Journal of Multivariate Analysis 180 (2020) 104663

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Locally optimal designs for multivariate generalized linear models

Osama Idais

Institute for Mathematical Stochastics, Otto-von-Guericke University Magdeburg, PF 4120, 39016 Magdeburg, Germany

A R T I C L E I N F O

Article history:
Received 1 November 2019
Received in revised form 20 July 2020
Accepted 20 July 2020
Available online 6 August 2020

AMS 2000 subject classifications:
62K05
62H12

Keywords:
Correlation matrix
Generalized linear model
Locally optimal design
Multivariate response
Saturated design

A B S T R A C T

The multivariate generalized linear model is considered. Each univariate response follows a generalized linear model. In this situation, the linear predictors and the link functions are not necessarily the same. The quasi-Fisher information matrix is obtained by using the method of generalized estimating equations. Then locally optimal designs for multivariate generalized linear models are investigated under the D- and A-optimality criteria. It turns out that under certain assumptions the optimality problem can be reduced to the marginal models. More precisely, a locally optimal saturated design for the univariate generalized linear models remains optimal for the multivariate structure in the set of all saturated designs. Moreover, the general equivalence theorem provides a necessary and sufficient condition under which the saturated design is locally D-optimal in the set of all designs. The results are applied for multivariate models with gamma-distributed responses. Furthermore, we consider a multivariate model with univariate gamma models having seemingly unrelated linear predictors. Under this constraint, locally D- and A-optimal designs are found as product of all D- and A-optimal designs, respectively for the marginal counterparts.

1. Introduction

Optimal designs of experiments allow to observe the response at certain settings of the control variables in order to achieve most precise estimates of the parameters in the statistical model. In many experiments, it is quite natural to observe more than one response at each setting of the control variables. This situation can be described by multivariate statistical models. Although deriving optimal designs for these models is more complicated than in univariate models, there are considerable efforts to find solutions of optimal designs for multivariate linear models (see Fedorov [6], Krafft and Schaefer [16], Chang [2], Kurotschka and Schwabe [17], Schwabe [25], Imhof [14], Huang et al. [12], Liu et al. [20], Yue et al. [32]). Recently, Rodríguez-Díaz and Sánchez-León [23] introduced analogous results to those in Kurotschka and Schwabe [17] for multiresponse models assuming double covariance structure (intra-correlation and inter-correlation). On the other hand, under multivariate nonlinear models the solutions of optimal designs depend on a prior knowledge of the model parameters. In this context, the research contributions in optimal design are limited (Heise and Myers [11], Zocchi and Atkinson [34], Fedorov and Leonov [8], Liu and Colditz [19]). However, numerical solutions were proposed in Wong et al. [29] to find optimal designs for multivariate linear and nonlinear models by using semi-definite programming (SDP).

In this paper, we consider the multivariate generalized linear model (MGLM). Here, each marginal model is addressed within the GLM framework. In practice, such a situation appears for example in a thermal spraying process where the in-flight properties (responses) follow gamma distributions and thus GLMs can be fitted. In this example of thermal spraying,

E-mail address: osama.idais@ovgu.de.

https://doi.org/10.1016/j.jmva.2020.104663
0047-259X/

optimal designs were derived by Dette et al. [5] based on maximizing the weighted sum of D-efficiencies under the marginal GLMs. Throughout, we will focus on local optimality approach (Chernoff [3]) to derive optimal designs for the MGLM by using the method of generalized estimating equations (GEEs) which was developed by Liang and Zeger [18]. The method of GEEs has been frequently employed to obtain optimal designs particularly, by making use of local optimality approach (Liu and Colditz [19]) and the Bayesian optimality approach (Woods and Van de Ven [30] and van de Ven and Woods [28]). Recently, Jankar et al. [15] employed GEEs to identify locally D-optimal crossover designs for GLMs in the presence of multiple treatments.

The purpose of this paper is to study local optimality of designs for multivariate generalized linear models under the D- and A-criteria. In Section 2, we introduce the model assumptions and optimality of designs. In Section 3, analytic solutions of optimal designs for MGLMs with emphasis on the impact of the marginal models are developed. The results are applied for gamma-distributed outcomes in Section 4. In Section 5, we concentrate on MGLMs with marginal univariate gamma models having seemingly unrelated linear predictors. Then optimal designs of product type are derived. Further applications and conclusions are provided in Section 6.

2. Preliminaries

Let $\boldsymbol{Y}(\boldsymbol{x}_1), \ldots, \boldsymbol{Y}(\boldsymbol{x}_n)$ be independent m-dimensional response variables for n experimental units taken at the experimental conditions $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ which belong to the experimental region $\mathcal{X} \subseteq \mathbb{R}^\nu$, $\nu \geq 1$. Here, $\boldsymbol{x}_i = (x_{1i}, \ldots, x_{\nu i})^\mathsf{T}$ is the ith value of the vector $\boldsymbol{x}$ of ν explanatory variables $x_1, \ldots, x_\nu$. Accordingly, $\boldsymbol{Y}(\boldsymbol{x}_i) = \big(Y_1(\boldsymbol{x}_i), \ldots, Y_m(\boldsymbol{x}_i)\big)^\mathsf{T}$ denotes the vector of responses for unit i at the point $\boldsymbol{x}_i$. That is, an m-dimensional real valued vector is observed instead of a single real valued random variable at each point $\boldsymbol{x}_i$, $i \in \{1, \ldots, n\}$. Otherwise speaking, there are n observations $Y_j(\boldsymbol{x}_1), \ldots, Y_j(\boldsymbol{x}_n)$ taken for each component $j \in \{1, \ldots, m\}$.

We assume that a single response $Y_{ij} := Y_j(\boldsymbol{x}_i)$ belongs to a one-parameter exponential family distributions in the canonical form

$$p(y_{ij}; \theta_{ij}) = \exp\Big(y_{ij}\theta_{ij} - b_j(\theta_{ij}) + c_j(y_{ij})\Big),$$

where $b_j(\cdot)$ and $c_j(\cdot)$ are known functions while θ_{ij} is a canonical parameter. Here, $\theta_{ij} := \theta_j(\boldsymbol{x}_i, \boldsymbol{\beta}_j)$ depends on $\boldsymbol{x}_i$ and the vector of model parameters $\boldsymbol{\beta}_j \in \mathbb{R}^{p_j}$, where $\boldsymbol{\beta}_j = \big(\beta_{j1}, \ldots, \beta_{jp_j}\big)^\mathsf{T}$. The expected mean is given by $\mathrm{E}(Y_{ij}) = \mu_j(\boldsymbol{x}_i, \boldsymbol{\beta}_j) = b_j'(\theta_{ij})$ with the variance function $V_j\big(\mu_j(\boldsymbol{x}_i, \boldsymbol{\beta}_j)\big) = b_j''(\theta_{ij})$ (see McCullagh and Nelder [21], Section 2.2.2).

Define $\boldsymbol{f}_j : \mathcal{X} \to \mathbb{R}^{p_j}$. The components of $\boldsymbol{f}_j$ are given by the functions $f_{j1}, \ldots, f_{jp_j}$ which are assumed to be linearly independent. In each component j the expected mean $\mu_j(\boldsymbol{x}_i, \boldsymbol{\beta}_j)$ is assumed to be related to a linear predictor

$$\eta_j(\boldsymbol{x}_i, \boldsymbol{\beta}_j) = \boldsymbol{f}_j^\mathsf{T}(\boldsymbol{x}_i)\boldsymbol{\beta}_j = \sum_{l=1}^{p_j} f_{jl}(\boldsymbol{x}_i)\beta_{jl}$$

via a one-to-one and differentiable link function g_j, i.e.,

$$\eta_j(\boldsymbol{x}_i, \boldsymbol{\beta}_j) = g_j(\mu_j(\boldsymbol{x}_i, \boldsymbol{\beta}_j)).$$

For each j the intensity function can be defined as

$$u_j(\boldsymbol{x}_i, \boldsymbol{\beta}_j) = \Big(V_j\big(g_j^{-1}\big(\boldsymbol{f}_j^\mathsf{T}(\boldsymbol{x}_i)\boldsymbol{\beta}_j\big)\big)\Big)^{-1}\Big(g_j'\big(g_j^{-1}\big(\boldsymbol{f}_j^\mathsf{T}(\boldsymbol{x}_i)\boldsymbol{\beta}_j\big)\big)\Big)^{-2}, \tag{1}$$

which is positive and depends on the value of the linear predictor $\boldsymbol{f}_j^\mathsf{T}(\boldsymbol{x}_i)\boldsymbol{\beta}_j$.

The total number of MGLM parameters is denoted by p, i.e., $p = \sum_{j=1}^m p_j$. The link functions g_j, $j \in \{1, \ldots, m\}$, are not necessarily similar and the single responses $Y_j(\boldsymbol{x}_i), j \in \{1, \ldots, m\}$, may belong to distinct one-parameter probability distributions. Let $\mathbf{f}(\boldsymbol{x}_i) = \mathrm{diag}\big(\boldsymbol{f}_1(\boldsymbol{x}_i), \ldots, \boldsymbol{f}_m(\boldsymbol{x}_i)\big)$ denote the $p \times m$ block diagonal multivariate regression function and $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\mathsf{T}, \ldots, \boldsymbol{\beta}_m^\mathsf{T})^\mathsf{T}$ be the stacked parameter vector of dimension $p \times 1$. Note that the components $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m$ of $\mathbf{f}$ are linearly independent functions. Denote by $\boldsymbol{\mu}(\boldsymbol{x}_i, \boldsymbol{\beta}) = (\mu_1(\boldsymbol{x}_i, \boldsymbol{\beta}_1), \ldots, \mu_m(\boldsymbol{x}_i, \boldsymbol{\beta}_m))^\mathsf{T}$ the vector of the expected means of unit i. Let $\boldsymbol{g}\big(\boldsymbol{\mu}(\boldsymbol{x}_i, \boldsymbol{\beta})\big) = \Big(g_1(\mu_1(\boldsymbol{x}_i, \boldsymbol{\beta}_1)), \ldots, g_m(\mu_m(\boldsymbol{x}_i, \boldsymbol{\beta}_m))\Big)^\mathsf{T}$ and $\mathbf{f}^\mathsf{T}(\boldsymbol{x}_i)\boldsymbol{\beta} = \Big(\boldsymbol{f}_1^\mathsf{T}(\boldsymbol{x}_i)\boldsymbol{\beta}_1, \ldots, \boldsymbol{f}_m^\mathsf{T}(\boldsymbol{x}_i)\boldsymbol{\beta}_m\Big)^\mathsf{T}$ be the vectors of the link functions and the linear predictors, respectively, of unit i. Then the MGLM is defined by

$$\boldsymbol{\eta}(\boldsymbol{x}_i, \boldsymbol{\beta}) = \boldsymbol{g}\big(\boldsymbol{\mu}(\boldsymbol{x}_i, \boldsymbol{\beta})\big) \quad \text{where} \quad \boldsymbol{\eta}(\boldsymbol{x}_i, \boldsymbol{\beta}) = \mathbf{f}^\mathsf{T}(\boldsymbol{x}_i)\boldsymbol{\beta}. \tag{2}$$

The simplest situation can be taken under identity links, i.e., $\boldsymbol{g}\big(\boldsymbol{\mu}\big) = \boldsymbol{\mu}$ for which the intensities $u_j(\boldsymbol{x}_i, \boldsymbol{\beta}_j)$, $j \in \{1, \ldots, m\}$, are constantly equal to 1 for all $i \in \{1, \ldots, n\}$. Therefore, the design problems can be addressed under the multivariate linear model. However, Liang and Zeger [18] mentioned that there is a lack of a rich class of distributions for the multivariate non-normal outcomes. Therefore, they proposed the method of GEEs to estimate the model parameters. It is noted that GEEs are considered as an extension of the score function for the GLM.

To employ the GEEs method we assume that the observations $\boldsymbol{Y}(\boldsymbol{x}_i)$, $i \in \{1, \ldots, n\}$, are uncorrelated across the n units while the components are correlated within each unit. Define $\boldsymbol{R}$ to be the $m \times m$ true correlation matrix of the components of each $\boldsymbol{Y}(\boldsymbol{x}_i)$ (see Crowder [4]). Here, $\boldsymbol{R} = \left(\rho_{jh}\right)_{j=1,\ldots,m}^{h=1,\ldots,m}$, $\rho_{jj} = 1 (1 \le j \le m)$, $-1 < \rho_{jh} < 1$ $(1 \le j < h \le m)$ which are assumed to be independent of $\boldsymbol{x}_i$ and $\boldsymbol{\beta}$. Throughout, $\boldsymbol{R}$ is assumed to be positive definite and its inverse is denoted by $\boldsymbol{R}^{-1} = \left(\rho^{(jh)}\right)_{j=1,\ldots,m}^{h=1,\ldots,m}$.

Remark 1. For a square matrix $\boldsymbol{B}$ if there exists a square matrix $\boldsymbol{C}$ such that $\boldsymbol{C}\boldsymbol{C}^{\mathsf{T}} = \boldsymbol{B}$, then we call $\boldsymbol{C}$ a square root of the matrix $\boldsymbol{B}$. If $\boldsymbol{B}$ is a diagonal matrix given by $\boldsymbol{B} = \mathrm{diag}(b_1, \ldots, b_m)$ then we can define its square root as $\boldsymbol{C} = \mathrm{diag}(b_1^{\frac{1}{2}}, \ldots, b_m^{\frac{1}{2}})$ and we denote $\boldsymbol{B}^{\frac{1}{2}} = \boldsymbol{C}$.

Define $\boldsymbol{A}(\boldsymbol{x}_i, \boldsymbol{\beta}) = \mathrm{diag}\Big(V_j\big(g_j^{-1}\big(\boldsymbol{f}_j^{\mathsf{T}}(\boldsymbol{x}_i)\boldsymbol{\beta}_j\big)\big)\Big)_{j=1}^m$ and $\boldsymbol{\Delta}(\boldsymbol{x}_i, \boldsymbol{\beta}) = \mathrm{diag}\left(1/g_j'\big(g_j^{-1}\big(\boldsymbol{f}_j^{\mathsf{T}}(\boldsymbol{x}_i)\boldsymbol{\beta}_j\big)\big)\right)_{j=1}^m$ for all i. Liang and Zeger [18] showed that the observation $\boldsymbol{Y}(\boldsymbol{x}_i)$ has the covariance structure $\mathrm{Cov}(\boldsymbol{Y}(\boldsymbol{x}_i)) = \boldsymbol{\Sigma}(\boldsymbol{x}_i, \boldsymbol{\beta})$ where

$$\boldsymbol{\Sigma}(\boldsymbol{x}_i, \boldsymbol{\beta}) = \boldsymbol{A}^{\frac{1}{2}}(\boldsymbol{x}_i, \boldsymbol{\beta})\boldsymbol{R}\boldsymbol{A}^{\frac{1}{2}}(\boldsymbol{x}_i, \boldsymbol{\beta}).$$

We define the p-vector of quasi-score functions by $\boldsymbol{U}\big(\boldsymbol{\beta}\big) = \sum_{i=1}^n \mathbf{f}(\boldsymbol{x}_i)\boldsymbol{\Delta}(\boldsymbol{x}_i, \boldsymbol{\beta})\,\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}_i, \boldsymbol{\beta})\,\big(\boldsymbol{Y}(\boldsymbol{x}_i) - \boldsymbol{\mu}(\boldsymbol{x}_i, \boldsymbol{\beta})\big)$. The maximum quasi-likelihood estimates $\hat{\boldsymbol{\beta}}$ is the solution of the generalized estimating equations $\boldsymbol{U}(\boldsymbol{\beta}) = \mathbf{0}_p$, where $\mathbf{0}_p$ is a p-vector of zeros, see Crowder [4]. The quasi-Fisher information matrix for the MGLM at the point $\boldsymbol{x}_i$ is given by

$$\boldsymbol{M}(\boldsymbol{x}_i, \boldsymbol{\beta}) = \mathbf{f}(\boldsymbol{x}_i)\boldsymbol{\Delta}(\boldsymbol{x}_i, \boldsymbol{\beta})\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}_i, \boldsymbol{\beta})\,\boldsymbol{\Delta}(\boldsymbol{x}_i, \boldsymbol{\beta})\mathbf{f}^{\mathsf{T}}(\boldsymbol{x}_i).$$

