



On local optimality of vertex type designs in generalized linear models

Osama Idais¹ 

Received: 5 June 2019 / Revised: 4 November 2019 / Published online: 7 February 2020

© The Author(s) 2020

Abstract

Locally optimal designs are derived for generalized linear models with first order linear predictors. We consider models including a single factor, two factors and multiple factors. Mainly, the experimental region is assumed to be a unit cube. In particular, models without intercept are considered on arbitrary experimental regions. Analytic solutions for optimal designs are developed under the D- and A-criteria, and more generally, for Kiefer's Φ_k -criteria. The focus is on the vertex type designs. That is, the designs are only supported by the vertices of the respective experimental regions. By the equivalence theorem, necessary and sufficient conditions are developed for the local optimality of these designs. The derived results are applied to gamma and Poisson models.

Keywords Generalized linear model · Approximate design · General equivalence theorem · Intercept term · Locally optimal design

1 Introduction

The generalized linear model (GLM) was developed by Nelder and Wedderburn (1972). It is viewed as a generalization of the ordinary linear regression which allows continuous or discrete observations from one-parameter exponential family distributions to be combined with explanatory variables (factors) via proper link functions. Generalized linear models include several types such as Poisson, gamma, logistic models among others. Therefore, wide applications can be addressed by GLMs such as social and educational sciences, clinical trials, insurance, industry (Walker and Duncan 1967; Myers and Montgomery 1997; Fox 2015; Goldburd et al. 2016).

✉ Osama Idais
osama.idais@ovgu.de

¹ Institute for Mathematical Stochastics, Otto-von-Guericke University Magdeburg, PF 4120, 39016 Magdeburg, Germany

While deriving optimal designs is obtained by minimizing the variance-covariance matrix there is no loss of generality to concentrate on maximizing the Fisher information matrix. For generalized linear models the Fisher information matrix depends on the model parameters. Therefore, the optimal design cannot be found without a prior knowledge of the parameters. One approach is the so-called local optimality, which was proposed by Chernoff (1953). This approach aims at deriving an optimal design at a given parameter value (best guess).

As in many research works the results on optimal designs in particular, on a continuous experimental region are influenced by the type of models used. For example, Ford et al. (1992) used single-factor GLMs. Moreover, Gaffke et al. (2019) and Russell et al. (2009) used the gamma and the Poisson models, respectively, while the logistic model was employed by Yang et al. (2011) and Atkinson and Haines (1996).

In this paper we focus on the problem of finding locally optimal designs for a class of generalized linear models, which is motivated by the work of Yang and Stufken (2009) and Tong et al. (2014), who provided analytic results for a general setup of the GLM with binary factors. Here, we are also interested in deriving locally optimal designs for a general setup of generalized linear models on continuous and discrete experimental regions. Schmidt and Schwabe (2017) showed that the support points of the optimal designs for GLMs on an experimental region given by a polytope are located at the edges of the experimental region. In particular, in Gaffke et al. (2019) we proved that the optimal designs for gamma models are supported by the vertices of the experimental region (polytope). In this paper, we will restrict our attention to the vertices of the experimental region by which optimal designs can be supported for the corresponding generalized linear models. Throughout the sequel, we confine ourselves to the general equivalence theorem to establish a necessary and sufficient condition for a design to be locally optimal.

The remainder of the paper is organized as follows. In Sect. 2 we introduce the generalized linear model and optimality of designs. Approaches to determine the optimal weights for some particular designs under D-, A- and Kiefer Φ_k -criteria are characterized in Sect. 3. Then optimal designs are derived under the single-factor and the two-factor models in Sects. 4 and 5, respectively. First order models of multiple factors are presented in Sect. 6, and optimal designs are derived for such models with and without intercept. Applications of the results are discussed under gamma and Poisson model in Sect. 7.

2 Preliminary

In this section, we introduce the generalized linear model and give a characterization of optimal designs. Let the univariate observation (response) Y belongs to a one-parameter exponential family distribution in the canonical form

$$p(y; \theta) = \exp \left(y\theta - b(\theta) + c(y) \right), \quad (2.1)$$

where $b(\cdot)$ and $c(\cdot)$ are known functions while θ is a canonical parameter. In the generalized linear model each response Y of a statistical unit is observed at a certain value of a covariate $\mathbf{x} = (x_1, \dots, x_v)^T$ that belongs to an experimental region $\mathcal{X} \subseteq$

\mathbb{R}^v , $v \geq 1$. Here, $\theta := \theta(\mathbf{x}, \boldsymbol{\beta})$ varies with the value of $\mathbf{x} \in \mathcal{X}$ at a fixed value of the vector of model parameters $\boldsymbol{\beta} \in \mathbb{R}^p$. The expected mean is given by $E(Y) = \mu(\mathbf{x}, \boldsymbol{\beta}) = b'(\theta)$ with the variance function $V(\mu(\mathbf{x}, \boldsymbol{\beta})) = b''(\theta)$ [see McCullagh and Nelder (1989, Sect. 2.2.2)]. Let $\mathbf{f}(\mathbf{x}) : \mathcal{X} \rightarrow \mathbb{R}^p$ be a vector of continuous regression functions $f_1(\mathbf{x}), \dots, f_p(\mathbf{x})$ which are assumed to be linearly independent. Denote the linear predictor by $\eta = \mathbf{f}^T(\mathbf{x})\boldsymbol{\beta}$. In the generalized linear model it is assumed that $\eta = g(\mu(\mathbf{x}, \boldsymbol{\beta}))$, where g is a link function and assumed to be one-to-one and differentiable. We can define the intensity function at a point $\mathbf{x} \in \mathcal{X}$ as

$$u(\mathbf{x}, \boldsymbol{\beta}) = V^{-1}(\mu(\mathbf{x}, \boldsymbol{\beta})) \left(\frac{d\mu(\mathbf{x}, \boldsymbol{\beta})}{d\eta} \right)^2 \quad (2.2)$$

where $d\mu(\mathbf{x}, \boldsymbol{\beta})/d\eta = 1/g'(\mu(\mathbf{x}, \boldsymbol{\beta}))$. Obviously, $u(\mathbf{x}, \boldsymbol{\beta})$ is positive for all $\mathbf{x} \in \mathcal{X}$ and may be regarded as a weight for the corresponding unit at the point \mathbf{x} (Atkinson and Woods 2015). The Fisher information matrix at $\mathbf{x} \in \mathcal{X}$ [see Fedorov and Leonov (2013, Sect. 1.3.2)] is given by

$$\mathbf{M}(\mathbf{x}, \boldsymbol{\beta}) = u(\mathbf{x}, \boldsymbol{\beta}) \mathbf{f}(\mathbf{x}) \mathbf{f}^T(\mathbf{x}). \quad (2.3)$$

An information matrix of the form (2.3) is appropriate for other nonlinear models, e.g., model with survival times observations employing the proportional hazards (Schmidt and Schwabe 2017). Moreover, under homoscedastic regression models the intensity function is constantly equal to 1 whereas, under heteroscedastic regression models we get intensity that is equal to $1/\text{var}(Y)$, which depends on \mathbf{x} only and thus we have information matrix of the form $\mathbf{M}(\mathbf{x}) = u(\mathbf{x}) \mathbf{f}(\mathbf{x}) \mathbf{f}^T(\mathbf{x})$ that does not depend on the model parameters. The latter case was discussed in Graßhoff et al. (2007) and in the book by Fedorov and Leonov (2013, p.13).

Throughout the present work we will deal with the approximate (continuous) design theory. An approximate design ξ can be defined as a probability measure with finite support on the experimental region \mathcal{X} ,

$$\xi = \begin{pmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_r \\ \omega_1 & \omega_2 & \dots & \omega_r \end{pmatrix}, \quad (2.4)$$

where $r \in \mathbb{N}$, $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_r \in \mathcal{X}$ are pairwise distinct points and $\omega_1, \omega_2, \dots, \omega_r > 0$ with $\sum_{i=1}^r \omega_i = 1$. The set $\text{supp}(\xi) = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_r\}$ is called the support of ξ and $\omega_1, \dots, \omega_r$ are called the weights of ξ [see Silvey (1980, p.15)]. The information matrix of a design ξ from (2.4) at a parameter point $\boldsymbol{\beta}$ is defined by

$$\mathbf{M}(\xi, \boldsymbol{\beta}) = \int_{\mathcal{X}} \mathbf{M}(\mathbf{x}, \boldsymbol{\beta}) \xi(d\mathbf{x}) = \sum_{i=1}^r \omega_i \mathbf{M}(\mathbf{x}_i, \boldsymbol{\beta}). \quad (2.5)$$

One might recognize $\mathbf{M}(\xi, \boldsymbol{\beta})$ as a convex combination of all information matrices for all support points of ξ . Another representation of the information matrix (2.5) can

be utilized based on the $r \times p$ design matrix $F = [f(x_1), \dots, f(x_r)]^T$ and the $r \times r$ weight matrix $V = \text{diag}(\omega_i u(x_i, \beta))_{i=1}^r$ and hence we can write

$$M(\xi, \beta) = F^T V F.$$

Remark A particular type of designs appears frequently when the support size equals the dimension of f , i.e., $r = p$. In such a case the design is minimally supported and it is often called a minimal-support or a saturated design.

This paper focuses on optimal designs within the family of Kiefer's Φ_k -criteria (Kiefer 1975). These criteria aim at minimizing the k -norm of the eigenvalues of the variance-covariance matrix and include the most common criteria for D-, A- and E- optimality. Denote by $\lambda_i(\xi, \beta)$ ($1 \leq i \leq p$) the eigenvalues of a nonsingular information matrix $M(\xi, \beta)$. Denote by "det" and "tr" the determinant and the trace of a matrix, respectively. The Kiefer's Φ_k -criteria are defined by

$$\Phi_k(M(\xi, \beta)) = \left(\frac{1}{p} \text{tr}(M^{-k}(\xi, \beta)) \right)^{\frac{1}{k}} = \left(\frac{1}{p} \sum_{i=1}^p \lambda_i^{-k}(\xi, \beta) \right)^{\frac{1}{k}}, \quad 0 < k < \infty,$$

$$\Phi_0(M(\xi, \beta)) = \lim_{k \rightarrow 0+} \Phi_k(M(\xi, \beta)) = \left(\det(M^{-1}(\xi, \beta)) \right)^{\frac{1}{p}},$$

$$\Phi_\infty(M(\xi, \beta)) = \lim_{k \rightarrow \infty} \Phi_k(M(\xi, \beta)) = \max_{1 \leq i \leq p} (\lambda_i^{-1}(\xi, \beta)).$$

Note that $\Phi_0(M(\xi, \beta))$, $\Phi_1(M(\xi, \beta))$ and $\Phi_\infty(M(\xi, \beta))$ are the D-, A- and E-criteria, respectively. Since $M(\xi, \beta)$ depends on the values of the parameters, a best guess of β is adopted here and locally D-optimal designs are constructed (Chernoff 1953). A locally Φ_k -optimal design ξ^* (at β) minimizes the function $\Phi_k(M(\xi, \beta))$ over all designs ξ whose information matrix $M(\xi, \beta)$ is nonsingular. For $0 \leq k < \infty$ the strict convexity of $\Phi_k(M(\xi, \beta))$ implies that the information matrix of a locally Φ_k -optimal design (at β) is unique. That is, if ξ^* and ξ^{**} are two locally Φ_k -optimal designs (at β) then $M(\xi^*, \beta) = M(\xi^{**}, \beta)$ (Kiefer 1975). In particular, D-optimal designs are constructed to minimize the determinant of the variance-covariance matrix of the estimates or equivalently to maximize the determinant of the information matrix. The D-criterion is typically defined by the convex function $\Phi_D(M(\xi, \beta)) = -\log \det(M(\xi, \beta))$. A-optimal designs are constructed to minimize the trace of the variance-covariance matrix of the estimates, i.e., to minimize the average variance of the estimates. The A-criterion is typically defined by $\Phi_A(M(\xi, \beta)) = \text{tr}(M^{-1}(\xi, \beta))$. Moreover, E-optimal designs maximize the smallest eigenvalue of $M(\xi, \beta)$.

In order to verify the local optimality of a design the general equivalence theorem is commonly employed [see Atkinson et al. (2007, p.137)]. It provides necessary and sufficient conditions for a design to be optimal and thus the optimality of a suggested design can be easily verified or disproved. The design ξ^* is Φ_k -optimal if and only if

$$u(x, \beta) f^T(x) M^{-k-1}(\xi^*, \beta) f(x) \leq \text{tr}(M^{-k}(\xi^*, \beta)) \quad \text{for all } x \in \mathcal{X}, \quad (0 \leq k < \infty). \quad (2.6)$$

Furthermore, if the design ξ^* is Φ_k -optimal then inequality (2.6) becomes equality at its support.

Remark The left hand side of condition (2.6) of the general equivalence theorem is called the sensitivity function.

