



Statistical Theory and Related Fields

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tstf20

Locally *R*-optimal designs for a class of nonlinear multiple regression models

Lei He & Rong-Xian Yue

To cite this article: Lei He & Rong-Xian Yue (2023) Locally *R*-optimal designs for a class of nonlinear multiple regression models, Statistical Theory and Related Fields, 7:2, 107-120, DOI: 10.1080/24754269.2022.2153540

To link to this article: https://doi.org/10.1080/24754269.2022.2153540

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



6

Published online: 12 Dec 2022.

ſ	
C	2

Submit your article to this journal 🗹

Article views: 240



View related articles 🖸

🕨 View Crossmark data 🗹



Taylor & Francis

OPEN ACCESS Check for updates

Locally *R*-optimal designs for a class of nonlinear multiple regression models

Lei He 🔎 and Rong-Xian Yue 🕬

^aDepartment of Statistics, Anhui Normal University, Wuhu, People's Republic of China; ^bDepartment of Mathematics, Shanghai Normal University, Shanghai, People's Republic of China

ABSTRACT

This paper concerns with optimal designs for a wide class of nonlinear models with information driven by the linear predictor. The aim of this study is to generate an *R*-optimal design which minimizes the product of the main diagonal entries of the inverse of the Fisher information matrix at certain values of the parameters. An equivalence theorem for the locally *R*-optimal designs is provided in terms of the intensity function. Analytic solutions for the locally saturated *R*-optimal designs are derived for the models having linear predictors with and without intercept, respectively. The particle swarm optimization method has been employed to generate locally non-saturated *R*-optimal designs. Numerical examples are presented for illustration of the locally *R*-optimal designs for Poisson regression models and proportional hazards regression models.

ARTICLE HISTORY

Received 9 June 2022 Revised 24 November 2022 Accepted 27 November 2022

KEYWORDS

Poisson regression models; proportional hazards models; *R*-optimality; particle swarm optimization

1. Introduction

Generalized linear models (GLMs) have been used quite effectively in statistical modelling but the associated design issues are undoubtedly challenging, since the intensity function within their information matrices depends on the value of the linear predictor, which means that the optimal designs are related to the unknown parameters. Following Konstantinou et al. (2014), we note that the information matrix of the proportional hazards regression models used in survival analysis also has the same feature when type I and random censoring are considered, and their intensity functions, which are similar to Poisson regression models and negative binomial regression models, are strictly monotonic rather than a symmetric structure that appears in the logistic and probit models. In this paper, we focus on such models with monotonic intensity functions.

In this direction, there is increasing interest in determining optimal designs under various criteria, especially for models with multiple covariates. Initial work was done by Konstantinou et al. (2014) who provided analytical results of *D*- and *c*-optimal designs for this class of models with only one covariate. Subsequently, Schmidt and Schwabe (2015) extended the results concerning *D*-optimality to one-dimensional discrete design space. For multiple regression, Schmidt and Schwabe (2017) determined *D*-optimal designs by identifying a complete subclass, which contains the results of Russell et al. (2009) for the Poisson regression model as a special case. Radloff and Schwabe (2019) gave a construction method of *D*-optimal designs when the design region is a *k*-dimensional ball. Recently, Schmidt (2019) systematically characterized *c*-, *L*- and Φ_k -optimal designs for models with a single covariate and for multiple regression with an arbitrary number of covariates.

It should be noted that the linear predictor in the previously mentioned literature always includes an intercept term. The intercept in GLMs and proportional hazards regression models with censoring, respectively, characterizes the expected mean and expected survival time when all the explanatory variables are equal to zero. In this case, the linear predictor reflects the influence of all the unobserved fixed variables in these models. As pointed out by Idais and Schwabe (2021), when the intercept is significantly zero, i.e., the average impact of all the unobserved fixed variables is significantly zero, one may claim that the model includes probably most variables which explain the outcome. For the gamma models without intercept, Idais and Schwabe (2021) obtained some explicit solutions of *D*- and *A*-optimal designs in some multi-linear cases, including the two-factor model with interaction.

This paper aims to provide a characterization of *R*-optimality for multiple regression models with and without intercept. It is well known that the *R*-optimality criterion proposed by Dette (1997) has a nice statistical interpretation, namely minimizing the volume of the Bonferroni rectangular confidence region of the regression parameters. Moreover, it satisfies an extremely useful invariance property which allows an easy calculation of optimal designs on many linearly transformed design spaces. This optimality has been frequently applied to the cases of multi-response

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

CONTACT Rong-Xian Yue Syue2@shnu.edu.cn Department of Mathematics, Shanghai Normal University, Shanghai 200234, People's Republic of China

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

experiments, multi-factor experiments and mixture experiments, see, e.g., X. Liu and Yue (2020), P. Liu et al. (2022) and Hao et al. (2021) for some recent references.

In general, the dependence of designs on parameters whose values are unknown *a priori* occurs in nonlinear regression models. Hence, utilizing a pre-specified parameter value we can obtain the so-called locally optimal designs in accordance with Chernoff (1953). In this paper, we concentrate on the construction of locally optimal designs and take *R*-optimal designs to replace locally *R*-optimal designs for simplicity. Some general notations and a brief introduction of the *R*-criterion are presented in Section 2. In Sections 3 and 4, we analytically and numerically determine *R*-optimal designs for models having an intensity function which only depends on the value of the linear predictor with and without intercept, respectively. A brief discussion is given in Section 5. All proofs are included in the Appendix.

2. Model specification and *R*-optimality criterion

Throughout the paper, we focus on a class of nonlinear multiple regression models with information driven by the linear predictor defined on a given design region \mathcal{X} , and consider the approximate designs ξ of the form

$$\xi = \left\{ \begin{array}{ccc} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_m \\ \omega_1 & \omega_2 & \cdots & \omega_m \end{array} \right\}, \quad \mathbf{x}_i \in \mathcal{X}, \quad 0 \leq \omega_i \leq 1, \quad \sum_{i=1}^m \omega_i = 1$$

or simply $\xi = {x_i; \omega_i}_{i=1}^m$ (see Silvey, 1980, p. 15). The Fisher information matrix of ξ with independent observations is assumed to be

$$M(\xi,\boldsymbol{\beta}) = \int_{\mathcal{X}} Q(\boldsymbol{f}(\boldsymbol{x})^{\top}\boldsymbol{\beta})\boldsymbol{f}(\boldsymbol{x})\boldsymbol{f}(\boldsymbol{x})^{\top} \mathrm{d}\xi(\boldsymbol{x}) = \sum_{i=1}^{m} \omega_i Q(\boldsymbol{f}(\boldsymbol{x}_i)^{\top}\boldsymbol{\beta})\boldsymbol{f}(\boldsymbol{x}_i)\boldsymbol{f}(\boldsymbol{x}_i)^{\top},$$
(1)

where $Q \neq 0$ is the intensity/efficiency function (see Fedorov, 1972, p. 39) which only depends on the value of the linear predictor, f is a $p \times 1$ vector of known regression functions, and $\beta \in \mathbb{R}^p$ denotes the vector of p unknown parameters.

This kind of information matrix is common in the widely used generalized linear models, while it may also arise in other models such as the exponential regression models in proportional hazards parametrization with various censoring, including type I and random censoring (see Konstantinou et al., 2014) as well as other censoring distributions (see Schmidt & Schwabe, 2017). Following Konstantinou et al. (2014), we further assume that the intensity function *Q* satisfies the following conditions.

- (A1) $Q(\theta)$ is positive for all $\theta \in \mathbb{R}$ and twice continuously differentiable.
- (A2) The derivative $Q'(\theta)$ is positive for all $\theta \in \mathbb{R}$.
- (A3) The second derivative $g''(\theta)$ of the function $g(\theta) = 1/Q(\theta)$ is injective.
- (A4) The function $Q(\theta)/Q'(\theta)$ is an increasing function.

It is clear that the intensity functions induced by Poisson regression models, negative binomial regression models and proportional hazards regression models with type I and random censoring abide by all the conditions above.

In what follows, we concentrate on the *R*-optimality of designs which minimizes the product of the diagonal entries of the inverse of the Fisher information matrix. A design $\xi^* \in \Xi$ is called *R*-optimal if it minimizes

$$\psi(\boldsymbol{\xi},\boldsymbol{\beta}) = \prod_{j=1}^{p} \left(\boldsymbol{M}(\boldsymbol{\xi},\boldsymbol{\beta})^{-1} \right)_{jj} = \prod_{j=1}^{p} \boldsymbol{e}_{j,p}^{\top} \boldsymbol{M}(\boldsymbol{\xi},\boldsymbol{\beta})^{-1} \boldsymbol{e}_{j,p}$$
(2)

over Ξ , where Ξ is the set of all designs with a non-singular information matrix on \mathcal{X} , and $e_{j,p}$ denotes the *j*th unit vector in \mathbb{R}^p . An important tool in optimal design theory is equivalence theorems that not only provide a characterization of the optimal design but also are the basis of many algorithms for their numerical construction (see, e.g., Yang et al., 2013; Freise et al., 2021). The following result gives the equivalence theorem for *R*-optimality.

Theorem 2.1: A design $\xi^* \in \Xi$ is *R*-optimal if and only if

$$\phi(\mathbf{x},\boldsymbol{\beta}) = Q(f(\mathbf{x})^{\top}\boldsymbol{\beta})f(\mathbf{x})^{\top}M(\xi^*,\boldsymbol{\beta})^{-1}\left(\sum_{j=1}^{p} \frac{\boldsymbol{e}_{j,p}\boldsymbol{e}_{j,p}^{\top}}{\boldsymbol{e}_{j,p}^{\top}M(\xi^*,\boldsymbol{\beta})^{-1}\boldsymbol{e}_{j,p}}\right)M(\xi^*,\boldsymbol{\beta})^{-1}f(\mathbf{x}) \leqslant p$$
(3)

holds for all $x \in \mathcal{X}$. Moreover, it is equal at the support points of the design ξ^* .