For each component j, given a parameter vector $\boldsymbol{\beta}_j$, we define the function $\boldsymbol{f}_{j,\boldsymbol{\beta}_j}(\boldsymbol{x}_i) = u_j^{\frac{1}{2}}(\boldsymbol{x}_i, \boldsymbol{\beta}_j)\boldsymbol{f}_j(\boldsymbol{x}_i)$, $j \in \{1, \ldots, m\}$ which constitute the $p \times m$ matrix $\mathbf{f}_{\boldsymbol{\beta}}(\boldsymbol{x}_i) = \mathrm{diag}\Big(\boldsymbol{f}_{1,\boldsymbol{\beta}_1}(\boldsymbol{x}_i), \ldots, \boldsymbol{f}_{m,\boldsymbol{\beta}_m}(\boldsymbol{x}_i)\Big)$. According to (1), it is straightforward to obtain

$$\boldsymbol{\Delta}(\boldsymbol{x}_i, \boldsymbol{\beta})\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}_i, \boldsymbol{\beta})\boldsymbol{\Delta}(\boldsymbol{x}_i, \boldsymbol{\beta}) = \mathrm{diag}\Big(u_j^{\frac{1}{2}}(\boldsymbol{x}_i, \boldsymbol{\beta}_j)\Big)_{j=1}^m \boldsymbol{R}^{-1}\,\mathrm{diag}\Big(u_j^{\frac{1}{2}}(\boldsymbol{x}_i, \boldsymbol{\beta}_j)\Big)_{j=1}^m, \tag{3}$$

and thus we have $\boldsymbol{U}\big(\boldsymbol{\beta}\big) = \sum_{i=1}^n \mathbf{f}_{\boldsymbol{\beta}}(\boldsymbol{x}_i)\boldsymbol{R}^{-1}\big(\boldsymbol{Y}(\boldsymbol{x}_i) - \boldsymbol{\mu}(\boldsymbol{x}_i, \boldsymbol{\beta})\big)$. For the whole experiment we introduce the quasi-Fisher information matrix $\boldsymbol{M}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{\beta}) = \sum_{i=1}^n \boldsymbol{M}(\boldsymbol{x}_i, \boldsymbol{\beta}) = \sum_{i=1}^n \mathbf{f}_{\boldsymbol{\beta}}(\boldsymbol{x}_i)\boldsymbol{R}^{-1}\mathbf{f}_{\boldsymbol{\beta}}^{\mathsf{T}}(\boldsymbol{x}_i)$ which can be described in the block representation

$$\boldsymbol{M}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{\beta}) = \left(\rho^{(jh)} \sum_{i=1}^n \boldsymbol{f}_{j,\boldsymbol{\beta}_j}(\boldsymbol{x}_i)\boldsymbol{f}_{h,\boldsymbol{\beta}_h}^{\mathsf{T}}(\boldsymbol{x}_i)\right)_{j=1,\ldots,m}^{h=1,\ldots,m}. \tag{4}$$

Multi-dimensional observations are rearranged in matrix form in different ways. For the design point of view, we are to emphasize the relation of MGLM to its univariate GLM as for the linear case in Zellner [33], Krafft and Schaefer [16] and Kurotschka and Schwabe [17]. The observational vector of the whole experiment is obtained by vectorization of the design matrix, i.e., by stacking all m column vectors on top of each other starting with the vector of the n observations of the first component. To see that, let $\boldsymbol{Y}_j = \big(Y_j(\boldsymbol{x}_1), \ldots, Y_j(\boldsymbol{x}_n)\big)^{\mathsf{T}}$ be the observations of the jth component of the whole experiment $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$. The stacked vector of responses for all the n units at the whole experiment is thus denoted by $\boldsymbol{Y} = \big(\boldsymbol{Y}_1^{\mathsf{T}}, \ldots, \boldsymbol{Y}_m^{\mathsf{T}}\big)^{\mathsf{T}}$. Accordingly, the design matrix $\boldsymbol{F}$ for the multivariate model is written in component wise. So let $\boldsymbol{F}_j = [\boldsymbol{f}_j(\boldsymbol{x}_1), \ldots, \boldsymbol{f}_j(\boldsymbol{x}_n)]^{\mathsf{T}}$ be the $n \times p_j$ design matrix for the jth marginal model, then we obtain $\boldsymbol{F} = \mathrm{diag}\big(\boldsymbol{F}_1, \ldots, \boldsymbol{F}_m\big)$ which represents the stacked $mn \times p$ design matrix for the MGLM. As a result, the stacked vector of the linear predictors is given by

$$\boldsymbol{H} = [\boldsymbol{\eta}_1^{\mathsf{T}}, \ldots, \boldsymbol{\eta}_m^{\mathsf{T}}]^{\mathsf{T}} = \boldsymbol{F}\boldsymbol{\beta},\ \ \boldsymbol{\eta}_j = \big(\eta_j(\boldsymbol{x}_1, \boldsymbol{\beta}_j), \ldots, \eta_j(\boldsymbol{x}_n, \boldsymbol{\beta}_j)\big)^{\mathsf{T}},\ j \in \{1, \ldots, m\}.$$

For each component j, define the following $n \times n$ diagonal matrices

$$\boldsymbol{D}_j = \mathrm{diag}\Big(V_j\big(g_j^{-1}\big(\boldsymbol{f}_j^{\mathsf{T}}(\boldsymbol{x}_i)\boldsymbol{\beta}_j\big)\big)\Big)_{i=1}^n,\ \ \boldsymbol{E}_j = \mathrm{diag}\left(\Big(1/g_j'\big(g_j^{-1}\big(\boldsymbol{f}_j^{\mathsf{T}}(\boldsymbol{x}_i)\boldsymbol{\beta}_j\big)\big)\Big)^2\right)_{i=1}^n.$$

It can be seen that $\boldsymbol{D}_j^{-1}\boldsymbol{E}_j = \mathrm{diag}\Big(u_j(\boldsymbol{x}_i, \boldsymbol{\beta}_j)\Big)_{i=1}^n$ for all $j \in \{1, \ldots, m\}$. We then denote the $mn \times mn$ diagonal matrices $\boldsymbol{D} = \mathrm{diag}\Big(\boldsymbol{D}_j\Big)_{j=1}^m$ and $\boldsymbol{E} = \mathrm{diag}\Big(\boldsymbol{E}_j\Big)_{j=1}^m$. By the Kronecker product "$\otimes$" the $mn \times mn$ variance–covariance matrix of $\boldsymbol{Y}$ is obtained by

$$\mathrm{Cov}(\boldsymbol{Y}) = \boldsymbol{D}^{\frac{1}{2}}(\boldsymbol{R} \otimes \boldsymbol{I}_n)\boldsymbol{D}^{\frac{1}{2}} = \begin{pmatrix} \rho_{11}\boldsymbol{D}_1 & \rho_{12}\boldsymbol{D}_1^{\frac{1}{2}}\boldsymbol{D}_2^{\frac{1}{2}} & \ldots & \rho_{1m}\boldsymbol{D}_1^{\frac{1}{2}}\boldsymbol{D}_m^{\frac{1}{2}} \\ \rho_{21}\boldsymbol{D}_2^{\frac{1}{2}}\boldsymbol{D}_1^{\frac{1}{2}} & \rho_{22}\boldsymbol{D}_2 & \ldots & \rho_{2m}\boldsymbol{D}_2^{\frac{1}{2}}\boldsymbol{D}_m^{\frac{1}{2}} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{m1}\boldsymbol{D}_m^{\frac{1}{2}}\boldsymbol{D}_1^{\frac{1}{2}} & \rho_{m2}\boldsymbol{D}_m^{\frac{1}{2}}\boldsymbol{D}_2^{\frac{1}{2}} & \ldots & \rho_{mm}\boldsymbol{D}_m \end{pmatrix},$$

where $\boldsymbol{I}_n$ is an $n \times n$ identity matrix. The overall $mn \times mn$ weight matrix $\boldsymbol{W}$ is defined as

$$\boldsymbol{W} = \boldsymbol{E}^{\frac{1}{2}}\left(\mathrm{Cov}(\boldsymbol{Y})\right)^{-1}\boldsymbol{E}^{\frac{1}{2}} = \boldsymbol{E}^{\frac{1}{2}}\boldsymbol{D}^{-\frac{1}{2}}(\boldsymbol{R}\otimes\boldsymbol{I}_n)^{-1}\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{E}^{\frac{1}{2}}.$$

Hence, the quasi-Fisher information matrix from (4) can be represented in the form

$$\boldsymbol{M}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n,\boldsymbol{\beta}) = \boldsymbol{F}^{\mathsf{T}}\boldsymbol{W}\boldsymbol{F}. \tag{5}$$

Lemma 1. *Consider the MGLM (2) on the whole experimental conditions $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n$. Let a parameter point $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^{\mathsf{T}},\ldots,\boldsymbol{\beta}_m^{\mathsf{T}})^{\mathsf{T}}$ be given. For all $j \in \{1,\ldots,m\}$ let $\boldsymbol{F}_j = [\boldsymbol{f}_j(\boldsymbol{x}_1),\ldots,\boldsymbol{f}_j(\boldsymbol{x}_n)]^{\mathsf{T}}$ and define $\boldsymbol{F}_{j,\boldsymbol{\beta}_j} = \boldsymbol{D}_j^{-\frac{1}{2}}\boldsymbol{E}_j^{\frac{1}{2}}\boldsymbol{F}_j = [\boldsymbol{f}_{j,\boldsymbol{\beta}_j}(\boldsymbol{x}_1),\ldots,\boldsymbol{f}_{j,\boldsymbol{\beta}_j}(\boldsymbol{x}_n)]^{\mathsf{T}}$. Denote $\boldsymbol{F}_{\boldsymbol{\beta}} = \boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{E}^{\frac{1}{2}}\boldsymbol{F} = \mathrm{diag}\big(\boldsymbol{F}_{1,\boldsymbol{\beta}_1},\ldots,\boldsymbol{F}_{m,\boldsymbol{\beta}_m}\big)$. Then the quasi-Fisher information matrix from (5) has the form*

$$\boldsymbol{M}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n,\boldsymbol{\beta}) = \boldsymbol{F}_{\boldsymbol{\beta}}^{\mathsf{T}}\big(\boldsymbol{R}^{-1}\otimes\boldsymbol{I}_n\big)\boldsymbol{F}_{\boldsymbol{\beta}}.$$

Proof. In view of (5), straightforward steps imply that

$$\boldsymbol{M}(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n,\boldsymbol{\beta}) = \boldsymbol{F}^{\mathsf{T}}\boldsymbol{W}\boldsymbol{F} = \boldsymbol{F}^{\mathsf{T}}\boldsymbol{E}^{\frac{1}{2}}\left(\mathrm{Cov}(\boldsymbol{Y})\right)^{-1}\boldsymbol{E}^{\frac{1}{2}}\boldsymbol{F} = \boldsymbol{F}^{\mathsf{T}}\boldsymbol{E}^{\frac{1}{2}}\boldsymbol{D}^{-\frac{1}{2}}(\boldsymbol{R}\otimes\boldsymbol{I}_n)^{-1}\boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{E}^{\frac{1}{2}}\boldsymbol{F} = \boldsymbol{F}_{\boldsymbol{\beta}}^{\mathsf{T}}\big(\boldsymbol{R}^{-1}\otimes\boldsymbol{I}_n\big)\boldsymbol{F}_{\boldsymbol{\beta}}. \quad \square$$

In this paper we are interested to find a finite set of experimental conditions at which the observations achieve the best estimates of the model parameters in the sense of the minimum variance–covariance matrix of the quasi-likelihood estimates $\hat{\boldsymbol{\beta}}$, or equivalently the maximum quasi-Fisher information matrix. Here, we will deal with the approximate (continuous) design theory, i.e., a design ξ is a probability measure with finite support on the experimental region $\mathcal{X}$,

$$\xi = \begin{pmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \ldots & \boldsymbol{x}_r \\ \omega_1 & \omega_2 & \ldots & \omega_r \end{pmatrix}, \tag{6}$$

where $r \in \mathbb{N}$, $\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_r \in \mathcal{X}$ are mutually distinct support points and $\omega_1,\omega_2,\ldots,\omega_r > 0$ with $\sum_{i=1}^{r}\omega_i = 1$. The set $\mathrm{supp}(\xi) = \{\boldsymbol{x}_1,\boldsymbol{x}_2,\ldots,\boldsymbol{x}_r\}$ is called the support of ξ and $\omega_1,\ldots,\omega_r$ are called the weights of ξ, see Silvey [26], p. 15. We denote by Ξ the set of all designs ξ on $\mathcal{X}$. The quasi-Fisher information matrix of ξ is given by

$$\boldsymbol{M}(\xi,\boldsymbol{\beta}) = \int_{\mathcal{X}}\boldsymbol{M}(\boldsymbol{x},\boldsymbol{\beta})\,\xi(\mathrm{d}\boldsymbol{x}) = \sum_{i=1}^{r}\omega_i\boldsymbol{M}(\boldsymbol{x}_i,\boldsymbol{\beta}). \tag{7}$$

Throughout we focus on locally optimal designs with respect to the D- and A-criteria. Denote by $\det(\boldsymbol{A})$ and $\mathrm{tr}(\boldsymbol{A})$ the determinant and the trace of a $p\times p$ matrix $\boldsymbol{A}$, respectively. D-optimal designs are constructed to minimize the determinant of the variance–covariance matrix of $\hat{\boldsymbol{\beta}}$ or equivalently to maximize the determinant of the quasi-Fisher information matrix. The D-criterion is defined by the convex function $\Phi_{\mathrm{D}}(\boldsymbol{M}(\xi,\boldsymbol{\beta})) = -\log\det\big(\boldsymbol{M}(\xi,\boldsymbol{\beta})\big)$. A-optimal designs are constructed to minimize the trace of the variance–covariance matrix of $\hat{\boldsymbol{\beta}}$, i.e., to minimize the average variance of the estimates. The A-criterion is defined by the function $\Phi_{\mathrm{A}}\big(\boldsymbol{M}(\xi,\boldsymbol{\beta})\big) = \mathrm{tr}\big(\boldsymbol{M}^{-1}(\xi,\boldsymbol{\beta})\big)$. The multivariate version of the equivalence theorem (see Fedorov et al. [7]) for checking the D- and A-optimality of a given design can be used.

Theorem 1. *Let $\boldsymbol{\beta}$ be a given parameter point and let ξ^* be a design with nonsingular quasi-Fisher information matrix $\boldsymbol{M}(\xi^*,\boldsymbol{\beta})$.*

(i) *A design ξ^* is locally D-optimal (at $\boldsymbol{\beta}$) for the MGLM if and only if*

$$\mathrm{tr}\Big(\boldsymbol{R}^{-1}\mathbf{f}_{\boldsymbol{\beta}}^{\mathsf{T}}(\boldsymbol{x})\boldsymbol{M}^{-1}(\xi^*,\boldsymbol{\beta})\mathbf{f}_{\boldsymbol{\beta}}(\boldsymbol{x})\Big) \le p \quad \forall \boldsymbol{x}\in\mathcal{X}. \tag{8}$$

(ii) *A design ξ^* is locally A-optimal (at $\boldsymbol{\beta}$) for the MGLM if and only if*

$$\mathrm{tr}\Big(\boldsymbol{R}^{-1}\mathbf{f}_{\boldsymbol{\beta}}^{\mathsf{T}}(\boldsymbol{x})\boldsymbol{M}^{-2}(\xi^*,\boldsymbol{\beta})\mathbf{f}_{\boldsymbol{\beta}}(\boldsymbol{x})\Big) \le \mathrm{tr}\Big(\boldsymbol{M}^{-1}(\xi^*,\boldsymbol{\beta})\Big) \quad \forall \boldsymbol{x}\in\mathcal{X}. \tag{9}$$

At the support points of ξ^* both inequalities (8) and (9) are equations.