3 Determination of locally optimal weights

In this section we provide the optimal weights of particular types of designs that will appear throughout the paper with respect to Kiefer's Φ_k -criteria. Particular emphasis will be on the A-criterion ($k = 1$) and the D-criterion ($k = 0$). This work mostly deals with saturated designs (i.e., $r = p$). Let the support points be given by $\mathbf{x}_1^*, \dots, \mathbf{x}_p^*$ such that $\mathbf{f}(\mathbf{x}_1^*), \dots, \mathbf{f}(\mathbf{x}_p^*)$ are linearly independent. For the A-criterion ($k = 1$) the optimal weights are given according to Pukelsheim (1993, Sect. 8.8), which has been modified in Gaffke et al. (2019). The design ξ^* which achieves the minimum value of $\text{tr}(\mathbf{M}^{-1}(\xi, \boldsymbol{\beta}))$ over all designs ξ with $\text{supp}(\xi) = \{\mathbf{x}_1^*, \dots, \mathbf{x}_p^*\}$ is given by

$$\xi^* = \left(\begin{matrix} \mathbf{x}_1^* & \dots & \mathbf{x}_p^* \\ \omega_1^* & \dots & \omega_p^* \end{matrix} \right), \text{ with } \omega_i^* = c^{-1} \left(\frac{c_{ii}}{u_i} \right)^{1/2} \quad (1 \leq i \leq p), \quad c = \sum_{k=1}^p \left(\frac{c_{kk}}{u_k} \right)^{1/2}, \quad (3.1)$$

where $u_i = u(\mathbf{x}_i^*, \boldsymbol{\beta})$ ($1 \leq i \leq p$) and c_{ii} ($1 \leq i \leq p$) are the diagonal entries of the matrix $\mathbf{C} = (\mathbf{F}^{-1})^T \mathbf{F}^{-1}$ and $\mathbf{F} = [\mathbf{f}(\mathbf{x}_1^*), \dots, \mathbf{f}(\mathbf{x}_p^*)]^T$.

For the D-criterion ($k = 0$) the optimal weights are given by $\omega_i^* = 1/p$ ($1 \leq i \leq p$), see Lemma 5.1.3 of Silvey (1980). This means that the locally D-optimal saturated design assigns equal weights to the support points. On the other hand, there is no unified formulas for the optimal weights of a non-saturated design specifically, with respect to D-criterion. However, for $p = 3$ the following lemma provides the optimal weights of a design with four support points $\xi^* = \{\mathbf{x}_i^*, \omega_i^*, i = 1, 2, 3, 4\}$ under certain conditions.

Lemma 3.1 *Let $p = 3$. Let the design points $\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*, \mathbf{x}_4^* \in \mathcal{X}$ be given such that any three of the four vectors $\mathbf{f}(\mathbf{x}_1^*), \mathbf{f}(\mathbf{x}_2^*), \mathbf{f}(\mathbf{x}_3^*), \mathbf{f}(\mathbf{x}_4^*)$ are linearly independent. Denote*

$$\begin{aligned} d_1 &= \det[\mathbf{f}(\mathbf{x}_2^*), \mathbf{f}(\mathbf{x}_3^*), \mathbf{f}(\mathbf{x}_4^*)], \quad d_2 = \det[\mathbf{f}(\mathbf{x}_1^*), \mathbf{f}(\mathbf{x}_3^*), \mathbf{f}(\mathbf{x}_4^*)], \\ d_3 &= \det[\mathbf{f}(\mathbf{x}_1^*), \mathbf{f}(\mathbf{x}_2^*), \mathbf{f}(\mathbf{x}_4^*)], \quad d_4 = \det[\mathbf{f}(\mathbf{x}_1^*), \mathbf{f}(\mathbf{x}_2^*), \mathbf{f}(\mathbf{x}_3^*)] \end{aligned}$$

such that $d_i \neq 0$, $i = 1, 2, 3, 4$. For a given parameter point $\boldsymbol{\beta}$ denote $u_i = u(\mathbf{x}_i^, \boldsymbol{\beta})$, $i = 1, 2, 3, 4$. Assume that $u_2 = u_3$ and $d_2^2 = d_3^2$ and let*

$$\begin{aligned} \omega_1^* &= \frac{3}{8} + \frac{1}{4} \left(1 + \frac{d_1^2 u_1}{d_4^2 u_4} - 4 \frac{d_2^2 u_1}{d_4^2 u_2} \right)^{-1}, \\ \omega_2^* &= \omega_3^* = \frac{1}{2} \left(4 - \frac{d_4^2 u_2}{d_2^2 u_1} - \frac{d_1^2 u_2}{d_2^2 u_4} \right)^{-1}, \end{aligned}$$

$$\omega_4^* = \frac{3}{8} + \frac{1}{4} \left(1 + \frac{d_4^2 u_4}{d_1^2 u_1} - 4 \frac{d_2^2 u_4}{d_1^2 u_2} \right)^{-1}.$$

Assume that $\omega_i^* > 0$, $i = 1, 2, 3, 4$. Then the design ξ^* which achieves the minimum value of $-\log \det(\mathbf{M}(\xi, \boldsymbol{\beta}))$ over all designs ξ with $\text{supp}(\xi) = \{\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*, \mathbf{x}_4^*\}$ is given by $\xi^* = \{(\mathbf{x}_i^*, \omega_i^*), i = 1, 2, 3, 4\}$.

Proof Let $\mathbf{f}_\ell = \mathbf{f}(\mathbf{x}_\ell^*)$ ($1 \leq \ell \leq 4$). The 4×3 design matrix is given by $\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3, \mathbf{f}_4]^T$. Denote $\mathbf{V} = \text{diag}(\omega_\ell u_\ell)_{\ell=1}^4$. Then $\mathbf{M}(\xi, \boldsymbol{\beta}) = \mathbf{F}^T \mathbf{V} \mathbf{F}$ and by the Cauchy–Binet formula the determinant of $\mathbf{M}(\xi, \boldsymbol{\beta})$ is given by the function $\varphi(\omega_1, \omega_2, \omega_3, \omega_4)$ where

$$\varphi(\omega_1, \omega_2, \omega_3, \omega_4) = \sum_{\substack{1 \leq i < j < k \leq 4 \\ h \in \{1, 2, 3, 4\} \setminus \{i, j, k\}}} d_h^2 u_i u_j u_k \omega_i \omega_j \omega_k. \quad (3.2)$$

By assumptions $u_2 = u_3$, $d_2^2 = d_3^2$ the function $\varphi(\omega_1, \omega_2, \omega_3, \omega_4)$ is invariant w.r.t. permuting ω_2 and ω_3 , i.e., $\varphi(\omega_1, \omega_2, \omega_3, \omega_4) = \varphi(\omega_1, \omega_3, \omega_2, \omega_4)$ and thus minimizing (3.2) has the same solutions for ω_2 and ω_3 . Thus we can write $\omega_4 = 1 - \omega_1 - 2\omega_2$ and (3.2) reduces to

$$\varphi(\omega_1, \omega_2) = \alpha_1 \omega_2^3 + \alpha_2 \omega_2^2 + \alpha_3 \omega_1^2 \omega_2 + \alpha_4 \omega_2^2 \omega_1 + \alpha_5 \omega_1 \omega_2,$$

where $\alpha_1 = -2\alpha_2 = -2d_4^2 u_2^2 u_4$, $\alpha_3 = -\alpha_5 = -4d_2^2 u_1 u_2 u_4$, $\alpha_4 = u_2^2 (d_1^2 u_1 - d_4^2 u_4) - 4d_2^2 u_1 u_2 u_4$. Thus we obtain the system of two equations $\partial\varphi/\partial\omega_1 = 0$, $\partial\varphi/\partial\omega_2 = 0$. Straightforward computations show that the solution of the above system is the optimal weights ω_ℓ^* ($1 \leq \ell \leq 4$) presented by the lemma. Hence, these optimal weights minimizing $\varphi(\omega_1, \omega_2)$. \square

Moreover, the choice of optimal weights of the saturated design under Kiefer Φ_k -criteria was given in Pukelsheim et al. (1991). It was stated in Schmidt (2019), Sect. 5, that the method of Pukelsheim et al. (1991) provides a system of equations that must be solved numerically. In the following, explicit optimal weights of the Φ_k -optimal saturated design are derived for a GLM without intercept specifically, under the first order model $\mathbf{f}(\mathbf{x}) = (x_1, \dots, x_\nu)^T$ and a parameter vector $\boldsymbol{\beta} = (\beta_1, \dots, \beta_\nu)^T$. The choice of locally Φ_k -optimal weights which yields the minimum value of $\Phi_k(\mathbf{M}(\xi, \boldsymbol{\beta}))$ over all saturated designs with the same support is given by the following lemma.

Lemma 3.2 Consider a GLM without intercept with $\mathbf{f}(\mathbf{x}) = (x_1, \dots, x_\nu)^T$ on an experimental region \mathcal{X} . Denote by \mathbf{e}_i for all ($1 \leq i \leq \nu$) the ν -dimensional unit vectors. Let $\mathbf{x}_i^* = a_i \mathbf{e}_i$, $a_i > 0$ for all ($1 \leq i \leq \nu$) be design points in \mathcal{X} . For a given parameter point $\boldsymbol{\beta} = (\beta_1, \dots, \beta_\nu)^T$ let $u_i = u(\mathbf{x}_i^*, \boldsymbol{\beta})$ for all ($1 \leq i \leq \nu$). Let a vector $\mathbf{a} = (a_1, \dots, a_\nu)^T$ be given with positive components. Then the design ξ_a^* which achieves the minimum value of $\Phi_k(\mathbf{M}(\xi_a, \boldsymbol{\beta}))$ over all designs ξ_a with $\text{supp}(\xi_a) = \{\mathbf{x}_1^*, \dots, \mathbf{x}_\nu^*\}$ assigns weights

$$\omega_i^* = \frac{(a_i^2 u_i)^{-\frac{k}{k+1}}}{\sum_{j=1}^v (a_j^2 u_j)^{-\frac{k}{k+1}}} \quad (1 \leq i \leq v)$$

to the corresponding design points $\mathbf{x}_1^*, \dots, \mathbf{x}_v^*$. Hence, for D -optimality ($k = 0$), $\omega_i^* = 1/v$ ($1 \leq i \leq v$).

for A -optimality ($k = 1$), $\omega_i^* = \frac{(a_i^2 u_i)^{-1/2}}{\sum_{j=1}^v (a_j^2 u_j)^{-1/2}}$ ($1 \leq i \leq v$).

for E -optimality ($k \rightarrow \infty$), $\omega_i^* = \frac{(a_i^2 u_i)^{-1}}{\sum_{j=1}^v (a_j^2 u_j)^{-1}}$ ($1 \leq i \leq v$).

Proof Define the $v \times v$ design matrix $\mathbf{F} = \text{diag}(a_i)_{i=1}^v$ with the $v \times v$ weight matrix $\mathbf{V} = \text{diag}(u_i \omega_i)_{i=1}^v$. Then we have $\mathbf{M}(\xi_a, \boldsymbol{\beta}) = \mathbf{F}^T \mathbf{V} \mathbf{F} = \text{diag}(a_i^2 u_i \omega_i)_{i=1}^v$ and $\mathbf{M}^{-k}(\xi_a, \boldsymbol{\beta}) = \text{diag}((a_i^2 u_i \omega_i)^{-k})_{i=1}^v$ with $\text{tr}(\mathbf{M}^{-k}(\xi_a, \boldsymbol{\beta})) = \sum_{i=1}^v (a_i^2 u_i \omega_i)^{-k}$. Thus

$$\Phi_k(\mathbf{M}(\xi_a, \boldsymbol{\beta})) = \left(\frac{1}{v} \sum_{i=1}^v (a_i^2 u_i \omega_i)^{-k} \right)^{\frac{1}{k}}, \quad (0 < k < \infty). \quad (3.3)$$

Now we aim at minimizing $\Phi_k(\mathbf{M}(\xi_a, \boldsymbol{\beta}))$ such that $\omega_i > 0$ and $\sum_{i=1}^v \omega_i = 1$. We write $\omega_v = 1 - \sum_{i=1}^{v-1} \omega_i$ then (3.3) becomes

$$\Phi_k(\mathbf{M}(\xi_a, \boldsymbol{\beta})) = \frac{1}{v^{1/k}} \left((a_v^2 u_v)^{-k} (1 - \sum_{i=1}^{v-1} \omega_i)^{-k} + \sum_{i=1}^{v-1} (a_i^2 u_i \omega_i)^{-k} \right)^{\frac{1}{k}}.$$

It is straightforward to see that the equation $\frac{\partial \Phi_k(\mathbf{M}(\xi_a, \boldsymbol{\beta}))}{\partial \omega_i} = 0$ is equivalent to

$$(a_i^2 u_i)^k \omega_i^{k+1} - (a_v^2 u_v)^k \left(1 - \sum_{i=1}^{v-1} \omega_i \right)^{k+1} = 0$$

which gives $\omega_i = \left(a_v^2 u_v / (a_i^2 u_i) \right)^{\frac{k}{k+1}} \omega_v$ ($1 \leq i \leq v-1$), thus $\omega_i (a_i^2 u_i)^{\frac{k}{k+1}} = \omega_v (a_v^2 u_v)^{\frac{k}{k+1}}$ ($1 \leq i \leq v-1$). This means $\omega_i (a_i^2 u_i)^{\frac{k}{k+1}}$ ($1 \leq i \leq v$) are all equal, i.e., $\omega_i (a_i^2 u_i)^{\frac{k}{k+1}} = c$ ($1 \leq i \leq v$), where $c > 0$. It implies that $\omega_i = c (a_i^2 u_i)^{-\frac{k}{k+1}}$ ($1 \leq i \leq v$). Due to $\sum_{i=1}^v \omega_i = 1$ we get $\sum_{i=1}^v c (a_i^2 u_i)^{-\frac{k}{k+1}} = c \sum_{i=1}^v (a_i^2 u_i)^{-\frac{k}{k+1}} = 1$, and thus $c = \left(\sum_{i=1}^v (a_i^2 u_i)^{-\frac{k}{k+1}} \right)^{-1}$. So we finally obtain $\omega_i = (a_i^2 u_i)^{-\frac{k}{k+1}} / \left(\sum_{j=1}^v (a_j^2 u_j)^{-\frac{k}{k+1}} \right)$ for all ($1 \leq i \leq v$) which are the optimal weights given by the lemma. \square

4 Single-factor model

In this section we deal with the simplest case under a model with a single factor

$$\eta(x, \boldsymbol{\beta}) = \mathbf{f}^T(x) \boldsymbol{\beta} = \beta_0 + \beta_1 x \quad \text{where } x \in \mathcal{X}.$$

Let the experimental region is taken to be the continuous unit interval $\mathcal{X} = [0, 1]$. In the following we introduce, for a fixed $\boldsymbol{\beta} = (\beta_0, \beta_1)^T$, the function

$$h(x) = \frac{1}{u(x, \boldsymbol{\beta})}, \quad x \in [0, 1],$$

which will be utilized for the characterization of the optimal designs. Consider the following conditions:

- (i) $u(x, \boldsymbol{\beta})$ is positive and twice continuously differentiable on $[0, 1]$.
- (ii) $u(x, \boldsymbol{\beta})$ is strictly increasing on $[0, 1]$.
- (iii) $h''(x)$ is an injective (one-to-one) function on $[0, 1]$.

