An equivalence theorem for design optimality with respect to a multi-objective criterion

Chiara Tommasi[1] · Juan M. Rodríguez-Díaz[2] · Jesús F. López-Fidalgo[3,4]

Abstract
Maxi-min efficiency criteria are a kind of multi-objective criteria, since they enable us to take into consideration several tasks expressed by different component-wise criteria. However, they are difficult to manage because of their lack of differentiability. As a consequence, maxi-min efficiency designs are frequently built through heuristic and ad hoc algorithms, without the possibility of checking for their optimality. The main contribution of this study is to prove that the maxi-min efficiency optimality is equivalent to a Bayesian criterion, which is differentiable. In addition, we provide an analytic method to find the prior probability associated with a maxi-min efficient design, making feasible the application of the equivalence theorem. Two illustrative examples show how the proposed theory works.

1 Introduction

In this study, we aim at solving a multi-objective optimization problem that consists in the maximization of a minimum design-efficiency. In optimal design literature,

✉ Juan M. Rodríguez-Díaz
juanmrod@usal.es

Chiara Tommasi
chiara.tommasi@unimi.it

Jesús F. López-Fidalgo
fidalgo@unav.es

[1] Department of Economics, Management and Quantitative Methods, University of Milan, via Conservatorio, 7, 20122 Milan, Italy

[2] Department of Statistics, Faculty of Science, IUFFyM, University of Salamanca, Pl. de los Caidos s/n, 37008 Salamanca, Spain

[3] Institute of Data Science and Artificial Intelligence, University of Navarra, Pamplona, Spain

[4] Tecnun Escuela de Ingeniería, University of Navarra, Pamplona, Spain

different approaches can be classified as maxi-min efficiency criteria. The standardized max-min criterion, introduced to tackle the problem of parameter uncertainty, is the most common. This last issue, however, is not considered in this work because it has been already extensively studied (see for instance, Chen et al. 2015; Dette and Biedermann 2003; Nyquist 2013; Fackle-Fornius et al. 2015; Dette et al. 2007, among others); furthermore, parameter uncertainty is not easily interpretable as a multi-task problem. Differently, examples of maxi-min efficiency criteria that can be interpreted as multi-objective problems are: the SMV-criterion (proposed by Dette (1997)), which aims to obtain an accurate estimation of each one of the model parameters, taking into account their different scale (see also López-Fidalgo and Tommasi 2004 and the references therein) and the extensions of T- and KL-criteria (proposed by Atkinson and Fedorov (1975) and Tommasi et al. (2016), respectively) to handle the problem of model uncertainty. Another interesting application might be the identification of an optimal design for model identification, precise parameter estimation and accurate predictions. This multiple objective could be achieved by maximizing the minimum efficiency of three criteria reflecting these three distinct goals.

The maxi-min approach arises naturally when we wish to protect against the worst case scenario; however, it is difficult to compute the corresponding optimal design (the maxi-min efficiency design) because this criterion is not differentiable. Consequently, a standard directional derivative argument cannot be applied to check whether a given design is optimal because unfortunately, the directional derivative involves an unknown measure; see for instance Wong (1992) and Atkinson and Fedorov (1975).

In addition, the construction of the maxi-min efficiency design is not straightforward at all. Frequently, it is found numerically by the application of some algorithm, but there is no way to prove that it is really the optimum.

The main contribution of this study is to prove the equivalence between the maxi-min efficiency approach and the Bayesian criterion for a specific prior, which is differentiable. Hence, the directional derivative of the Bayesian criterion can be used to check for the minimum efficiency optimality. Let us note that the Bayesian criterion is another kind of multi-objective optimality function, being a convex combination of different quantities. The connection between maxi-min efficiency and Bayesian optimum designs has been already explored by other authors, see for instance Schervish (1995), Müller and Pazman (1998) and Dette et al. (2007); other versions of the equivalence theorem can be found but they are specialized for specific problems; for instance, Dette and Biedermann (2003) or Berger et al. (2000) consider parameter uncertainty in a non-linear model and the D-criterion.