In Sections 3 and 4 the minimally supported designs (i.e., the so-called saturated designs) will appear as candidates for determining *R*-optimal designs. A design $\xi = \{x_i; \omega_i\}_{i=1}^m$ has minimal support if the number of support points is equal to the number of parameters, i.e., m = p. Let $D_{\omega} = \text{diag}(\omega_1, \ldots, \omega_m)$, $\tilde{X} = Q^{1/2}X$ with $X = (f(x_1), \ldots, f(x_m))^\top$ and $Q^{1/2} = \text{diag}(\sqrt{Q(f(x_1)^\top \beta)}, \cdots, \sqrt{Q(f(x_m)^\top \beta)}))$. Accordingly, the information matrix (1) for the design ξ can be decomposed as $M(\xi, \beta) = \tilde{X}^\top D_{\omega} \tilde{X}$. Furthermore, if the design ξ is saturated and the regression vectors located at the rows of X are linearly independent, the following result exhibits the optimal weights of a design with minimal support for the *R*-optimality criterion.

Lemma 2.1 (Pukelsheim & Torsney, 1993): The *R*-optimal weights for a saturated design ξ are given by

$$\omega_j^* = \sqrt{\frac{s_{jj}}{p}}, \quad j = 1, \dots, p, \tag{4}$$

where s_{11}, \ldots, s_{pp} are the diagonal entries of the matrix

$$\boldsymbol{S} = \left(\tilde{\boldsymbol{X}}^{-1}\right)^{\top} \left(\sum_{j=1}^{p} \frac{\boldsymbol{e}_{j,p} \boldsymbol{e}_{j,p}^{\top}}{\boldsymbol{e}_{j,p}^{\top} \boldsymbol{M}(\boldsymbol{\xi}, \boldsymbol{\beta})^{-1} \boldsymbol{e}_{j,p}}\right) \tilde{\boldsymbol{X}}^{-1}.$$
(5)

Remark 2.1: Since the diagonal entries s_{jj} of the matrix S depend on the weights, a fixed point iteration procedure is available to determine a solution of optimal weights. In step r + 1, the weight $\omega_j^{(r+1)}$ is equivalent to $\sqrt{s_{jj}(\omega_1^{(r)}, \ldots, \omega_p^{(r)})/p}$ under a given initial weight vector $(\omega_1^{(0)}, \ldots, \omega_p^{(0)})^{\top}$.

3. *R*-optimal designs for models with intercept

In this section, we will describe how to construct *R*-optimal designs for multiple regression models with information matrices of the form (1) and satisfying the assumptions (A1)-(A4). More precisely, multiple regression with additive linear effects of the covariates in the linear predictor that contains an intercept term will be considered.

We consider the multi-linear case which has $f = (1, \mathbf{x}^{\top})^{\top}$ with $\mathbf{x} = (x_1, \dots, x_{p-1})^{\top} \in \mathcal{X} \subset \mathbb{R}^{p-1}$ and denote the parameter vector $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_{p-1})^{\top}$ for convenience. Suppose that the design region \mathcal{X} is a multidimensional polyhedron. By applying the complete class results described in Theorem 2 and Lemma 1 of Schmidt (2019), *R*-optimal designs can be found from the complete subclass, which has at most two support points on each edge of \mathcal{X} when the hyperplanes $H_{\eta} = \{\mathbf{x} \in \mathbb{R}^{p-1} : f(\mathbf{x})^{\top} \boldsymbol{\beta} = \eta\}$ are assumed to be bounded on \mathcal{X} for all $\eta \in \mathbb{R}$,

The following theorem provides an approach to generate *R*-optimal designs on the rectangular design region for the model under consideration, the proof of which can be established by using the similar arguments as in Theorem 9 in Schmidt (2019) via the fact $(\sum_{j=1}^{p} \sqrt{s_{jj}})^2 = \sqrt{p}$. It should be pointed out that using the same reasoning as in the proof of Theorems 8 and 9 in Schmidt (2019) the uniqueness of the solution of the common system of equations described in Theorem 3.1 follows. Here the case p = 2 with one covariate is also satisfied, so $p \ge 2$.

Theorem 3.1: Let $\mathcal{X} = [u_1, v_1] \times \cdots \times [u_{p-1}, v_{p-1}]$ and let the assumptions (A1)–(A4) be satisfied for a model with information matrices of the form (1). Let $\beta_i \neq 0$ for $i = 1, \ldots, p-1$. Define $a_i = v_i$ if $\beta_i > 0$, and $a_i = u_i$ if $\beta_i < 0$. Let s_{ij} $(i, j = 1, \ldots, p)$ denote the elements of the matrix S in (5) evaluated at a design with support points $\mathbf{a} - (x_1/\beta_1)\mathbf{e}_{1,p-1}, \ldots, \mathbf{a} - (x_{p-1}/\beta_{p-1})\mathbf{e}_{p-1,p-1}$ and $\mathbf{a} = (a_1, \ldots, a_{p-1})^{\top}$. If a solution exists, let $x_i^* \in (0, \infty)$ and the weights ω_i^* for $i = 1, \ldots, p-1$ be the unique solutions of the common system of Equations (4) and (6) given by

$$x_{i} - 2\frac{Q(f(\boldsymbol{a})^{\top}\boldsymbol{\beta} - x_{i})}{Q'(f(\boldsymbol{a})^{\top}\boldsymbol{\beta} - x_{i})} \left[1 - \sqrt{\frac{Q(f(\boldsymbol{a})^{\top}\boldsymbol{\beta} - x_{i})}{Q(f(\boldsymbol{a})^{\top}\boldsymbol{\beta})}} \frac{s_{1,i+1}}{\sqrt{s_{11}s_{i+1,i+1}}} \right] = 0.$$
(6)

If $x_i^* \leq |\beta_i|(v_i - u_i)$ holds for i = 1, ..., p - 1, then the design

$$\xi^* = \begin{cases} a - (x_1^*/\beta_1)e_{1,p-1} & \cdots & a - (x_{p-1}^*/\beta_{p-1})e_{p-1,p-1} & a \\ \omega_1^* & \cdots & \omega_{p-1}^* & \omega_p^* \end{cases}$$

is a unique R-optimal design.

Remark 3.1: Let $\mathcal{X} = [u_1, v_1] \times \cdots \times [u_{p-1}, v_{p-1}]$, when some of the parameters $\beta_1, \ldots, \beta_{p-1}$ are equal to zero. It follows from Lemma 1 of Schmidt (2019) that the two endpoints of the corresponding edges must be the support

points of optimal design, and the number of support points is possibly over *p*. This means that we need to use some numerical algorithms to generate the *R*-optimal design, such as the commonly used particle swarm optimization (PSO) method, see Section 4.2 for details.

The following two examples illustrate the results in Theorem 3.1 and Remark 3.1, and exhibit the performance of different designs in terms of the Bonferroni confidence intervals and the relative efficiencies of designs. In Example 3.1 the first-order Poisson regression model with intercept will be considered, and the difference of the averaged Bonferroni confidence intervals derived by *R*- and *D*-optimal designs as well as the balanced design ξ_b will be investigated by a numerical simulation. Example 3.2 considers the Poisson regression model with two covariants which was discussed in Schmidt (2019), and the relative efficiency of *R*-optimal designs will be compared with other designs. In both of the examples, the intensity function *Q* is given by $Q(\theta) = \exp(\theta)$ for $\theta \in \mathbb{R}$.

Example 3.1: Consider the first-order Poisson regression model with intercept and let $\mathcal{X} = [0, 5]$. If we set $\boldsymbol{\beta} = (6, -1)^{\top}$ and $\boldsymbol{\beta} = (1, 1)^{\top}$, the *R*- and *D*-optimal designs can be calculated numerically by Theorem 3.1 and Theorem 2 of Konstantinou et al. (2014). For example, under the *R*-optimality criterion with $\boldsymbol{\beta} = (6, -1)^{\top}$ we obtain $x_1^* = 2.1886$ and $\omega_2^* = 0.5431$ by solving the Equations (4) and (6). In Table 1, the results on optimal designs for both cases are reported. Furthermore, we carry out a simulation study in order to assess the difference of the Bonferroni confidence intervals derived by different designs. The averaged Bonferroni confidence intervals under different sample sizes are calculated by 10,000 simulation runs, which are shown in Table 1. We find from Table 1 that for both cases the coverage probabilities perform in a similar manner and much close to 0.95 as the sample size increases, and the width of the averaged Bonferroni confidence intervals designs is narrower comparing with *D*-optimal designs and the balanced design ξ_b when there exists a serious loss of *R*-efficiency.

Example 3.2: Assume that the linear predictor in the considered Poisson regression model includes an intercept term. Let $\mathcal{X} = [0, 5]^2$ and $\boldsymbol{\beta} = (0, -1, -1)^{\top}$. A numerical calculation yields immediately $x_1^* = x_2^* = 2.1785$, $\omega_1^* = \omega_2^* = 0.3060$ and $\omega_3^* = 0.3880$ by the Equations (4) and (6). The *R*-optimal design is given in Table 2. The PSO method was used to find the corresponding *D*-, *R*-, and *A*-optimal designs when $\boldsymbol{\beta} = (0, -1, 0)^{\top}$. The resulting designs are listed in Table 2. In this case, we find that the endpoints (0,0) and (0,5) must be the support points of optimal designs. Figure 1 displays the plot of the function $\phi(\mathbf{x}, \boldsymbol{\beta})$ defined in (3) for the cases $\boldsymbol{\beta} = (0, -1, -1)^{\top}$ and $\boldsymbol{\beta} = (0, -1, 0)^{\top}$, which indicates that the function ϕ attains its maximum 3 at each support point.

To compare the performance of different designs, such as the common *D*- and *A*-optimality, we may calculate the related efficiency that usually defines the value of the criterion function for the optimal design relative to the value of the criterion function of a design and can thus take values between 0 and 1. For instance, the *R*-efficiency is defined as $\text{Eff}_R(\xi) = \psi(\xi^*, \beta)/\psi(\xi, \beta)$. The results are summarized in Table 2. For the case $\beta = (0, -1, -1)^{\top}$, all the designs have relatively high efficiencies regarding the other optimality criteria. This occurs primarily because the designs have a similar structure. For the case $\beta = (0, -1, 0)^{\top}$, we observe that the *R*-optimal design still has relatively high *D*- and *A*-efficiencies, but the *A*-optimal design has a large loss of *R*-efficiency.