Recently, Lemma 1 in Konstantinou et al. (2014) showed that under the above conditions (i)-(iii) a locally D-optimal design on $[0, 1]$ is only supported by two points a and b where $0 \leq a < b \leq 1$. In what follows an analogous result is presented for locally optimal designs under various optimality criteria.

Lemma 4.1 *Consider a GLM with $\mathbf{f}(x) = (1, x)^T$ and the experimental region $\mathcal{X} = [0, 1]$. Let a parameter point $\boldsymbol{\beta} = (\beta_0, \beta)^T$ be given. Let conditions (i)-(iii) be satisfied. Denote by \mathbf{A} a positive definite matrix and let c be constant. Then if the condition of the general equivalence theorem is of the form*

$$u(x, \boldsymbol{\beta}) \mathbf{f}^T(x) \mathbf{A} \mathbf{f}(x) \leq c$$

then the support points of a locally optimal design ξ^ is concentrated on exactly two points a and b where $0 \leq a < b \leq 1$.*

Proof Let $\mathbf{A} = [a_{ij}]_{i,j=1,2}$. Then let $p(x) = \mathbf{f}^T(x) \mathbf{A} \mathbf{f}(x) = a_{22}x^2 + 2a_{12}x + a_{11}$ which is a polynomial in x of degree 2 where $x \in \mathcal{X}$. Hence, by the general equivalence theorem ξ^* is locally optimal (at $\boldsymbol{\beta}$) if and only if

$$p(x) \leq c h(x) \quad \text{for all } x \in [0, 1].$$

The above inequality is similar to that obtained in the proof of Lemma 1 in Konstantinou et al. (2014) and thus the rest of our proof is analogous to that. \square

Accordingly, for D-optimality we have $c = 2$ and $\mathbf{A} = \mathbf{M}^{-1}(\xi^*, \boldsymbol{\beta})$. For A-optimality we have $c = \text{tr}(\mathbf{M}^{-1}(\xi^*, \boldsymbol{\beta})) = (\sqrt{(a^2 + 1)/u_b} + \sqrt{(b^2 + 1)/u_a})/(b - a)^2$ where $u_a = u(a, \boldsymbol{\beta})$ and $u_b = u(b, \boldsymbol{\beta})$ with $\mathbf{A} = \mathbf{M}^{-2}(\xi^*, \boldsymbol{\beta})$. In general, under Kiefer's Φ_k -criteria we denote $c = \text{tr}(\mathbf{M}^{-k}(\xi^*, \boldsymbol{\beta}))$ and $\mathbf{A} = \mathbf{M}^{-k-1}(\xi^*, \boldsymbol{\beta})$. Moreover, the

Generalized D-criterion and L-criterion can be applied (Atkinson and Woods 2015, Chapter 10).

As a consequence of Lemma 4.1, we next provide sufficient conditions for a design supported by the boundary points 0 and 1 to be locally D- or A-optimal on $\mathcal{X} = [0, 1]$ at a given β . Let $q(x) = 1/u(x, \beta)$ and denote $q_0 = q^{\frac{1}{2}}(0)$ and $q_1 = q^{\frac{1}{2}}(1)$.

Theorem 4.1 Consider a GLM with $f(x) = (1, x)^T$ and the experimental region $\mathcal{X} = [0, 1]$. Let a parameter point $\beta = (\beta_0, \beta)^T$ be given. Let $q(x)$ be positive and twice continuously differentiable. Then:

- (i) The unique locally D-optimal design ξ^* (at β) is the two-point design supported by 0 and 1 with equal weights $1/2$ if

$$q_0^2 + q_1^2 > q''(x)/2 \text{ for all } x \in (0, 1). \quad (4.1)$$

- (ii) The unique locally A-optimal design ξ^* (at β) is the two-point design supported by 0 and 1 with weights

$$\omega_0^* = \frac{\sqrt{2}q_0}{\sqrt{2}q_0 + q_1} \text{ and } \omega_1^* = \frac{q_1}{\sqrt{2}q_0 + q_1}$$

if

$$q_0^2 + q_1^2 + \sqrt{2}q_0q_1 > q''(x)/2 \text{ for all } x \in (0, 1). \quad (4.2)$$

Proof Part (i): Condition (2.6) of the general equivalence theorem for $k = 0$ implies that ξ^* is locally D-optimal if and only if

$$(1-x)^2q_0^2 + x^2q_1^2 - q(x) \leq 0 \quad \forall x \in [0, 1]. \quad (4.3)$$

Since the support points are $\{0, 1\}$, the l.h.s. of the above inequality equals zero at the boundaries of $[0, 1]$. Then it is sufficient to show that the aforementioned l.h.s. is convex on the interior $(0, 1)$ and this convexity realizes under condition (4.1) asserted in the theorem. Now to show that ξ^* is unique at β assume that ξ^{**} is locally D-optimal at β . Then $M(\xi^*, \beta) = M(\xi^{**}, \beta)$ and therefore, the condition of the equivalence theorem under ξ^{**} is equivalent to (4.3) and this is an equation only at the support of ξ^* , i.e., 0 and 1.

Part (ii): This case can be shown in analogy to Part (i) by employing condition (2.6) of the general equivalence theorem for $k = 1$ with $\text{tr}(M^{-1}(\xi^*, \beta)) = (\sqrt{2}q_0 + q_1)^2$. The optimal weights ω_0^* and ω_1^* are derived according to (3.1) in Sect. 3. \square

5 Two-factor model

In this section we consider a first order model of two factors

$$f(x) = (1, x_1, x_2)^T \text{ where } x = (x_1, x_2)^T \in \mathcal{X}. \quad (5.1)$$

5.1 Continuous factors

Let the experimental region be given by the unit rectangle $\mathcal{X} = [0, 1]^2$. Denote the vertices of \mathcal{X} by $\mathbf{x}_1^* = (0, 0)^T$, $\mathbf{x}_2^* = (1, 0)^T$, $\mathbf{x}_3^* = (0, 1)^T$ and $\mathbf{x}_4^* = (1, 1)^T$. In the following we provide necessary and sufficient conditions for the designs that are supported by the vertices \mathbf{x}_1^* , \mathbf{x}_2^* , \mathbf{x}_3^* , \mathbf{x}_4^* to be locally D- and A-optimal.

Theorem 5.1 Consider a GLM with $f(\mathbf{x}) = (1, x_1, x_2)^T$ and the experimental region $\mathcal{X} = [0, 1]^2$. For a given parameter point $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$ let $u_i = u(\mathbf{x}_i^*, \boldsymbol{\beta})$ ($1 \leq i \leq 4$). Then:

(o) The locally D-optimal design ξ^* (at $\boldsymbol{\beta}$) is unique.

$$(1) \quad \xi^* = \begin{pmatrix} \mathbf{x}_1^* & \mathbf{x}_2^* & \mathbf{x}_3^* \\ 1/3 & 1/3 & 1/3 \end{pmatrix} \text{ if and only if}$$

$$(1 - x_1 - x_2)^2 u_1^{-1} + x_1^2 u_2^{-1} + x_2^2 u_3^{-1} \leq u^{-1}(\mathbf{x}, \boldsymbol{\beta}) \quad \forall \mathbf{x} \in [0, 1]^2.$$

$$(2) \quad \xi^* = \begin{pmatrix} \mathbf{x}_1^* & \mathbf{x}_2^* & \mathbf{x}_4^* \\ 1/3 & 1/3 & 1/3 \end{pmatrix} \text{ if and only if}$$

$$(1 - x_1)^2 u_1^{-1} + (x_1 - x_2)^2 u_2^{-1} + x_2^2 u_4^{-1} \leq u^{-1}(\mathbf{x}, \boldsymbol{\beta}) \quad \forall \mathbf{x} \in [0, 1]^2.$$

$$(3) \quad \xi^* = \begin{pmatrix} \mathbf{x}_1^* & \mathbf{x}_3^* & \mathbf{x}_4^* \\ 1/3 & 1/3 & 1/3 \end{pmatrix} \text{ if and only if}$$

$$(1 - x_2)^2 u_1^{-1} + (x_2 - x_1)^2 u_3^{-1} + x_1^2 u_4^{-1} \leq u^{-1}(\mathbf{x}, \boldsymbol{\beta}) \quad \forall \mathbf{x} \in [0, 1]^2.$$

$$(4) \quad \xi^* = \begin{pmatrix} \mathbf{x}_2^* & \mathbf{x}_3^* & \mathbf{x}_4^* \\ 1/3 & 1/3 & 1/3 \end{pmatrix} \text{ if and only if}$$

$$(1 - x_2)^2 u_2^{-1} + (1 - x_1)^2 u_3^{-1} + (x_1 + x_2 - 1)^2 u_4^{-1} \leq u^{-1}(\mathbf{x}, \boldsymbol{\beta}) \quad \forall \mathbf{x} \in [0, 1]^2.$$

(5) Otherwise, ξ^* is supported by the four design points \mathbf{x}_1^* , \mathbf{x}_2^* , \mathbf{x}_3^* , \mathbf{x}_4^* .

Proof The proof of cases (1)–(4) is demonstrated by making use of condition (2.6) for $k = 0$ of the general equivalence theorem. For case (ℓ) ($1 \leq \ell \leq 4$) denote the design matrix $\mathbf{F} = [f(\mathbf{x}_i^*), f(\mathbf{x}_j^*), f(\mathbf{x}_k^*)]^T$ and the weight matrix $\mathbf{U} = \text{diag}(u_i, u_j, u_k)$ such that $1 \leq i < j < k \leq 4$ and $i, j, k \neq 4 - \ell + 1$. We will show that the condition in each case (1)–(4) is equivalent to

$$(1, x_1, x_2) \mathbf{F}^{-1} \mathbf{U}^{-1} (\mathbf{F}^T)^{-1} (1, x_1, x_2)^T \leq u^{-1}(\mathbf{x}, \boldsymbol{\beta}) \quad \forall \mathbf{x} = (x_1, x_2)^T \in [0, 1]^2.$$

To this end, for each case (1) – (4), we report the matrices F , F^{-1} and U

$$\begin{aligned}
 (1) : \quad F &= \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \quad F^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \quad U = \text{diag}(u_1, u_2, u_3). \\
 (2) : \quad F &= \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad F^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}, \quad U = \text{diag}(u_1, u_2, u_4). \\
 (3) : \quad F &= \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad F^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ -1 & 1 & 0 \end{pmatrix}, \quad U = \text{diag}(u_1, u_3, u_4). \\
 (4) : \quad F &= \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad F^{-1} = \begin{pmatrix} 1 & 1 & -1 \\ 0 & -1 & 1 \\ -1 & 0 & 1 \end{pmatrix}, \quad U = \text{diag}(u_2, u_3, u_4).
 \end{aligned}$$

It remains to show that the design ξ^* is unique at β . Suppose that ξ^* and ξ^{**} are locally D-optimal at β . Then by the strict convexity of the D-criterion we have $M(\xi^*, \beta) = M(\xi^{**}, \beta)$. Thus $M(\xi^*, \beta) - M(\xi^{**}, \beta) = \sum_{i=1}^4 (\omega_i^* - \omega_i^{**}) u_i f(x_i^*) f^T(x_i^*) = 0$. The intensities u_i ($1 \leq i \leq 4$) are positive and $f(x_i^*) f^T(x_i^*)$ ($1 \leq i \leq 4$) are linearly independent. It follows that $\omega_i^* - \omega_i^{**} = 0$ ($1 \leq i \leq 4$). \square

In analogy to Theorem 5.1 we introduce locally A-optimal designs in the next theorem.