In this study, we prove a more general version of the equivalence theorem, because it covers any multi-objective problem that can be expressed as a minimum design-efficiency (for any component-wise criteria). Furthermore, following similar ideas as in Chen et al. (2017), we provide a method to determine the prior probability that matches the maxi-min efficiency criterion and the Bayesian optimality; this makes possible the application of the equivalence theorem.

The paper is organized as follows. In Sect. 2, we recall some background information and the used notation. In Sect. 3, we state the equivalence theorem and the rule to determine the prior probability that makes the minimum efficiency and the Bayesian

criteria equivalent. Section 4 concerns a pair of illustrative examples. Section 5 provides some conclusions, and finally Appendix A includes the proofs of the theoretical results.

2 Background and notation

In this section, we introduce the main ideas of optimal experimental design and the notation used in what follows.

Let us assume that $f(y, x, \theta)$ is a statistical model that describes the response Y at the experimental condition x, which may be chosen in a compact set $\mathcal{X}$ and $\theta \in \Theta \subseteq \mathbb{R}^p$ denotes a $p \times 1$ parameter vector.

An approximate design is a probability measure on the *design space* $\mathcal{X}$ with a finite support, i.e.

$$\xi = \left\{ \begin{array}{cccc} x_1 & x_2 & \cdots & x_r \\ \xi(x_1) & \xi(x_2) & \cdots & \xi(x_r) \end{array} \right\},$$

where $\xi(x_i) \approx n_i/n$, and n_i is the number of observations to be taken at the experimental condition $x_i, i = 1, \ldots, r$.

The aim is to find a design ξ_θ^* maximizing (minimizing) a concave (convex) optimality criterion function $\Phi(\xi; \theta)$ defined on the space of all designs Ξ to the real line. This means that an optimal design ξ_θ^* may be found according to several criteria reflecting different inferential goals: parameter estimation, prediction or model discrimination. Many optimality criteria for the precise estimation of θ are concave (or convex) functions of the information matrix of a design $\xi \in \Xi$, i.e. $\Phi(\xi; \theta) = \Phi[M(\xi, \theta)]$, where

$$M(\xi, \theta) = \int_{\mathcal{X}} \mathrm{E}_Y \left\{ \frac{\partial \log f(y, x, \theta)}{\partial \theta} \frac{\partial \log f(y, x, \theta)}{\partial \theta^T} \right\} d\xi(x). \qquad (1)$$

If $\Phi(\xi; \theta)$ is a non-negative concave function, then a measure of the goodness of a design ξ with respect to the optimal design ξ_θ^*, is the following efficiency function:

$$0 \leq \mathrm{Eff}(\xi, \theta) = \frac{\Phi(\xi, \theta)}{\Phi(\xi_\theta^*, \theta)} \leq 1. \qquad (2)$$

If $\Phi(\xi; \theta)$ is convex, then the ratio on the right-hand side of Eq. (2) should be reversed.

3 Minimum efficiency and pseudo-Bayesian criteria

Let $\Phi_i(\xi; \theta_i)$ with $i = 1, \ldots, k$ be k different concave optimality criteria, that reflect distinct goals and possibly depend on some unknown parameter vector θ_i. Let θ_{0i} be a guessed value for θ_i; thus, $\xi_i^* = \xi_{i;\theta_{0i}}^* = \arg \max_{\xi \in \Xi} \Phi_i(\xi; \theta_{0i})$ are local optimum designs. When we are interested in a compromise design that is 'good' for all the different criteria, we need to combine $\Phi_i(\xi; \theta_i)$, for $i = 1, \ldots, k$, in a multi-objective criterion. To this aim, as suggested by Dette (1997), we should

first standardize the criteria $\Phi_i(\xi) = \Phi_i(\xi; \theta_{0i})$, obtaining their efficiency functions: $\mathrm{Eff}_i(\xi) = \Phi_i(\xi)/\Phi_i(\xi_i^*)$, $i = 1, \ldots, k$.