Theorem 3.2: Let $\mathcal{X} = [0, \infty)^{p-1}$ and let the assumptions (A1)–(A4) be satisfied for models with information matrices of the form (1). If $\beta_i < 0$ for all i = 1, ..., p - 1, then the solutions x_i^* and ω_i^* from the system of Equations (4) and (6) do not depend on the parameter vector $(\beta_1, ..., \beta_{p-1})^\top$, that is, the *R*-optimal weights are unchanged for different $(\beta_1, ..., \beta_{p-1})^\top$ when β_0 is fixed.

Remark 3.2: The optimal weights of saturated designs for *L*- and Φ_k -optimality criteria, except for *D*-optimality, depend on the parameters $(\beta_1, \ldots, \beta_{p-1})^{\top}$ in the same settings of Theorem 3.2.

To further illustrate Theorem 3.2, we consider the proportional hazards regression models with two types of censoring. Under type I censoring with a fixed censoring time *c* the intensity function *Q* is given by $Q(\theta) = 1 - \exp(-c \exp(\theta))$. For random censoring the intensity function $Q(\theta)$ equals to $1 - [1 - \exp(-c \exp(\theta))]/[c \exp(\theta)]$ if the censoring times are assumed to follow a uniform distribution U(0, c) (see Konstantinou et al., 2014). For both the above-mentioned intensity functions the variable θ belongs to \mathbb{R} .

Example 3.3: For the proportional hazards regression models with type I and random censoring Schmidt (2019) discussed the behaviour of *c*- and Φ_k -optimal designs for different parameter values. Their results are consistent with Remark 3.2. Analogous to Schmidt (2019), we investigate the performance of *R*-optimal designs on the design region $\mathcal{X} = [0, 3]^{p-1}$ by comparing with a balanced design, say ξ_b , which is supported at all vertices of \mathcal{X} with equal weights. First, the censoring time *c* as an unknown parameter should be determined previously, which is closely

Sample size	Desigr	า	<i>R</i> -efficiency	BCI [†]	CP [‡]
Case A: $\beta_0 = 6$, β	$\beta_1 = -1$	o (00)			
	$\xi_R^* = \begin{cases} 0\\ 0.5431 \end{cases}$	2.1886	1	[5.9506, 6.0491]	0.9709
	(,		[-1.0209, -0.9792]	0.9821
10	$\xi_D^* = \begin{cases} 0\\ 0.5 \end{cases}$	0.5	0.9792	[5.9485, 6.0512]	0.9803
		,		[-1.0756, -0.9266]	0.9783
	$\xi_b = \begin{cases} 0\\ 0.5 \end{cases}$	5 0.5	0.3436	[5.9482, 6.0509]	0.9790
		, 2.1006)		[—1.1357 , —0.8776]	0.9821
	$\xi_R^* = \begin{cases} 0\\ 0.5431 \end{cases}$	0.4569	1	[5.9866, 6.0134]	0.9510
	(,		[-1.0209, -0.9792]	0.9536
100	$\xi_D^* = \begin{cases} 0\\ 0.5 \end{cases}$	0.5	0.9792	[5.9860, 6.0139]	0.9538
		,		[-1.0203, -0.9799]	0.9523
	$\xi_b = \begin{cases} 0\\ 0.5 \end{cases}$	5 0.5	0.3436	[5.9860, 6.0140]	0.9540
	C C	,		[-1.0351, -0.9665]	0.9561
Case B: $\beta_0 = 1, \beta_1$	= 1	5)			
	$\xi_R^* = \begin{cases} 2.4078 \\ 0.8234 \end{cases}$	0.1766	1	[0.7045, 1.2904]	0.9765
	(3	5)		[0.9346, 1.0660]	0.9792
10	$\xi_D^* = \begin{cases} 0.5 \\ 0.5 \end{cases}$	0.5	0.5221	[0.6404, 1.3557]	0.9800
	(o	E)		[0.9258, 1.0747]	0.9792
	$\xi_b = \begin{cases} 0\\ 0.5 \end{cases}$	0.5	0.0598	[0.3139, 1.6046]	0.9853
	(24679	5)		[0.8785, 1.1375]	0.9813
	$\xi_R^* = \begin{cases} 2.4078\\ 0.8234 \end{cases}$	0.1766	1	[0.9205, 1.0797]	0.9532
	(2	ر د ا		[0.9821, 1.0178]	0.9547
100	$\xi_D^* = \begin{cases} 3\\ 0.5 \end{cases}$	0.5	0.5221	[0.9031, 1.0976]	0.9523
	()	, ,		[0.9796, 1.0202]	0.9531
	$\xi_b = \begin{cases} 0\\ 0.5 \end{cases}$	5 0.5	0.0598	[0.8258, 1.1671]	0.9621
	ť	,		[0.9664, 1.0349]	0.9603

Table 1. The simulation results obtained from the locally *R*-, *D*-optimal designs and the balanced design on $\mathcal{X} = [0, 5]$ for the first-order Poisson regression model with intercept.

[†] The abbreviation 'BCI' represents the Bonferroni confidence intervals of the parameters β_0 and β_1 , where the first row under each design is the simulation results of β_0 , and the second row is β_1 .

[‡] The abbreviation 'CP' represents the coverage probabilities of 95% confidence intervals for the parameters β_0 and β_1 .

Table 2. Comparison of *R*-optimal design with *D*- and *A*-optimal designs on $\mathcal{X} = [0, 5]^2$ for the Poisson regression model discussed in Example 3.2.

				Efficiencies	
β	Criterion	Optimal design	$\operatorname{Eff}_{D}(\xi)^{\dagger}$	$\operatorname{Eff}_R(\xi)$	$\operatorname{Eff}_{A}(\xi)^{\dagger}$
(0, -1, -1)	D	$\xi_D^* = \begin{cases} (2,0) & (0,2) & (0,0) \\ 1/3 & 1/3 & 1/3 \end{cases}$	1	0.9526	0.9856
	R	$\xi_R^* = \begin{cases} (2.1785, 0) & (0, 2.1785) & (0, 0) \\ 0.3060 & 0.3060 & 0.3880 \end{cases}$	0.9886	1	0.9704
	Α	$\xi_A^* = \begin{cases} (2.2453, 0) & (0, 2.2453) & (0, 0) \\ 0.3492 & 0.3492 & 0.3016 \end{cases}$	0.9884	0.9409	1
	D	$\xi_D^* = \begin{cases} (1.8493, 5) & (1.8493, 0) & (0, 5) & (0, 0) \\ 0.3197 & 0.3197 & 0.1802 & 0.1802 \end{cases}$	1	0.8454	0.7905
(0, -1, 0)	R	$\xi_R^* = \begin{cases} (1.9449, 5) & (1.9449, 0) & (0, 5) & (0, 0) \\ 0.1476 & 0.1951 & 0.2185 & 0.4388 \end{cases}$	0.9622	1	0.8435
	Α	$\xi_A^* = \begin{cases} (2.1798, 5) & (2.1798, 0) & (0, 5) & (0, 0) \\ 0.1198 & 0.4054 & 0.0757 & 0.3991 \end{cases}$	0.7717	0.6183	1

[†]The *D*- and *A*-efficiencies are defined as $\text{Eff}_D(\xi) = \{\det(M(\xi, \beta))/\det(M(\xi^*, \beta))\}^{1/p} \text{ and } \text{Eff}_A(\xi) = \operatorname{tr}(M^{-1}(\xi^*, \beta))/\operatorname{tr}(M^{-1}(\xi, \beta)), \text{ respectively, where } \det(\cdot) \text{ and } \operatorname{tr}(\cdot) \text{ are the matrix determinant and trace functions.}$

related to the amount of censoring q. Here q is also called the overall probability of censoring and given by $q = 1 - \sum \omega_i P(Y_j < C)$, in which Y_i is the survival time and C is the censoring distribution (see Kalish & Harrington, 1988). For the two-covariate case, we choose c = 32 for type I censoring and c = 69 for random censoring such that q is, respectively, equal to 60% for $\beta = (3, -2.5, -2.5)^{\top}$, 71% for $\beta = (3, -3, -3)^{\top}$ and 75% for $\beta = (3, -4, -4)^{\top}$



Figure 1. Plot of the function ϕ in (3) for the *R*-optimal design on $\mathcal{X} = [0, 5]^2$ for the Poisson regression model discussed in Example 3.2: (a) for $\boldsymbol{\beta} = (0, -1, -1)^{\top}$ and (b) for $\boldsymbol{\beta} = (0, -1, 0)^{\top}$.

β	<i>R</i> -optimal de	R-optimal design for type I censoring				
(3, -2.5, -2.5)	{(2.4406, 0) 0.2443	(0, 2.4406) 0.2443	(0,0) 0.5114	0.1986		
(3, -3, -3)	(2.0338, 0) 0.2443	(0, 2.0338) 0.2443	(0, 0) 0.5114	0.0389		
(3, -4, -4)	(1.5254, 0) 0.2443	(0, 1.5254) 0.2443	(0, 0) 0.5114	0.0004		
	<i>R</i> -optimal desi	gn for random	censoring			
(3, -2.5, -2.5)	{(2.4972, 0) 0.2508	(0, 2.4972) 0.2508	(0, 0) 0.4984	0.2307		
(3, -3, -3)	{(2.0810, 0) 0.2508	(0, 2.0810) 0.2508	(0, 0) 0.4984	0.0497		
(3, -4, -4)	(1.5607, 0) 0.2508	(0, 1.5607) 0.2508	(0, 0) 0.4984	0.0005		

Table 3. *R*-optimal designs on $\mathcal{X} = [0, 3]^2$ for proportional hazards regression models for different $\boldsymbol{\beta}$ and *R*-efficiencies of the balanced design ξ_b .

when the balanced design ξ_b on $\mathcal{X} = [0, 3]^2$ had been used. The *R*-optimal designs on $\mathcal{X} = [0, 3]^2$ and *R*-efficiencies of ξ_b are summarized in Table 3. We observe from Table 3 that the *R*-optimal designs shift the support points on the edges towards the vertex $(0, 0)^{\top}$ and the *R*-efficiencies of the balanced design ξ_b decrease with increasing amounts of censoring. Note also that the *R*-efficiency of ξ_b is quite low for these censoring scenarios. In addition, the *R*-optimal weights are the same for each model by Theorem 3.2.