Theorem 5.2 Consider the assumptions and notations of Theorem 5.1. Denote $q_i = u_i^{-1/2}$ ($1 \leq i \leq 4$). Then:

(o) The locally A-optimal design ξ^* (at β) is unique.

$$(1) \quad \xi^* = \begin{pmatrix} \mathbf{x}_1^* & \mathbf{x}_2^* & \mathbf{x}_3^* \\ \sqrt{3}q_1/c & q_2/c & q_3/c \end{pmatrix} \text{ if and only if}$$

$$\begin{aligned}
 & (1 - x_1 - x_2)^2 q_1^2 + x_1^2 q_2^2 + x_2^2 q_3^2 - \frac{2}{\sqrt{3}} (1 - x_1 - x_2) (x_1 q_2 + x_2 q_3) q_1 \\
 & \leq u^{-1}(\mathbf{x}, \beta) \quad \forall \mathbf{x} \in [0, 1]^2.
 \end{aligned}$$

$$(2) \quad \xi^* = \begin{pmatrix} \mathbf{x}_1^* & \mathbf{x}_2^* & \mathbf{x}_4^* \\ \sqrt{2}q_1/c & \sqrt{2}q_2/c & q_4/c \end{pmatrix} \text{ if and only if}$$

$$\begin{aligned}
 & (1 - x_1)^2 q_1^2 + (x_1 - x_2)^2 q_2^2 + x_2^2 q_4^2 - (x_1 - x_2) ((1 - x_1) q_1 + \sqrt{2} x_2 q_4) q_2 \\
 & \leq u^{-1}(\mathbf{x}, \beta) \quad \forall \mathbf{x} \in [0, 1]^2.
 \end{aligned}$$

$$(3) \xi^* = \begin{pmatrix} \mathbf{x}_1^* & \mathbf{x}_3^* & \mathbf{x}_4^* \\ \sqrt{2}q_1/c & \sqrt{2}q_3/c & q_4/c \end{pmatrix} \text{ if and only if}$$

$$(1 - x_2)^2 q_1^2 + (x_2 - x_1)^2 q_3^2 + x_1^2 q_4^2 - (x_2 - x_1)((1 - x_2)q_1 + \sqrt{2}x_1 q_4)q_3 \leq u^{-1}(\mathbf{x}, \boldsymbol{\beta}) \quad \forall \mathbf{x} \in [0, 1]^2.$$

$$(4) \xi^* = \begin{pmatrix} \mathbf{x}_2^* & \mathbf{x}_3^* & \mathbf{x}_4^* \\ \sqrt{2}q_2/c & \sqrt{2}q_3/c & \sqrt{3}q_4/c \end{pmatrix} \text{ if and only if}$$

$$(1 - x_2)^2 q_2^2 + (1 - x_1)^2 q_3^2 + (x_2 + x_1 - 1)^2 q_4^2 + (1 - x_1)(1 - x_2)q_2 q_3 - 2\sqrt{\frac{2}{3}}(x_1 + x_2 - 1)((1 - x_2)q_2 - (1 - x_1)q_3)q_4 \leq u^{-1}(\mathbf{x}, \boldsymbol{\beta}) \quad \forall \mathbf{x} \in [0, 1]^2.$$

For each case (1)–(4), the constant c appearing in the weights equals the sum of the numerators of the three ratios.

(5) Otherwise, ξ^* is supported by the four design points $\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*, \mathbf{x}_4^*$.

Proof We make use of condition (2.6) for $k = 1$ of the general equivalence theorem. In analogy to the proof of Theorem 5.1 for case (ℓ) ($1 \leq \ell \leq 4$) denote $\mathbf{F} = [\mathbf{f}(\mathbf{x}_i^*), \mathbf{f}(\mathbf{x}_j^*), \mathbf{f}(\mathbf{x}_k^*)]^T$, $\mathbf{U} = \text{diag}(u_i, u_j, u_k)$ and $\boldsymbol{\Omega} = \text{diag}(\omega_i^*, \omega_j^*, \omega_k^*)$ such that $1 \leq i < j < k \leq 4$ and $i, j, k \neq 4 - \ell + 1$. Then we obtain $\mathbf{C} = (\mathbf{F}^{-1})^T \mathbf{F}^{-1}$. An elementary calculation shows that the weights given by (3.1) for an A-optimal design coincide with the ω_i^* ($1 \leq i \leq 3$) as stated in the theorem. Now we show that the design ξ^* is locally A-optimal if and only if the corresponding condition holds. We have

$$\begin{aligned} \mathbf{M}(\xi^*, \boldsymbol{\beta}) &= \mathbf{F}^T \mathbf{V} \mathbf{F} = \mathbf{F}^T \boldsymbol{\Omega} \mathbf{U} \mathbf{F}, \\ \text{tr}(\mathbf{M}^{-1}(\xi^*, \boldsymbol{\beta})) &= \text{tr}(\mathbf{F}^{-1} \mathbf{U}^{-1} \boldsymbol{\Omega}^{-1} (\mathbf{F}^{-1})^T) \\ &= c \left(\left(\frac{c_{11}}{u_1} \right)^{1/2} + \left(\frac{c_{22}}{u_2} \right)^{1/2} + \left(\frac{c_{33}}{u_3} \right)^{1/2} \right) = c^2. \end{aligned}$$

Since $\mathbf{U}^{-1/2} \boldsymbol{\Omega}^{-1} = c \text{diag}(c_{11}^{-1/2}, c_{22}^{-1/2}, c_{33}^{-1/2})$, we obtain

$$\begin{aligned} \mathbf{M}^{-2}(\xi^*, \boldsymbol{\beta}) &= \mathbf{F}^{-1} \mathbf{U}^{-1} \boldsymbol{\Omega}^{-1} (\mathbf{F}^{-1})^T \mathbf{F}^{-1} \mathbf{U}^{-1} \boldsymbol{\Omega}^{-1} (\mathbf{F}^{-1})^T \\ &= c^2 \mathbf{F}^{-1} \mathbf{U}^{-1/2} \mathbf{C}^* \mathbf{U}^{-1/2} (\mathbf{F}^{-1})^T \end{aligned}$$

where $\mathbf{C}^* = \text{diag}(c_{11}^{-1/2}, c_{22}^{-1/2}, c_{33}^{-1/2}) \mathbf{C} \text{diag}(c_{11}^{-1/2}, c_{22}^{-1/2}, c_{33}^{-1/2})$. So, together with condition (2.6) of the general equivalence theorem for $k = 1$ the design ξ^* is locally A-optimal (at $\boldsymbol{\beta}$) if and only if

$$\begin{aligned} & (1, x_1, x_2) \mathbf{F}^{-1} \mathbf{U}^{-1/2} \mathbf{C}^* \mathbf{U}^{-1/2} (\mathbf{F}^{-1})^T (1, x_1, x_2)^T \\ & \leq u^{-1}(\mathbf{x}, \boldsymbol{\beta}) \quad \forall \mathbf{x} \in [0, 1]^2 \end{aligned} \quad (5.2)$$

Straightforward calculation shows that condition (5.2) is equivalent to the respective condition in Case (ℓ). \square

Remark Yang et al. (2011) developed a method to find locally optimal designs for logistic models of multiple factors. It was assumed that one factor is defined on the whole real line while the other factors belong to a compact region which seems in conflict with the experimental region given in Theorem 5.1. Then a subclass of designs was established by Loewner semi ordering of nonnegative definite matrices and so, one could focus on this subclass to derive optimal designs. A similar strategy was used in Gaffke et al. (2019) for gamma models on the experimental region $[0, 1]^v$, $v \geq 1$. Nevertheless, it seems that this strategy may not work for a general setup of the generalized linear model. However, consider a logistic model of two factors with $\mathbf{f}(\mathbf{x}) = (1, x_1, x_2)^T$ and intensity function $u(\mathbf{x}, \boldsymbol{\beta}) = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2) / (1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2))^2$. According to Yang et al. (2011) the experimental region is assumed to be $\mathcal{X} = [0, 1] \times \mathbb{R}$, i.e., $x_2 \in (-\infty, \infty)$. From Yang et al. (2011), Corollary 1, a locally D-optimal design is given by

$$\xi^* = \left(\begin{array}{cccc} (0, \frac{c^* - \beta_0}{\beta_2})^T & (0, \frac{-c^* - \beta_0}{\beta_2})^T & (1, \frac{c^* - \beta_0 - \beta_1}{\beta_2})^T & (1, \frac{-c^* - \beta_0 - \beta_1}{\beta_2})^T \\ 1/4 & 1/4 & 1/4 & 1/4 \end{array} \right)$$

where c^* is the maximizer of $c^2 (\exp(c)/(1 + \exp(c)))^3$. In general, ξ^* is not covered by Theorem 5.1. In contrast to that, for a particular parameter point $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$ such that $\beta_1 = 0$, $\beta_2 = -2\beta_0$ and $\beta_0 = c^*$ the design ξ^* is supported by the vertices of $[0, 1]^2$.

5.2 Discrete factors

Here, we assume two factors each at two levels, i.e., 0 and 1. The experimental region is given by $\tilde{\mathcal{X}} = \{0, 1\}^2$ which consists of the vertices of the unit rectangle $[0, 1]^2$. So we write $\tilde{\mathcal{X}} = \{\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*, \mathbf{x}_4^*\}$.

Corollary 5.1 Consider a GLM with $\mathbf{f}(\mathbf{x}) = (1, x_1, x_2)^T$ and the experimental region $\tilde{\mathcal{X}} = \{0, 1\}^2$. For a given parameter point $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$ let $u_i = u(\mathbf{x}_i^*, \boldsymbol{\beta})$ ($1 \leq i \leq 4$). Denote by $u_{(1)} \leq u_{(2)} \leq u_{(3)} \leq u_{(4)}$ the intensity values u_1, u_2, u_3, u_4 rearranged in ascending order. Then:

- (i) The design ξ^* is supported by the three design points whose intensity values are given by $u_{(2)}, u_{(3)}, u_{(4)}$, with equal weights $1/3$ if and only if

$$u_{(2)}^{-1} + u_{(3)}^{-1} + u_{(4)}^{-1} \leq u_{(1)}^{-1}.$$

(ii) The design ξ^* is supported by the four design points $\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*, \mathbf{x}_4^*$ with weights $\omega_1^*, \omega_2^*, \omega_3^*, \omega_4^*$ which are uniquely determined by the condition

$$\omega_i^* > 0 \ (1 \leq i \leq 4), \ \sum_{i=1}^4 \omega_i^* = 1, \ \text{and} \ u_i \omega_i^* \left(\frac{1}{3} - \omega_i^*\right) \ (1 \leq i \leq 4) \ \text{are equal} \quad (5.3)$$

if and only if $u_{(2)}^{-1} + u_{(3)}^{-1} + u_{(4)}^{-1} > u_{(1)}^{-1}$.

Proof The proof is demonstrated by Theorem 5.1. The condition of ξ^* in part (i) comes by the corresponding inequality in cases (1)–(4) of Theorem 5.1 which arises at the point that is not a support of the respective ξ^* . The equal values of the identity $u_i \omega_i^* (\frac{1}{3} - \omega_i^*)$ for $i = 1, 2, 3, 4$ were proved in Gaffke et al. (2019). \square

Remark In part (i) of Corollary 5.1 the design points with highest intensities perform as a support of a locally D-optimal design.

Theorem 5.3 Under the assumptions of Corollary 5.1 let the parameter point $\beta = (\beta_0, \beta_1, \beta_2)^T$ be given with $\beta_1 = \beta_2 = \beta$ which fulfills assumption (ii) of Corollary 5.1. Then the locally D-optimal design (at β) is supported by the four design points $\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*, \mathbf{x}_4^*$ with positive weights

$$\begin{aligned} \omega_1^* &= \frac{3}{8} + \frac{1}{4} \left(1 + \frac{u_1}{u_4} - 4 \frac{u_1}{u_2} \right)^{-1}, \\ \omega_2^* &= \omega_3^* = \frac{1}{2} \left(4 - \frac{u_2}{u_1} - \frac{u_2}{u_4} \right)^{-1}, \\ \omega_4^* &= \frac{3}{8} + \frac{1}{4} \left(1 + \frac{u_4}{u_1} - 4 \frac{u_4}{u_2} \right)^{-1}. \end{aligned}$$

Proof Since assumption (ii) of Corollary 5.1 is fulfilled by a point β the design is supported by all points $\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*, \mathbf{x}_4^*$. Then the optimal weights are obtained according to Lemma 3.1 where we have $d_i^2 = 1$ ($1 \leq i \leq 4$) and $u_2 = u_3$. Hence, the results follow. \square

Now we restrict to A-optimal designs on the set of vertices $\tilde{\mathcal{X}} = \{0, 1\}^2$. It can also be noted that the design points with highest intensities perform as a support of a locally A-optimal design at a given parameter value.

Corollary 5.2 Consider the assumptions and notations of Corollary 5.1. Denote $q_i = u_i^{-1/2}$ ($1 \leq i \leq 4$). Then the unique locally A-optimal design ξ^* is as follows.

$$(1) \ \xi^* = \left(\begin{array}{ccc} \mathbf{x}_1^* & \mathbf{x}_2^* & \mathbf{x}_3^* \\ \sqrt{3}q_1/c & q_2/c & q_3/c \end{array} \right) \text{ if and only if}$$

$$q_1^2 + q_2^2 + q_3^2 + \frac{2}{\sqrt{3}}q_1(q_2 + q_3) \leq q_4^2.$$

$$(2) \xi^* = \begin{pmatrix} \mathbf{x}_1^* & \mathbf{x}_2^* & \mathbf{x}_4^* \\ \sqrt{2}q_1/c & \sqrt{2}q_2/c & q_4/c \end{pmatrix} \text{ if and only if}$$

$$q_1^2 + q_2^2 + q_4^2 + q_1q_2 + \sqrt{2}q_2q_4 \leq q_3^2.$$

$$(3) \xi^* = \begin{pmatrix} \mathbf{x}_1^* & \mathbf{x}_3^* & \mathbf{x}_4^* \\ \sqrt{2}q_1/c & \sqrt{2}q_3/c & q_4/c \end{pmatrix} \text{ if and only if}$$

$$q_1^2 + q_3^2 + q_4^2 + q_1q_3 + \sqrt{2}q_3q_4 \leq q_2^2.$$

$$(4) \xi^* = \begin{pmatrix} \mathbf{x}_2^* & \mathbf{x}_3^* & \mathbf{x}_4^* \\ \sqrt{2}q_2/c & \sqrt{2}q_3/c & \sqrt{3}q_4/c \end{pmatrix} \text{ if and only if}$$

$$q_2^2 + q_3^2 + q_4^2 + q_2q_3 + 2\sqrt{\frac{2}{3}}q_4(q_2 + q_3) \leq q_1^2.$$

For each case (i) – (iv), the constant c appearing in the weights equals the sum of the numerators of the three ratios.