An easy way of combining the standardized criteria is through a linear combination. If we have same prior knowledge about criteria $\Phi_i(\xi)$ for $i = 1, \ldots, k$, we might compute a Bayesian optimum design maximizing the following criterion:

$$\Phi_B(\xi; \pi) = \sum_{i=1}^{k} \pi_i \cdot \mathrm{Eff}_i(\xi), \quad 0 \le \pi_i \le 1, \quad \sum_{i=1}^{k} \pi_i = 1, \tag{3}$$

where $\pi^T = (\pi_1, \ldots, \pi_k)$ is a prior probability on the set $\{1, \ldots, k\}$. For an application of this criterion, see for instance Tommasi and López-Fidalgo (2010).

A design ξ_π^* is Bayesian optimal if and only if $\partial \Phi_B(\xi_\pi^*, \bar{\xi}; \pi) \le 0$ for any $\bar{\xi}$, where

$$\partial \Phi_B(\xi_\pi^*, \bar{\xi}; \pi) = \int \sum_{i=1}^{k} \pi_i \frac{\partial \Phi_i(\xi_\pi^*, \xi_x)}{\Phi_i(\xi_i^*)} d\bar{\xi}(dx)$$

is the directional derivative of criterion (3) at ξ_π^* in the direction of $\bar{\xi} - \xi_\pi^*$, and $\partial \Phi_i(\xi_\pi^*, \xi_x)$ denotes the directional derivative of the component-wise criterion $\Phi_i(\cdot)$ at ξ_π^* in the direction of $\xi_x - \xi_\pi^*$. It is easy to prove that ξ_π^* is a Bayesian optimal design if and only if it satisfies the following inequality:

$$\sum_{i=1}^{k} \pi_i \frac{\partial \Phi_i(\xi_\pi^*, \xi_x)}{\Phi_i(\xi_i^*)} \le 0, \quad x \in \mathcal{X}, \tag{4}$$

and that

$$\sum_{i=1}^{k} \pi_i \frac{\partial \Phi_i(\xi_\pi^*, \xi_x)}{\Phi_i(\xi_i^*)} = 0, \quad \text{at the support points of } \xi_\pi^*. \tag{5}$$

When we are unable to provide a prior distribution π, another possibility to takes into consideration all the objectives represented by the k different criteria is the following minimum efficiency criterion:

$$\Phi(\xi) = \min_{i \in \{1, \ldots, k\}} \mathrm{Eff}_i(\xi) = \left[\max_{i \in \{1, \ldots, k\}} \frac{1}{\mathrm{Eff}_i(\xi)} \right]^{-1}.$$

This multi-objective optimality function, differently from the previous one, is not differentiable, and thus the computation of Φ-optimal designs is not straightforward at all.

A design ξ^* is a *maxi-min efficiency design* if and only if

$$\xi^* = \arg\max_{\xi} \min_{i \in \{1, \ldots, k\}} \mathrm{Eff}_i(\xi) = \arg\min_{\xi} \max_{i \in \{1, \ldots, k\}} \frac{1}{\mathrm{Eff}_i(\xi)}.$$

From the last equation, ξ^* is also the design that minimizes the maximum inefficiency optimality criterion:

$$\Phi^{-1}(\xi) = \max_{i \in \{1,...,k\}} \frac{1}{\text{Eff}_i(\xi)}. \tag{6}$$

We find maxi-min efficiency designs by minimizing $\Phi^{-1}(\xi)$, for which we can state the following propositions:

Proposition 1 *The maximum inefficiency criterion $\Phi^{-1}(\xi)$ is a convex function.*

The proof is straightforward.

Proposition 2 *The directional derivative of $\Phi^{-1}(\xi)$ at ξ in the direction of $\bar{\xi} - \xi$ is*

$$\partial \Phi^{-1}(\xi; \bar{\xi}) = \max_{e_i \in \mathcal{C}(\xi)} \int_{\mathcal{X}} \psi(x, e_i, \xi) \bar{\xi}(dx),$$

where e_i denotes the canonical vector of the Euclidean space,

$$\mathcal{C}(\xi) = \left\{ e_i : i = \arg \max_{j \in \{1,...k\}} \frac{1}{\text{Eff}_j(\xi)} \right\} = \left\{ e_i : i = \arg \min_{j \in \{1,...k\}} \text{Eff}_j(\xi) \right\},$$

and $\psi(x, e_i, \xi) = -\Phi_i(\xi_i^*) \dfrac{\partial \Phi_i(\xi, \xi_x)}{\Phi_i^2(\xi)}$.