3. Optimal designs for MGLMs

The locally optimal design for a MGLM is derived at a given parameter point $\boldsymbol{\beta}$ under known correlation matrix $\boldsymbol{R}$. In each jth component under a design $\xi \in \Xi$ with support $\{\boldsymbol{x}_1,\ldots,\boldsymbol{x}_r\}$ we obtain the corresponding $r\times p_j$ design matrix $\boldsymbol{F}_j = [\boldsymbol{f}_j(\boldsymbol{x}_1),\ldots,\boldsymbol{f}_j(\boldsymbol{x}_r)]^{\mathsf{T}}$. In view of Lemma 1, let $n = r$ and denote by $\boldsymbol{\Omega} = \mathrm{diag}(\omega_1,\ldots,\omega_r)$ the diagonal matrix of the design weights. Then by Lemma 1 the quasi-Fisher information matrix (7) of a design ξ rewrites

$$\boldsymbol{M}(\xi,\boldsymbol{\beta}) = \boldsymbol{F}_{\boldsymbol{\beta}}^{\mathsf{T}}\big(\boldsymbol{R}^{-1}\otimes\boldsymbol{\Omega}\big)\boldsymbol{F}_{\boldsymbol{\beta}}. \tag{10}$$

Furthermore, when $\boldsymbol{M}(\xi,\boldsymbol{\beta})$ is positive definite the inverse $\boldsymbol{M}^{-1}(\xi,\boldsymbol{\beta}) = \left(\boldsymbol{F}_{\boldsymbol{\beta}}^{\top}\left(\boldsymbol{R}^{-1}\otimes\boldsymbol{\Omega}\right)\boldsymbol{F}_{\boldsymbol{\beta}}\right)^{-1}$ factorizes if $\boldsymbol{F}_{\boldsymbol{\beta}}$ is square, and thus we obtain

$$\boldsymbol{M}^{-1}(\xi,\boldsymbol{\beta}) = \boldsymbol{F}_{\boldsymbol{\beta}}^{-1}\left(\boldsymbol{R}\otimes\boldsymbol{\Omega}^{-1}\right)\left(\boldsymbol{F}_{\boldsymbol{\beta}}^{\top}\right)^{-1}.$$

A block representation of $\boldsymbol{M}(\xi,\boldsymbol{\beta})$ can be given in the form

$$\boldsymbol{M}(\xi,\boldsymbol{\beta}) = \left(\rho^{(jh)}\boldsymbol{M}_{jh}(\xi,\boldsymbol{\beta}_j,\boldsymbol{\beta}_h)\right)_{j=1,\ldots,m}^{h=1,\ldots,m}, \tag{11}$$

where $\boldsymbol{M}_j(\xi,\boldsymbol{\beta}_j) = \boldsymbol{M}_{jj}(\xi,\boldsymbol{\beta}_j) = \boldsymbol{F}_{j,\boldsymbol{\beta}_j}^{\top}\boldsymbol{\Omega}\boldsymbol{F}_{j,\boldsymbol{\beta}_j} = \sum_{i=1}^{r}\omega_i\boldsymbol{f}_{j,\boldsymbol{\beta}_j}(\boldsymbol{x}_i)\boldsymbol{f}_{j,\boldsymbol{\beta}_j}^{\top}(\boldsymbol{x}_i)$ is the $p_j\times p_j$ information matrix for the jth marginal model $(1\le j\le m)$, whereas the $p_j\times p_h$ submatrices $\boldsymbol{M}_{jh}(\xi,\boldsymbol{\beta}_j,\boldsymbol{\beta}_h) = \boldsymbol{F}_{j,\boldsymbol{\beta}_j}^{\top}\boldsymbol{\Omega}\boldsymbol{F}_{h,\boldsymbol{\beta}_h} = \sum_{i=1}^{r}\omega_i\boldsymbol{f}_{j,\boldsymbol{\beta}_j}(\boldsymbol{x}_i)\boldsymbol{f}_{h,\boldsymbol{\beta}_h}^{\top}(\boldsymbol{x}_i)$ $(1\le j\ne h\le m)$ which are not necessarily square.

In the following lemma we present a block structure of $\boldsymbol{M}^{-1}(\xi,\boldsymbol{\beta})$ under specific restrictions. This is useful to derive optimal designs, in particular, under the A-criterion.

Lemma 2. *Consider a design ξ defined on $\mathcal{X}$ with quasi-Fisher information matrix $\boldsymbol{M}(\xi,\boldsymbol{\beta})$ given by (11). Let a parameter point $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^{\top},\ldots,\boldsymbol{\beta}_m^{\top})^{\top}$ be given. Assume that all submatrices $\boldsymbol{M}_{jh}(\xi,\boldsymbol{\beta}_j,\boldsymbol{\beta}_h)$, $j,h\in\{1,\ldots,m\}$ are square, i.e., $p_1=\cdots=p_m=p_0$, and nonsingular. If $\sum_{k=1}^{m}\rho_{hk}\rho^{(jk)}\boldsymbol{M}_{jk}(\xi,\boldsymbol{\beta}_j,\boldsymbol{\beta}_k)\boldsymbol{M}_{hk}^{-1}(\xi,\boldsymbol{\beta}_h,\boldsymbol{\beta}_k)=0$ for all $(1\le j\ne h\le m)$ then $\boldsymbol{M}(\xi,\boldsymbol{\beta})$ is nonsingular and a block representation of its inverse is given by*

$$\boldsymbol{M}^{-1}(\xi,\boldsymbol{\beta}) = \left(\rho_{hj}\boldsymbol{M}_{hj}^{-1}(\xi,\boldsymbol{\beta}_h,\boldsymbol{\beta}_j)\right)_{j=1,\ldots,m}^{h=1,\ldots,m}. \tag{12}$$

Proof. The assumption $\sum_{k=1}^{m}\rho_{hk}\rho^{(jk)}\boldsymbol{M}_{jk}(\xi,\boldsymbol{\beta}_j,\boldsymbol{\beta}_k)\boldsymbol{M}_{hk}^{-1}(\xi,\boldsymbol{\beta}_h,\boldsymbol{\beta}_k)=0$ for all $(1\le j\ne h\le m)$ describes explicitly the multiplication of the off-diagonal submatrices of $\boldsymbol{M}(\xi,\boldsymbol{\beta})$ and $\boldsymbol{M}^{-1}(\xi,\boldsymbol{\beta})$. Thus under this assumption $\boldsymbol{M}(\xi,\boldsymbol{\beta})\boldsymbol{M}^{-1}(\xi,\boldsymbol{\beta})$ is an identity matrix. □

3.1. Reduction to the marginal GLMs

The previous situation can be simplified under saturated designs, i.e., when the number of the support points of a design ξ is equal to the number of the parameters of the marginal model ($r=p_0$). Let $\Xi_{p_0}=\{\xi:\mathrm{supp}(\xi)\subseteq\mathcal{X}, r=p_0\}$ denote the set of all saturated designs under each jth univariate GLM $(1\le j\le m)$. Clearly, for any design $\xi\in\Xi_{p_0}$ the design matrix $\boldsymbol{F}$ (or $\boldsymbol{F}_{\boldsymbol{\beta}}$) of the MGLM is square. In particular, for $\xi\in\Xi_{p_0}$ the design matrices $\boldsymbol{F}_{\boldsymbol{\beta}_j}$, $j\in\{1,\ldots,m\}$ are square. Hence, the submatrices $\boldsymbol{M}_j(\xi,\boldsymbol{\beta}_j), j\in\{1,\ldots,m\}$ and $\boldsymbol{M}_{jh}(\xi,\boldsymbol{\beta})$ $(1\le j\ne h\le m)$ factorize and thus, the assumptions $\sum_{k=1}^{m}\rho_{hk}\rho^{(jk)}\boldsymbol{M}_{jk}(\xi,\boldsymbol{\beta}_j,\boldsymbol{\beta}_k)\boldsymbol{M}_{hk}^{-1}(\xi,\boldsymbol{\beta}_h,\boldsymbol{\beta}_k)=0$ for all $(1\le j\ne h\le m)$ in Lemma 2 will be implicitly satisfied as it is clarified by the next corollary.

Corollary 1. *Consider the notations presented in Lemma 2. Then for any $\xi\in\Xi_{p_0}$ the assumption given in Lemma 2 is satisfied, i.e., $\sum_{k=1}^{m}\rho_{hk}\rho^{(jk)}\boldsymbol{M}_{jk}(\xi,\boldsymbol{\beta}_j,\boldsymbol{\beta}_k)\boldsymbol{M}_{hk}^{-1}(\xi,\boldsymbol{\beta}_h,\boldsymbol{\beta}_k)=0$ for all $(1\le j\ne h\le m)$.*

Proof. Note that for all $\xi\in\Xi_{p_0}$ we can write $\boldsymbol{M}_j(\xi,\boldsymbol{\beta}_j)=\boldsymbol{F}_{j,\boldsymbol{\beta}_j}^{\top}\boldsymbol{\Omega}\boldsymbol{F}_{j,\boldsymbol{\beta}_j}, j\in\{1,\ldots,m\}$ and $\boldsymbol{M}_{jh}(\xi,\boldsymbol{\beta})=\boldsymbol{F}_{j,\boldsymbol{\beta}_j}^{\top}\boldsymbol{\Omega}\boldsymbol{F}_{h,\boldsymbol{\beta}_h}$ $(1\le j\ne h\le m)$. Then $\boldsymbol{M}_{jk}(\xi,\boldsymbol{\beta}_j,\boldsymbol{\beta}_k)\boldsymbol{M}_{hk}^{-1}(\xi,\boldsymbol{\beta}_h,\boldsymbol{\beta}_k)=\boldsymbol{F}_{j,\boldsymbol{\beta}_j}^{\top}(\boldsymbol{F}_{h,\boldsymbol{\beta}_h}^{\top})^{-1}$ for all $k\in\{1,\ldots,m\}$. Since $\sum_{k=1}^{m}\rho_{hk}\rho^{(jk)}=0$ we conclude $\sum_{k=1}^{m}\rho_{hk}\rho^{(jk)}\boldsymbol{M}_{jk}(\xi,\boldsymbol{\beta}_j,\boldsymbol{\beta}_k)\boldsymbol{M}_{hk}^{-1}(\xi,\boldsymbol{\beta}_h,\boldsymbol{\beta}_k)=\boldsymbol{F}_{j,\boldsymbol{\beta}_j}^{\top}(\boldsymbol{F}_{h,\boldsymbol{\beta}_h}^{\top})^{-1}\sum_{k=1}^{m}\rho_{hk}\rho^{(jk)}=0$ for all $(1\le j\ne h\le m)$. □

Remark 2. The locally D-optimal saturated design assigns equal weights $\omega_i=1/p_0\,\forall i$ to its support points (see Silvey [26], Lemma 5.3.1).

Lemma 3. *The locally D-optimal saturated design ξ^* in Ξ_{p_0} at a given parameter point $\boldsymbol{\beta}$ for a MGLM (2) is independent of the correlation matrix $\boldsymbol{R}$.*

Proof. Let $\xi\in\Xi_{p_0}$. The determinant of the quasi-Fisher information matrix $\boldsymbol{M}(\xi,\boldsymbol{\beta})$ from (10) is given by

$$\det\boldsymbol{M}(\xi,\boldsymbol{\beta}) = \det\boldsymbol{F}_{\boldsymbol{\beta}}^{\top}\left(\boldsymbol{R}^{-1}\otimes\boldsymbol{\Omega}\right)\boldsymbol{F}_{\boldsymbol{\beta}} = \det\left(\boldsymbol{F}_{\boldsymbol{\beta}}^{\top}\boldsymbol{F}_{\boldsymbol{\beta}}\right)\det\left(\boldsymbol{R}^{-1}\otimes\boldsymbol{\Omega}\right) = \det\left(\boldsymbol{F}_{\boldsymbol{\beta}}^{\top}\boldsymbol{F}_{\boldsymbol{\beta}}\right)\left(\det\boldsymbol{\Omega}\right)^m\left(\det\boldsymbol{R}^{-1}\right)^r.$$

It follows that $\det\boldsymbol{M}(\xi,\boldsymbol{\beta})$ is proportional to $\det\left(\boldsymbol{F}_{\boldsymbol{\beta}}^{\top}\boldsymbol{F}_{\boldsymbol{\beta}}\right)\left(\det\boldsymbol{\Omega}\right)^m$. Thus the optimization w.r.t. to the D-criterion in Ξ_{p_0} is independent of $\boldsymbol{R}$. □

Theorem 2. *Consider the MGLM (2) with a positive definite correlation matrix $\boldsymbol{R}$. For a given parameter point $\boldsymbol{\beta}=(\boldsymbol{\beta}_1^{\top},\ldots,\boldsymbol{\beta}_m^{\top})^{\top}$ let the design $\xi^*\in\Xi_{p_0}$ be locally D-optimal (at $\boldsymbol{\beta}_j$) for each jth marginal model $(1\le j\le m)$. Then ξ^* is locally D-optimal (at $\boldsymbol{\beta}$) for the MGLM (2) within the set Ξ_{p_0}.*

Proof. Under the design ξ^* we get $\boldsymbol{\Omega} = (1/p_0)\boldsymbol{I}_{p_0}$ where $\boldsymbol{I}_{p_0}$ is a $p_0 \times p_0$ identity matrix. From the proof of Lemma 3 we have

$$\begin{aligned}\det \boldsymbol{M}(\xi^*, \boldsymbol{\beta}) &= \det\big(\boldsymbol{F}_{\boldsymbol{\beta}}^{\mathsf{T}}\boldsymbol{F}_{\boldsymbol{\beta}}\big)\big(\det \boldsymbol{\Omega}\big)^m \big(\det \boldsymbol{R}^{-1}\big)^r = p_0^{-rm}(\det \boldsymbol{F})^2\Big(\det \boldsymbol{E}^{\frac{1}{2}}\boldsymbol{D}^{-\frac{1}{2}}\Big)^2\big(\det \boldsymbol{R}^{-1}\big)^r\\ &= p_0^{-rm}(\det \boldsymbol{F})^2\Big(\prod_{j=1}^{m}\prod_{i=1}^{r} u_j(\boldsymbol{x}_i, \boldsymbol{\beta}_j)\Big)\big(\det \boldsymbol{R}^{-1}\big)^r,\end{aligned}$$

where $\det \boldsymbol{F} = \prod_{j=1}^{m} \det \boldsymbol{F}_j$. Moreover, the determinant of the information matrix for the jth marginal models is $\det \boldsymbol{M}_j(\xi^*, \boldsymbol{\beta}_j) = p_0^{-r}(\det \boldsymbol{F}_j)^2 \prod_{i=1}^{r} u_j(\boldsymbol{x}_i, \boldsymbol{\beta}_j)$, $j \in \{1, \ldots, m\}$. Thus $\prod_{i=1}^{r} u_j(\boldsymbol{x}_i, \boldsymbol{\beta}_j) = p_0^r \left(\det \boldsymbol{F}_j\right)^{-2} \det \boldsymbol{M}_j(\xi^*, \boldsymbol{\beta}_j)$. It follows that

$$\begin{aligned}\det \boldsymbol{M}(\xi^*, \boldsymbol{\beta}) &= p_0^{-rm}\Big(\prod_{j=1}^{m}\det \boldsymbol{F}_j\Big)^2\Big(\prod_{j=1}^{m} p_0^r \left(\det \boldsymbol{F}_j\right)^{-2} \det \boldsymbol{M}_j(\xi^*, \boldsymbol{\beta}_j)\Big)\big(\det \boldsymbol{R}^{-1}\big)^r\\ &= p_0^{-rm}p_0^{rm}\Big(\prod_{j=1}^{m}\det \boldsymbol{F}_j\Big)^2\Big(\prod_{j=1}^{m}\det \boldsymbol{F}_j\Big)^{-2}\prod_{j=1}^{m}\det \boldsymbol{M}_j(\xi^*, \boldsymbol{\beta}_j)\big(\det \boldsymbol{R}^{-1}\big)^r = \big(\det \boldsymbol{R}^{-1}\big)^r\prod_{j=1}^{m}\det \boldsymbol{M}_j(\xi^*, \boldsymbol{\beta}_j).\end{aligned}$$

Since ξ^* is locally D-optimal for the jth marginal model it maximizes $\det \boldsymbol{M}_j(\xi, \boldsymbol{\beta}_j)$ on Ξ_{p_0}. Thus $\prod_{j=1}^{m} \det \boldsymbol{M}_j(\xi^*, \boldsymbol{\beta}_j) \geq \prod_{j=1}^{m} \det \boldsymbol{M}_j(\xi, \boldsymbol{\beta}_j)$ for all $\xi \in \Xi_{p_0}$. As a result, ξ^* maximizes $\det \boldsymbol{M}(\xi, \boldsymbol{\beta})$ on Ξ_{p_0}. Hence, ξ^* is locally D-optimal (at $\boldsymbol{\beta}$) for a MGLM within the set Ξ_{p_0}. □

Next we will deal with local A-optimality. The following lemma is immediate.