(5) Otherwise, ξ^* is supported by the four design points $\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*, \mathbf{x}_4^*$.

As the optimal weights of the A-optimal designs depend on the model parameters each condition provided in the theorem characterizes a subregion of the parameter space where the corresponding designs with the same support are A-optimal.

6 Multiple regression model

6.1 Model with intercept

Consider a first order model of multiple factors

$$f(\mathbf{x}) = (1, \mathbf{x}^T)^T \text{ where } \mathbf{x} = (x_1, \dots, x_v)^T \in \mathcal{X} = [0, 1]^v. \quad (6.1)$$

Here, we are interested in providing an extension of locally D- and A-optimal designs with support $(0, 0)^T$, $(1, 0)^T$, $(0, 1)^T$ that are given in part (1) of Theorems 5.1 and 5.2.

Theorem 6.1 Consider model (6.1) with experimental region $\mathcal{X} = [0, 1]^v$, where $v \geq 2$. Define particular design points by

$$\mathbf{x}_1^* = (0, \dots, 0)^T, \mathbf{x}_2^* = (1, \dots, 0)^T, \dots, \mathbf{x}_{v+1}^* = (0, \dots, 1)^T.$$

For a given parameter point $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_v)^T$ let $u_i = u(\mathbf{x}_i^*, \boldsymbol{\beta})$ ($1 \leq i \leq v+1$). Then the design ξ^* which assigns equal weights $1/(v+1)$ to the design points \mathbf{x}_i^* for

all $(1 \leq i \leq v+1)$ is locally D -optimal (at β) if and only if

$$u_1^{-1} \left(1 - \sum_{j=1}^v x_j \right)^2 + \sum_{i=1}^v u_{i+1}^{-1} x_i^2 \leq u^{-1}(\mathbf{x}, \beta) \text{ for all } \mathbf{x} \in [0, 1]^v. \quad (6.2)$$

Proof Define the $(v+1) \times (v+1)$ design matrix $\mathbf{F} = [\mathbf{f}(\mathbf{x}_1^*), \dots, \mathbf{f}(\mathbf{x}_{v+1}^*)]^T$, then

$$\mathbf{M}(\xi^*, \beta) = \frac{1}{v+1} \mathbf{F}^T \mathbf{U} \mathbf{F}, \text{ where } \mathbf{U} = \text{diag}(u_i)_{i=1}^{v+1}.$$

We have

$$\mathbf{F} = \begin{bmatrix} 1 & \mathbf{0}_{1 \times v} \\ \mathbf{1}_{v \times 1} & \mathbf{I}_v \end{bmatrix}, \text{ hence } \mathbf{F}^{-1} = \begin{bmatrix} 1 & \mathbf{0}_{1 \times v} \\ -\mathbf{1}_{v \times 1} & \mathbf{I}_v \end{bmatrix}, \quad (6.3)$$

where $\mathbf{0}_{1 \times v}$, $\mathbf{1}_{v \times 1}$, and \mathbf{I}_v denote the v -dimensional row vector of zeros, the v -dimensional column vector of ones, and the $v \times v$ unit matrix, respectively. So, by condition (2.6) of the general equivalence theorem for $k = 0$ the design is locally D -optimal if and only if

$$u(\mathbf{x}, \beta) \mathbf{f}^T(\mathbf{x}) \mathbf{M}^{-1}(\xi^*, \beta) \mathbf{f}(\mathbf{x}) \leq v+1 \quad \forall \mathbf{x} \in [0, 1]^v. \quad (6.4)$$

The l.h.s. of (6.4) reads as

$$\begin{aligned} & u(\mathbf{x}, \beta) (v+1) \mathbf{f}^T(\mathbf{x}) \mathbf{F}^{-1} \mathbf{U}^{-1} (\mathbf{F}^{-1})^T \mathbf{f}(\mathbf{x}) \\ &= (v+1) u(\mathbf{x}, \beta) \left(u_1^{-1} \left(1 - \sum_{j=1}^v x_j \right)^2 + \sum_{i=1}^v u_{i+1}^{-1} x_i^2 \right), \end{aligned}$$

and hence it is obvious that (6.4) is equivalent to (6.2). \square

Remark The D -optimal design under a two-factor model with support $(0, 0)^T$, $(1, 0)^T$, $(0, 1)^T$ from Theorem 5.1, part (1) is covered by Theorem 6.1 for $v = 2$. It is clear that condition (6.2) for $v = 2$ is equivalent to the inequality $(1 - x_1 - x_2)^2 u_1^{-1} + x_1^2 u_2^{-1} + x_2^2 u_3^{-1} \leq u^{-1}(\mathbf{x}, \beta) \quad \forall \mathbf{x} \in [0, 1]^2$.

In analogy to Theorem 6.1 we present locally A -optimal designs in the next theorem.

Theorem 6.2 Consider the assumptions and notations of Theorem 6.1. Denote $q_i = u_i^{-1/2}$ ($1 \leq i \leq v+1$). Then the design ξ^* which is supported by \mathbf{x}_i^* ($1 \leq i \leq v+1$) with weights

$$\omega_1^* = \sqrt{v+1} q_1 / c \text{ and } \omega_{i+1}^* = q_{i+1} / c, \quad i = 1, \dots, v \text{ where } c = \sqrt{v+1} q_1 + \sum_{j=2}^v q_j$$

is locally A -optimal (at β) if and only if for all $\mathbf{x} = (x_1, \dots, x_v)^T \in [0, 1]^v$

$$\begin{aligned} q_1^2 \left(1 - \sum_{j=1}^v x_j\right)^2 + \sum_{i=1}^v q_{i+1}^2 x_i^2 + \frac{2q_1}{\sqrt{v+1}} \left(\sum_{j=1}^v x_j - 1\right) \sum_{i=1}^v q_{i+1} x_i \\ \leq u^{-1}(\mathbf{x}, \beta). \end{aligned} \quad (6.5)$$

Proof As in the proof of Theorem 6.1 the design matrix \mathbf{F} and its inverse are given by (6.3) and we obtain

$$\mathbf{C} = (\mathbf{F}^{-1})^T \mathbf{F}^{-1} = \begin{bmatrix} v+1 & -\mathbf{1}_{1 \times v} \\ -\mathbf{1}_{v \times 1} & \mathbf{I}_v \end{bmatrix}.$$

This yields $\sqrt{c_{11}/u_1} = \sqrt{v+1}q_1$ and $\sqrt{c_{ii}/u_i} = q_i$ for $i = 2, \dots, v+1$ according to (3.1) in Sect. 3 with $p = v+1$. An elementary calculation shows that the weights given by (3.1) for an A -optimal design coincide with the ω_i^* ($1 \leq i \leq p$) as stated in the theorem. Now we show that the design ξ^* is locally A -optimal if and only if (6.5) holds. Let $\mathbf{U} = \text{diag}(u_1, \dots, u_p)$, $\mathbf{\Omega} = \text{diag}(\omega_1^*, \dots, \omega_p^*)$ and $\mathbf{V} = \mathbf{\Omega}\mathbf{U}$. Then we have

$$\begin{aligned} \mathbf{M}(\xi^*, \beta) &= \mathbf{F}^T \mathbf{V} \mathbf{F} = \mathbf{F}^T \mathbf{\Omega} \mathbf{U} \mathbf{F}, \\ \text{tr}(\mathbf{M}^{-1}(\xi^*, \beta)) &= \text{tr}(\mathbf{F}^{-1} \mathbf{U}^{-1} \mathbf{\Omega}^{-1} (\mathbf{F}^{-1})^T) = c \sum_{i=1}^p \left(\frac{c_{ii}}{u_i}\right)^{1/2} = c^2. \end{aligned}$$

Since $\mathbf{U}^{-1/2} \mathbf{\Omega}^{-1} = c \text{diag}(c_{11}^{-1/2}, \dots, c_{pp}^{-1/2})$, we obtain

$$\begin{aligned} \mathbf{M}^{-2}(\xi^*, \beta) &= \mathbf{F}^{-1} \mathbf{U}^{-1} \mathbf{\Omega}^{-1} (\mathbf{F}^{-1})^T \mathbf{F}^{-1} \mathbf{U}^{-1} \mathbf{\Omega}^{-1} (\mathbf{F}^{-1})^T \\ &= c^2 \mathbf{F}^{-1} \mathbf{U}^{-1/2} \mathbf{C}^* \mathbf{U}^{-1/2} (\mathbf{F}^{-1})^T \end{aligned}$$

where $\mathbf{C}^* = \text{diag}(c_{11}^{-1/2}, \dots, c_{pp}^{-1/2}) \mathbf{C} \text{diag}(c_{11}^{-1/2}, \dots, c_{pp}^{-1/2})$. So, together with condition (2.6) of the general equivalence theorem for $k = 1$ the design ξ^* is locally A -optimal (at β) if and only if

$$\left(\mathbf{U}^{-1/2} (\mathbf{F}^{-1})^T \mathbf{f}(\mathbf{x})\right)^T \mathbf{C}^* \left(\mathbf{U}^{-1/2} (\mathbf{F}^{-1})^T \mathbf{f}(\mathbf{x})\right) \leq u^{-1}(\mathbf{x}, \beta) \quad \forall \mathbf{x} \in [0, 1]^v \quad (6.6)$$

Straightforward calculation shows that condition (6.6) is equivalent to condition (6.5). \square

Remark Theorem 6.2 with $v = 2$ covers the result stated in part (1) of Theorem 5.2. It can be seen that with the notations of Theorem 5.2, the inequality $q_1^2 + q_2^2 + q_3^2 + \frac{2}{\sqrt{3}}q_1q_2 + \frac{2}{\sqrt{3}}q_1q_3 \leq q_4^2$ is equivalent to condition (6.5) of Theorem 6.2 for $v = 2$.

6.2 Model without intercept

Consider a model of multiple factors and without intercept. We assume a first order model

$$\mathbf{f}(\mathbf{x}) = (x_1, \dots, x_v)^T \quad \text{where } \mathbf{x} \in \mathcal{X}. \quad (6.7)$$

The experimental region \mathcal{X} has an arbitrary form. Locally optimal designs will be derived under Kiefer's Φ_k -criteria. The support points are located at the boundary of \mathcal{X} and the optimal weights are obtained according to Lemma 3.2.

Theorem 6.3 Consider model (6.7) on an experimental region \mathcal{X} . Let a vector $\mathbf{a} = (a_1, \dots, a_v)^T$ be given such that $a_i > 0$ ($1 \leq i \leq v$). Denote the design points by $\mathbf{x}_i^* = a_i \mathbf{e}_i$ ($1 \leq i \leq v$) that are assumed to belong to \mathcal{X} . For a given parameter point $\boldsymbol{\beta}$ let $u_i = u(\mathbf{x}_i^*, \boldsymbol{\beta})$ ($1 \leq i \leq v$). Let k with $0 \leq k < \infty$ be given. Let ξ_a^* be the saturated design whose support consists of the points \mathbf{x}_i^* ($1 \leq i \leq v$) with the corresponding weights

$$\omega_i^* = \frac{(a_i^2 u_i)^{\frac{-k}{k+1}}}{\sum_{j=1}^v (a_j^2 u_j)^{\frac{-k}{k+1}}} \quad (1 \leq i \leq v).$$

Then ξ_a^* is locally Φ_k -optimal (at $\boldsymbol{\beta}$) if and only if

$$u(\mathbf{x}, \boldsymbol{\beta}) \sum_{i=1}^v u_i^{-1} a_i^{-2} x_i^2 \leq 1 \quad \text{for all } \mathbf{x} = (x_1, \dots, x_v)^T \in \mathcal{X}. \quad (6.8)$$

Proof Define the $v \times v$ design matrix $\mathbf{F} = \text{diag}(a_i)_{i=1}^v$ with the $v \times v$ weight matrix

$$\mathbf{V} = \text{diag}(u_i \omega_i^*)_{i=1}^v = \left(\sum_{j=1}^v (a_j^2 u_j)^{\frac{-k}{k+1}} \right)^{-1} \text{diag} \left((a_i^{-2k} u_i)^{\frac{1}{k+1}} \right)_{i=1}^v.$$

Then we have

$$\begin{aligned} \mathbf{M}(\xi_a^*, \boldsymbol{\beta}) &= \mathbf{F}^T \mathbf{V} \mathbf{F} = \left(\sum_{j=1}^v (a_j^2 u_j)^{\frac{-k}{k+1}} \right)^{-1} \text{diag} \left((a_i^2 u_i)^{\frac{1}{k+1}} \right)_{i=1}^v, \\ \mathbf{M}^{-k-1}(\xi_a^*, \boldsymbol{\beta}) &= \left(\sum_{j=1}^v (a_j^2 u_j)^{\frac{-k}{k+1}} \right)^{k+1} \text{diag} \left(a_i^{-2} u_i^{-1} \right)_{i=1}^v, \text{ and} \\ \text{tr}(\mathbf{M}^{-k}(\xi_a^*, \boldsymbol{\beta})) &= \left(\sum_{j=1}^v (a_j^2 u_j)^{\frac{-k}{k+1}} \right)^{k+1}. \end{aligned}$$

Adopting these formulas simplifies the l.h.s. of condition (2.6) of the general equivalence theorem to $u(\mathbf{x}, \boldsymbol{\beta}) \left(\sum_{j=1}^v (a_j^2 u_j)^{\frac{-k}{k+1}} \right)^{k+1} \sum_{i=1}^v u_i^{-1} a_i^{-2} x_i^2$ which is bounded by $\left(\sum_{j=1}^v (a_j^2 u_j)^{\frac{-k}{k+1}} \right)^{k+1}$ if and only if condition (6.8) holds true. \square

The optimality condition (6.8) does not depend on the value of k . However, from Theorem 6.3 the locally D-optimal design ($k = 0$) has weights $\omega_i^* = 1/v$ ($1 \leq i \leq v$) and the locally A-optimal design ($k = 1$) has weights $\omega_i^* = (a_i^2 u_i)^{-1/2} / \sum_{j=1}^v (a_j^2 u_j)^{-1/2}$ ($1 \leq i \leq v$).