The proof is deferred to Appendix A.

3.1 Equivalence theorem

Bayesian optimum designs are usually found by applying standard algorithms, because the equivalence inequality (4) is completely known. Maxi-min efficiency designs are difficult to determine because they are not differentiable (their equivalence inequality depends on an unknown measure); see for instance, Wong (1992). See also Chen et al. (2017) and Dette and Biedermann (2003) for an equivalence theorem for the standard-ized max-min D-optimal design criterion. In this section, we provide a new formulation of the equivalence theorem, which establishes a connection between $\Phi_B(\xi; \pi)$ and $\Phi(\xi)$.

Theorem 3 *(Equivalence Theorem) A design ξ^* is a maxi-min efficiency design if and only if there exists a probability distribution π^* on the index set*

$$\mathcal{I}(\xi^*) = \left\{ i : i = \arg \min_{j \in \{1,...,k\}} \text{Eff}_j(\xi^*) \right\}, \tag{7}$$

such that ξ^* is a Bayesian optimum design for the prior distribution π^*, that is, if and only if ξ^* fulfils the following inequality,

$$\sum_{i \in \mathcal{I}(\xi^*)} \pi_i^* \frac{\partial \Phi_i(\xi^*, \xi_x)}{\Phi_i(\xi_i^*)} \leq 0, \qquad x \in \mathcal{X}. \tag{8}$$

The detailed proof of the equivalence theorem is deferred to Appendix A. In addition, from (5), we can state the following corollary:

Corollary 3.1 *The quantity $\sum_{i \in \mathcal{I}(\xi^*)} \pi_i^* \frac{\partial \Phi_i(\xi^*, \xi_x)}{\Phi_i(\xi_i^*)}$ attains its maximum value of zero at every support point of ξ^*.*

The equivalence between the minimum efficiency and the Bayesian optimality criteria can be used to check whether a design is optimal with respect to criterion (6). Recently, several algorithms have been applied to construct optimal designs numerically; see for instance, Dette et al. (2003) who apply the Nedler–Mead algorithm, (Chen et al. 2015, 2020), where the authors use particle swarm optimization, or (Belmiro et al. 2015), where a semi-infinite programming based algorithm is considered. These algorithms provide a solution based on a suitable stopping rule; however, it is necessary to check the equivalence inequality to prove that an 'optimum' has been reached. We follow the same idea as in Chen et al. (2017) (page 87). Given a solution of a numerical procedure ξ_s^*, from the equivalence inequality (8) with ξ_s^* instead of ξ^*, we can compute the prior distribution π^* solving the minimization problem

$$\min_{\pi_i \in [0;1], \sum_{i \in \mathcal{I}(\xi_s^*)} \pi_i = 1} \sum_{x \in \mathcal{S}_{\xi_s^*}} \left[\sum_{i=1}^k \pi_i \frac{\partial \Phi_i(\xi_s^*, \xi_x)}{\Phi_i(\xi_i^*)} \right]^2, \tag{9}$$

where $\mathcal{S}_{\xi_s^*}$ denotes the support of ξ_s^* and $\mathcal{I}(\xi_s^*)$ is the set defined in (7) with ξ^* replaced by ξ_s^*. Equation (9) comes out from the equivalence theorem, for which the weighted sum of the component-wise criteria's derivatives must be zero at each support point of the optimal design, and thus, the weights can be chosen by minimizing the sum of squares of these expressions for all the support points.

Given a design ξ_s^*, using the solutions of (9) we can check whether ξ_s^* really is an optimal design by computing the equivalence inequality (8).

Remark 1 At the optimal design, the value of (9) should be zero (except for rounding approximations).

4 Illustrative examples

The first example of this section underlines the difficulty in finding out a maxi-min efficiency design, when the search is done step-by-step by comparing the k efficiencies. This leads to the conclusion that suitable optimization algorithms should be applied, and then their numerical solutions should be checked for their optimality through Equivalence inequality (8). This procedure is followed in Example 4.2.