Lemma 4. *The locally A-optimal saturated design ξ^* in Ξ_{p_0} at a given parameter point $\boldsymbol{\beta}$ for a MGLM (2) is independent of correlation matrix $\boldsymbol{R}$.*

Proof. Let $\xi \in \Xi_{p_0}$. According to Lemma 2 and Corollary 1, the inverse of the quasi-Fisher information matrix of ξ is given by the block representation (12). Thus $\mathrm{tr}(\boldsymbol{M}^{-1}(\xi, \boldsymbol{\beta})) = \sum_{j=1}^{m} \mathrm{tr}(\boldsymbol{M}_j^{-1}(\xi, \boldsymbol{\beta}_j))$. It is clear that $\mathrm{tr}(\boldsymbol{M}^{-1}(\xi, \boldsymbol{\beta}))$ does not depend on $\boldsymbol{R}$. □

Theorem 3. *Consider the MGLM (2) with a positive definite correlation matrix $\boldsymbol{R}$. For a given parameter point $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^{\mathsf{T}}, \ldots, \boldsymbol{\beta}_m^{\mathsf{T}})^{\mathsf{T}}$ let the design $\xi^* \in \Xi_{p_0}$ be locally A-optimal (at $\boldsymbol{\beta}_j$) for each jth marginal model ($1 \leq j \leq m$). Then ξ^* is locally A-optimal (at $\boldsymbol{\beta}$) for the MGLM (2) within the set Ξ_{p_0}.*

Proof. For the design $\xi^* \in \Xi_{p_0}$ we have $\mathrm{tr}(\boldsymbol{M}^{-1}(\xi^*, \boldsymbol{\beta})) = \sum_{j=1}^{m} \mathrm{tr}(\boldsymbol{M}_j^{-1}(\xi^*, \boldsymbol{\beta}_j))$. As ξ^* is locally A-optimal for the jth marginal model then $\mathrm{tr}(\boldsymbol{M}_j^{-1}(\xi^*, \boldsymbol{\beta}_j)) \leq \mathrm{tr}(\boldsymbol{M}_j^{-1}(\xi, \boldsymbol{\beta}_j))$ for all $\xi \in \Xi_{p_0}$. Thus $\sum_{j=1}^{m} \mathrm{tr}(\boldsymbol{M}_j^{-1}(\xi^*, \boldsymbol{\beta}_j)) \leq \sum_{j=1}^{m} \mathrm{tr}(\boldsymbol{M}_j^{-1}(\xi, \boldsymbol{\beta}_j))$ for all $\xi \in \Xi_{p_0}$. As a result, ξ^* minimizes $\mathrm{tr}(\boldsymbol{M}^{-1}(\xi, \boldsymbol{\beta}))$ on Ξ_{p_0}. Hence, ξ^* is locally A-optimal (at $\boldsymbol{\beta}$) for a MGLM within the set Ξ_{p_0}. □

Remark 3. It is well known that the optimal weights of a locally A-optimal saturated design under a univariate generalized linear model depend on the model parameters through the intensity functions (see Gaffke et al. [9], Lemma 2.1). Therefore, the locally A-optimal saturated design for marginal univariate models is A-optimal for the MGLM if all intensities in all components have the same form and the A-optimality is derived at equal parameter points. This guarantees that the A-optimal saturated design is the same for all marginal models.

In the following theorem we consider a multivariate GLM with identical components. Locally optimal designs are derived within the set Ξ of all possible designs on the experimental region.

Theorem 4. *Consider the MGLM (2) such that $\boldsymbol{f}_1(\boldsymbol{x}) = \cdots = \boldsymbol{f}_m(\boldsymbol{x}) = \boldsymbol{f}_0(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathcal{X}$. Let a positive definite correlation matrix $\boldsymbol{R}$ be given. Let the parameter point $\boldsymbol{\beta}$ be given such that $\boldsymbol{\beta}_1 = \cdots = \boldsymbol{\beta}_m = \boldsymbol{\beta}_0$, i.e., $\boldsymbol{\beta} = \boldsymbol{1} \otimes \boldsymbol{\beta}_0$. Assume that $u_1(\boldsymbol{x}, \boldsymbol{\beta}_1) = \cdots = u_m(\boldsymbol{x}, \boldsymbol{\beta}_m) = u_0(\boldsymbol{x}, \boldsymbol{\beta}_0)$ for all $\boldsymbol{x} \in \mathcal{X}$. Thus $\boldsymbol{F}_{1,\boldsymbol{\beta}_1} = \cdots = \boldsymbol{F}_{m,\boldsymbol{\beta}_m} = \boldsymbol{F}_{0,\boldsymbol{\beta}_0}$. Let the design ξ^* be locally D- or A-optimal (at $\boldsymbol{\beta}_0$) for each jth marginal model ($1 \leq j \leq m$). Then the design ξ^* is locally D- or A-optimal (at $\boldsymbol{\beta} = \boldsymbol{1} \otimes \boldsymbol{\beta}_0$) for the MGLM (2), respectively.*

Proof. Denote $\boldsymbol{M}_0(\xi, \boldsymbol{\beta}_0) = \boldsymbol{F}_{0,\boldsymbol{\beta}_0}^{\mathsf{T}}\boldsymbol{\Omega}\boldsymbol{F}_{0,\boldsymbol{\beta}_0}$. Under the assumptions proposed in the theorem it can be seen that $\boldsymbol{M}_{jh}(\xi, \boldsymbol{\beta}_j, \boldsymbol{\beta}_h) = \boldsymbol{M}_0(\xi, \boldsymbol{\beta}_0)$ for all $j, h \in \{1, \ldots, m\}$. As a result, the quasi-Fisher information matrix (11) and its inverse factorize as $\boldsymbol{M}(\xi, \boldsymbol{\beta}_0) = \boldsymbol{R}^{-1} \otimes \boldsymbol{M}_0(\xi, \boldsymbol{\beta}_0)$ and $\boldsymbol{M}^{-1}(\xi, \boldsymbol{\beta}_0) = \boldsymbol{R} \otimes \boldsymbol{M}_0^{-1}(\xi, \boldsymbol{\beta}_0)$, respectively. Therefore, we obtain $\det \boldsymbol{M}(\xi, \boldsymbol{\beta}_0) = \big(\det \boldsymbol{R}^{-1}\big)^{p_0}\big(\det \boldsymbol{M}_0(\xi, \boldsymbol{\beta}_0)\big)^m$, $\mathrm{tr}(\boldsymbol{M}^{-1}(\xi, \boldsymbol{\beta}_0)) = \mathrm{tr}(\boldsymbol{R})\mathrm{tr}(\boldsymbol{M}_0^{-1}(\xi, \boldsymbol{\beta}_0))$. Hence, the optimization with respect to the D- and A-criteria is independent of $\boldsymbol{R}$ and reduces to the corresponding univariate optimization problem. □

3.2. D-optimality under exchangeable correlation matrix

In what follows we restrict the correlation matrix $\boldsymbol{R}$ to the exchangeable structure

$$\boldsymbol{R} = \big(1-\rho\big)\boldsymbol{I}_m + \rho\mathbf{1}\mathbf{1}^\top \quad \text{with} \quad \boldsymbol{R}^{-1} = \big(1-\rho\big)^{-1}\Big(\boldsymbol{I}_m - \frac{\rho}{1+(m-1)\rho}\mathbf{1}\mathbf{1}^\top\Big), \tag{13}$$

where $-1 < \rho < 1$, $\boldsymbol{I}_m$ is an $m \times m$ identity matrix and $\mathbf{1}$ is an m-dimensional vector of ones. As a result, expressions (3), (7) and (12), respectively become

$$\boldsymbol{\Delta}(\boldsymbol{x}_i, \boldsymbol{\beta})\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}_i, \boldsymbol{\beta})\boldsymbol{\Delta}(\boldsymbol{x}_i, \boldsymbol{\beta}) = (1-\rho)^{-1}\Big(\mathrm{diag}\big(u_j(\boldsymbol{x}_i, \boldsymbol{\beta}_j)\big)_{j=1}^m - \frac{\rho}{1+(m-1)\rho}\ \Big(u_j^{\frac{1}{2}}(\boldsymbol{x}_i, \boldsymbol{\beta}_j)u_h^{\frac{1}{2}}(\boldsymbol{x}_i, \boldsymbol{\beta}_h)\Big)_{j=1,\dots,m}^{h=1,\dots,m}\Big),$$

$$\boldsymbol{M}(\xi, \boldsymbol{\beta}) = \frac{1}{(1-\rho)(1+(m-1)\rho)}\Big((1+(m-1)\rho)\mathrm{diag}(\boldsymbol{M}_j(\xi, \boldsymbol{\beta}_j))_{j=1}^m - \rho(\boldsymbol{M}_{jh})_{j,h=1}^m\Big),$$

$$\boldsymbol{M}^{-1}(\xi, \boldsymbol{\beta}) = (1-\rho)\mathrm{diag}\Big(\boldsymbol{M}_j^{-1}(\xi, \boldsymbol{\beta}_j)\Big)_{j=1}^m + \rho\Big(\boldsymbol{M}_{jh}^{-1}(\xi, \boldsymbol{\beta}_j, \boldsymbol{\beta}_h)\Big)_{j,h=1}^m.$$

In the following theorem we provide a sufficient and necessary condition for a saturated design $\xi^* \in \Xi_{p_0}$ to be locally D-optimal for the MGLM in the set Ξ of all possible designs by employing condition (8) of the general equivalence theorem. The proof of the theorem can be found in Appendix.

Theorem 5. *Consider the MGLM (2) on the experimental region $\mathcal{X}$. Let a positive definite correlation matrix $\boldsymbol{R}$ from (13) be given. Let $\xi^* \in \Xi_{p_0}$. For a given parameter point $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \dots, \boldsymbol{\beta}_m^\top)^\top$ define*

$$d_j(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j) = \boldsymbol{f}_{\boldsymbol{\beta}_j}^\top(\boldsymbol{x})\boldsymbol{M}_j^{-1}(\xi^*, \boldsymbol{\beta}_j)\boldsymbol{f}_{\boldsymbol{\beta}_j}(\boldsymbol{x}), j \in \{1, \dots, m\},\ \text{for all } \boldsymbol{x} \in \mathcal{X},$$

$$d_{jh}(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j, \boldsymbol{\beta}_h) = \boldsymbol{f}_{\boldsymbol{\beta}_j}^\top(\boldsymbol{x})\boldsymbol{M}_{jh}^{-1}(\xi^*, \boldsymbol{\beta}_j, \boldsymbol{\beta}_h)\boldsymbol{f}_{\boldsymbol{\beta}_h}(\boldsymbol{x}),\ 1 \le j < h \le m,\ \text{for all } \boldsymbol{x} \in \mathcal{X}.$$

Then the design ξ^* with non-singular quasi-Fisher information matrix $\boldsymbol{M}(\xi^*, \boldsymbol{\beta})$ is locally D-optimal (at $\boldsymbol{\beta}$) for the MGLM if and only if

$$\frac{1}{(1-\rho)(1+(m-1)\rho)}\Big((1+(m-2)\rho)\sum_{j=1}^m d_j(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j) - 2\,\rho^2 \sum_{j<h=1}^m d_{jh}(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j, \boldsymbol{\beta}_h)\Big) \le mp_0 \text{ for all } \boldsymbol{x} \in \mathcal{X}. \tag{14}$$

Remark 4. According to the general equivalence theorem for a univariate generalized linear model (see Silvey [26], p. 40, p. 48 and p. 54), the design ξ^* is locally D-optimal for the jth marginal model if and only if $d_j(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j) \le p_0$ for all $\boldsymbol{x} \in \mathcal{X}$. The function $d_j(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j)$ is commonly called the sensitivity function.

Remark 5. Condition (14) for a saturated design $\xi^* \in \Xi_{p_0}$ is useful to approve the local D-optimality of ξ^* in Ξ. It can be seen that this condition depends on the correlation ρ and is equivalent to

$$a(\boldsymbol{x})\,\rho^2 + b(\boldsymbol{x})\,\rho + c(\boldsymbol{x}) \le 0 \ \text{ for all } \boldsymbol{x} \in \mathcal{X}, \tag{15}$$

where the coefficients $a(\boldsymbol{x}) = \Big(m(m-1)mp_0 - 2\sum_{j<h=1}^m d_{jh}(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j, \boldsymbol{\beta}_h)\Big)$, $b(\boldsymbol{x}) = (m-2)\Big(\sum_{j=1}^m d_j(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j) - mp_0\Big)$, and $c(\boldsymbol{x}) = \sum_{j=1}^m d_j(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j) - mp_0$. The l.h.s. of (15) is a polynomial in ρ of degree 2. The set of solutions of the system of inequalities characterized by (15) is given by

$$\cap_{\boldsymbol{x}\in\mathcal{X}\setminus\mathrm{supp}(\xi^*)}\Big[\Big(-b(\boldsymbol{x}) \pm \sqrt{b^2(\boldsymbol{x}) - 4a(\boldsymbol{x})c(\boldsymbol{x})}\Big)/2a(\boldsymbol{x})\Big] \cap (-1, 1). \tag{16}$$

Accordingly, the sensitivity functions for the locally D-optimal saturated design ξ^* under all marginal GLMs can be used to determine the range of ρ at which the design ξ^* can be locally D-optimal in Ξ. Hence, the value of ρ is chosen from the set (16).

Lemma 5. *In view of Theorem 5 assume that $\boldsymbol{f}_1(\boldsymbol{x}) = \cdots = \boldsymbol{f}_m(\boldsymbol{x}) = \boldsymbol{f}_0(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathcal{X}$. Then for the given parameter vectors $\boldsymbol{\beta}_j, j \in \{1, \dots, m\}$ we have*

$$d_{jh}(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j, \boldsymbol{\beta}_h) \le \max\{d_j(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j), d_h(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_h)\} \ \text{ for all } \boldsymbol{x} \in \mathcal{X} \text{ and for all } (1 \le j < h \le m).$$

If $\max\{d_j(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j), d_h(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_h)\} \le p_0$ for all $\boldsymbol{x} \in \mathcal{X}$ then $d_{jh}(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}) \le p_0$ for all $\boldsymbol{x} \in \mathcal{X}$ and for all $(1 \le j < h \le m)$.

The proof of the above lemma can be found in Appendix.

4. Applications to gamma-distributed outcomes

We consider a MGLM which consists of first order univariate generalized linear models defined on the standardized unit cube $\mathcal{X} = [0, 1]^\nu$, $\nu \ge 1$ as an experimental region. Here, for each j we have $p_j = \nu + 1 = p_0$ with linear predictor

$$\eta_j(\boldsymbol{x}, \boldsymbol{\beta}_j) = \boldsymbol{f}_j^\top(\boldsymbol{x})\boldsymbol{\beta}_j,\ \boldsymbol{f}_j(\boldsymbol{x}) = (1, x_1, \dots, x_\nu)^\top,\ \boldsymbol{\beta}_j = (\beta_{j0}, \beta_{j1}, \dots, \beta_{j\nu})^\top,\ j \in \{1, \dots, m\}.$$

Table 1
Design points of locally D- and A-optimal saturated designs on $\mathcal{X} = [0,1]^{\nu}$, $\nu \geq 1$, for the MGLM with marginal gamma models at restricted parameter values.

	Design points				
Run	x_1	x_2	x_3	$\cdots$	x_ν
1	0	0	0	$\cdots$	0
2	1	0	0	$\cdots$	0
3	0	1	0	$\cdots$	0
4	0	0	1	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\nu+1$	0	0	0	$\cdots$	1

Assume that each marginal observation Y_j follows a gamma distribution. In this situation, Y_j is typically non-negative and continuous with a variance that is proportional to the square of the expected mean. The canonical link is the inverse, i.e.,

$$\eta_j(\boldsymbol{x}, \boldsymbol{\beta}_j) = \alpha_j/\mu_j(\boldsymbol{x}, \boldsymbol{\beta}_j) = \boldsymbol{f}_j^{\mathsf{T}}(\boldsymbol{x})\boldsymbol{\beta}_j \text{ for all } \boldsymbol{x} \in \mathcal{X} \text{ and for all } j,$$

where $\alpha_j > 0$ is the shape parameter of the gamma distribution. The intensity function is given by

$$u_j(\boldsymbol{x}, \boldsymbol{\beta}_j) = \alpha_j\big(\boldsymbol{f}_j^{\mathsf{T}}(\boldsymbol{x})\boldsymbol{\beta}_j\big)^{-2} \text{ for all } \boldsymbol{x} \in \mathcal{X} \text{ and for all } j.$$

According to Burridge and Sebastiani [1] the saturated design ξ^*_{1/p_0} that assigns equal weights $\omega_k = 1/p_0$ for all $k \in \{1,\ldots,p_0\}$ to the design points presented in Table 1 is locally D-optimal for the jth univariate gamma model at $\boldsymbol{\beta}_j = (\beta_{j0}, \beta_{j1}, \ldots, \beta_{j\nu})^{\mathsf{T}}$ such that $\beta_{j0}^2 \leq \beta_{jk}\beta_{jk'}$ for all $(1 \leq k < k' \leq \nu)$ and for all $j \in \{1,\ldots,m\}$. By Theorem 4 design ξ^*_{1/p_0} is locally D-optimal for the MGLM at $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^{\mathsf{T}}, \ldots, \boldsymbol{\beta}_m^{\mathsf{T}})^{\mathsf{T}}$ in Ξ_{p_0} under the same parameter constraints, i.e., $\beta_{j0}^2 \leq \beta_{jk}\beta_{jk'}$ for all $(1 \leq k < k' \leq \nu)$ and for all $j \in \{1,\ldots,m\}$.