7 Applications

In this section, we give a discussion on the application of the previous results for the generalized linear models. Here, emphasis will be laid on gamma and Poisson models. However, it is known that the linear regression model is a GLM. Therefore, to begin with, we briefly focus on the Φ_k -optimality under a non-intercept linear model with $\mathbf{f}(\mathbf{x}) = (x_1, \dots, x_v)^T$ on the continuous experimental region $\mathcal{X} = [0, 1]^v$, $v \geq 2$. Here, $u(\mathbf{x}, \boldsymbol{\beta}) = 1$ for all $\mathbf{x} \in \mathcal{X}$ so the information matrices in a linear model are independent of $\boldsymbol{\beta}$. Note that Theorem 6.3 does not cover a non-intercept linear model on \mathcal{X} since condition (6.8) does not hold true for $v \geq 2$. However, the l.h.s. of condition (2.6) of the general equivalence theorem under linear models, i.e., when $u(\mathbf{x}, \boldsymbol{\beta}) = 1$, is strictly convex and it attains its maximum at some vertices of \mathcal{X} . Thus the support of any Φ_k (or D, A)-optimal design is a subset of $\{0, 1\}^v$. As a result, in particular for D- and A-optimality, one might apply the results of Theorem 3.1 in Huda and Mukerjee (1988), which were obtained under linear models on $\{0, 1\}^v$.

- For odd numbers of factors $v = 2q + 1$, $q \in \mathbb{N}$, the equally weighted designs ξ^* supported by all $\mathbf{x}^* = (x_1, \dots, x_v) \in \{0, 1\}^v$ such that $\sum_{i=1}^v x_i = q + 1$ is either D- or A-optimal.
- For even numbers of factors $v = 2q$, $q \in \mathbb{N}$, the equally weighted design ξ^* supported by all $\mathbf{x}^* = (x_1, \dots, x_v) \in \{0, 1\}^v$ such that $\sum_{i=1}^v x_i = q$ or $\sum_{i=1}^v x_i = q + 1$ is D-optimal. Moreover, the design ξ^* which assigns equal weights to all points $\mathbf{x}^* = (x_1, \dots, x_v) \in \{0, 1\}^v$ such that $\sum_{i=1}^v x_i = q$ is A-optimal.

7.1 Gamma model

A gamma model is given by

$$\mathbf{f}^T(\mathbf{x})\boldsymbol{\beta} = \frac{\kappa}{\mu(\mathbf{x}, \boldsymbol{\beta})} \quad \text{with intensity } u(\mathbf{x}, \boldsymbol{\beta}) = (\mathbf{f}^T(\mathbf{x})\boldsymbol{\beta})^{-2} \quad \forall \mathbf{x} \in \mathcal{X}.$$

Here, κ is the shape parameter of the gamma distribution which is assumed to be fixed and positive. The expected mean $\mu(\mathbf{x}, \boldsymbol{\beta})$ for the gamma distribution is positive for all $\mathbf{x} \in \mathcal{X}$. The parameter space including all possible parameter vector $\boldsymbol{\beta}$ is determined by the assumption $\mathbf{f}^T(\mathbf{x})\boldsymbol{\beta} > 0$ for all $\mathbf{x} \in \mathcal{X}$.

Let the experimental region be the cube $\mathcal{X} = [0, 1]^v$, $v > 1$. In Gaffke et al. (2019) we showed that the locally optimal designs under a first order gamma model $f(\mathbf{x}) = (1, x_1, \dots, x_v)^T$ are only supported by the vertices of the cube $[0, 1]^v$. Therefore, in the following we focus on the set of vertices $\{0, 1\}^v$.

In what follows, firstly we consider a two-factor gamma model with the linear predictor $\eta(\mathbf{x}, \boldsymbol{\beta}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ and experimental region $\mathcal{X} = [0, 1]^2$. Denote $\mathbf{x}_1^* = (0, 0)^T$, $\mathbf{x}_2^* = (1, 0)^T$, $\mathbf{x}_3^* = (0, 1)^T$ and $\mathbf{x}_4^* = (1, 1)^T$. Let $u_k = u(\mathbf{x}_k^*, \boldsymbol{\beta})$ ($1 \leq k \leq 4$), i.e.,

$$u_1 = \beta_0^{-2}, \quad u_2 = (\beta_0 + \beta_1)^{-2}, \quad u_3 = (\beta_0 + \beta_2)^{-2}, \quad u_4 = (\beta_0 + \beta_1 + \beta_2)^{-2}.$$

In view of Corollaries 5.1 and 5.2 the following explicit results are immediate.

Corollary 7.1 Consider a gamma model with $f(\mathbf{x}) = (1, x_1, x_2)^T$ and the experimental region $\mathcal{X} = [0, 1]^2$. Let $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$ be a parameter point such that $\beta_0 > 0$, $\beta_0 + \beta_1 > 0$, $\beta_0 + \beta_2 > 0$, and $\beta_0 + \beta_1 + \beta_2 > 0$. Then the unique locally D-optimal design ξ^* (at $\boldsymbol{\beta}$) is as follows.

- (1) ξ^* assigns equal weights $1/3$ to $\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*$ if and only if $\beta_0^2 - \beta_1\beta_2 \leq 0$.
- (2) ξ^* assigns equal weights $1/3$ to $\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_4^*$ if and only if $(\beta_0 + \beta_1)^2 + \beta_1\beta_2 \leq 0$.
- (3) ξ^* assigns equal weights $1/3$ to $\mathbf{x}_1^*, \mathbf{x}_3^*, \mathbf{x}_4^*$ if and only if $(\beta_0 + \beta_2)^2 + \beta_1\beta_2 \leq 0$.
- (4) ξ^* assigns equal weights $1/3$ to $\mathbf{x}_2^*, \mathbf{x}_3^*, \mathbf{x}_4^*$ if and only if $\beta_0^2 + \beta_1^2 + \beta_2^2 + \beta_1\beta_2 + 2\beta_0(\beta_1 + \beta_2) \leq 0$.
- (5) Otherwise, ξ^* is supported by the four points $\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*, \mathbf{x}_4^*$.

Proof In view of Corollary 5.1, part (i), straightforward computations show that the corresponding conditions of parts (1)–(4) are equivalent to $u_1^{-1} + u_2^{-1} + u_3^{-1} \leq u_4^{-1}$, $u_1^{-1} + u_2^{-1} + u_4^{-1} \leq u_3^{-1}$, $u_1^{-1} + u_3^{-1} + u_4^{-1} \leq u_2^{-1}$, and $u_2^{-1} + u_3^{-1} + u_4^{-1} \leq u_1^{-1}$, respectively.

Remark According to Corollary 7.1, part (5), the subregion where a four-point design is D-optimal has been determined by computer algebra and is given below.

- $-\beta_0 < \beta_1 < 0$ and $\frac{1}{2}(\sqrt{-(3\beta_1^2 + 4\beta_0\beta_1)} - (\beta_1 + 2\beta_0)) < \beta_2 < -(\beta_1 + \beta_0)^2/\beta_1$.
- $\beta_1 = 0$ and $\beta_2 > -\beta_0$.
- $\beta_1 > 0$ and $\frac{1}{2}(\sqrt{4\beta_0\beta_1 + \beta_1^2} - (\beta_1 + 2\beta_0)) < \beta_2 < \beta_0^2/\beta_1$.

On each subregion the optimal weights of a D-optimal design depend on the parameter values.

If $\beta_0 > 0$ define the ratios $\gamma_1 = \beta_1/\beta_0$ and $\gamma_2 = \beta_2/\beta_0$. So we have $\gamma_1 > -1$, $\gamma_2 > -1$ and $\gamma_1 + \gamma_2 > -1$. Without loss of generality the conditions of the D-optimal designs given in Corollary 7.1 can be written in terms of γ_1 and γ_2 . In the left panel of Fig. 1 the parameter subregions of γ_1 and γ_2 are depicted where the designs given by Corollary 7.1 are locally D-optimal. In particular, the design with support $\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*$ is locally D-optimal over the larger subregion for positive larger values of γ_1 and γ_2 . The diagonal line represents the case of equal effects $\beta_1 = \beta_2 = \beta$ where

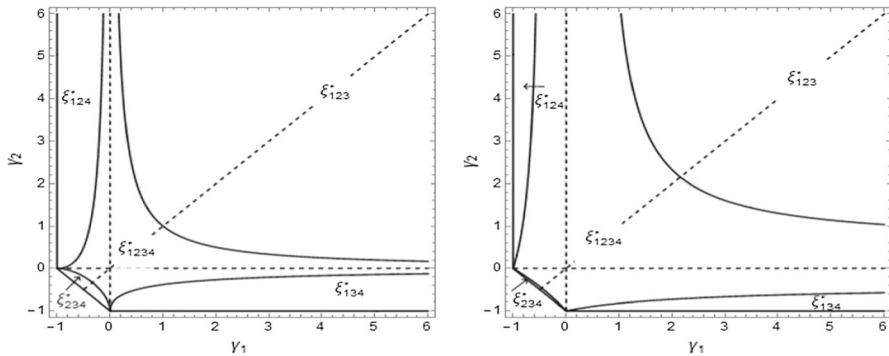


Fig. 1 Dependence of optimal designs under gamma models on β ; Left panel: D-optimal designs. Right panel: A-optimal designs. $\text{supp}(\xi_{ijk}^*) = \{x_i^*, x_j^*, x_k^*\} \subset \{x_1^*, x_2^*, x_3^*, x_4^*\}$ and $\text{supp}(\xi_{1234}^*) = \{x_1^*, x_2^*, x_3^*, x_4^*\}$. The diagonal dashed line is $\gamma_2 = \gamma_1$ where $\gamma_i = \beta_i/\beta_0$, $i = 1, 2$

$\beta > -(1/2)\beta_0$. In particular, in the case $-(1/3)\beta_0 < \beta < \beta_0$ the design is supported by the four design points with optimal weights

$$\omega_1^* = \frac{3\gamma + 1}{4(2\gamma + 1)}, \quad \omega_2^* = \omega_3^* = \frac{(\gamma + 1)^2}{4(2\gamma + 1)}, \quad \omega_4^* = \frac{1 - \gamma}{4}, \quad \text{where } \gamma = \frac{\beta}{\beta_0}.$$

These weights as functions of γ are exhibited in Fig. 2. Obviously, the weights are positive over the respective domain $\gamma \in (-1/3, 1)$ and $1/4 \leq \omega_2^* = \omega_3^* \leq 1/3$. The design ξ^* at $\gamma = 0$ assigns uniform weights $1/4$ to the set of points $\{x_1^*, x_2^*, x_3^*, x_4^*\}$. This case is equivalent to an ordinary linear regression model with two binary factors. At the limits of $(-1/3, 1)$ the D-optimal four-point design becomes a D-optimal saturated design. This means that at $\gamma = -1/3$ we have $\omega_1^* = 0$ and at $\gamma = 1$ we have $\omega_4^* = 0$.