4.1 SMV-optimum designs in biology immunoassays

In biology, immunoassays are usually performed to quantify the concentration of an analyte. In this example, the SMV-optimality criterion is applied to the four-parameter

logistic model, which is the most frequently used model for symmetric immunoassay data,

$$y = \theta_1 + \frac{\theta_2 - \theta_1}{1 + \left(\dfrac{x}{\theta_4}\right)^{\theta_3}} + \varepsilon, \quad x \in \mathcal{X} = [0, \infty), \tag{10}$$

where y is the response at the concentration x, $\varepsilon \sim N(0; \sigma^2)$ is a random error, and $\theta_1 > 0$, $\theta_2 > 0$, $\theta_3 \in \mathbb{R}$, and $\theta_4 > 0$ are unknown parameters.

The SMV-optimality criterion, proposed by Dette (1997),

$$\Phi_{SMV}(\xi) = \max_{i \in \{1,\dots,4\}} \frac{e_i^T M^{-1}(\xi, \theta_0) e_i}{e_i^T M^-(\xi_i^*, \theta_0) e_i}$$

is an example of maximum inefficiency criterion (6), where $k = 4$ is the dimension of $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$; θ_0 is a guessed value for θ; $\Phi_i(\xi)$ is given by

$$\Phi_i(\xi) = \begin{cases} [e_i^T M^-(\xi, \theta_0) e_i]^{-1} & \text{if } e_i \in \text{Range}[M(\xi, \theta_0)] \\ 0 & \text{otherwise} \end{cases}, \quad i = 1, \dots, 4$$

where $M(\xi, \theta)$ is the information matrix (1) for model (10), and e_i, $i = 1, ..., 4$ are the canonical basis of $\mathbb{R}^4$.

In this example, $\mathcal{X} = [0, 5]$, $\theta_0 = (1, 2, 1, 1)$ and the gradient in (1) is

$$\frac{\partial \log f(y, x, \theta)}{\partial \theta} = \left(1 - \frac{1}{1+x}, \frac{1}{1+x}, -\frac{x \log[x + 10^{-6}]}{(1+x)^2}, \frac{x}{(1+x)^2}\right)^T,$$

where the third component has been slightly modified for computational reasons.

The procedure followed to find out the optimal design is quite cumbersome, but the prior probabilities which solve (9) enable us to identify the right maxi-min efficiency design. At first we search for designs that have the same efficiencies for any pair of the indices. Let $I = \{i_1, \dots, i_l\}$, with $l = 2, \dots, k$, be an index set. For instance, for $I = \{2, 4\}$ we obtain the design

$$\xi_1^{(2,4)} = \left\{ \begin{matrix} 0 & 1.321 & 2.756 & 5 \\ 0.212 & 0.221 & 0.395 & 0.173 \end{matrix} \right\},$$

which has the same efficiency, 0.2116, for both 2 and 4 standardized criteria, but the efficiencies of the other criteria are smaller (0.1139 and 0.0525 for standardized criteria 1 and 3, respectively). Thus, $\xi_1^{(2,4)}$ is not a minimum efficiency design and is discarded. In particular, one of the other efficiencies is very low, and thus after a new search the design

$$\xi_2^{(2,4)} = \left\{ \begin{matrix} 0 & 0.649 & 1.192 & 5 \\ 0.139 & 0.574 & 0.115 & 0.172 \end{matrix} \right\},$$

Table 1 Designs giving the same efficiency for $i \in \{1, 2, 4\}$

Design ξ	Eff$_i(\xi)$, $i \in \{1, 2, 4\}$	Eff$_3(\xi)$
$\left\{ \begin{array}{cccc} 0 & 0.764 & 1.162 & 5 \\ 0.154 & 0.388 & 0.233 & 0.225 \end{array} \right\}$	0.1546	0.0864
$\left\{ \begin{array}{cccc} 0 & 0.815 & 1.927 & 5 \\ 0.167 & 0.551 & 0.114 & 0.168 \end{array} \right\}$	0.1670	0.1058
$\left\{ \begin{array}{cccc} 0 & 0.771 & 2.532 & 5 \\ 0.224 & 0.299 & 0.313 & 0.164 \end{array} \right\}$	0.2238	0.1745