4. R-optimal designs for models without intercept

In the present section we turn to discuss the multi-linear case without intercept, where the vector of the regressor functions is specified as $f = (x_1, \ldots, x_p)^\top \in \mathcal{X}, p \ge 2$, and the parameter vector is denoted by $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$. The regression model without an intercept is common and it usually arises from the physical characteristics of the variables measured. The design issue for this model has attracted considerable attention in the literature (see, e.g., Idais & Schwabe, 2021; Li et al., 2005 and the references cited there).

Let $\mathcal{X} = [u_1, v_1] \times \cdots \times [u_p, v_p]$. Then the support points of *R*-optimal designs are also at the edges of \mathcal{X} by Theorem 2 in Schmidt (2019). Moreover, from the proof of Theorem 8 in Schmidt (2019) the support points of an *R*-optimal design ξ^* must be given by $\mathbf{a} - (x_1/\beta_1)\mathbf{e}_{1,p-1}, \ldots, \mathbf{a} - (x_p/\beta_p)\mathbf{e}_{p,p}$ and $\mathbf{a} = (a_1, \ldots, a_p)^\top$ with $a_i = v_i$ if $\beta_i > 0$ and $a_i = u_i$ if $\beta_i < 0$ for $i = 1, \ldots, p$. As a result, the optimality condition for approximate designs in Theorem 2.1 can be simplified to verify whether it is satisfied at the boundary of the design region, i.e., the design ξ^* is *R*-optimal if and only if

$$h_i(x_i,\xi^*) = \ell_i(x_i,\xi^*) - \frac{p}{Q(\boldsymbol{a}^\top\boldsymbol{\beta} + \beta_i(x_i - a_i))} \leqslant 0$$
(7)

holds for all x_i , i = 1, ..., p, where the function $\ell_i(x_i, \xi^*)$ is given by

$$\ell_i(x_i,\xi^*) = (\boldsymbol{a} + (x_i - a_i)\boldsymbol{e}_{i,p})^\top \boldsymbol{M}(\xi^*,\boldsymbol{\beta})^{-1} \left(\sum_{j=1}^p \frac{\boldsymbol{e}_{j,p}\boldsymbol{e}_{j,p}^\top}{\boldsymbol{e}_{j,p}^\top \boldsymbol{M}(\xi^*,\boldsymbol{\beta})^{-1}\boldsymbol{e}_{j,p}} \right) \boldsymbol{M}(\xi^*,\boldsymbol{\beta})^{-1} (\boldsymbol{a} + (x_i - a_i)\boldsymbol{e}_{i,p})^\top \boldsymbol{h}(\xi^*,\boldsymbol{\beta})^{-1} (\boldsymbol{a} + (x_i - a_i)\boldsymbol{e}_{i,p})^\top \boldsymbol{h}(\xi^*,\boldsymbol{\beta})^\top \boldsymbol{h}(\xi^*,\boldsymbol{\beta})^{-1} (\boldsymbol{a} + (x_i - a_i)\boldsymbol{e}_{i,p})^\top \boldsymbol{h}(\xi^*,\boldsymbol{\beta})^\top \boldsymbol{h$$

4.1. Some theoretical results on saturated designs

In this subsection, we will provide two types of saturated *R*-optimal designs for the model considered by distinguishing whether the vertex *a* is a support point. The first result reveals designs which include the support point *a*, where the condition $a_{\tau} \neq 0, \tau \in \{1, ..., p\}$, described in Theorem 4.1 is to guarantee the design ξ_{τ}^* in (9) with a non-singular information matrix.

Theorem 4.1: Let the assumptions (A1)–(A4) be satisfied for models with information matrices of the form (1) and without intercept. Let s_{ij} (i, j = 1, ..., p) be the elements of the matrix **S** in (5) evaluated at a design of the form $\{a - (x_i/\beta_i)e_{i,p}; \omega_i\}_{i=1}^p$. For a given index τ ($\tau \in \{1, ..., p\}$) with $a_\tau \neq 0$, if a solution exists, let $x_i^* \in (0, \infty)$ and the weights ω_i^* be the unique solutions of the common system of Equations (4) and (8) given by

$$x_{i} - 2\frac{Q(\boldsymbol{a}^{\top}\boldsymbol{\beta} - x_{i})}{Q'(\boldsymbol{a}^{\top}\boldsymbol{\beta} - x_{i})} \left[1 - \sqrt{\frac{Q(\boldsymbol{a}^{\top}\boldsymbol{\beta} - x_{i})}{Q(\boldsymbol{a}^{\top}\boldsymbol{\beta})}} \frac{s_{\tau i}}{\sqrt{s_{\tau \tau} s_{ii}}} \right] = 0$$
(8)

for all $i \neq \tau$. If $0 < x_i^* \leq |\beta_i|(v_i - u_i)$ holds, then the design

$$\xi_{\tau}^{*} = \begin{cases} \boldsymbol{a} - (x_{1}^{*}/\beta_{1})\boldsymbol{e}_{1,p} & \cdots & \boldsymbol{a} - (x_{\tau}^{*}/\beta_{\tau})\boldsymbol{e}_{\tau,p} & \cdots & \boldsymbol{a} - (x_{p}^{*}/\beta_{p})\boldsymbol{e}_{p,p} \\ \omega_{1}^{*} & \cdots & \omega_{\tau}^{*} & \cdots & \omega_{p}^{*} \end{cases}$$
(9)

with $x_{\tau}^* = 0$ will be R-optimal on \mathcal{X} provided that the condition $h_{\tau}(x_{\tau}, \xi_{\tau}^*) < 0$ (or $\phi(\mathbf{a} + (x_{\tau} - a_{\tau})\mathbf{e}_{\tau,p}, \boldsymbol{\beta}) < p$) for $x_{\tau} \in [u_{\tau}, v_{\tau}] \setminus \{a_{\tau}\}$ is satisfied, where $h_{\tau}(\cdot, \cdot)$ is defined as in (7) and $\phi(\cdot, \cdot)$ as in (3).

Example 4.1: Consider the Poisson regression model with two covariates and without intercept, where the design region is $\mathcal{X} = [0, 5]^2$. To find the *R*-optimal design we first fix $\boldsymbol{\beta} = (-0.5, 0.5)^{\top}$. According to Theorem 4.1 we can only specify $\tau = 2$, i.e., $x_2^* = 0$. By solving the common system of Equations (4) and (10) we obtain $x_1^* = 2.1886$ and the design ξ_2^* is of the form

$$\xi_2^* = \begin{cases} (4.3772, 5) & (0,5) \\ 0.4569 & 0.5431 \end{cases}.$$

It is easily verified that the condition $h_2(x_2, \xi_2^*) < 0$ is satisfied for $x_2 \in [0, 5)$ and the design ξ_2^* is then *R*-optimal (see also Figure 2(a)). If we choose $\boldsymbol{\beta} = (1, 1)^{\top}$, however, it follows from Theorem 4.1 that the designs ξ_1^* and ξ_2^* are given by

$$\xi_1^* = \begin{cases} (2.4678, 5) & (5, 5) \\ 0.8234 & 0.1766 \end{cases}, \quad \xi_2^* = \begin{cases} (5, 2.4678) & (0, 5) \\ 0.8234 & 0.1766 \end{cases},$$

respectively. In this case, the conditions $h_i(x_i, \xi_i^*) < 0$, i = 1, 2, are not satisfied, which means that both designs are not *R*-optimal. Accordingly, finding a saturated *R*-optimal design for this model that does not contain the support point *a* may be an alternative scheme. This type of *R*-optimal design will be elaborated in Theorem 4.2.

Theorem 4.2: Let the assumptions (A1)–(A4) be satisfied for models with information matrices of the form (1) and without intercept. Let s_{ij} (i, j = 1, ..., p) be the elements of the matrix **S** in (5) evaluated at a design of the form $\{a - (x_i/\beta_i)e_{i,p}; \omega_i\}_{i=1}^p$. If a solution exists, let $x_i^* \in (0, \infty)$ and the weights ω_i^* be the unique solutions of the common system of Equations (4) and (10) given by

$$x_i - 2\frac{Q(\boldsymbol{a}^{\top}\boldsymbol{\beta} - x_i)}{Q'(\boldsymbol{a}^{\top}\boldsymbol{\beta} - x_i)} \left[\frac{1}{1 - \boldsymbol{z}^{\top}\boldsymbol{1}_p} \sum_{j=1}^p \frac{z_j \sqrt{Q(\boldsymbol{a}^{\top}\boldsymbol{\beta} - x_i)}}{\sqrt{Q(\boldsymbol{a}^{\top}\boldsymbol{\beta} - x_j)}} \frac{s_{ji}}{\sqrt{s_{ii}s_{jj}}} + 1 \right] = 0$$
(10)

for all *i*, where $\mathbf{z} = (z_1, \ldots, z_p)^\top$ with $z_j = a_j \beta_j / x_j$, $\mathbf{1}_p = (1, \ldots, 1)^\top \in \mathbb{R}^p$. If $0 < x_i^* \leq |\beta_i| (v_i - u_i)$ holds, then the design

$$\xi^{*} = \begin{cases} a - (x_{1}^{*}/\beta_{1})e_{1,p} & \cdots & a - (x_{p}^{*}/\beta_{p})e_{p,p} \\ \omega_{1}^{*} & \cdots & \omega_{p}^{*} \end{cases}$$
(11)

will be R-optimal on \mathcal{X} provided that the condition $h_i(a_i, \xi^*) < 0$ (or $\phi(\mathbf{a}, \boldsymbol{\beta}) < p$) is satisfied.



Figure 2. Plots of the functions ϕ in (3) for the *R*-optimal designs on $\mathcal{X} = [0, 5]^2$ for the Poisson regression models discussed in Examples 4.1 and 4.2: (a) for $\beta = (-0.5, 0.5)^{\top}$ and (b) for $\beta = (1, 1)^{\top}$.

Example 4.2: Consider the same model given in Example 4.1. Here we consider $\boldsymbol{\beta} = (1, 1)^{\top}$. The common system of Equations (4) and (10) has the solutions $x_1^* = x_2^* = 1.8755$ and $\omega_1^* = \omega_2^* = 0.5$. Then the design

$$\xi^* = \begin{cases} (3.1245, 5) & (5, 3.1245) \\ 0.5 & 0.5 \end{cases}$$

is *R*-optimal due to $\phi(a, \beta) < 2$ (see also Figure 2(b)).