Let the correlation matrix be given with the exchangeable structure (13), i.e., $\boldsymbol{R} = \big(1-\rho\big)\boldsymbol{I}_m + \rho\boldsymbol{1}\boldsymbol{1}^{\mathsf{T}}$, where ρ belongs the range determined by (16). Since ξ^*_{1/p_0} is D-optimal for the marginal models we have $d_j(\boldsymbol{x}, \xi^*_{1/p_0}, \boldsymbol{\beta}_j) \leq p_0$ for all $\boldsymbol{x} \in [0,1]^{\nu}$ and for all j (see Remark 4). From Lemma 5 each $d_{jh}(\boldsymbol{x}, \xi^*_{1/p_0}, \boldsymbol{\beta}_j, \boldsymbol{\beta}_h) \leq p_0$ for all $\boldsymbol{x} \in [0,1]^{\nu}$ and for all $j < h$ with equality at each support point of ξ^*_{1/p_0}. As a result, for correlations ρ from (16) condition (14) of Theorem 5 holds and hence ξ^*_{1/p_0} is locally D-optimal for the MGLM in Ξ.

For illustration, when $m = 2$ and $\nu = 2$ with the set of vertices $\{(0,0)^{\mathsf{T}}, (1,0)^{\mathsf{T}}, (0,1)^{\mathsf{T}}, (1,1)^{\mathsf{T}}\}$ as an experimental region, the equally weighted design $\xi^*_{1/3} = \xi^*_{1/p_0}$ with support points $(0,0)^{\mathsf{T}}$, $(1,0)^{\mathsf{T}}$, $(0,1)^{\mathsf{T}}$ is locally D-optimal at $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^{\mathsf{T}}, \boldsymbol{\beta}_2^{\mathsf{T}})^{\mathsf{T}} = (\beta_{10}, \beta_{11}, \beta_{12}, \beta_{20}, \beta_{21}, \beta_{22})^{\mathsf{T}}$ with $\beta_{j0}^2 \leq \beta_{j1}\beta_{j2}, j = 1,2$ for correlation $\rho \in [\pm\rho(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)]$ where

$$\rho(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2) = \sqrt{\frac{6 - \big(d_1((1,1)^{\mathsf{T}}, \xi^*_{1/3}, \boldsymbol{\beta}_1) + d_2((1,1)^{\mathsf{T}}, \xi^*_{1/3}, \boldsymbol{\beta}_2)\big)}{6 - 2d_{12}((1,1)^{\mathsf{T}}, \xi^*_{1/3}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2)}}. \tag{17}$$

The function $\rho(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$ given by (17) determines the range of correlation ρ at which the design $\xi^*_{1/3}$ is D-optimal. Note that $\rho(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$ depends on the values of the sensitivity functions at the vertex $(1,1)^{\mathsf{T}}$ given the parameter vectors $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$. Fig. 1 displays the curve of $\rho(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$ for $\boldsymbol{\beta}_1 = (1,1,\beta)^{\mathsf{T}}$, $\boldsymbol{\beta}_2 = (1,3,\beta)^{\mathsf{T}}$, $1 \leq \beta \leq 10$. It can be seen that the values of $\rho(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$ increase as β increases. The minimum value is given by 0.77 at $\beta = 1$ and thus $\rho \in [-0.77, 0.77]$. The maximum of $\rho(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$ is given by 0.96 at $\beta = 10$ and thus $\rho \in [-0.96, 0.96]$.

For A-optimality let the parameter vector $\boldsymbol{\beta}$ be given such that $\boldsymbol{\beta}_j = \boldsymbol{\beta}_0 = (\beta_0, \beta_1, \ldots, \beta_\nu)^{\mathsf{T}}$ for all $j \in \{1,\ldots,m\}$ (see Remark 3). Gaffke et al. [9] showed that the saturated design $\xi^*_{\boldsymbol{\beta}_0}$ with support points given by Table 1 and weights

$$\omega_1^* = \frac{\sqrt{p_0}}{\sqrt{p_0} + \nu + \sum_{k=1}^{\nu}\gamma_k}, \quad \omega_k^* = \frac{1 + \gamma_{k-1}}{\sqrt{p_0} + \nu + \sum_{k=1}^{\nu}\gamma_k}, \quad k \in \{2,\ldots,p_0\}, \gamma_k = \frac{\beta_k}{\beta_0}, k \in \{1,\ldots,\nu\}$$

is locally A-optimal for the jth univariate gamma model at $\boldsymbol{\beta}_0$ under the condition

$$\gamma_k\gamma_{k'} - \frac{1}{\sqrt{p_0}}(\gamma_k + \gamma_{k'}) \geq \Big(1 + \frac{2}{\sqrt{p_0}}\Big) \text{ for all } (1 \leq k < k' \leq \nu). \tag{18}$$

By Theorem 3 the design $\xi^*_{\boldsymbol{\beta}_0}$ is locally A-optimal in Ξ_{p_0} for the MGLM at $\boldsymbol{\beta} = \boldsymbol{1} \otimes \boldsymbol{\beta}_0$ if condition (18) holds.

In particular, let $\nu = 2$ so $p_0 = 3$. Let the jth the parameter vector $\boldsymbol{\beta}_j = \boldsymbol{\beta}_0 = (\beta_0, \beta, \beta)^{\mathsf{T}}$ be given for all $j \in \{1,\ldots,m\}$. Consequently, condition (18) is equivalent to $\gamma^2 - (2/\sqrt{3})\gamma - (\sqrt{3}+2)/\sqrt{3} \geq 0$ where $\gamma = \beta/\beta_0$. E.g., at $\gamma = 3$ the design $\xi^*_{\boldsymbol{\beta}_0}$ has the form

$$\xi^*_{\boldsymbol{\beta}_0} = \begin{pmatrix} (0,0)^{\mathsf{T}} & (1,0)^{\mathsf{T}} & (0,1)^{\mathsf{T}} \\ \frac{\sqrt{3}}{\sqrt{3}+8} & \frac{4}{\sqrt{3}+8} & \frac{4}{\sqrt{3}+8} \end{pmatrix}. \tag{19}$$

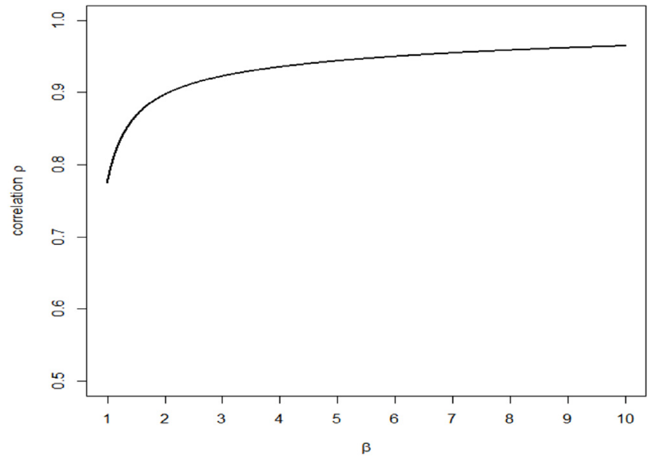

Fig. 1. The curve of the function $\rho(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2)$ from (17) at the parameter vector $\boldsymbol{\beta} = (1, 1, \beta, 1, 3, \beta)^\mathsf{T}$, $1 \le \beta \le 10$.

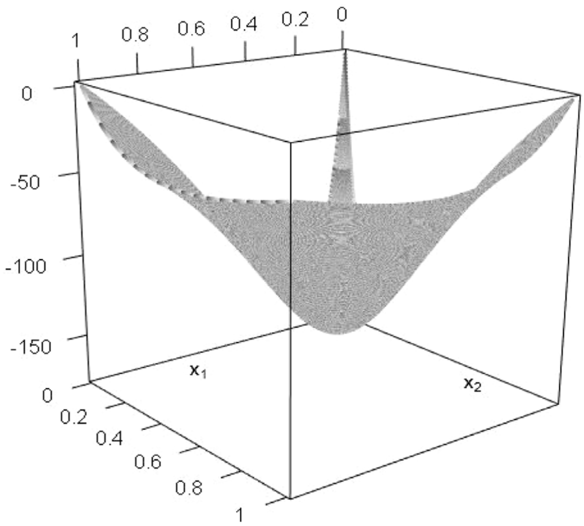

Fig. 2. The function $\psi(\boldsymbol{x}, \xi^*_{\boldsymbol{\beta}_0}, \boldsymbol{\beta}_0)$ from (20) for the locally A-optimal saturated design $\xi^*_{\boldsymbol{\beta}_0}$ given by (19), at the parameter vector $\boldsymbol{\beta}_0 = (1, 3, 3)^\mathsf{T}$, on the experimental region $[0, 1]^2$ for any correlation $\rho \in (-1, 1)$.

We can approve the local A-optimality of $\xi^*_{\boldsymbol{\beta}_0}$ from (19) in Ξ using condition (9) of the general equivalence theorem. To this end, define the function

$$\psi(\boldsymbol{x}, \xi^*_{\boldsymbol{\beta}_0}, \boldsymbol{\beta}_0) = \mathrm{tr}\Big(\boldsymbol{R}^{-1}\mathbf{f}^\mathsf{T}_{\beta}(\boldsymbol{x})\boldsymbol{M}^{-2}(\xi^*_{\boldsymbol{\beta}_0}, \boldsymbol{\beta})\mathbf{f}_{\beta}(\boldsymbol{x})\Big) - \mathrm{tr}\Big(\boldsymbol{M}^{-1}(\xi^*_{\boldsymbol{\beta}_0}, \boldsymbol{\beta})\Big),\ \boldsymbol{x} \in \mathcal{X}. \tag{20}$$

Then $\xi^*_{\boldsymbol{\beta}_0}$ from (19) is A-optimal if and only if $\psi(\boldsymbol{x}, \xi^*_{\boldsymbol{\beta}_0}, \boldsymbol{\beta}) \le 0$ for all $\boldsymbol{x} \in [0, 1]^2$. The function (20) is plotted in Fig. 2 at $\boldsymbol{\beta} = \mathbf{1} \otimes \boldsymbol{\beta}_0$ where $\boldsymbol{\beta}_0 = (1, 3, 3)^\mathsf{T}$ and $-1 < \rho < 1$. It can be seen that $\psi(\boldsymbol{x}, \xi^*_{\boldsymbol{\beta}_0}, \boldsymbol{\beta})$ is bounded by 0 for all $\boldsymbol{x} \in [0, 1]^2$ and $\psi(\boldsymbol{x}, \xi^*_{\boldsymbol{\beta}_0}, \boldsymbol{\beta}) = 0$ for $\boldsymbol{x} \in \{(0, 0)^\mathsf{T}, (1, 0)^\mathsf{T}, (0, 1)^\mathsf{T}\}$, i.e., at the support points of $\xi^*_{\boldsymbol{\beta}_0}$. From calculations, the A-optimality of $\xi^*_{\boldsymbol{\beta}_0}$ in Ξ seems to be independent of ρ.

In addition to the previous examples provided under particular restrictions, it is worthy to propose certain algorithms to find unrestricted locally optimal designs numerically with emphasis on the D-criterion. For simplification purposes we concentrate on bivariate generalized linear models, i.e., $m = 2$. We propose the multiplicative algorithm (see Yu [31] and Harman and Trnovská [10]) and the general non-linear optimization using augmented Lagrange multiplier method via software `R` (see R Core Team [22]), package `Rsolnp`.

Let us begin with a model with linear predictors of one factor

$$\begin{cases} \eta_1(x, \boldsymbol{\beta}_1) = \beta_{10} + \beta_{11}x \\ \eta_2(x, \boldsymbol{\beta}_2) = \beta_{20} + \beta_{21}x \end{cases},\ x \in [0, 1].$$

Table 2
Numerical solutions for locally D-optimal designs for a one-factor bivariate gamma model on the experimental region $\mathcal{X} = [0, 1]$ at the correlation $\rho = 0.99$.

Parameter values				Design point x^*			
				0	0.240	0.415	1
β_{10}	β_{11}	β_{20}	β_{21}	Optimal weights			
1	0	1	0	0.5000			0.5000
1	1	1	1	0.5000			0.5000
1	1	1	0	0.3900		0.2195	0.3905
1	4	1	1	0.3663	0.2673		0.3664

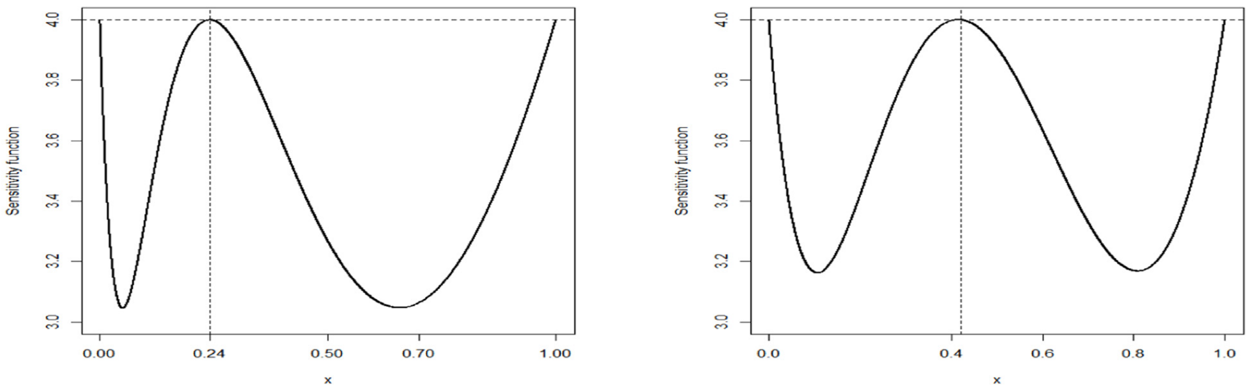

Fig. 3. The sensitivity function under numerically D-optimal designs ξ^{**} given by (22) for a one-factor bivariate GLM at the parameter vectors $\boldsymbol{\beta} = (1, 1, 1, 0)^{\mathsf{T}}$ (Left panel) and $\boldsymbol{\beta} = (1, 4, 1, 1)^{\mathsf{T}}$ (Right panel) for the correlation $\rho = 0.99$.