Table 2 Designs giving the same efficiency for $i \in \{2, 3, 4\}$ and a larger one for $i = 1$

Design ξ	Eff$_i(\xi)$, $i \in \{2, 3, 4\}$	Eff$_1(\xi)$
$\left\{ \begin{array}{cccc} 0 & 0.642 & 2.662 & 5 \\ 0.207 & 0.327 & 0.268 & 0.198 \end{array} \right\}$	0.2075	0.2441
$\left\{ \begin{array}{cccc} 0 & 0.527 & 1.499 & 5 \\ 0.231 & 0.284 & 0.109 & 0.376 \end{array} \right\}$	0.23047	0.3220
$\left\{ \begin{array}{cccc} 0 & 0.470 & 0.821 & 5 \\ 0.177 & 0.313 & 0.262 & 0.248 \end{array} \right\}$	0.1772	0.2477

gets a common efficiency of 0.1386 for indices 2 and 4, and efficiencies for the other criteria which are not so bad as in the previous case (0.1696 and 0.1140 for indices 1 and 3, respectively). However, once again one of the efficiencies is smaller than that for the indices in I, and $\xi_2^{(2,4)}$ is discarded as well.

After some attempts, finally we find out the design

$$\xi_3^{(2,4)} = \left\{ \begin{array}{cccc} 0 & 0.101 & 1.244 & 5 \\ 0.478 & 0.171 & 0.242 & 0.109 \end{array} \right\},$$

giving the same efficiency, 0.4778, for indices 2 and 4, which is smaller than the other efficiencies, 0.5734 and 0.5879 (for indices 1 and 3, respectively); therefore, $\xi_3^{(2,4)}$ is a candidate design for the Bayesian optimality. To prove that, it is necessary to identify a prior distribution in I, $\pi = \{\pi_2, \pi_4\}$, such that $\xi_3^{(2,4)}$ is Bayesian optimal for π. To find such a distribution, we employ condition (9). The weights minimizing this expression are $\pi_2 = 0.608$ and $\pi_4 = 1 - \pi_2$, but the minimum value obtained from these weights is 3.547, which is far from zero. Thus, this design cannot be Bayesian optimal (and neither maxi-min efficiency optimal).

We obtain similar results with every pair (i_1, i_2) of indices. Thus, the search proceeds among designs producing equal efficiencies for three of the component-wise criteria. At first we look for designs that produce a common efficiency for the indices in $I = \{1, 2, 4\}$; Table 1 lists some designs verifying this condition, however, none of them has the remaining efficiency larger than this common value.

The same happens with the triplets of indices $\{1, 3, 4\}$ and $\{1, 2, 3\}$. Differently, for the set $I = \{2, 3, 4\}$, we find out some designs with the same common efficiency, which is smaller than that for $i = 1$; see Table 2.

However, none of them can be Bayesian optimal because no distribution of weights in I gets a minimum value zero in (9). Finally, for the same index set, we obtain the design

$$\xi^* = \left\{ \begin{array}{cccc} 0 & 0.126 & 1.279 & 5 \\ 0.497 & 0.114 & 0.241 & 0.148 \end{array} \right\},$$

with efficiencies $\{0.5963, 0.4970, 0.4970, 0.4970\}$. Setting $\xi_s^* = \xi^*$ in (9), we get the solution $\pi^* = \{0, 0.493, 0.054, 0.453\}$ with a minimum value of 6.644 x 10^{-4} and hence, ξ^* turns out to be Bayesian optimal for π^*.

