1)}(x_{1u}^*, \beta)}} & 0_{p \times 1} \\ 0_{p \times 1} & \sqrt{\tilde{\omega}_{(2)}} \frac{f(x_{2v}^*)}{\sqrt{\sigma_{(2)}(x_{2v}^*, \beta)}} \end{pmatrix} \varepsilon_{uv} \right)^2 \\ &\leq \left(\sum_{u=1}^{k_1} \sum_{v=1}^{k_2} \omega_{uv}^* d^T \bar{K}^T \begin{pmatrix} \tilde{\omega}_{(1)} \frac{f(x_{1u}^*) f^T(x_{1u}^*)}{\sigma_{(1)}(x_{1u}^*, \beta)} & 0_{p \times p} \\ 0_{p \times p} & \tilde{\omega}_{(2)} \frac{f(x_{2v}^*) f^T(x_{2v}^*)}{\sigma_{(2)}(x_{2v}^*, \beta)} \end{pmatrix} \bar{K} d \right) \left(\sum_{u=1}^{k_1} \sum_{v=1}^{k_2} \omega_{uv}^* \|\varepsilon_{uv}\|^2 \right) \\ &= \left(d^T \bar{K}^T \begin{pmatrix} \tilde{\omega}_{(1)} \sum_{u=1}^{k_1} \omega_{1u}^* \frac{f(x_{1u}^*) f^T(x_{1u}^*)}{\sigma_{(1)}(x_{1u}^*, \beta)} & 0_{p \times p} \\ 0_{p \times p} & \tilde{\omega}_{(2)} \sum_{v=1}^{k_2} \omega_{2v}^* \frac{f(x_{2v}^*) f^T(x_{2v}^*)}{\sigma_{(2)}(x_{2v}^*, \beta)} \end{pmatrix} \bar{K} d \right) \\ &= \gamma^2 c^T M_{all}^{-1}(\xi^*, \beta) c \\ &= 1, \end{aligned}$$

and the above equality holds if and only if

$$\varepsilon_{uv} = \kappa_{uv} \begin{pmatrix} \sqrt{\tilde{\omega}_{(1)}} \frac{f^T(x_{1u}^*)}{\sqrt{\sigma_{(1)}(x_{1u}^*, \beta)}} & 0_{1 \times p} \\ 0_{1 \times p} & \sqrt{\tilde{\omega}_{(2)}} \frac{f^T(x_{2v}^*)}{\sqrt{\sigma_{(2)}(x_{2v}^*, \beta)}} \end{pmatrix} \bar{K} d, \quad u = 1, \dots, k_1, \quad v = 1, \dots, k_2. \quad (\text{A7})$$

Combining (A5) and (11), we obtain (A8)

$$\kappa_{uv} = \frac{1}{\sqrt{d^T \left(\tilde{\omega}_{(1)} K_{(1)}^T \frac{f(x_{1u}^*) f^T(x_{1u}^*)}{\sigma_{(1)}(x_{1u}^*, \beta)} K_{(1)} + \tilde{\omega}_{(2)} K_{(2)}^T \frac{f(x_{2v}^*) f^T(x_{2v}^*)}{\sigma_{(2)}(x_{2v}^*, \beta)} K_{(2)} \right) d}} = 1, \quad (\text{A8})$$

where $u = 1, \dots, k_1, v = 1, \dots, k_2$.

Consequently, (14) gives

$$\begin{aligned} \gamma c &= \sum_{u=1}^{k_1} \sum_{v=1}^{k_2} \omega_{uv}^* \bar{K}^T \begin{pmatrix} \tilde{\omega}_{(1)} \frac{f(x_{1u}^*) f^T(x_{1u}^*)}{\sigma_{(1)}(x_{1u}^*, \beta)} & 0_{p \times p} \\ 0_{p \times p} & \tilde{\omega}_{(2)} \frac{f(x_{2v}^*) f^T(x_{2v}^*)}{\sigma_{(2)}(x_{2v}^*, \beta)} \end{pmatrix} \bar{K} d \\ &= M_{all}(\xi^*, \beta) d. \end{aligned}$$

It follows from [27] that there exists a generalized inverse G of the matrix $M_{all}(\xi^*, \beta)$ such that $d = \gamma Gc$, and from (A6), we have for all $x_1, x_2 \in \mathcal{X}$

$$\gamma^2 c^T G \left(\sum_{l=1}^2 \tilde{\omega}_{(l)} K_{(l)}^T \frac{f(x_l) f^T(x_l)}{\sigma_{(l)}(x_l, \beta)} K_{(l)} \right) Gc \leq 1. \quad (\text{A9})$$