Corollary 4.1: Let $\mathcal{X} = [0, \infty)^p$ and let the assumptions (A1)–(A4) be satisfied for models with information matrices of the form (1) and without intercept. Then the following design ξ^* is R-optimal for each $\beta_i < 0, i = 1, ..., p$,

$$\xi^* = \begin{cases} -(\psi^{-1}(0)/\beta_1)\boldsymbol{e}_{1,p} & \cdots & -(\psi^{-1}(0)/\beta_p)\boldsymbol{e}_{p,p} \\ 1/p & \cdots & 1/p \end{cases},$$
(12)

where the function $\psi(\cdot)$ is defined as $\psi(x) = x - 2Q(-x)/Q'(-x)$ for x > 0.

Remark 4.1: It is worthwhile mentioning that the design (12) shown in Corollary 4.1 is also D- and A-optimal.

Remark 4.2: For the case of p = 1, the one point design $\xi^* = \{a, 1\}$ that simultaneously satisfies the additional conditions in Theorem 4.1 is saturated *R*-optimal. However, if the design $\xi^* = \{a, 1\}$ is not *R*-optimal, the proposed method in Theorem 4.2 can be used to search for a saturated *R*-optimal design.

4.2. PSO-generated R-optimal designs

Only finding saturated designs may not be enough for determining the *R*-optimality of a design in the design class Ξ , since the optimal designs depend on the unknown parameters. For example, let $\boldsymbol{\beta} = (0.5, 0.5)^{\top}$ for the Poisson regression model discussed in Example 4.1. The *R*-optimal design (see Figure 3) is then given by

 $\xi^* = \begin{cases} (1.4321, 5) & (5, 1.4321) & (5, 5) \\ 0.4666 & 0.4666 & 0.0668 \end{cases}.$

This means that the number of support points of *R*-optimal designs for models without intercept may exceed the number of regression parameters. Thereby, the aforementioned method by solving equations is unable to determine an *R*-optimal design in Ξ and we require an effective algorithm to generate an *R*-optimal design. Here we employ the PSO algorithm to find optimal designs for the models under consideration and a pseudo code of PSO is described as below.

In Section 4.2, the support points of each position ξ_i must be given by $\mathbf{a} - (x_1/\beta_1)\mathbf{e}_{1,p-1}, \dots, \mathbf{a} - (x_p/\beta_p)\mathbf{e}_{p,p}$ and \mathbf{a} .

The index *t* is equal to 0, 1, . . ., and two criteria can be used to end iteration, achieving a maximum number of iterations or verifying whether the equivalence condition attains a pre-specified threshold.

The notations $v_i^{(t)}$ and $\xi_i^{(t)}$ are, respectively, the current velocity and position for the *i*th particle. θ_t is the inertia weight that modulates the influence of the former velocity, which can be a constant or a decreasing function with



Figure 3. Plot of the function ϕ in (3) for the *R*-optimal design on $\mathcal{X} = [0, 5]^2$ for $\boldsymbol{\beta} = (0.5, 0.5)^{\top}$ in the Poisson regression model without intercept.

Algorithm 1 The pseudo code for PSO

Randomly generate *n* initial positions (designs), $\xi_1^{(0)}, \ldots, \xi_n^{(0)}$, and initial velocities, $v_1^{(0)}, \ldots, v_n^{(0)}$, for *n* particles. Determine personal and global best positions, say ξ_i^{pbest} (*i* = 1, ..., *n*) and ξ^{gbest} , by comparing the criterion function values for *n* particles.

while not stopping criteria do

Update each particle velocity using $v_i^{(t+1)} = \theta_t v_i^{(t)} + \gamma_1 \alpha_1 (\xi_i^{\text{pbest}} - \xi_i^{(t)}) + \gamma_2 \alpha_2 (\xi^{\text{gbest}} - \xi_i^{(t)}).$ Update each particle position using $\xi_i^{(t+1)} = \xi_i^{(t)} + v_i^{(t+1)}.$

Update ξ_i^{pbest} (*i* = 1,...,*n*) and personal best criterion values as well as ξ^{gbest} and global best criterion value. **return** ξ^{gbest} and global best criterion value.

values between 0 and 1. α_1 and α_2 are both random variables from U(0, 1). γ_1 and γ_2 are two constants reflecting the cognitive learning level and social learning level, respectively.

Example 4.3: Consider the proportional hazards regression models with three covariates, in which the linear predictor does not include an intercept term. Let $\mathcal{X} = [0,3]^3$ and $\boldsymbol{\beta} = (-2.5, 0.5, 0.5)^\top$. In order to assess the effect of the amount of censoring *q* on *R*-optimal designs, we adjust the censoring time *c* to achieve overall censoring probabilities of 20%, 40%, 60% and 80% for the balanced design ξ_b . For instance, we choose c = 60 for type I censoring and c = 133 for random censoring when q = 0.4. In this example, the PSO algorithm with 150 particles and 100 iterations is able to find the *R*-optimal design with the required accuracy, which can be implemented by **R** software in less than 20 seconds on a standard PC. Table 4 summarizes the numerical results, including *R*-efficiencies of the balanced design ξ_b and the amount of censoring q^* under the corresponding *R*-optimal design. Some of the points are clear from the numerical results.

- For the three-covariate case, the support points of *R*-optimal designs for type I censoring and random censoring exceed the number of parameters.
- With increasing amounts of censoring, the *R*-optimal designs shift the support point on the edge $(x_1, 3, 3)^{\top}$ towards the vertex $\mathbf{a} = (0, 3, 3)^{\top}$, and the *R*-efficiency of the balanced design ξ_b is reduced gradually.
- The overall probability of censoring q^* under the *R*-optimal design is less than for the balanced design ξ_b .

5. Discussion

The present paper investigates the construction of locally *R*-optimal designs for a large class of nonlinear multiple regression models. For the case of models with intercept, the *R*-optimal designs on a rectangular design region can be determined by utilizing the similar arguments in Schmidt (2019) but finding its optimal weight is different. We notice that the structure of the *R*-optimal designs is similar to those criteria reported in Schmidt (2019), especially in terms of the location of support points. For the case of models without intercept, however, with the same method we can determine the saturated *R*-optimal designs only from two design subclasses addressed in Section 4.1. Moreover,

Table 4. The locally *R*-optimal designs on $\mathcal{X} = [0, 3]^3$ for $\boldsymbol{\beta} = (-2.5, 0.5, 0.5)^\top$ in the proportional hazards regression models, the *R*-efficiencies of the balanced design ξ_b and the overall probability of censoring under the *R*-optimal design.

q	<i>R</i> -optima	al design fo	r type l cen	soring	$\operatorname{Eff}_R(\xi_b)$	<i>q</i> *
0.2	{(3,3,3) 0.2071	(0, 0, 3) 0.3053	(0, 3, 0) 0.3053	(0, 3, 3) 0.1823	0.6286	0.0029
0.4	{(2.6970, 3, 3) 0.2200	(0, 0, 3) 0.3032	(0, 3, 0) 0.3032	(0, 3, 3) 0.1736	0.2170	0.0525
0.6	{(1.2854, 3, 3) 0.2564	(0, 0, 3) 0.2943	(0, 3, 0) 0.2943	(0, 3, 3) 0.1550	0.0193	0.2309
0.8	(1.0354, 3, 3) 0.2762	(0, 0, 3) 0.2887	(0, 3, 0) 0.2887	(0, 3, 3) 0.1464	0.0122	0.6477
	<i>R</i> -optimal	design for	random ce	nsoring		
0.2	{(3, 3, 3) 0.2112	(0, 0, 3) 0.3046	(0, 3, 0) 0.3046	(0, 3, 3) 0.1796	0.6941	0.0192
0.4	{(2.7596, 3, 3) 0.2274	(0, 0, 3) 0.3019	(0, 3, 0) 0.3019	(0, 3, 3) 0.1688	0.2367	0.0796
0.6	(1.4216, 3, 3) 0.2538	(0, 0, 3) 0.2946	(0, 3, 0) 0.2946	(0, 3, 3) 0.1570	0.0244	0.2541
0.8	(1.0763, 3, 3) 0.2735	(0, 0, 3) 0.2882	(0, 3, 0) 0.2882	(0, 3, 3) 0.1501	0.0132	0.6387

Table 5. The *R*-efficiencies of the locally *R*-optimal designs on $\mathcal{X} = [0, 5]^{p-1}$ for various misspecified β for the first-order Poisson regression models with intercept.

<i>p</i> = 2		<i>p</i> = 3		<i>p</i> = 4		
β	$\operatorname{Eff}_{R}(\xi)$	β	$\operatorname{Eff}_{R}(\xi)$	β	$\operatorname{Eff}_{R}(\xi)$	
(0, -0.5)	0.6376	(0, -0.5, -0.5)	0.4091	(0, -0.5, -0.5, -0.5)	0.2649	
(0, -0.8)	0.9467	(0, -0.8, -0.8)	0.8970	(0, -0.8, -0.8, -0.8)	0.8519	
(0, -1.0)	1	(0, -1.0, -1.0)	1	(0, -1.0, -1.0, -1.0)	1	
(0, -1.2)	0.9588	(0, -1.2, -1.2)	0.9200	(0, -1.2, -1.2, -1.2)	0.8816	
(0, -1.5)	0.7990	(0, -1.5, -1.5)	0.6409	(0, -1.5, -1.5, -1.5)	0.5132	
(0, -2.0)	0.4854	(0, -2.0, -2.0)	0.2384	(0, -2.0, -2.0, -2.0)	0.1171	

the PSO algorithm has been used to generate non-saturated *R*-optimal designs for both cases. Some conditions in Theorems 3.1, 4.1 and 4.2 are required, which ensure that the support points are located within the design region. If these conditions are not satisfied, optimal designs may then be having more support points. In addition, a nonlinear system of equations must be solved numerically in order to search for the saturated *R*-optimal designs. Although the existence of the solution to these equations is not proved theoretically, the numerical exploration shows that the solution in all considered examples always exists.