For the parameter vector $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^{\mathsf{T}}, \boldsymbol{\beta}_2^{\mathsf{T}})^{\mathsf{T}} = (\beta_{10}, \beta_{11}, \beta_{20}, \beta_{21})^{\mathsf{T}}$, the locally D-optimal saturated design (see Idais [13]) under each marginal model is given by

$$\xi^*_{1/2} = \begin{pmatrix} 0 & 1 \\ 0.5 & 0.5 \end{pmatrix}. \tag{21}$$

The correlation range (16) is given by $[\pm \min_{x\in(0,1)} \rho(x, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2)]$, where

$$\rho(x, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2) = \sqrt{\frac{4 - (d_1(x,\ \xi^*_{1/2}, \boldsymbol{\beta}_1) + d_2(x,\ \xi^*_{1/2}, \boldsymbol{\beta}_2))}{4 - 2d_{12}(x,\ \xi^*_{1/2}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2)}},\ x \in (0, 1).$$

For $\rho \in [\pm \min_{x\in(0,1)} \rho(x, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2)]$ numerical solutions are expected to be coincide with the analytic results from Theorem 5. In contrast to that, at the parameter vector $\boldsymbol{\beta} = (1, 1, 1, 0)^{\mathsf{T}}$ or $\boldsymbol{\beta} = (1, 4, 1, 1)^{\mathsf{T}}$ and ρ outside the respective range, specifically at $\rho = 0.99$, numerical solutions showed that a locally D-optimal deign ξ^{**} may contain one additional point

$$\xi^{**} = \begin{pmatrix} 0 & x^* & 1 \\ \omega^* & 1 - 2\omega^* & \omega^* \end{pmatrix}, \tag{22}$$

where $x^* \in (0, 1)$. Table 2 presents the design points of ξ^{**} with corresponding weights derived numerically at certain values of $\boldsymbol{\beta}$ and $\rho = 0.99$. As it is shown, at $\boldsymbol{\beta} = (1, 0, 1, 0)^{\mathsf{T}}$ and $\boldsymbol{\beta} = (1, 1, 1, 1)^{\mathsf{T}}$ the design is given by $\xi^*_{1/2}$ since the correlation $\rho = 0.99$ belongs to the respective range. Otherwise, at $\boldsymbol{\beta} = (1, 1, 1, 0)^{\mathsf{T}}$ and $\boldsymbol{\beta} = (1, 4, 1, 1)^{\mathsf{T}}$ although the resulting designs may not saturated, the weights of both 0 and 1 are equal. The equivalence theorem (8) were employed to approve the local D-optimality of designs in Ξ for $\rho = 0.99$ at the parameter vectors $\boldsymbol{\beta} = (1, 1, 1, 0)^{\mathsf{T}}$ and $\boldsymbol{\beta} = (1, 4, 1, 1)^{\mathsf{T}}$ as it shown in Fig. 3. Under both parameter vectors the sensitivity function $d(x, \xi^{**}, \boldsymbol{\beta}) = \mathrm{tr}\Big(\boldsymbol{R}^{-1}\mathbf{f}_{\boldsymbol{\beta}}^{\mathsf{T}}(x)\boldsymbol{M}^{-1}(\xi^{**}, \boldsymbol{\beta})\mathbf{f}_{\boldsymbol{\beta}}(x)\Big)$ is bounded by 4 for all $x \in [0, 1]$ and equality holds at the support points of the design ξ^{**}.

The D-efficiency of the saturated design $\xi = \xi^*_{1/2}$ from (21) at $\rho = 0.99$ can be examined on a specific region of parameter values determined specifically by $\boldsymbol{\beta} = (1, \beta_{11}, 1, 1)^{\mathsf{T}}$, $0 \le \beta_{11} \le 5$. For calculating the efficiency we use locally D-optimal designs ξ^{**} from (22) which is obtained numerically at $\rho = 0.99$. The D-efficiencies of ξ, as a function of $\boldsymbol{\beta}$, are calculated by

$$\mathrm{Eff}(\xi, \boldsymbol{\beta}) = \left(\frac{\det \boldsymbol{M}(\xi, \boldsymbol{\beta})}{\det \boldsymbol{M}(\xi^*_{\boldsymbol{\beta}}, \boldsymbol{\beta})}\right)^{1/p}.$$

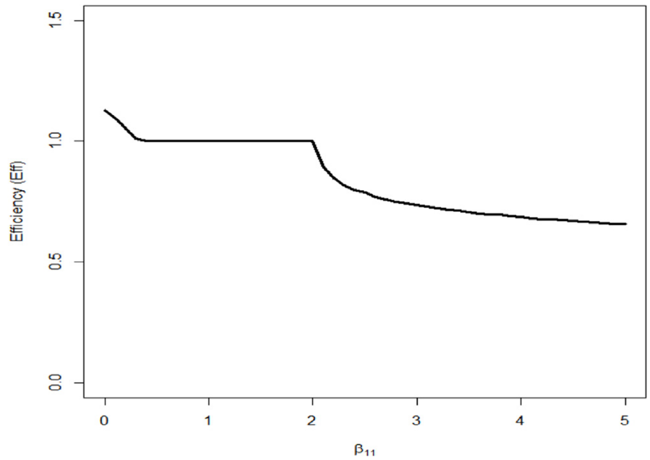

Fig. 4. Efficiencies of ξ from (21) at the parameter vectors $\boldsymbol{\beta} = (1, \beta_{11}, 1, 1)^{\mathsf{T}}$, $0 \le \beta_{11} \le 5$ for the correlation $\rho = 0.99$.

Table 3
Numerical solutions for locally D-optimal designs for a two-factor bivariate gamma model on the experimental region $\mathcal{X} = [0, 1]^2$ at the parameter vector $\boldsymbol{\beta} = (1, 3, 3, 1, 1, -0.8)^{\mathsf{T}}$ with the corresponding efficiencies under some values of the correlation ρ.

Design points		Correlation ρ					
		0	0.3	0.5	0.7	0.9	0.99
x_1	x_2	Optimal weights					
0	0	0.3157	0.3152	0.3141	0.3041	0.2337	0.2035
0	1	0.3116	0.3107	0.3085	0.2970	0.2302	0.2027
1	0	0.1899	0.1901	0.1905	0.1918	0.1942	0.1997
1	1	0.1828	0.1840	0.1869	0.1916	0.1949	
0	0.50				0.0155	0.0877	0.0381
0	0.55					0.0593	0.1570
0.85	1						0.1753
0.90	1						0.0237
D-Efficiencies		0.9729	0.9737	0.9752	0.9780	0.9163	0.6638

Here, $p = 4$ and $\xi^*_{\boldsymbol{\beta}} = \xi^{**}$. The D-efficiencies of ξ are depicted in Fig. 4. It can be seen that for β_{11} very close to zero the saturated design $\xi = \xi^*_{1/2}$ performs better than ξ^{**}. Then we have the perfect performance when β_{11} moves towards 2 because we get $\xi = \xi^{**}$. Otherwise, as β_{11} increases the efficiencies curve tends to have small values. The overall performance of ξ does not show satisfactory at $\rho = 0.99$. This is of course because the optimality of the saturated design in Ξ depends on ρ.

Next we focus on a bivariate GLM with two factors in the experimental region $[0, 1]^2$ and linear predictors

$$\begin{cases} \eta_1(\boldsymbol{x}, \boldsymbol{\beta}_1) = \beta_{10} + \beta_{11}x_1 + \beta_{21}x_2, \\ \eta_2(\boldsymbol{x}, \boldsymbol{\beta}_2) = \beta_{20} + \beta_{21}x_1 + \beta_{22}x_2, \end{cases}$$

where $\boldsymbol{x} = (x_1, x_2)^{\mathsf{T}} \in [0, 1]^2$ and $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^{\mathsf{T}}, \boldsymbol{\beta}_2^{\mathsf{T}})^{\mathsf{T}} = (\beta_{10}, \beta_{11}, \beta_{12}, \beta_{20}, \beta_{21}, \beta_{22})^{\mathsf{T}}$. We consider $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ having different types. For example; at $\boldsymbol{\beta} = (1, 3, 3, 1, 1, -0.8)^{\mathsf{T}}$ numerical solutions cannot provide D-optimal saturated designs as it is shown by Table 3. It can be seen that the results depend on the value of ρ. It may be desirable to study the potential merits of the locally D-optimal designs $\xi^*_{\boldsymbol{\beta}}$ derived numerically for $\rho \in \{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.99\}$ at $\boldsymbol{\beta} = (1, 3, 3, 1, 1, -0.8)^{\mathsf{T}}$ by employing the function of D-efficiency with $p = 6$. For examination of the efficiency we select the design $\xi_{1/4}$ which is uniform on $(0, 0)^{\mathsf{T}}, (1, 0)^{\mathsf{T}}, (0, 1)^{\mathsf{T}}, (1, 1)^{\mathsf{T}}$. Note that $\xi_{1/4}$ is locally D-optimal at $\boldsymbol{\beta} = (1, 0, 0, 1, 0, 0)^{\mathsf{T}}$ for the bivariate model as well as it is locally D-optimal for the jth marginal model on $[0, 1]^2$ at $\boldsymbol{\beta}_j = (1, 0, 0)^{\mathsf{T}}$. The efficiencies of $\xi_{1/4}$ are depicted in Fig. 5 and some of them are presented in Table 3. It can been noted that for $\rho \le 0.9$ the design $\xi_{1/4}$ performs quite well. Otherwise, $\xi_{1/4}$ shows the worst performance for $\rho > 0.9$ where the minimum efficiency is equal to 0.6638 at $\rho = 0.99$. The maximum efficiency is given by 0.9780 at $\rho = 0.7$.

5. Product type designs

Another situation of the multivariate structure of GLMs can be considered when all components follow specifically univariate gamma models assuming seemingly unrelated linear predictors. That means, we allow different factors which

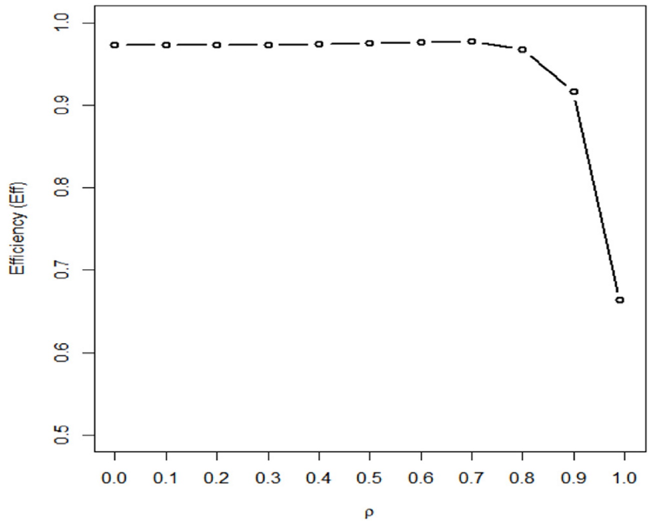

Fig. 5. D-Efficiencies of the uniform design $\xi_{1/4}$ at the parameter vector $\boldsymbol{\beta} = (1, 3, 3, 1, 1, -0.8)^\mathsf{T}$ for a two-factor bivariate gamma model on the experimental region $\mathcal{X} = [0, 1]^2$.

belong to different experimental regions and different regression functions in the linear predictors. Let ν_j denote the number of the factors associated with the component j and ν denote the total number of all factors in the MGLM, i.e., $\nu = \sum_{j=1}^{m} \nu_j$. For any unit i the univariate responses $Y_j(\boldsymbol{x}_{ij})$, $j \in \{1, \ldots, m\}$, are assumed to be correlated through the correlation matrix $\boldsymbol{R}$. Here, all $Y_j(\boldsymbol{x}_{ij})$, $j \in \{1, \ldots, m\}$, come from gamma distributions where the experimental condition $\boldsymbol{x}_{ij} = (x_{ij1}, \ldots, x_{ij\nu_j})^\mathsf{T}$ may differ across the components of unit i and is chosen from an experimental region $\mathcal{X}_j \subseteq \mathbb{R}^{\nu_j}$, $\nu_j \geq 1$, $j \in \{1, \ldots, m\}$. As in Gaffke et al. [9], the univariate gamma models with power links can be employed

$$\eta_j(\boldsymbol{x}_{ij}, \boldsymbol{\beta}_j) = \mu_j^{\kappa_j}(\boldsymbol{x}_{ij}, \boldsymbol{\beta}_j),\ \eta_j(\boldsymbol{x}_{ij}, \boldsymbol{\beta}_j) = \boldsymbol{f}_j^\mathsf{T}(\boldsymbol{x}_{ij})\boldsymbol{\beta}_j = \sum_{l=1}^{p_j} f_{jl}(\boldsymbol{x}_{ij})\beta_{jl}.$$

Note that κ_j is a given nonzero real number and denotes the exponent of the power link in the jth model. Furthermore, $\boldsymbol{f}_j$ is the vector of p_j linearly independent regression functions $f_{j1}, \ldots, f_{jp_j}$ and $\boldsymbol{\beta}_j = \left(\beta_{j1}, \ldots, \beta_{jp_j}\right)^\mathsf{T} \in \mathbb{R}^{p_j}$. The intensity function is given by $u_j(\boldsymbol{x}_{ij}, \boldsymbol{\beta}_j) = \alpha_j\kappa_j\left(\boldsymbol{f}_j^\mathsf{T}(\boldsymbol{x}_{ij})\boldsymbol{\beta}_j\right)^{-2}$. Here, α_j is the shape parameter of a gamma distribution. Note that $\alpha_j\kappa_j$ is a positive constant and can be ignored. The expected mean $\mu_j(\boldsymbol{x}_{ij}, \boldsymbol{\beta}_j)$ for the gamma distribution is positive and therefore, we assume $\boldsymbol{f}_j^\mathsf{T}(\boldsymbol{x}_{ij})\boldsymbol{\beta}_j > 0$ for all $i \in \{1, \ldots, n\}$.

The experimental region $\mathcal{X}$ for the MGLM with power links and seemingly unrelated linear predictors is given by the Cartesian product of all marginal experimental regions $\mathcal{X}_j$, $j \in \{1, \ldots, m\}$, i.e., $\mathcal{X} = \times_{j=1}^{m}\mathcal{X}_j$. The $p \times m$ block diagonal multivariate regression is given by $\mathbf{f}(\boldsymbol{x}_i) = \mathrm{diag}\big(\boldsymbol{f}_1(\boldsymbol{x}_{i1}), \ldots, \boldsymbol{f}_m(\boldsymbol{x}_{im})\big)$ where $\boldsymbol{x}_i = (\boldsymbol{x}_{i1}^\mathsf{T}, \ldots, \boldsymbol{x}_{im}^\mathsf{T})^\mathsf{T} \in \mathcal{X}$. Obviously, $\boldsymbol{x}_i = (x_{i11}, \ldots, x_{i1\nu_1}, \ldots, x_{im1}, \ldots, x_{im\nu_m})^\mathsf{T}$ is a ν-tuple. Let the stacked parameter p-vector $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\mathsf{T}, \ldots, \boldsymbol{\beta}_m^\mathsf{T})^\mathsf{T}$ be given. The MGLM with univariate gamma models is defined by

$$\boldsymbol{\eta}(\boldsymbol{x}_i, \boldsymbol{\beta}) = \left(\mu_1^{\kappa_1}(\boldsymbol{x}_{i1}, \boldsymbol{\beta}_1), \ldots, \mu_m^{\kappa_m}(\boldsymbol{x}_{im}, \boldsymbol{\beta}_m)\right)^\mathsf{T},\ \ \boldsymbol{\eta}(\boldsymbol{x}_i, \boldsymbol{\beta}) = \mathbf{f}^\mathsf{T}(\boldsymbol{x}_i)\boldsymbol{\beta}.$$

In particular, for multivariate models with seemingly unrelated linear models, i.e., $\eta_j(\boldsymbol{x}_{ij}, \boldsymbol{\beta}_j) = \mu_j(\boldsymbol{x}_{ij}, \boldsymbol{\beta}_j)$ with $u_j(\boldsymbol{x}_{ij}, \boldsymbol{\beta}_j) = 1$ for all $j \in \{1, \ldots, m\}$, Soumaya et al. [27] reduced the optimality problem to the marginal counterparts and product type designs were developed. In the following, we will extend their results under the MGLM with seemingly unrelated gamma models. The product type design is supported by the cross-product of the finite sets of design points of the designs under marginal ν_j-factor gamma models and the weights are given by the product of the weights of those designs. To be more specific, denote by ξ_j a design defined on $\mathcal{X}_j$ for a marginal ν_j-factor gamma model ($1 \leq j \leq m$). We present ξ_j as in (6) which assigns the weights $\omega_{j1}, \omega_{j2}, \ldots, \omega_{jr_j}$ to the support points $\boldsymbol{x}_{j1}, \boldsymbol{x}_{j2}, \ldots, \boldsymbol{x}_{jr_j}$. Then the product type design $\xi = \bigotimes_{j=1}^{m} \xi_j$ is defined on $\mathcal{X} = \mathcal{X}_1 \times \ldots \times \mathcal{X}_m$ and has $r = \prod_{j=1}^{m} r_j$ design points $\boldsymbol{x}_{i_1,\ldots,i_m} = (x_{1i_1}, \ldots, x_{ji_m})^\mathsf{T}$ with corresponding weights $\omega_{i_1,\ldots,i_m} = \prod_{j=1}^{m} \omega_{ji_j}$, $i_j \in \{1, \ldots, r_j\}$, $j \in \{1, \ldots, m\}$.