Corollary 7.2 *Under the assumptions and notations of Corollary 7.1. The unique locally A-optimal design (at β) is as follows.*

$$(1) \quad \xi^* = \begin{pmatrix} x_1^* & x_2^* & x_3^* \\ \sqrt{3}\beta_0/c & (\beta_0 + \beta_1)/c & (\beta_0 + \beta_2)/c \end{pmatrix} \text{ if and only if}$$

$$(1 + 2/\sqrt{3})\beta_0^2 + (1/\sqrt{3})\beta_0(\beta_1 + \beta_2) - \beta_1\beta_2 \leq 0.$$

$$(2) \quad \xi^* = \begin{pmatrix} x_1^* & x_2^* & x_4^* \\ \sqrt{2}\beta_0/c & \sqrt{2}(\beta_0 + \beta_1)/c & (\beta_0 + \beta_1 + \beta_2)/c \end{pmatrix} \text{ if and only if}$$

$$(3 + \sqrt{2})\beta_0^2 + (2 + \sqrt{2})(\beta_1^2 + \beta_1\beta_2) + (5 + 2\sqrt{2})\beta_0\beta_1 + \sqrt{2}\beta_0\beta_2 \leq 0.$$

$$(3) \quad \xi^* = \begin{pmatrix} x_1^* & x_3^* & x_4^* \\ \sqrt{2}\beta_0/c & \sqrt{2}(\beta_0 + \beta_2)/c & (\beta_0 + \beta_1 + \beta_2)/c \end{pmatrix} \text{ if and only if}$$

$$(3 + \sqrt{2})\beta_0^2 + (2 + \sqrt{2})(\beta_2^2 + \beta_1\beta_2) + (5 + 2\sqrt{2})\beta_0\beta_2 + \sqrt{2}\beta_0\beta_1 \leq 0.$$

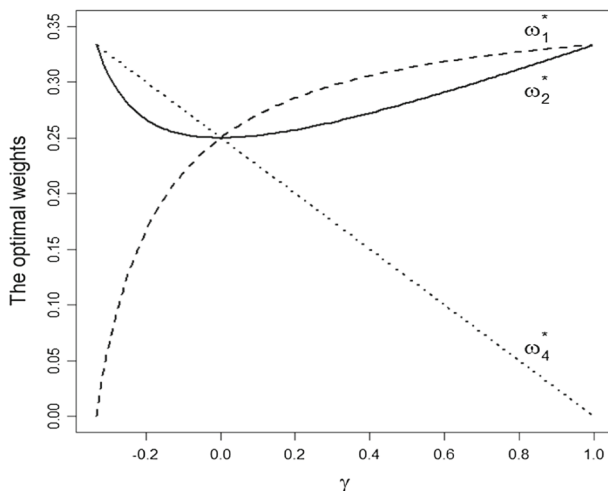


Fig. 2 Effect of γ on the optimal weights ω_1^* , ω_2^* and ω_4^* of the locally D-optimal four-point design under a two-factor gamma model, where $\gamma = \beta/\beta_0$ and $\beta_1 = \beta_2 = \beta$

$$(4) \quad \xi^* = \left(\frac{\mathbf{x}_2^*}{\sqrt{2}(\beta_0 + \beta_1)/c} \quad \frac{\mathbf{x}_3^*}{\sqrt{2}(\beta_0 + \beta_2)/c} \quad \frac{\mathbf{x}_4^*}{\sqrt{3}(\beta_0 + \beta_1 + \beta_2)/c} \right) \text{ if and only if}$$

$$(3 + 4\sqrt{2/3})(\beta_0^2 + \beta_1\beta_2) + 2(1 + \sqrt{2/3})(\beta_1^2 + \beta_2^2) + (5 + 6\sqrt{2/3})\beta_0(\beta_1 + \beta_2) \leq 0.$$

For each case (1) – (4), the constant c appearing in the weights equals the sum of the numerators of the three ratios.

$$(5) \quad \text{Otherwise, } \xi^* \text{ is supported by the four design points } \mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*, \mathbf{x}_4^*.$$

Proof The result follows from Corollary 5.2 by denoting $u_k = u(\mathbf{x}_k^*, \boldsymbol{\beta})$ ($1 \leq k \leq 4$) and $q_i = u_i^{-1/2}$ ($1 \leq i \leq 4$). \square

Again, the conditions of A-optimal designs can be written in terms of the ratios $\gamma_1 = \beta_1/\beta_0$ and $\gamma_2 = \beta_2/\beta_0$, $\beta_0 > 0$. In the right panel of Fig. 1 the parameter subregions of γ_1 and γ_2 are depicted where the designs given by Corollary 7.2 are locally A-optimal. Comparing to the left panel under D-optimality, similar interpretation might be observed. In particular, the largest subregion of the parameter points is for A-optimal designs with support $\{\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*\}$.

Remark For the multiple-factor gamma model $\mathbf{f}(\mathbf{x}) = (1, x_1, \dots, x_v)^T$ on the experimental region $\mathcal{X} = [0, 1]^v$, $v \geq 2$, the result of Burrige and Sebastiani (1994) can be applied. It was shown that the design which assigns equal weights $1/(v+1)$ to the design points $\mathbf{x}_1^* = (0, \dots, 0)^T$, $\mathbf{x}_2^* = (1, \dots, 0)^T$, \dots , $\mathbf{x}_{v+1}^* = (0, \dots, 1)^T$ is locally D-optimal for a given $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_v)^T$ if and only if $\beta_0^2 \leq \beta_i\beta_j$, $1 \leq i < j \leq v$. This result can be considered as a special case of Theorem 6.1. Note that for $v = 2$ this result is also covered by part (1) of Corollary 7.1.

Now consider a non-intercept gamma model with multiple factors. The linear predictor can be written as $\eta(\mathbf{x}, \boldsymbol{\beta}) = \mathbf{x}^T \boldsymbol{\beta} = \sum_{i=1}^v \beta_i x_i$. Let the experimental region

be given by $\mathcal{X} = [0, \infty)^v \setminus \{\mathbf{0}\}$ with intensity $u(\mathbf{x}, \boldsymbol{\beta}) = (\mathbf{x}^T \boldsymbol{\beta})^{-2}$ for all $\mathbf{x} \in \mathcal{X}$. The parameter space is determined by $\boldsymbol{\beta} \in (0, \infty)^v$, i.e., $\beta_i > 0$ for all $(1 \leq i \leq v)$. From Theorem 6.3 the following corollary is obtained for the gamma model without intercept.

Corollary 7.3 *Consider a non-intercept gamma model with $f(\mathbf{x}) = \mathbf{x}$ on the experimental region $\mathcal{X} = [0, \infty)^v \setminus \{\mathbf{0}\}$ and intensity $u(\mathbf{x}, \boldsymbol{\beta}) = (\mathbf{x}^T \boldsymbol{\beta})^{-2}$ for all $\mathbf{x} \in \mathcal{X}$. For a given vector $\mathbf{a} = (a_1, \dots, a_v)^T$ where $a_i > 0$ ($1 \leq i \leq v$) let $\mathbf{x}_i^* = a_i \mathbf{e}_i$ for all $i = 1, \dots, v$. Let k with $0 \leq k < \infty$ be given. For a given parameter point $\boldsymbol{\beta} \in (0, \infty)^v$ let $\xi_{\mathbf{a}}^*$ be the saturated design whose support consists of the points \mathbf{x}_i^* ($1 \leq i \leq v$) with the corresponding weights*

$$\omega_i^* = \frac{\beta_i^{\frac{2k}{k+1}}}{\sum_{j=1}^v \beta_j^{\frac{2k}{k+1}}} \quad (1 \leq i \leq v).$$

Then $\xi_{\mathbf{a}}^$ is locally Φ_k -optimal (at $\boldsymbol{\beta}$).*

Proof Let $u_i = u(\mathbf{x}_i^*, \boldsymbol{\beta})$ ($1 \leq i \leq v$). Thus $u_i = (a_i \beta_i)^{-2}$ ($1 \leq i \leq v$). Then condition (6.8) of Theorem 6.3 is equivalent to $-2 \sum_{i < j=1}^v \beta_i \beta_j x_i x_j \leq 0$ for all $\mathbf{x} \in \mathcal{X}$. Since $\beta_i > 0$, $x_i > 0$ ($1 \leq i \leq v$) the condition holds true. \square

The optimal weights given in Corollary 7.3 are the same irrespective of the values a_1, \dots, a_v . The reason is that the information matrix under a gamma model without intercept is invariant with respect to simultaneous scaling of the factors. Therefore, we get $\mathbf{M}(a_i \mathbf{e}_i, \boldsymbol{\beta}) = \mathbf{M}(\mathbf{e}_i, \boldsymbol{\beta})$, $i = 1, \dots, v$. Note also that Corollary 7.3 covers Theorem 3.1 in Idais and Schwabe (2020) who provided locally D- and A-optimal designs for non-intercept gamma models.

7.2 Poisson model

A Poisson model is given by

$$f^T(\mathbf{x})\boldsymbol{\beta} = \log(\mu(\mathbf{x}, \boldsymbol{\beta})) \quad \text{with intensity } u(\mathbf{x}, \boldsymbol{\beta}) = \exp(f^T(\mathbf{x})\boldsymbol{\beta}) \quad \forall \mathbf{x} \in \mathcal{X}.$$

Here, the expected mean $\mu(\mathbf{x}, \boldsymbol{\beta})$ for the Poisson distribution is positive for all $\mathbf{x} \in \mathcal{X}$. The parameter vector $\boldsymbol{\beta} \in \mathbb{R}^p$ is a real-valued vector.

For a two-factor model with linear predictor $\eta(\mathbf{x}, \boldsymbol{\beta}) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ and experimental region $\mathcal{X} = \{0, 1\}^2$ the next corollary presents the locally D-optimal designs. The same results for D-optimality were obtained in Graßhoff et al. (2013) under the Rasch Poisson counts model in item response theory. Denote $\mathbf{x}_1^* = (0, 0)^T$, $\mathbf{x}_2^* = (1, 0)^T$, $\mathbf{x}_3^* = (0, 1)^T$ and $\mathbf{x}_4^* = (1, 1)^T$. Let $u_k = u(\mathbf{x}_k^*, \boldsymbol{\beta})$ ($1 \leq k \leq 4$), i.e.,

$$u_1 = \exp(\beta_0), \quad u_2 = \exp(\beta_0 + \beta_1), \quad u_3 = \exp(\beta_0 + \beta_2), \quad u_4 = \exp(\beta_0 + \beta_1 + \beta_2).$$

Corollary 7.4 Consider a Poisson model with $f(\mathbf{x}) = (1, x_1, x_2)^T$ and the experimental region $\mathcal{X} = \{0, 1\}^2$. Let $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$ be a given parameter point. Then the unique locally D-optimal design ξ^* (at $\boldsymbol{\beta}$) is as follows.

- (1) ξ^* assigns equal weights $1/3$ to $\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*$ if and only if $\beta_2 \leq \log((1 + \exp(\beta_1))/(1 - \exp(\beta_1)))$.
- (2) ξ^* assigns equal weights $1/3$ to $\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_4^*$ if and only if $\beta_2 \leq \log((1 - \exp(-\beta_1))/(1 + \exp(-\beta_1)))$.
- (3) ξ^* assigns equal weights $1/3$ to $\mathbf{x}_1^*, \mathbf{x}_3^*, \mathbf{x}_4^*$ if and only if $\beta_2 \leq -\log((1 - \exp(\beta_1))/(1 + \exp(\beta_1)))$.
- (4) ξ^* assigns equal weights $1/3$ to $\mathbf{x}_2^*, \mathbf{x}_3^*, \mathbf{x}_4^*$ if and only if $\beta_2 \leq -\log((1 - \exp(-\beta_1))/(1 + \exp(-\beta_1)))$.
- (5) Otherwise, ξ^* is supported by the four points $\mathbf{x}_1^*, \mathbf{x}_2^*, \mathbf{x}_3^*, \mathbf{x}_4^*$.

Proof The proof is analogous to that of Corollary 7.1. \square

Remark Graßhoff et al. (2013) showed that a four-point design from part (5) of Corollary 7.4 is locally D-optimal in the subregion

$$|\beta_2| < \log\left(\frac{\exp(|\beta_1|) + 1}{\exp(|\beta_1|) - 1}\right).$$

Next we provide the locally A-optimal designs.

Corollary 7.5 Under the assumptions and notations of Corollary 7.4 the unique locally A-optimal design ξ^* (at $\boldsymbol{\beta}$) is as follows.

- (1) $\xi^* = \begin{pmatrix} \mathbf{x}_1^* & \mathbf{x}_2^* & \mathbf{x}_3^* \\ \sqrt{3}/c & \exp(-\beta_1)/c & \exp(-\beta_2)/c \end{pmatrix}$ if and only if $\exp(-\beta_1) + \exp(-\beta_2) + \exp(-(\beta_1 + \beta_2)) + \exp(-(\beta_1 + \beta_2)/2) + 2\sqrt{2/3}(\exp(-(2\beta_1 + \beta_2)/2)) + \exp(-(2\beta_2 + \beta_1)/2)) - 1 \leq 0$.
- (2) $\xi^* = \begin{pmatrix} \mathbf{x}_1^* & \mathbf{x}_2^* & \mathbf{x}_4^* \\ \sqrt{2}/c & \sqrt{2} \exp(-\beta_1)/c & \exp(-(\beta_1 + \beta_2)/2)/c \end{pmatrix}$ if and only if $\exp(\beta_1) + \exp(-\beta_2) + \exp(\beta_1 - \beta_2) + \exp(\beta_1 - \beta_2/2) + \sqrt{2} \exp(\beta_1/2 - \beta_2) - 1 \leq 0$.
- (3) $\xi^* = \begin{pmatrix} \mathbf{x}_1^* & \mathbf{x}_3^* & \mathbf{x}_4^* \\ \sqrt{2}/c & \sqrt{2} \exp(-\beta_2)/c & \exp(-(\beta_1 + \beta_2)/2)/c \end{pmatrix}$ if and only if $\exp(\beta_2) + \exp(-\beta_1) + \exp(\beta_2 - \beta_1) + \exp(\beta_2 - \beta_1/2) + \sqrt{2} \exp(\beta_2/2 - \beta_1) - 1 \leq 0$.
- (4) $\xi^* = \begin{pmatrix} \mathbf{x}_2^* & \mathbf{x}_3^* & \mathbf{x}_4^* \\ \sqrt{2} \exp(-\beta_1)/c & \sqrt{2} \exp(-\beta_2)/c & \sqrt{3} \exp(-(\beta_1 + \beta_2)/2)/c \end{pmatrix}$ if and only if $\exp(\beta_1) + \exp(\beta_2) + \exp(\beta_1 + \beta_2) + 2\sqrt{2/3}(\exp(\beta_2 + \beta_1/2) + \exp(\beta_1 + \beta_2/2)) - 1 \leq 0$.