It is worthwhile mentioning that the locally optimal designs discussed so far are derived for a given value of the model parameter vector $\boldsymbol{\beta}$. One might choose such a value of $\boldsymbol{\beta}$ from an initial guess or estimation when some historical observations can be obtained. It might be of interest to study how the design will be affected by wrongly specified parameters. For illustration, we consider the locally *R*-optimal designs for the first-order Poisson regression models with an intercept on the design region $\mathcal{X} = [0, 5]^{p-1}$ for p = 2, 3, 4, respectively. In specific, we assume that the true parameter vector is $\boldsymbol{\beta} = (0, -1)^{\top}$ for $p = 2, \boldsymbol{\beta} = (0, -1, -1)^{\top}$ for p = 3, and $\boldsymbol{\beta} = (0, -1, -1, -1)^{\top}$ for p = 4. We generate the locally *R*-optimal designs for various misspecified values of $\boldsymbol{\beta}$, and calculate their *R*-efficiencies with respect to the locally *R*-optimal design for the true parameter vector for each p = 2, 3, 4, which are shown in Table 5. It is observed from Table 5 that the efficiency of the locally *R*-optimal design for the misspecified value of $\boldsymbol{\beta}$ diverges from its true value, and the loss of *R*-efficiency is disastrous when the difference between the true value and the misspecified value of $\boldsymbol{\beta}$ is relatively serious. Numerical results with other examples yield similar conclusions, which are not reported here for the sake of saving space.

To overcome the parameter dependence of the locally optimal design, a commonly used approach to the computation of locally optimal designs is weighted designs (see, e.g., Atkinson et al., 2007, Chap. 18), where a prior distribution for β , which may be either discrete or continuous, is assumed in advance. Another method is the computation of maximin efficient designs, i.e., maximizing the minimal efficiency with respect to the parameters (see Dette, 1997; Konstantinou et al., 2014).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Lei He's work is supported by the National Natural Science Foundation of China [Grant Number 12101013] and the Natural Science Foundation of Anhui Province [Grant Number 2008085QA15]. Rong-Xian Yue's work is supported by the National Natural Science Foundation of China [Grant Numbers 11971318, 11871143].

ORCID

Lei He http://orcid.org/0000-0001-6943-5123 *Rong-Xian Yue* http://orcid.org/0000-0001-6899-4636

References

- Atkinson, A. C., Donev, A. N., & Tobias, R. D. (2007). Optimum experimental designs, with SAS. Oxford University Press.
- Chernoff, H. (1953). Locally optimal designs for estimating parameters. *The Annals of Mathematical Statistics*, 24(4), 586–602. https://doi.org/10.1214/aoms/1177728915
- Dette, H. (1997). Designing experiments with respect to 'standardized' optimality criteria. *Journal of the Royal Statistical Society:* Series B (Statistical Methodology), 59(1), 97–110. https://doi.org/10.1111/rssb.1997.59.issue-1
- Fedorov, V. V. (1972). Theory of optimal experiments. Academic Press.
- Freise, F., Gaffke, N., & Schwabe, R. (2021). The adaptive Wynn algorithm in generalized linear models with univariate response. *The Annals of Statistics*, 49(2), 702–722. https://doi.org/10.1214/20-AOS1974
- Hao, H., Zhu, X., Zhang, X., & Zhang, C. (2021). *R*-optimal design of the second-order Scheffé mixture model. *Statistics & Probability Letters*, *173*(*C*), 109069. https://doi.org/10.1016/j.spl.2021.109069
- Idais, O., & Schwabe, R. (2021). Analytic solutions for locally optimal designs for gamma models having linear predictors without intercept. *Metrika*, 84(1), 1–26. https://doi.org/10.1007/s00184-019-00760-3
- Kalish, L. A., & Harrington, D. P. (1988). Efficiency of balanced treatment allocation for survival analysis. *Biometrics*, 44(3), 815–821. https://doi.org/10.2307/2531593
- Konstantinou, M., Biedermann, S., & Kimber, A. (2014). Optimal designs for two-parameter nonlinear models with application to survival models. *Statistica Sinica*, 24(1), 415–428. https://doi.org/10.5705/ss.2011.271
- Li, K. H., Lau, T. S., & Zhang, C. (2005). A note on D-optimal designs for models with and without an intercept. *Statistical Papers*, *46*(3),451–458. https://doi.org/10.1007/BF02762844
- Liu, P., Gao, L. L., & Zhou, J. (2022). R-optimal designs for multi-response regression models with multi-factors. Communications in Statistics – Theory and Methods, 51(2), 340–355. https://doi.org/10.1080/03610926.2020.1748655
- Liu, X., & Yue, R.-X. (2020). Elfving's theorem for *R*-optimality of experimental designs. *Metrika*, 83(4), 485–498. https://doi.org/10.1007/s00184-019-00728-3
- Pukelsheim, F., & Torsney, B. (1993). Optimal weights for experimental designs on linearly independent support points. *The Annals of Statistics*, *19*(*3*), 1614–1625. https://doi.org/10.2307/2241966
- Radloff, M., & Schwabe, R. (2019). Locally *D*-optimal designs for non-linear models on the *k*-dimensional ball. *Journal of Statistical Planning and Inference*, 203, 106–116. https://doi.org/10.1016/j.jspi.2019.03.004
- Russell, K. G., Woods, D. C., Lewis, S. M., & Eccleston, E. C. (2009). *D*-optimal designs for Poisson regression models. *Statistica Sinica*, 19(2), 721–730. https://doi.org/10.2307/24308852
- Schmidt, D. (2019). Characterization of *c*-, *L* and ϕ_k -optimal designs for a class of non-linear multiple-regression models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 81(1), 101–120. https://doi.org/10.1111/rssb.12292
- Schmidt, D., & Schwabe, R. (2015). On optimal designs for censored data. *Metrika*, 78(3), 237–257. https://doi.org/10.1007/s00184-014-0500-1
- Schmidt, D., & Schwabe, R. (2017). Optimal design for multiple regression with information driven by the linear predictor. *Statistica Sinica*, *27*(*3*), 1371–1384. https://doi.org/10.5705/ss.202015.0385

Silvey, S. D. (1980). Optimal design. Chapman and Hall.

- Whittle, P. (1973). Some general points in the theory of optimal experimental designs. *Journal of the Royal Statistical Society: Series B (Methodological)*, 35(1), 123–130. https://doi.org/10.1111/j.2517-6161.1973.tb00944.x
- Yang, M., Biedermann, S., & Tang, E. (2013). On optimal designs for nonlinear models: A general and efficient algorithm. Journal of the American Statistical Association, 108(504), 1411–1420. https://doi.org/10.1080/01621459.2013.806268

Appendix

Proof of Theorem 2.1: Let $f_{\beta}(\mathbf{x}) = \sqrt{Q(f(\mathbf{x})^{\top}\boldsymbol{\beta})f(\mathbf{x})}$. The Fisher information matrix of approximate design ξ given by (1) can then be written as

$$\boldsymbol{M}(\boldsymbol{\xi},\boldsymbol{\beta}) = \int_{\mathcal{X}} \boldsymbol{f}_{\boldsymbol{\beta}}(\boldsymbol{x}) \boldsymbol{f}_{\boldsymbol{\beta}}(\boldsymbol{x})^{\top} \mathrm{d}\boldsymbol{\xi}(\boldsymbol{x}).$$

118 😉 L. HE AND R.-X. YUE

For any designs $\xi, \bar{\xi} \in \Xi$ and $\alpha \in (0, 1)$, define $\xi_{\alpha} = (1 - \alpha)\xi + \alpha \bar{\xi}$. We have $M(\xi_{\alpha}, \beta) = (1 - \alpha)M(\xi, \beta) + \alpha M(\bar{\xi}, \beta)$ and

$$\frac{\partial \psi(\xi_{\alpha}, \boldsymbol{\beta})}{\partial \alpha} = \frac{\partial}{\partial \alpha} \left[\prod_{j=1}^{p} \boldsymbol{e}_{j,p}^{\top} \boldsymbol{M}^{-1}(\xi_{\alpha}, \boldsymbol{\beta}) \boldsymbol{e}_{j,p} \right]$$
$$= \sum_{j=1}^{p} \left(\prod_{\substack{i=1\\i\neq j}}^{p} \boldsymbol{e}_{i,p}^{\top} \boldsymbol{M}^{-1}(\xi_{\alpha}, \boldsymbol{\beta}) \boldsymbol{e}_{i,p} \right) \frac{\partial}{\partial \alpha} \left(\boldsymbol{e}_{j,p}^{\top} \boldsymbol{M}^{-1}(\xi_{\alpha}, \boldsymbol{\beta}) \boldsymbol{e}_{j,p} \right),$$

where

$$\frac{\partial}{\partial \alpha} \left(\boldsymbol{e}_{j,p}^{\top} \boldsymbol{M}^{-1}(\xi_{\alpha}, \boldsymbol{\beta}) \boldsymbol{e}_{j,p} \right) = \boldsymbol{e}_{j,p}^{\top} \left[\boldsymbol{M}^{-1}(\xi_{\alpha}, \boldsymbol{\beta}) \left(\boldsymbol{M}(\xi, \boldsymbol{\beta}) - \boldsymbol{M}(\bar{\xi}, \boldsymbol{\beta}) \right) \boldsymbol{M}^{-1}(\xi_{\alpha}, \boldsymbol{\beta}) \right] \boldsymbol{e}_{j,p}.$$