For each component j denote $\boldsymbol{f}_{j,\boldsymbol{\beta}_j}(\boldsymbol{x}_{ij}) = \left(\boldsymbol{f}_j^\mathsf{T}(\boldsymbol{x}_{ij})\boldsymbol{\beta}_j\right)^{-1}\boldsymbol{f}_j(\boldsymbol{x}_{ij})$ for all $i \in \{1, \ldots, n\}$. Note that $\boldsymbol{f}_{j,\boldsymbol{\beta}_j}(\boldsymbol{x}_{ij})$ involves implicitly the intercept term, i.e., there exists a constant vector $\boldsymbol{c}_j$ such that $\boldsymbol{c}_j^\mathsf{T}\boldsymbol{f}_{j,\boldsymbol{\beta}_j}(\boldsymbol{x}_{ij}) = 1$ for all $i \in \{1, \ldots, n\}$. Here, $\boldsymbol{c}_j = \boldsymbol{\beta}_j$, $j \in \{1, \ldots, m\}$. As a result, the quasi-Fisher information matrix for the MGLM factorizes analogously to Soumaya et al. [27] and the optimality problem at $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \ldots, \boldsymbol{\beta}_m^\top)^\top$ for the MGLM can be reduced to the marginal ν_j-factor models at $\boldsymbol{\beta}_j$, $j \in \{1, \ldots, m\}$.

The following theorem develops a product structure of a locally optimal design for our MGLM with respect to the D- and A-criteria. The proof is similar to the analogous results in Soumaya et al. [27].

Theorem 6. *Consider a MGLM with seemingly unrelated gamma models. Let* $\xi = \bigotimes_{j=1}^m \xi_j$ *be a product-type design on the experimental region* $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j$. *Let a parameter point* $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \ldots, \boldsymbol{\beta}_m^\top)^\top$ *be given. For each marginal design* ξ_j *denote by* $\boldsymbol{M}_j(\xi_j, \boldsymbol{\beta}_j) = \int_{\mathcal{X}_j} \boldsymbol{f}_{j,\boldsymbol{\beta}_j}(\boldsymbol{x}_j) \boldsymbol{f}_{j,\boldsymbol{\beta}_j}^\top(\boldsymbol{x}_j)\, \xi_j(\mathrm{d}\boldsymbol{x}_j)$ *and by* $\boldsymbol{m}_j(\xi_j, \boldsymbol{\beta}_j) = \int_{\mathcal{X}_j} \boldsymbol{f}_{j,\boldsymbol{\beta}_j}(\boldsymbol{x}_j)\, \xi_j(\mathrm{d}\boldsymbol{x}_j)$ *the information matrix and the moment vector, respectively. Then the quasi-Fisher information matrix of design* ξ *has the form*

$$\boldsymbol{M}(\xi, \boldsymbol{\beta}) = \mathrm{diag}\Big(\rho^{(jj)} \left(\boldsymbol{M}_j(\xi_j, \boldsymbol{\beta}_j) - \boldsymbol{m}_j(\xi_j, \boldsymbol{\beta}_j)\boldsymbol{m}_j^\top(\xi_j, \boldsymbol{\beta}_j)\right)\Big) + \boldsymbol{m}(\xi, \boldsymbol{\beta})\boldsymbol{R}^{-1}\boldsymbol{m}^\top(\xi, \boldsymbol{\beta}),$$

where $\boldsymbol{m}(\xi, \boldsymbol{\beta}) = \mathrm{diag}\Big(\boldsymbol{m}_j(\xi_j, \boldsymbol{\beta}_j)\Big)_{j=1}^m$. *If the information matrix* $\boldsymbol{M}_j(\xi_j, \boldsymbol{\beta}_j)$ *for all* $j \in \{1, \ldots, m\}$ *is nonsingular then the quasi-Fisher information matrix* $\boldsymbol{M}(\xi, \boldsymbol{\beta})$ *for the MGLM with seemingly unrelated gamma models is nonsingular and*

$$\boldsymbol{M}^{-1}(\xi, \boldsymbol{\beta}) = \mathrm{diag}\left(\frac{1}{\rho^{(jj)}}\left(\boldsymbol{M}_j^{-1}(\xi_j, \boldsymbol{\beta}_j) - \boldsymbol{\beta}_j\boldsymbol{\beta}_j^\top\right)\right)_{j=1}^m + \boldsymbol{B}\boldsymbol{R}\boldsymbol{B}^\top,$$

where $\boldsymbol{B} = \mathrm{diag}\left(\boldsymbol{\beta}_j\right)_{j=1}^m$. *Moreover, for each* $j \in \{1, \ldots, m\}$, *let* ξ_j^* *be a locally D- or A-optimal design (at* $\boldsymbol{\beta}_j$*) for the jth univariate gamma model on the experimental region* $\mathcal{X}_j$. *Then the product type design* $\xi^* = \bigotimes_{j=1}^m \xi_j^*$ *is a locally D- or A-optimal design (at* $\boldsymbol{\beta}$*) for the MGLM with seemingly unrelated gamma models on the experimental region* $\mathcal{X} = \times_{j=1}^m \mathcal{X}_j$, *respectively.*

Example 1. Consider a bivariate generalized linear model with two marginal univariate gamma models with seemingly unrelated linear predictors, i.e., $m = 2$. Let us begin with only one factor in each linear predictor where

$$\begin{cases} \eta_1(x_1, \boldsymbol{\beta}_1) = \beta_{10} + \beta_{11}x_1,\ x_1 \in \mathcal{X}_1 = [0, 1], \\ \eta_2(x, \boldsymbol{\beta}_2) = \beta_{20} + \beta_{21}x_2,\ x_2 \in \mathcal{X}_2 = [0, 1]. \end{cases}$$

According to Idais [13] the locally D-optimal design for each jth one-factor gamma model with the experimental region $\mathcal{X}_j = [0, 1]$ is $\xi_j^* = \xi_{1/2}^*$ given in (21) at any parameter vector $\boldsymbol{\beta}_j = (\beta_{j0}, \beta_{j1})^\top$ such that $\beta_{j0} > 0$ and $\beta_{j1} > -\beta_{j0}$ for $j = 1, 2$. So ξ_j^* assigns weights 1/2 to the boundaries 0 and 1. Based on Theorem 6 the product type design

$$\xi^* = \xi_1^* \otimes \xi_2^* = \begin{pmatrix} (0,0)^\top & (1,0)^\top & (0,1)^\top & (1,1)^\top \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}, \tag{23}$$

is locally D-optimal for the bivariate gamma model with seemingly unrelated linear predictors at $\boldsymbol{\beta} = \left(\beta_{10}, \beta_{11}, \beta_{20}, \beta_{21}\right)^\top$ such that $\beta_{j0} > 0$ and $\beta_{j1} > -\beta_{j0}$ for $j = 1, 2$.

Moreover, the multiplicative algorithm can be used to a find locally optimal design numerically for the proposed bivariate model. The resulting optimal designs are not necessarily of product type. For example; when $\rho = 0$ the numerical results show that as the ratio β_{j1}/β_{j0} becomes smaller the locally D-optimal design is given by the product type design of form (23). Otherwise, as β_{j1}/β_{j0} becomes larger the weights of $(0,0)^\top$ and $(1,1)^\top$ decrease whereas the weights of $(1,0)^\top$ and $(0,1)^\top$ increase. E.g., at $\boldsymbol{\beta} = (1, 7, 1, 7)^\top$ and $\rho = 0$ the locally D-optimal design is given numerically by

$$\xi^{**} = \begin{pmatrix} (0,0)^\top & (1,0)^\top & (0,1)^\top & (1,1)^\top \\ 0.185 & 0.315 & 0.315 & 0.185 \end{pmatrix}. \tag{24}$$

Although ξ^{**} is not a product type design, we have $\det \boldsymbol{M}(\xi^*, \boldsymbol{\beta}) = \det \boldsymbol{M}(\xi^{**}, \boldsymbol{\beta})$. As a result, neither ξ^* from (23) nor ξ^{**} is unique at $\boldsymbol{\beta} = (1, 7, 1, 7)^\top$. For another parameter point for instance, $\boldsymbol{\beta} = (1, 3, 1, 4)^\top$ and arbitrary $\rho \in \{0.3, 0.5, 0.7, 0.9, 0.99\}$ the locally D-optimal design on the experimental region $[0, 1]^2$ is uniformly supported by its vertices which is given by (23). So the product type design (23) is independent of $\rho \in \{0.3, 0.5, 0.7, 0.9, 0.99\}$.

Now we consider marginal gamma models with two-factor linear predictors

$$\begin{cases} \eta_1(x_1, \boldsymbol{\beta}_1) = \beta_{10} + \beta_{11}x_{11} + \beta_{12}x_{12},\ \boldsymbol{x}_1 = (x_{11}, x_{12})^\top \in \mathcal{X}_1 = [0, 1]^2, \\ \eta_2(x, \boldsymbol{\beta}_2) = \beta_{20} + \beta_{21}x_{21} + \beta_{22}x_{22},\ \boldsymbol{x}_2 = (x_{21}, x_{22})^\top \in \mathcal{X}_2 = [0, 1]^2. \end{cases}$$

Idais [13] provided a complete solution of the local D-optimality in the two-factor gamma model with a linear predictor as given above. In particular, for the first marginal gamma model on $\mathcal{X}_1$ the saturated design

$$\xi_1^* = \begin{pmatrix} (0,0)^\top & (1,0)^\top & (1,1)^\top \\ 1/3 & 1/3 & 1/3 \end{pmatrix},$$

is locally D-optimal at any $\boldsymbol{\beta}_1 = (\beta_{10}, \beta_{11}, \beta_{12})^\top$ with $(\beta_{10} + \beta_{11})^2 + \beta_{11}\beta_{12} \le 0$ such as $\boldsymbol{\beta}_1 = (1, -0.8, 1)^\top$. For the second marginal gamma model on $\mathcal{X}_2$ the uniform design ξ_2^* given by (24) is locally D-optimal at any $\boldsymbol{\beta}_2 = (\beta_{20}, \beta_{21}, \beta_{22})^\top$

with $\beta_{20} > 0, \beta_{21} = \beta_{22} = 0$. Then by Theorem 6 the product type design $\xi^* = \xi_1^* \otimes \xi_2^*$ which assigns the weights $\omega_i = 1/12, i \in \{1, \dots, 12\}$ to the support points

$$(0,0,0,0)^\mathsf{T}, (0,0,1,0)^\mathsf{T}, (0,0,0,1)^\mathsf{T}, (0,0,1,1)^\mathsf{T}, (1,0,0,0)^\mathsf{T}, (1,0,1,0)^\mathsf{T},$$
$$(1,0,0,1)^\mathsf{T}, (1,0,1,1)^\mathsf{T}, (1,1,0,0)^\mathsf{T}, (1,1,1,0)^\mathsf{T}, (1,1,0,1)^\mathsf{T}, (1,1,1,1)^\mathsf{T}$$

is locally D-optimal for the bivariate gamma model with seemingly unrelated linear predictors at the parameter vector $\boldsymbol{\beta} = \big(\beta_{10}, \beta_{11}, \beta_{12}, \beta_{20}, 0, 0\big)^\mathsf{T}$ such that $(\beta_{10} + \beta_{11})^2 + \beta_{11}\beta_{12} \le 0$, $\beta_{10} > 0$, $\beta_{20} > 0$.

6. Conclusion

In the present paper we studied the local optimality of designs for the multivariate generalized linear model with respect to the D- and A-criteria. The model can deal with multiple responses observed for a unit where each response comes from a one-parameter exponential family distributions. We found that the focus on all saturated designs simplifies the optimality solutions. That is the optimality problem reduces to the marginal univariate models. Consequently, the locally optimal saturated design is still optimal for the MGLM in the set of all saturated designs and is independent of the correlation. For D-optimality and an exchangeable correlation matrix the general equivalence theorem provides a necessary and sufficient condition for the saturated design to be D-optimal for the MGLM in the set of all designs. Here, the correlation is bounded by certain limits determined by the sensitivity functions in the general equivalence theorem under the univariate marginal models. Our results were discussed for responses coming from gamma distributions. In this situation, locally D- and A-optimal saturated designs are provided analytically. Then we employed particular algorithms to find locally D-optimal non-saturated designs for different correlation values. We concluded this paper with the development of optimal designs for multivariate gamma models with seemingly unrelated linear predictors. The optimal design is given by the product of optimal designs obtained under the marginal models.

There are wide applications for the results introduced in Section 3. Besides gamma models presented in Section 4 the first order Poisson models can be considered. That is Y_j comes from a Poisson distribution and thus the canonical link is given by $\eta_j(\boldsymbol{x}, \boldsymbol{\beta}_j) = \log\big(\mu_j(\boldsymbol{x}, \boldsymbol{\beta}_j)\big) = \boldsymbol{f}_j^\mathsf{T}(\boldsymbol{x})\boldsymbol{\beta}_j = \beta_{j0} + \sum_{k=1}^{\nu} \beta_{jk}x_k$ for all $\boldsymbol{x} \in \mathcal{X}$ and for all j. The intensity function is given by $u_j(\boldsymbol{x}, \boldsymbol{\beta}_j) = \exp\big(\boldsymbol{f}_j^\mathsf{T}(\boldsymbol{x})\boldsymbol{\beta}_j\big)$ for all $\boldsymbol{x} \in \mathcal{X}$ and for all j.

According to Russell et al. [24] the equally weighted design ξ^*_{1/p_0} on $\mathcal{X} = [0,1]^\nu$ with support given by Table 1 is locally D-optimal for the jth univariate Poisson model at $\boldsymbol{\beta}_j = (\beta_{j0}, \beta_{j1}, \dots, \beta_{j\nu})^\mathsf{T}$ with $\beta_{jk} = -2$ for all $k \in \{1, \dots, \nu\}$ and for all $j \in \{1, \dots, m\}$. By Theorem 2 the design ξ^*_{1/p_0} is locally D-optimal in Ξ_{p_0} for the MGLM at $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\mathsf{T}, \dots, \boldsymbol{\beta}_m^\mathsf{T})^\mathsf{T}$ such that $\beta_{jk} = -2$ for all $k \in \{1, \dots, p_0\}$ and for all $j \in \{1, \dots, m\}$. In this situation even Theorem 4 can approve the local D-optimality of ξ^*_{1/p_0} in Ξ.

Theorem 2 covers MGLMs that combine multiple responses from distinct probability distributions. As an example; let $m = 2$ so we only get $\boldsymbol{Y} = (Y_1, Y_2)^\mathsf{T}$. Assume that Y_1 comes from a gamma distribution and Y_2 is from a Poisson distribution. Let $\mathcal{X} = [0,1]^2$ and $\boldsymbol{f}_j(\boldsymbol{x}) = (1, x_1, x_2)^\mathsf{T}$ with $\boldsymbol{\beta}_j = (\beta_{j0}, \beta_{j1}, \beta_{j2})^\mathsf{T}$, $j = 1, 2$. Thus by Theorem 2, the equally weighted design $\xi^*_{1/3}$ with support $(0,0)^\mathsf{T}, (1,0)^\mathsf{T}, (0,1)^\mathsf{T}$ is locally D-optimal in Ξ_3 for the bivariate GLM at $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\mathsf{T}, \boldsymbol{\beta}_2^\mathsf{T})^\mathsf{T}$ such that $\beta_{10}^2 \le \beta_{11}\beta_{12}$ and $\beta_{21} = \beta_{22} = -2$.

Acknowledgments

The author acknowledges the support by a scholarship of the Landesgraduiertenförderung Sachsen-Anhalt, Germany. This paper has been completed based on the final scholarship for doctoral students with family responsibility and funding gap from Otto-von-Guericke University Magdeburg, Germany.