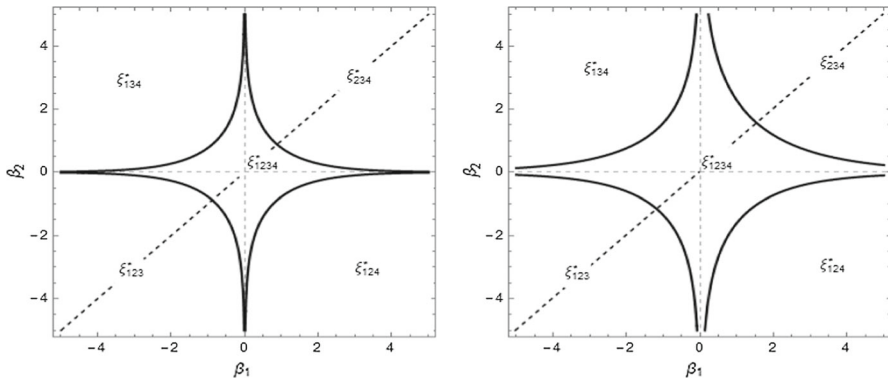


Fig. 3 Dependence of optimal designs under Poisson models on β ; Left panel: D-optimal designs. Right panel: A-optimal designs. $\text{supp}(\xi_{ijk}^*) = \{x_i^*, x_j^*, x_k^*\} \subset \{x_1^*, x_2^*, x_3^*, x_4^*\}$ and $\text{supp}(\xi_{1234}^*) = \{x_1^*, x_2^*, x_3^*, x_4^*\}$. The diagonal dashed line is $\beta_2 = \beta_1$

For each case (1) – (4), the constant c appearing in the weights equals the sum of the numerators of the three ratios.

(5) Otherwise, ξ^* is supported by the four design points $x_1^*, x_2^*, x_3^*, x_4^*$.

Figure 3 shows the dependence of the locally D- and A-optimal designs from Corollaries 7.4 and 7.5, respectively on the parameters β_1 and β_2 .

Remark For the multiple-factor Poisson model $f(\mathbf{x}) = (1, x_1, \dots, x_v)^T$ on the experimental region $\mathcal{X} = [0, 1]^v$, $v \geq 2$, the result of Russell et al. (2009) can be applied. It was shown that the design which assigns equal weights $1/(v+1)$ to the design points $\mathbf{x}_1^* = (0, \dots, 0)^T$, $\mathbf{x}_2^* = (1, \dots, 0)^T$, \dots , $\mathbf{x}_{v+1}^* = (0, \dots, 1)^T$ is locally D-optimal at a given $\beta = (\beta_0, \beta_1, \dots, \beta_v)^T$ such that $\beta_i = -2$, $i = 1, \dots, v$.

Now consider a non-intercept Poisson model with multiple factors. The linear predictor can be written as $\eta(\mathbf{x}, \beta) = \mathbf{x}^T \beta = \sum_{i=1}^v \beta_i x_i$. Let the experimental region be given by $\mathcal{X} = \{0, 1\}^v$ with intensity $u(\mathbf{x}, \beta) = \exp(\mathbf{x}^T \beta)$ for all $\mathbf{x} \in \mathcal{X}$. In the following we will apply Theorem 6.3 for a Poisson model without intercept. Let us restrict to the case of $a_i = 1$ ($1 \leq i \leq v$), i.e., the design points are the unit vectors \mathbf{e}_i ($1 \leq i \leq v$). As a result, condition (6.8) is simplified as presented in the following corollary.

Corollary 7.6 Consider a non-intercept Poisson model with $f(\mathbf{x}) = \mathbf{x}$ on the experimental region $\mathcal{X} = \{0, 1\}^v$, $v \geq 2$ and intensity $u(\mathbf{x}, \beta) = \exp(\mathbf{x}^T \beta)$ for all $\mathbf{x} \in \mathcal{X}$. For a given parameter point $\beta = (\beta_1, \dots, \beta_v)^T$ define $u_i = \exp(\beta_i)$ ($1 \leq i \leq v$) and denote by $u_{[1]} \geq u_{[2]} \geq \dots \geq u_{[v]}$ the descending order of u_1, u_2, \dots, u_v . Let k be given with $0 \leq k < \infty$. Let ξ^* be the saturated design supported by the unit vectors \mathbf{e}_i ($1 \leq i \leq v$) with weights $\omega_i^* = u_i^{-\frac{k}{k+1}} / \sum_{j=1}^v u_j^{-\frac{k}{k+1}}$ ($1 \leq i \leq v$). Then ξ^* is locally Φ_k -optimal (at β) if and only if

$$u_{[1]} + u_{[2]} \leq 1. \quad (7.1)$$

Proof Condition (6.8) of Theorem 6.3 reduces to

$$\exp\left(\sum_{i=1}^v \beta_i x_i\right) \sum_{i=1}^v \exp(-\beta_i) x_i^2 \leq 1 \quad \forall \mathbf{x} \in \mathcal{X}. \quad (7.2)$$

For any $\mathbf{x} = (x_1, \dots, x_v) \in \{0, 1\}^v$, $v \geq 2$ define the index set $S \subseteq \{1, \dots, v\}$ such that $x_i = 1$ if $i \in S$ and $x_i = 0$ else. So for \mathbf{x} described by $S \subseteq \{1, \dots, v\}$ and $s = \#S$, if $s = 0$ (i.e., $S = \emptyset$) then the l.h.s. of (7.2) is zero. If $s = 1$, inequality (7.2) becomes an equality. Let $s \geq 2$. Then the l.h.s. of (7.2) is equal to $\exp(\sum_{i \in S} \beta_i) \sum_{i \in S} \exp(-\beta_i)$ which thus rewrites as $\prod_{i \in S} u_i \sum_{i \in S} u_i^{-1}$ or equivalently as $\sum_{i \in S} \prod_{j \in S \setminus \{i\}} u_j$. By the the descending order $u_{[1]} \geq u_{[2]} \geq \dots \geq u_{[v]}$ of u_1, u_2, \dots, u_v we obtain for all subsets $S \subseteq \{1, \dots, v\}$ of the same size $s \geq 2$,

$$\sum_{i=1}^s u_{[i]}^{-1} \prod_{i=1}^s u_{[i]} = \sum_{i=1}^s \prod_{i \neq j=1}^s u_{[j]} \geq \sum_{i \in S} \prod_{j \in S \setminus \{i\}} u_j.$$

Denote $T_s = \sum_{i=1}^s u_{[i]}^{-1} \prod_{i=1}^s u_{[i]}$. Hence, inequality (7.2) is equivalent to $T_s \leq 1$ for all $s = 2, \dots, v$. Then it is sufficient to show that

$$u_{[1]} + u_{[2]} \leq 1 \iff T_s \leq 1 \quad \forall s = 2, \dots, v.$$

For “ \Leftarrow ”, $T_2 = u_{[1]} + u_{[2]} \leq 1$. For “ \Rightarrow ”, we use induction. Firstly, note that $T_2 = u_{[1]} + u_{[2]}$ thus $T_s \leq 1$ is true for $s = 2$. Now assume $T_s \leq 1$ is true for some $s = q$ where $2 \leq q \leq v$. We want to show that it is true for $s = q + 1$. We can write

$$\begin{aligned} T_{q+1} &= \left(\sum_{i=1}^q u_{[i]}^{-1} + u_{[q+1]}^{-1} \right) \left(\prod_{i=1}^q u_{[i]} \right) u_{[q+1]} \\ &= T_q u_{[q+1]} + \prod_{i=1}^q u_{[i]} = T_q u_{[q+1]} + T_q \left(\sum_{i=1}^q u_{[i]}^{-1} \right)^{-1} \\ &= T_q \left(u_{[q+1]} + \left(\sum_{i=1}^q u_{[i]}^{-1} \right)^{-1} \right) \\ &\text{since } \left(\sum_{i=1}^q u_{[i]}^{-1} \right)^{-1} \leq \frac{1}{q} u_{[1]} \text{ and } u_{[q+1]} + \frac{1}{q} u_{[1]} \leq T_2 = u_{[1]} + u_{[2]} \\ &\leq 1 \text{ we have} \\ T_{q+1} &\leq T_q \left(u_{[q+1]} + \frac{1}{q} u_{[1]} \right) \leq T_q T_2 \leq 1. \end{aligned}$$

Acknowledgements Open Access funding provided by Projekt DEAL. The author acknowledges the support by a scholarship of the Landesgraduiertenförderung Sachsen-Anhalt [GradFVO (GVBl. LSA 1992, 402)].

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

References

- Atkinson A, Haines L (1996) 14 designs for nonlinear and generalized linear models. In: Ghosh S, Rao C (eds) Design and analysis of experiments. Handbook of statistics, vol 13. Elsevier, Amsterdam, pp 437–475
- Atkinson AC, Woods DC (2015) Designs for generalized linear models. In: Dean A, Morris JSM, Bingha D (eds) Handbook of design and analysis of experiments. Chapman & Hall/CRC Press, Boca Raton, pp 471–514
- Atkinson A, Donev A, Tobias R (2007) Optimum experimental designs, with SAS, vol 34. Oxford University Press, Oxford
- Burridge J, Sebastiani P (1994) D-optimal designs for generalised linear models with variance proportional to the square of the mean. *Biometrika* 81:295–304
- Chernoff H (1953) Locally optimal designs for estimating parameters. *Ann Math Stat* 24:586–602
- Fedorov VV, Leonov SL (2013) Optimal design for nonlinear response models. CRC Press, Boca Raton
- Ford I, Torsney B, Wu CFJ (1992) The use of a canonical form in the construction of locally optimal designs for non-linear problems. *J R Stat Soc Ser B (Methodol)* 54:569–583
- Fox J (2015) Applied regression analysis and generalized linear models, 3rd edn. SAGE Publications, Los Angeles
- Gaffke N, Idais O, Schwabe R (2019) Locally optimal designs for gamma models. *J Stat Plan Inference* 203:199–214
- Goldburd M, Khare A, Tevet D (2016) Generalized linear models for insurance rating. Casualty Actuarial Society, Virginia
- Graßhoff U, Großmann H, Holling H, Schwabe R (2007) Design optimality in multi-factor generalized linear models in the presence of an unrestricted quantitative factor. *J Stat Plan Inference* 137:3882–3893
- Graßhoff U, Holling H, Schwabe R (2013) Optimal design for count data with binary predictors in item response theory. In: Ucinski D, Atkinson AC, Patan M (eds) mODa 10-advances in model-oriented design and analysis. Springer, Heidelberg, pp 117–124
- Huda S, Mukerjee R (1988) Optimal weighing designs: approximate theory. *Statistics* 19:513–517
- Idais O, Schwabe R (2020) Analytic solutions for locally optimal designs for gamma models having linear predictors without intercept. *Metrika*. <https://doi.org/10.1007/s00184-019-00760-3>
- Kiefer J (1975) Optimal design: variation in structure and performance under change of criterion. *Biometrika* 62:277–288
- Konstantinou M, Biedermann S, Kimber A (2014) Optimal designs for two-parameter nonlinear models with application to survival models. *Stat Sin* 24:415–428
- McCullagh P, Nelder J (1989) Generalized linear models, 2nd edn. Chapman & Hall, London
- Myers RH, Montgomery DC (1997) A tutorial on generalized linear models. *J Qual Technol* 29:274–291
- Nelder JA, Wedderburn RWM (1972) Generalized linear models. *J R Stat Soc Ser A (Gen)* 135:370–384
- Pukelsheim F (1993) Optimal design of experiments. Wiley, New York
- Pukelsheim F, Torsney B (1991) Optimal weights for experimental designs on linearly independent support points. *Ann Stat* 19:1614–1625
- Russell KG, Woods DC, Lewis SM, Eccleston JA (2009) D-optimal designs for poisson regression models. *Stat Sin* 19:721–730

- Schmidt D (2019) Characterization of c -, l - and ϕ_k -optimal designs for a class of non-linear multiple-regression models. *J R Stat Soc Ser B (Stat Methodol)* 81:101–120
- Schmidt D, Schwabe R (2017) Optimal design for multiple regression with information driven by the linear predictor. *Stat Sin* 27:1371–1384
- Silvey SD (1980) Optimal design. Chapman & Hall, London
- Tong L, Volkmer HW, Yang J (2014) Analytic solutions for D-optimal factorial designs under generalized linear models. *Electron J Stat* 8:1322–1344
- Walker SH, Duncan DB (1967) Estimation of the probability of an event as a function of several independent variables. *Biometrika* 54:167–179
- Yang M, Stufken J (2009) Support points of locally optimal designs for nonlinear models with two parameters. *Ann Stat* 37:518–541
- Yang M, Zhang B, Huang S (2011) Optimal designs for generalized linear models with multiple design variables. *Stat Sin* 21:1415–1430

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