The directional derivative of ψ at ξ in the direction of $\overline{\xi}$, denoted by $\nabla_{\psi}(\xi, \overline{\xi})$, is then given by

$$\begin{aligned} \nabla_{\psi}(\xi,\bar{\xi}) &= \lim_{\alpha \to 0^+} \frac{\partial \psi(\xi_{\alpha},\boldsymbol{\beta})}{\partial \alpha} \\ &= \sum_{j=1}^{p} \frac{\psi(\xi)}{\boldsymbol{e}_{j,p}^{\top} \boldsymbol{M}^{-1}(\xi,\boldsymbol{\beta}) \boldsymbol{e}_{j,p}} \boldsymbol{e}_{j,p}^{\top} \left[\boldsymbol{M}^{-1}(\xi,\boldsymbol{\beta}) \left(\boldsymbol{M}(\xi,\boldsymbol{\beta}) - \boldsymbol{M}(\bar{\xi},\boldsymbol{\beta}) \right) \boldsymbol{M}^{-1}(\xi,\boldsymbol{\beta}) \right] \boldsymbol{e}_{j,p} \\ &= \psi(\xi) \sum_{j=1}^{p} \left(1 - \frac{\boldsymbol{e}_{j,p}^{\top} \boldsymbol{M}^{-1}(\xi,\boldsymbol{\beta}) \boldsymbol{M}(\bar{\xi},\boldsymbol{\beta}) \boldsymbol{M}^{-1}(\xi,\boldsymbol{\beta}) \boldsymbol{e}_{j,p}}{\boldsymbol{e}_{j,p}^{\top} \boldsymbol{M}^{-1}(\xi,\boldsymbol{\beta}) \boldsymbol{e}_{j,p}} \right) \\ &= \psi(\xi) \left(p - \operatorname{tr} \left\{ \boldsymbol{M}^{-1}(\xi,\boldsymbol{\beta}) \boldsymbol{M}(\bar{\xi},\boldsymbol{\beta}) \boldsymbol{M}^{-1}(\xi,\boldsymbol{\beta}) \sum_{j=1}^{q} \frac{\boldsymbol{e}_{j,p} \boldsymbol{e}_{j,p}^{\top}}{\boldsymbol{e}_{j,p}^{\top} \boldsymbol{M}^{-1}(\xi,\boldsymbol{\beta}) \boldsymbol{e}_{j,p}} \right\} \right). \end{aligned}$$

Note that the directional derivative $\nabla_{\psi}(\xi, \overline{\xi})$ is linear in $\overline{\xi}$ for any fixed $\xi \in \Xi$, i.e.,

$$\nabla_{\psi}(\xi,\bar{\xi}) = \int_{\mathcal{X}} \nabla_{\psi}(\xi,\delta_{\mathbf{x}}) \,\mathrm{d}\bar{\xi}(\mathbf{x}),$$

where δ_x is the Dirac measure at x. Following Whittle (1973), the design $\xi^* \in \Xi$ is R-optimal if and only if $\inf_{x \in \mathcal{X}} \nabla_{\psi}(\xi, \delta_x) = 0$. As a consequence the assertion of Theorem 2.1 follows.

Proof of Theorem 3.2: From Theorem 3.1, *R*-optimal design on $\mathcal{X} = [0, \infty)^{p-1}$ for $\beta_i < 0, i = 1, \dots, p-1$, has the following form

$$\xi = \begin{cases} -(x_1/\beta_1)\boldsymbol{e}_{1,p-1} & \cdots & -(x_{p-1}/\beta_{p-1})\boldsymbol{e}_{p-1,p-1} & \boldsymbol{0}_{p-1} \\ \omega_1 & \cdots & \omega_{p-1} & \omega_p \end{cases},$$
(A1)

,

where $\mathbf{0}_{p-1}$ is a vector of (p-1) 0's, x_i and ω_i , i = 1, ..., p-1, can be determined by the common system of Equations (4) and (6) for convenience. We denote $Q_i = Q(\beta_0 - x_i)$ for i = 1, ..., p-1 and $Q_p = Q(\beta_0)$ according to the support points of the design (A1). Employing the previously mentioned decomposition approach and letting $\mathbf{y} = (\beta_1/x_1, ..., \beta_{p-1}/x_{p-1})^{\top}$ and $\mathbf{A} = \text{diag}((Q_p \omega_p)^{-1}, (Q_1 \omega_1)^{-1}, ..., (Q_{p-1} \omega_{p-1})^{-1})$, we can obtain that the inverse of the information matrix of ξ is given by

$$M(\xi, \boldsymbol{\beta})^{-1} = \boldsymbol{X}^{-1} \boldsymbol{Q}^{-1/2} \boldsymbol{D}_{\omega}^{-1} \boldsymbol{Q}^{-1/2} (\boldsymbol{X}^{\top})^{-1}$$
$$= \begin{pmatrix} 1 & \boldsymbol{0}_{p-1}^{\top} \\ \boldsymbol{y} & -\operatorname{diag}(\boldsymbol{y}) \end{pmatrix} \boldsymbol{\Lambda} \begin{pmatrix} 1 & \boldsymbol{y}^{\top} \\ \boldsymbol{0}_{p-1} & -\operatorname{diag}(\boldsymbol{y}) \end{pmatrix} = \begin{pmatrix} \frac{1}{Q_{p}\omega_{p}} & \frac{1}{Q_{p}\omega_{p}} \boldsymbol{y}^{\top} \\ \frac{1}{Q_{p}\omega_{p}} \boldsymbol{y} & -\operatorname{diag}(\tilde{\boldsymbol{y}}) \end{pmatrix},$$

where $\tilde{\boldsymbol{y}} = (\beta_1^2 \tilde{\lambda}_{1p} / x_1^2, \dots, \beta_{p-1}^2 \tilde{\lambda}_{p-1,p} / x_{p-1}^2)^\top$ with $\tilde{\lambda}_{ip} = (Q_i \omega_i)^{-1} + (Q_p \omega_p)^{-1}$, $i = 1, \dots, p-1$. Then the matrix \boldsymbol{S} defined in (5) is given by

$$\mathbf{S} = \begin{pmatrix} \omega_p + \frac{1}{Q_p} \sum_{i=1}^{p-1} \frac{1}{\tilde{\lambda}_{ip}} & -\frac{1}{\sqrt{Q_1 Q_p} \tilde{\lambda}_{1p}} & \cdots & -\frac{1}{\sqrt{Q_{p-1} Q_p} \tilde{\lambda}_{p-1,p}} \\ -\frac{1}{\sqrt{Q_1 Q_p} \tilde{\lambda}_{1p}} & \frac{1}{Q_1 \tilde{\lambda}_{1p}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{\sqrt{Q_{p-1} Q_p} \tilde{\lambda}_{p-1,p}} & 0 & \cdots & \frac{1}{Q_{p-1} \tilde{\lambda}_{p-1,p}} \end{pmatrix}$$

which is entirely unrelated to the parameters $(\beta_1, \ldots, \beta_{p-1})^{\top}$. Hence, the solutions of x_i and ω_i for $i = 1, \ldots, p-1$ do not depend on the parameters $(\beta_1, \ldots, \beta_{p-1})^{\top}$.

In order to prove the following Theorems 4.1 and 4.2, the extended design region with $x_i \in (-\infty, v_i]$ for $\beta_i > 0$ and $x_i \in [u_i, \infty)$ for $\beta_i < 0$ will be considered.

Proof of Theorem 4.1: We will only prove the case $\tau = 1$, and the others can be treated similarly. With the previous decomposition strategy the information matrix of ξ_1^* can be written as

$$\boldsymbol{M}(\boldsymbol{\xi}_1^*,\boldsymbol{\beta}) = \tilde{\boldsymbol{X}}^{\top} \boldsymbol{D}_{\boldsymbol{\omega}} \tilde{\boldsymbol{X}} = \boldsymbol{X}^{\top} \boldsymbol{Q}^{1/2} \boldsymbol{D}_{\boldsymbol{\omega}} \boldsymbol{Q}^{1/2} \boldsymbol{X},$$

where

$$\mathbf{X} = \begin{pmatrix} a_1 & \mathbf{a}_{-1}^{\mathsf{T}} \\ a_1 \mathbf{1}_{p-1} & \mathbf{1}_{p-1} \mathbf{a}_{-1}^{\mathsf{T}} - \operatorname{diag}(\mathbf{x}_{-1}^*) \end{pmatrix}$$

with $\boldsymbol{a}_{-1} = (a_2, \dots, a_p)^\top$, $\mathbf{1}_{p-1} = (1, \dots, 1)^\top$ and $\boldsymbol{x}_{-1}^* = (x_2^*/\beta_2, \dots, x_p^*/\beta_p)^\top$. Denote $\boldsymbol{y}_{-1}^* = (\beta_2/x_2^*, \dots, \beta_p/x_p^*)^\top$ and $\boldsymbol{z}_{-1}^* = (a_2\beta_2/x_2^*, \dots, a_p\beta_p/x_p^*)^\top$, and then we have

$$X^{-1} = \begin{pmatrix} \frac{1 - \boldsymbol{z}_{-1}^{*\top} \boldsymbol{1}_{p-1}}{a_1} & \frac{1}{a_1} \boldsymbol{z}_{-1}^{*\top} \\ \boldsymbol{y}_{-1}^{*} & -\operatorname{diag}(\boldsymbol{y}_{-1}^{*}) \end{pmatrix}.$$

It follows that $(X^{\top})^{-1}a = e_{1,p}$ and $(X^{\top})^{-1}(a - (x_i^*/\beta_i)e_{i,p}) = e_{i,p}, i = 2, ..., p$.

Now letting $\ell_i(x_i) = \ell_i(x_i, \xi^*)$ and $h_i(x_i) = h_i(x_i, \xi^*)$ which are defined in (7) to simplify writing, and using the formulas (4) and (5) we obtain

$$\ell_i(a_i) = \boldsymbol{e}_{1,p}^\top \boldsymbol{Q}^{-1/2} \boldsymbol{D}_{\omega}^{-1} \boldsymbol{S} \boldsymbol{D}_{\omega}^{-1} \boldsymbol{Q}^{-1/2} \boldsymbol{e}_{1,p} = \frac{s_{11}}{(\omega_1^*)^2 Q(\boldsymbol{a}^\top \boldsymbol{\beta})} = \frac{p}{Q(\boldsymbol{a}^\top \boldsymbol{\beta})} = p_1(a_1),$$

and $\ell_i(a_i - x_i^*/\beta_i) = p/Q(\mathbf{a}^\top \boldsymbol{\beta} - x_i^*)$ for i = 2, ..., p. Thus we have $h_1(a_1) = h_i(a_i) = h_i(a_i - x_i^*/\beta_i) = 0$ for i = 2, ..., p. As in the proof of Theorem 9 in Schmidt (2019), it is easily shown that $h'_i(x_i)$ has at most two roots, $\lim_{x_i \to \pm \infty} h_i(x_i) = \pm \infty$ for $\beta_i > 0$ and $\lim_{x_i \to \pm \infty} h_i(x_i) = \mp \infty$ for $\beta_i < 0$. To prove the *R*-optimality of ξ^* , it suffices here to show that $h'_i(a_i - x_i^*/\beta_i) = 0$ holds since $h_1(x_1) < 0$ for $x_1 \in [u_1, v_1] \setminus \{a_1\}$ and $h_i(a_i) = 0$ for all i = 2, ..., p. We have

$$\ell'_{i}(x) = 2e_{i,p}^{\top}M(\xi^{*}, \beta)^{-1}\left(\sum_{j=1}^{p} \frac{e_{j,p}e_{j,p}^{\top}}{e_{j,p}^{\top}M(\xi^{*}, \beta)^{-1}e_{j,p}}\right)M(\xi^{*}, \beta)^{-1}(a + (x - a_{i})e_{i,p}).$$