Appendix

Proof of Theorem 5. We employ condition (8) of the general equivalence theorem to approve the D-optimality of the saturated design ξ^* in Ξ. To this end, denote $\boldsymbol{A} = \mathbf{f}_\beta^\mathsf{T}(\boldsymbol{x})\boldsymbol{M}^{-1}(\xi^*, \boldsymbol{\beta})\mathbf{f}_\beta(\boldsymbol{x})$. Then the l.h.s. of condition (8) with $\boldsymbol{R}$ from (13) is equal to $\mathrm{tr}\Big(\boldsymbol{R}^{-1}\boldsymbol{A}\Big) = \frac{1}{1-\rho}\mathrm{tr}(\boldsymbol{A} + \frac{\rho}{1+(m-1)\rho}\mathbf{1}\mathbf{1}^\mathsf{T}\boldsymbol{A})$. We can write $\boldsymbol{A} = \boldsymbol{A}_1 + \boldsymbol{A}_2$ where $\boldsymbol{A}_1 = (1-\rho)\mathrm{diag}(\boldsymbol{f}_{\beta_j}^\mathsf{T}(\boldsymbol{x})\boldsymbol{M}_j^{-1}(\xi^*, \boldsymbol{\beta}_j)\boldsymbol{f}_{\beta_j}(\boldsymbol{x}))_{j=1}^m$ and $\boldsymbol{A}_2 = \rho\Big[\boldsymbol{f}_{\beta_j}^\mathsf{T}(\boldsymbol{x})\boldsymbol{M}_{jh}^{-1}(\xi^*, \boldsymbol{\beta}_j, \boldsymbol{\beta}_h)\boldsymbol{f}_{\beta_h}(\boldsymbol{x})\Big]_{j=1,\dots,m}^{h=1,\dots,m}$. It is straightforward to obtain $\mathrm{tr}(\boldsymbol{A}_1) = (1-\rho)\sum_{j=1}^m d_j(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j)$ and $\mathrm{tr}(\boldsymbol{A}_2) = \rho\sum_{j=1}^m d_j(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j)$. Thus $\mathrm{tr}(\boldsymbol{A}) = \sum_{j=1}^m d_j(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j)$. Note that $\frac{\rho}{1+(m-1)\rho}\mathbf{1}\mathbf{1}^\mathsf{T}\boldsymbol{A} = \frac{\rho}{1+(m-1)\rho}(\mathbf{1}\mathbf{1}^\mathsf{T}\boldsymbol{A}_1 + \mathbf{1}\mathbf{1}^\mathsf{T}\boldsymbol{A}_2)$ with $\mathrm{tr}(\mathbf{1}\mathbf{1}^\mathsf{T}\boldsymbol{A}_1) = \mathrm{tr}(\boldsymbol{A}_1)$ and $\mathrm{tr}(\mathbf{1}\mathbf{1}^\mathsf{T}\boldsymbol{A}_2) = \rho\sum_{j=1}^m d_j(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j) + 2\rho\sum_{j<h=1}^m d_{jh}(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j, \boldsymbol{\beta}_h)$. It follows that

$$\begin{aligned}
\mathrm{tr}\Big(\boldsymbol{R}^{-1}\boldsymbol{A}\Big) &= \mathrm{tr}\Big(\boldsymbol{R}^{-1}\mathbf{f}_\beta^\mathsf{T}(\boldsymbol{x})\boldsymbol{M}^{-1}(\xi^*, \boldsymbol{\beta})\mathbf{f}_\beta(\boldsymbol{x})\Big) = \frac{1}{1-\rho}\Big(\mathrm{tr}(\boldsymbol{A}) - \frac{\rho}{1+(m-1)\rho}(\mathrm{tr}(\mathbf{1}\mathbf{1}^\mathsf{T}\boldsymbol{A}_1) + \mathrm{tr}(\mathbf{1}\mathbf{1}^\mathsf{T}\boldsymbol{A}_2))\Big) \\
&= \frac{1}{1-\rho}\Big(\frac{1+(m-2)\rho}{1+(m-1)\rho}\sum_{j=1}^m d_j(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j) - \frac{2\,\rho^2}{1+(m-1)\rho}\sum_{j<h=1}^m d_{jh}(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j, \boldsymbol{\beta}_h)\Big),
\end{aligned}$$

which is less than or equal to mp_0 for all $\boldsymbol{x} \in \mathcal{X}$ by (14). □

Proof of Lemma 5. Denote by $\boldsymbol{x}_i^*$ for all $i \in \{1, \dots, p_0\}$ the support points of the design $\xi^* \in \Xi_{p_0}$. Since $\boldsymbol{f}_j(\boldsymbol{x}) = \boldsymbol{f}_0(\boldsymbol{x})$ for all $\boldsymbol{x} \in \mathcal{X}$ and for all j we have $\boldsymbol{F}_j = \boldsymbol{F}_0$ for all j. Define the vector $\boldsymbol{a}(\boldsymbol{x}) = \left(\boldsymbol{F}_0^{\mathsf{T}}\right)^{-1}\boldsymbol{f}_0(\boldsymbol{x}) = (a_1(\boldsymbol{x}), \dots, a_{p_0}(\boldsymbol{x}))^{\mathsf{T}}$ for all $\boldsymbol{x} \in \mathcal{X}$. Let $u_{ij} = u_j(\boldsymbol{x}_i^*, \boldsymbol{\beta}_j)$ for all $i \in \{1, \dots, p_0\}$. Denote $\tilde{u}_{ij}(\boldsymbol{x}, \boldsymbol{\beta}_j) = u_j(\boldsymbol{x}, \boldsymbol{\beta}_j)u_{ij}^{-1}$ for all $\boldsymbol{x} \in \mathcal{X}$. Let $\mathbf{Q}_j(\boldsymbol{x}, \boldsymbol{\beta}_j) = \mathrm{diag}\left(\tilde{u}_{ij}(\boldsymbol{x}, \boldsymbol{\beta}_j)\right)_{j=1}^{p_0}$ and $\mathbf{Q}_{jh}(\boldsymbol{x}, \boldsymbol{\beta}_j, \boldsymbol{\beta}_h) = \mathrm{diag}\left(\tilde{u}_{ij}^{1/2}(\boldsymbol{x}, \boldsymbol{\beta}_j)\tilde{u}_{ih}^{1/2}(\boldsymbol{x}, \boldsymbol{\beta}_h)\right)_{j=1}^{p_0}$ for all $\boldsymbol{x} \in \mathcal{X}$. Then we have

$$d_j(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j) = \boldsymbol{f}_0^{\mathsf{T}}(\boldsymbol{x})\boldsymbol{F}_0^{-1}\mathrm{diag}\left(u_j(\boldsymbol{x}, \boldsymbol{\beta}_j)u_{ij}^{-1}\right)_{j=1}^{p_0}\left(\boldsymbol{F}_0^{\mathsf{T}}\right)^{-1}\boldsymbol{f}_0(\boldsymbol{x}) = \boldsymbol{a}^{\mathsf{T}}(\boldsymbol{x})\mathbf{Q}_j(\boldsymbol{x}, \boldsymbol{\beta}_j)\boldsymbol{a}(\boldsymbol{x}) = \sum_{i=1}^{p_0} a_i^2(\boldsymbol{x})\tilde{u}_{ij}(\boldsymbol{x}, \boldsymbol{\beta}_j),\ \boldsymbol{x} \in \mathcal{X},$$

$$d_{jh}(\boldsymbol{x}, \xi^*, \boldsymbol{\beta}_j, \boldsymbol{\beta}_h) = \boldsymbol{a}^{\mathsf{T}}(\boldsymbol{x})\mathbf{Q}_{jh}(\boldsymbol{x}, \boldsymbol{\beta}_j, \boldsymbol{\beta}_h)\boldsymbol{a}(\boldsymbol{x}) = \sum_{i=1}^{p_0} a_i^2(\boldsymbol{x})\sqrt{\tilde{u}_{ij}(\boldsymbol{x}, \boldsymbol{\beta}_j)\tilde{u}_{ih}(\boldsymbol{x}, \boldsymbol{\beta}_j)},\ \boldsymbol{x} \in \mathcal{X}.$$

Since $\mathbf{Q}_j(\boldsymbol{x}, \boldsymbol{\beta}_j)$ and $\mathbf{Q}_{jh}(\boldsymbol{x}, \boldsymbol{\beta}_j, \boldsymbol{\beta}_h)$ are positive definite we have $\boldsymbol{a}^{\mathsf{T}}(\boldsymbol{x})\mathbf{Q}_j(\boldsymbol{x}, \boldsymbol{\beta}_j)\boldsymbol{a}(\boldsymbol{x}) > 0$ and $\boldsymbol{a}^{\mathsf{T}}(\boldsymbol{x})\mathbf{Q}_{jh}(\boldsymbol{x}, \boldsymbol{\beta}_j, \boldsymbol{\beta}_h)\boldsymbol{a}(\boldsymbol{x}) > 0$ for all $\boldsymbol{x} \in \mathcal{X}$. Let $r(\boldsymbol{x}) = \boldsymbol{a}^{\mathsf{T}}(\boldsymbol{x})\mathbf{Q}_j(\boldsymbol{x}, \boldsymbol{\beta}_j)\boldsymbol{a}(\boldsymbol{x}) - \boldsymbol{a}^{\mathsf{T}}(\boldsymbol{x})\mathbf{Q}_{jh}(\boldsymbol{x}, \boldsymbol{\beta}_j, \boldsymbol{\beta}_h)\boldsymbol{a}(\boldsymbol{x}) = \sum_{i=1}^{p_0} a_i^2(\boldsymbol{x})\left(\tilde{u}_{ij}(\boldsymbol{x}, \boldsymbol{\beta}_j) - \sqrt{\tilde{u}_{ij}(\boldsymbol{x}, \boldsymbol{\beta}_j)\tilde{u}_{ih}(\boldsymbol{x}, \boldsymbol{\beta}_j)}\right)$ for all $\boldsymbol{x} \in \mathcal{X}$. For an arbitrary point $\boldsymbol{x}_0 \in \mathcal{X}$, if $r(\boldsymbol{x}_0) \geq 0$, we get $\boldsymbol{a}^{\mathsf{T}}(\boldsymbol{x}_0)\mathbf{Q}_j(\boldsymbol{x}_0, \boldsymbol{\beta}_j)\boldsymbol{a}(\boldsymbol{x}_0) \geq \boldsymbol{a}^{\mathsf{T}}(\boldsymbol{x}_0)\mathbf{Q}_{jh}(\boldsymbol{x}_0, \boldsymbol{\beta}_j, \boldsymbol{\beta}_h)\boldsymbol{a}(\boldsymbol{x}_0) \geq \boldsymbol{a}^{\mathsf{T}}(\boldsymbol{x}_0)\mathbf{Q}_h(\boldsymbol{x}_0, \boldsymbol{\beta}_h)\boldsymbol{a}(\boldsymbol{x}_0)$. Otherwise, if $r(\boldsymbol{x}_0) \leq 0$, we get $\boldsymbol{a}^{\mathsf{T}}(\boldsymbol{x}_0)\mathbf{Q}_j(\boldsymbol{x}_0, \boldsymbol{\beta}_j)\boldsymbol{a}(\boldsymbol{x}_0) \leq \boldsymbol{a}^{\mathsf{T}}(\boldsymbol{x}_0)\mathbf{Q}_{jh}(\boldsymbol{x}_0, \boldsymbol{\beta}_j, \boldsymbol{\beta}_h)\boldsymbol{a}(\boldsymbol{x}_0) \leq \boldsymbol{a}^{\mathsf{T}}(\boldsymbol{x}_0)\mathbf{Q}_h(\boldsymbol{x}_0, \boldsymbol{\beta}_h)\boldsymbol{a}(\boldsymbol{x}_0)$. Hence, the lemma is proved. □

References

[1] J. Burridge, P. Sebastiani, D-optimal designs for generalised linear models with variance proportional to the square of the mean, Biometrika 81 (1994) 295–304.
[2] S.I. Chang, Some properties of multiresponse D-optimal designs, J. Math. Anal. Appl. 184 (1994) 256–262.
[3] H. Chernoff, Locally optimal designs for estimating parameters, Ann. Math. Stat. 24 (1953) 586–602.
[4] M. Crowder, On the use of a working correlation matrix in using generalised linear models for repeated measures, Biometrika 82 (1995) 407–410.
[5] H. Dette, L. Hoyden, S. Kuhnt, K. Schorning, Optimal designs for thermal spraying, J. R. Stat. Soc. Ser. C. Appl. Stat. 66 (2017) 53–72.
[6] V.V. Fedorov, The design of experiments in the multiresponse case, Theory Probab. Appl. 16 (1971) 323–332.
[7] V.V. Fedorov, R.C. Gagnon, S.L. Leonov, Design of experiments with unknown parameters in variance, Appl. Stoch. Models Bus. Ind. 18 (2002) 207–218.
[8] V.V. Fedorov, S.L. Leonov, Optimal Design for Nonlinear Response Models, CRC Press, Boca Raton, 2013.
[9] N. Gaffke, O. Idais, R. Schwabe, Locally optimal designs for gamma models, J. Statist. Plann. Inference 203 (2019) 199–214.
[10] R. Harman, M. Trnovská, Approximate D-optimal designs of experiments on the convex hull of a finite set of information matrices, Math. Slovaca 59 (2009) 693–704,.
[11] M.A. Heise, R.H. Myers, Optimal designs for bivariate logistic regression, Biometrics 52 (1996) 613–624.
[12] M.-N.L. Huang, R.-B. Chen, C.-S. Lin, W.K. Wong, Optimal designs for parallel models with correlated responses, Statist. Sinica 16 (2006) 121–133.
[13] O. Idais, On local optimality of vertex type designs in generalized linear models, Statist. Papers (2020) http://dx.doi.org/10.1007/s00362-020-01158-4.
[14] L. Imhof, Optimum designs for a multiresponse regression model, J. Multivariate Anal. 72 (2000) 120–131.
[15] J. Jankar, A. Mandal, J. Yang, Optimal crossover designs for generalized linear models, 2020, arXiv:2001.06168.
[16] O. Krafft, M. Schaefer, D-Optimal designs for a multivariate regression model, J. Multivariate Anal. 42 (1992) 130–140.
[17] V.G. Kurotschka, R. Schwabe, The reduction of design problems for multivariate experiments to univariate possibilities and their limitations, in: E. Brunner, M. Denker (Eds.), Research Developments in Probability and Statistics, in: Festschrift in Honor of Madan L. Puri on the Occasion of his 65th Birthday, VSP, Utrecht, 1996, pp. 193–204.
[18] K.-Y. Liang, S. Zeger, Longitudinal data analysis using generalized linear models, Biometrika 73 (1986) 13–22.
[19] J. Liu, G.A. Colditz, Optimal design of longitudinal data analysis using generalized estimating equation models, Biom. J. 59 (2017) 315–330.
[20] X. Liu, R.-X. Yue, F.J. Hickernell, Optimality criteria for multiresponse linear models based on predictive ellipsoids, Statist. Sinica 21 (2011) 421–432.
[21] P. McCullagh, J. Nelder, Generalized Linear Models, second ed., Chapman & Hall, London, 1989.
[22] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2020.
[23] J. Rodríguez-Díaz, G. Sánchez-León, Optimal designs for multiresponse models with double covariance structure, Chemometr. Intell. Lab. Syst. 189 (2019) 1–7.
[24] K.G. Russell, D.C. Woods, S.M. Lewis, J.A. Eccleston, D-optimal designs for Poisson regression models, Statist. Sinica 19 (2009) 721–730.
[25] R. Schwabe, Optimal design of experiments for multivariate response in two-factor linear models, in: A.K. Gupta, V.L. Girko (Eds.), Proceedings of the Sixth Eugene Lukacs Symposium, Bowling Green, OH, in: Multidimensional Statistical Analysis and Theory of Random Matrices, VSP, Utrecht, 1996, pp. 235–242.
[26] S.D. Silvey, Optimal Design, Chapman & Hall, London, 1980.
[27] M. Soumaya, N. Gaffke, R. Schwabe, Optimal design for multivariate observations in seemingly unrelated linear models, J. Multivariate Anal. 142 (2015) 48–56.
[28] P. van de Ven, D. Woods, Optimal blocked minimum-support designs for non-linear models, J. Statist. Plann. Inference 144 (2014) 152–159.
[29] W.K. Wong, Y. Yin, J. Zhou, Optimal designs for multi-response nonlinear regression models with several factors via semidefinite programming, J. Comput. Graph. Statist. 28 (2019) 61–73.
[30] D.C. Woods, P. Van de Ven, Blocked designs for experiments with correlated non-normal response, Technometrics 53 (2011) 173–182.
[31] Y. Yu, Monotonic convergence of a general algorithm for computing optimal designs, Ann. Statist. (2010) 1593–1606.
[32] R.-X. Yue, X. Liu, K. Chatterjee, D-optimal designs for multiresponse linear models with a qualitative factor, J. Multivariate Anal. 124 (2014) 57–69.
[33] A. Zellner, An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, J. Amer. Statist. Assoc. 57 (1962) 348–368.
[34] S.S. Zocchi, A.C. Atkinson, Optimum experimental designs for multinomial logistic models, Biometrics 55 (1999) 437–444.