By (A3), we have $\gamma^2 = (c^T Gc)^{-1} = (c^T M_{all}^{-1}(\xi^*, \beta) c)^{-1}$ and the inequality (A9) reduces to (11). Consequently, the c -optimality of the design ξ^* follows from Theorem 1, which proves the assertion. $\square$

Proof of Theorem 4. If $\bar{K}$ is an orthogonal matrix, then

$$\bar{K}^{-1} = \bar{K}^T, \quad K_{(l)} K_{(l)}^T = I_p, \quad K_{(l)} K_{(t)}^T = 0, \quad l \neq t, \quad l, t = 1, \dots, L, \quad (\text{A10})$$

and $\text{rank}(K_{(l)}) = p = q/L$ for $l = 1, \dots, L$, which means the number of unknown parameters in each group of models is p .

Since $\xi_{(l)}^*$ is the locally D -optimal design for the linear ME model in the l th group, it yields the inequality based on Theorem 1 in [15],

$$\frac{1}{\sigma_{(l)}(x_l, \beta)} f^T(x_l) M_{(l)}^{-1}(\xi_{(l)}^*, \beta) f(x_l) \leq p \quad (\text{A11})$$

for any $x_l \in \mathcal{X}, l = 1, \dots, L$.

Let $\xi^* = (\xi_{(1)}^*, \dots, \xi_{(L)}^*)$; from (21), (22), (A10) and (A11), we obtain

$$\begin{aligned} &\phi_D(x_1, \dots, x_L, \xi^*) \\ &= \sum_{l=1}^L \frac{\tilde{\omega}_{(l)}}{\sigma_{(l)}(x_l, \beta)} f^T(x_l) K_{(l)} M_{all}^{-1}(\xi^*, \beta) K_{(l)}^T f(x_l) \\ &= \sum_{l=1}^L \frac{\tilde{\omega}_{(l)}}{\sigma_{(l)}(x_l, \beta)} f^T(x_l) K_{(l)} \bar{K}^{-1} \bar{M}^{-1}(\xi^*, \beta) (\bar{K}^T)^{-1} K_{(l)}^T f(x_l) \\ &= \sum_{l=1}^L \frac{\tilde{\omega}_{(l)}}{\sigma_{(l)}(x_l, \beta)} f^T(x_l) K_{(l)} \bar{K}^T \bar{M}^{-1}(\xi^*, \beta) \bar{K} K_{(l)}^T f(x_l) \\ &= \sum_{l=1}^L \frac{1}{\sigma_{(l)}(x_l, \beta)} f^T(x_l) M_{(l)}^{-1}(\xi_{(l)}^*, \beta) f(x_l) \\ &\leq Lp \\ &= q \end{aligned}$$

for all $(x_1, \dots, x_L) \in \mathcal{X}^L$. Here

$$M_{all}(\xi^*, \beta) = \bar{K}^T \bar{M}(\xi^*, \beta) \bar{K}$$

and

$$\bar{M}(\xi^*, \beta) = \begin{pmatrix} \tilde{\omega}_{(1)} M_{(1)}(\xi_{(1)}^*, \beta) & 0_{p \times p} & \cdots & 0_{p \times p} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{p \times p} & 0_{p \times p} & \cdots & \tilde{\omega}_{(L)} M_{(L)}(\xi_{(L)}^*, \beta) \end{pmatrix},$$