For i = 2, ..., p, utilizing $\mathbf{e}_{i,p}^{\top} \mathbf{X}^{-1} = (\beta_i / x_i^*)(1, -\mathbf{e}_{i,p-1}^{\top})$ we obtain

$$\ell_i'\left(a_i - \frac{x_i^*}{\beta_i}\right) = \frac{2\beta_i}{x_i^*}(1, -e_{i,p-1}^\top)\mathbf{Q}^{-1/2}\mathbf{D}_{\omega}^{-1}\mathbf{S}\mathbf{D}_{\omega}^{-1}\mathbf{Q}^{-1/2}e_{i,p}$$

$$= \frac{2(\beta_i/x_i^*)s_{1i}}{\omega_1^*\omega_i^*\sqrt{Q(\mathbf{a}^\top\boldsymbol{\beta})}\sqrt{Q(\mathbf{a}^\top\boldsymbol{\beta} - x_i^*)}} - \frac{2(\beta_i/x_i^*)s_{ii}}{\omega_i^{*2}Q(\mathbf{a}^\top\boldsymbol{\beta} - x_i^*)}.$$

Hence

$$h_{i}'\left(a_{i}-\frac{x_{i}^{*}}{\beta_{i}}\right)=p\left[\frac{2(\beta_{i}/x_{i}^{*})}{\sqrt{Q(a^{\top}\beta)}\sqrt{Q(a^{\top}\beta-x_{i}^{*})}}\frac{s_{1i}}{\sqrt{s_{11}s_{ii}}}-\frac{2(\beta_{i}/x_{i}^{*})}{Q(a^{\top}\beta-x_{i}^{*})}+\frac{\beta_{i}Q'(a^{\top}\beta-x_{i}^{*})}{(Q(a^{\top}\beta-x_{i}^{*}))^{2}}\right].$$

The equations $h'_i(a_i - x_i^*/\beta_i) = 0$ are equivalent to the equations

$$x_i^* - 2\frac{Q(\boldsymbol{a}^\top \boldsymbol{\beta} - x_i^*)}{Q'(\boldsymbol{a}^\top \boldsymbol{\beta} - x_i^*)} \left[1 - \sqrt{\frac{Q(\boldsymbol{a}^\top \boldsymbol{\beta} - x_i^*)}{Q(\boldsymbol{a}^\top \boldsymbol{\beta})}} \frac{s_{1i}}{\sqrt{s_{11}s_{ii}}} \right] = 0$$

for i = 2, ..., p. If the x_i^* and ω_i^* are the solutions of this system of equations combined with Equation (4), then the design ξ_1^* is *R*-optimal.

Proof of Theorem 4.2: With the same decomposition method we have

$$\boldsymbol{M}(\boldsymbol{\xi}^*,\boldsymbol{\beta}) = \tilde{\boldsymbol{X}}^\top \boldsymbol{D}_{\boldsymbol{\omega}} \tilde{\boldsymbol{X}} = \boldsymbol{X}^\top \boldsymbol{Q}^{1/2} \boldsymbol{D}_{\boldsymbol{\omega}} \boldsymbol{Q}^{1/2} \boldsymbol{X},$$

where $X = \mathbf{1}_{p} a^{\top} - \text{diag}(x^{*})$ and $x^{*} = (x_{1}^{*}/\beta_{1}, \dots, x_{p}^{*}/\beta_{p})^{\top}$. Let $y^{*} = (\beta_{1}/x_{1}^{*}, \dots, \beta_{p}/x_{p}^{*})^{\top}$ and $z^{*} = (a_{1}\beta_{1}/x_{1}^{*}, \dots, a_{p}\beta_{p}/x_{p}^{*})^{\top}$, and then we get

$$X^{-1} = -\operatorname{diag}(\boldsymbol{y}^*) - \frac{\boldsymbol{y}^* \boldsymbol{z}^{*\top}}{1 - \boldsymbol{z}^{*\top} \boldsymbol{1}_p}$$

and $(\mathbf{X}^{\top})^{-1}(\mathbf{a} - (x_i^*/\beta_i)\mathbf{e}_{i,p}) = \mathbf{e}_{i,p}$. We still let $\ell_i(x_i) = \ell_i(x_i, \xi^*)$ and $h_i(x_i) = h_i(x_i, \xi^*)$ to simplify writing. It follows from the formulas (4) and (5) that

$$\ell_i(a_i - x_i^*/\beta_i) = \mathbf{e}_{i,p}^\top \mathbf{Q}^{-1/2} \mathbf{D}_{\omega}^{-1} \mathbf{S} \mathbf{D}_{\omega}^{-1/2} \mathbf{e}_{i,p} = \frac{s_{i,i}}{(\omega_i^*)^2 Q(\mathbf{a}^\top \boldsymbol{\beta} - x_i^*)} = \frac{p}{Q(\mathbf{a}^\top \boldsymbol{\beta} - x_i^*)}$$

Thus we have $h_i(a_i - x_i^*/\beta_i) = 0$ for i = 1, ..., p. Due to the fact that $h'_i(x_i)$ has at most two roots, $\lim_{x_i \to \pm \infty} h_i(x_i) = \pm \infty$ for $\beta_i > 0$ and $\lim_{x_i \to \pm \infty} h_i(x_i) = \mp \infty$ for $\beta_i < 0$. To prove the *R*-optimality of ξ^* , it suffices here to show that $h'_i(a_i - x_i^*/\beta_i) = 0$

120 🕒 L. HE AND R.-X. YUE

holds since $h_i(a_i) < 0$ for all $i = 1, \ldots, p$. We have

$$\ell'_{i}(x) = 2e_{i,p}M(\xi^{*}, \beta)^{-1}\left(\sum_{j=1}^{p} \frac{e_{j,p}e_{j,p}^{\top}}{e_{j,p}^{\top}M(\xi^{*}, \beta)^{-1}e_{j,p}}\right)M(\xi^{*}, \beta)^{-1}(a + (x - a_{i})e_{i,p}).$$

With $\boldsymbol{e}_{i,p}^{\top}\boldsymbol{X}^{-1} = -(\beta_i/x_i^*)(\boldsymbol{e}_{i,p}^{\top} + \boldsymbol{z}^{*\top}/(1 - \boldsymbol{z}^{*\top}\boldsymbol{1}_p))$ we have

$$\ell_{i}'\left(a_{i}-\frac{x_{i}^{*}}{\beta_{i}}\right) = -\frac{2\beta_{i}}{x_{i}^{*}}\left(\boldsymbol{e}_{i,p}^{\top}+\frac{\boldsymbol{z}^{*\top}}{1-\boldsymbol{z}^{*\top}\boldsymbol{1}_{p}}\right)\boldsymbol{Q}^{-1/2}\boldsymbol{D}_{\omega}^{-1}\boldsymbol{S}\boldsymbol{D}_{\omega}^{-1}\boldsymbol{Q}^{-1/2}\boldsymbol{e}_{i,p}$$
$$= -\frac{2(\beta_{i}/x_{i}^{*})s_{ii}}{\omega_{i}^{*2}\boldsymbol{Q}(\boldsymbol{a}^{\top}\boldsymbol{\beta}-x_{i}^{*})} - \frac{2(\beta_{i}/x_{i}^{*})}{(1-\boldsymbol{z}^{*T}\boldsymbol{1}_{p})\sqrt{\boldsymbol{Q}(\boldsymbol{a}^{\top}\boldsymbol{\beta}-x_{i}^{*})}}\sum_{j=1}^{p}\frac{z_{j}s_{ij}}{\omega_{i}^{*}\omega_{j}^{*}\sqrt{\boldsymbol{Q}(\boldsymbol{a}^{\top}\boldsymbol{\beta}-x_{j}^{*})}}$$

Hence

$$h_{i}'\left(a_{i} - \frac{x_{i}^{*}}{\beta_{i}}\right) = p \left[-\frac{2(\beta_{i}/x_{i}^{*})}{Q(\boldsymbol{a}^{\top}\boldsymbol{\beta} - x_{i}^{*})} - \frac{2(\beta_{i}/x_{i}^{*})}{(1 - \boldsymbol{z}^{*T}\boldsymbol{1}_{p})\sqrt{Q(\boldsymbol{a}^{\top}\boldsymbol{\beta} - x_{i}^{*})}} \sum_{j=1}^{p} \frac{z_{j}s_{ij}/\sqrt{s_{ii}s_{jj}}}{\sqrt{Q(\boldsymbol{a}^{\top}\boldsymbol{\beta} - x_{j}^{*})}} + \frac{\beta_{i}Q'(\boldsymbol{a}^{\top}\boldsymbol{\beta} - x_{i}^{*})}{(Q(\boldsymbol{a}^{\top}\boldsymbol{\beta} - x_{i}^{*}))^{2}} \right].$$

The equations $h_i'(a_i - x_i^*/\beta_i) = 0$ (i = 1, ..., p) are equivalent to

$$x_{i}^{*}-2\frac{Q(a^{\top}\beta-x_{i}^{*})}{Q'(a^{\top}\beta-x_{i}^{*})}\left[\frac{1}{1-z^{*\top}\mathbf{1}_{p}}\sum_{j=1}^{p}\frac{z_{j}\sqrt{Q(a^{\top}\beta-x_{i}^{*})}}{\sqrt{Q(a^{\top}\beta-x_{j}^{*})}}\frac{s_{ij}}{\sqrt{s_{ii}s_{jj}}}+1\right]=0, \quad i=1,\ldots,p.$$

If the x_i^* and ω_i^* are the solutions of this system of equations combined with Equation (4), then the design ξ^* is *R*-optimal.