which completes the proof by means of Corollary 1. $\square$

References

1. Masayo, O.S.; Shintaro, H.; Nobushige, M. A comparison of the pharmacokinetics and drug safety among east asian populations. *Ther. Innov. Regul. Sci.* **2014**, *48*, 393–403. [\[CrossRef\]](#)
2. Park, M.S. Multiple Comparison based on the Bootstrapping Methods. *J. Kor. Data Anal. Soc.* **2013**, *15*, 187–198.
3. Chen, S.; Wu, H.; Zhao, H.W. A comparison of causal inference methods for evaluating multiple treatment groups. *J. Nonparametr. Stat.* **2025**, *37*, 1317–1340. [\[CrossRef\]](#)
4. Budig, S.; Vogel, C.; Schaarschmidt, F. Multiple Comparisons with Overdispersed Multinomial Data: Methods, Properties and Application. *Pharm. Stat.* **2026**, *25*, e70073. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Sandoval-Moreno, G.; Galea, M.; Arellano-Valle, R. Inference in multivariate regression models with measurement errors. *J. Stat. Comput. Simul.* **2023**, *93*, 1997–2025. [\[CrossRef\]](#)
6. Le, Y.; Bai, Y.; Qin, G.Y. Subgroup analysis of linear models with measurement error. *Can. J. Stat.* **2023**, *52*, 26–42. [\[CrossRef\]](#)
7. van Leeuwen, C.C.E.; Mulder, V.L.; Batjes, N.H.; Heuvelink, G.B.M. Effect of measurement error in wet chemistry soil data on the calibration and model performance of pedotransfer functions. *Geoderma* **2024**, *442*, 116762. [\[CrossRef\]](#)
8. Fuller, W.A. *Measurement Error Models*; Wiley: New York, NY, USA, 1987.
9. Carroll, R.; Ruppert, D.; Stefanski, L. *Measurement Error in Nonlinear Models: A Modern Perspective*, 2nd ed.; Chapman and Hall: London, UK, 2006.
10. Buonaccorsi, J.P. *Measurement Error: Models, Methods, and Applications*; Chapman and Hall: London, UK, 2010.
11. Schwabe, R.; Schmelter, T. On optimal designs in random intercept models. *Tatra. Mt. Math. Publ.* **2008**, *39*, 145–153.
12. Prus, M. Optimal designs in multiple group random coefficient regression models. *Test* **2020**, *29*, 233–254. [\[CrossRef\]](#)
13. Liu, X.; Yue, R.X.; Wong, W.K. Equivalence theorems for c and D_A -optimality for linear mixed effects models with applications to multitreatment group assignments in health care. *Scand. J. Stat.* **2022**, *49*, 1842–1859. [\[CrossRef\]](#)
14. Schmelter, T. Consideration on group-wise identical designs for linear mixed models. *J. Stat. Plann. Inference* **2007**, *137*, 4003–4010. [\[CrossRef\]](#)
15. Zhang, M.J.; Yue, R.X. Optimal designs for homoscedastic functional polynomial measurement error models. *AStA Adv. Stat. Anal.* **2021**, *105*, 485–501. [\[CrossRef\]](#)
16. Silvey, S.D. *Optimal Design*; Chapman and Hall: London, UK, 1980.
17. Elfving, G. Optimum allocation in linear regression theory. *Ann. Math. Stat.* **1952**, *23*, 255–262. [\[CrossRef\]](#)
18. Fan, S.K.; Chaloner, K. A geometric method for singular c -optimal designs. *J. Stat. Plann. Inference* **2003**, *113*, 249–257.
19. Studden, W.J. Elfving's theorem revisited. *J. Stat. Plann. Inference* **2005**, *130*, 85–94. [\[CrossRef\]](#)
20. Dette, H.; Holland-Letz, T. A geometric characterization of c -optimal designs for heteroscedastic regression. *Ann. Stat.* **2009**, *37*, 4088–4103. [\[CrossRef\]](#)
21. Liski, E.P.; Mandal, N.K.; Shah, K.R.; Sinha, B.K. *Topics in Optimal Design*; Lecture Notes in Statistics; Springer: New York, NY, USA, 2002; Volume 163.
22. Grant, M.C.; Boyd, S.P. *The CVX Users Guide, Release 2.0 (Beta)*; CVX Research Inc.: Stanford, CA, USA, 2013.
23. Wong, W.K.; Zhou, J. CVX based algorithms for constructing various optimal regression designs. *Scand. J. Stat.* **2019**, *47*, 374–391. [\[CrossRef\]](#)
24. Khattree, R.; Naik, D.N. *Applied Multivariate Statistics with SAS Software*; SAS Institute Inc: Cary, NC, USA, 1999.
25. Zhang, S.Y. An exact method for the multiple comparison of several polynomial regression models with applications in dose-response study. *AStA Adv. Stat. Anal.* **2018**, *102*, 413–429.
26. Konstantinou, M.; Dette, H. Locally optimal designs for errors-in-variables models. *Biometrika* **2015**, *102*, 951–958. [\[CrossRef\]](#)
27. Searle, S.R. *Matrix Algebra Useful for Statistics*; Wiley: New York, NY, USA, 1982.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.