4.2 Maxi-min optimal discriminating designs in toxicology studies

In toxicology studies, we may have a continuous response and several possible models for the true mean response. As in Dette et al. (2010), we assume the following rival models for the mean response of the outcome Y:

$$\eta_1(x, \theta) = a e^{-bx}; \quad \theta = (a, b)^T, a > 0, b > 0,$$
$$\eta_2(x, \theta) = a e^{-bx^d}; \quad \theta = (a, b, d)^T, a > 0, b > 0, d \geq 1,$$
$$\eta_3(x, \theta) = a\left[c - (c-1)e^{-bx} \right]; \quad \theta = (a, b, c)^T, a > 0, b > 0, c \in [0, 1],$$
$$\eta_4(x, \theta) = a\left[c - (c-1)e^{-bx^d} \right]; \quad \theta = (a, b, c, d)^T, a > 0, b > 0, c \in [0, 1], d \geq 1,$$

and the following criterion for discriminating between pairs of models:

$$\min_{i \in \{1,2,3,4\}} \text{Eff}_i(\xi) = \min \left\{ \text{Eff}^{2-1}(\xi), \text{Eff}^{3-1}(\xi), \text{Eff}^{4-2}(\xi), \text{Eff}^{4-3}(\xi) \right\}, \quad (11)$$

where index i denotes 4 different pairwise comparisons: η_1 vs η_2, η_1 vs η_3, η_2 vs η_4, and η_3 vs η_4, respectively. In other terms, for a fixed value θ_0,

$$\text{Eff}_i(\xi) = \frac{\min_\xi \mathbf{e}_i^T \mathbf{M}_i^-(\xi, \theta_0)\mathbf{e}_i}{\mathbf{e}_i^T \mathbf{M}_i^{-1}(\xi, \theta_0)\mathbf{e}_i},$$

with

$$\mathbf{e}_i = \begin{cases} e_3 \in \mathbb{R}^3 \text{ for } i = 1, 2 \\ e_3 \in \mathbb{R}^4 \text{ for } i = 3 \\ e_4 \in \mathbb{R}^4 \text{ for } i = 4 \end{cases} \quad \text{and} \quad \mathbf{M}_i(\xi, \theta_0) = \begin{cases} M_2(\xi, \theta_0) \text{ for } i = 1 \\ M_3(\xi, \theta_0) \text{ for } i = 2 \\ M_4(\xi, \theta_0) \text{ for } i = 3, 4 \end{cases},$$

where e_j, $j = 3, 4$ is the j-th canonical basis of the Euclidean space and $M_j(\xi, \theta_0)$ is the information matrix (1) for the mean response $\eta_j(x, \theta)$, with $j = 1, 2, 3, 4$. Dette et al. (2003) found the maxi-min efficiency designs using a numerical procedure based on the Nedler–Mead algorithm. Setting $\theta_0 = (1, 3, 0, 1)^T$ (see Table 3 in Dette et al.
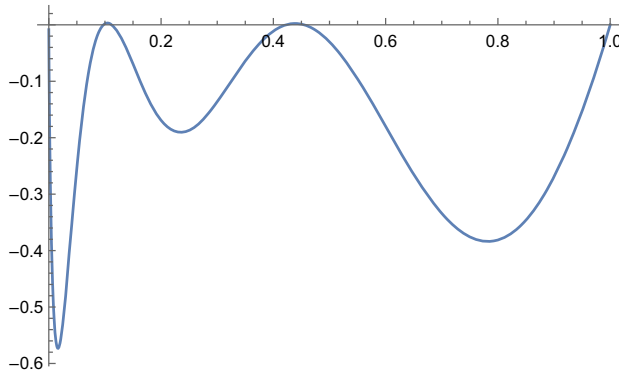
Fig. 1 Sensitivity function

2003), the authors found the following numerical solution:

$$\xi_s^* = \left\{ \begin{array}{cccc} 0 & .105 & .44 & 1 \\ .141 & .233 & .199 & .427 \end{array} \right\},$$

for which $\mathrm{Eff}_1(\xi_s^*) = .705$, $\mathrm{Eff}_2(\xi_s^*) = \mathrm{Eff}_4(\xi_s^*) = .682$ and $\mathrm{Eff}_3(\xi_s^*) = .871$, and hence $\mathcal{I}(\xi_s^*) = \{2; 4\}$ and $\pi_1^* = \pi_3^* = 0$. From (9), where

$$\Phi_i(\xi) = \begin{cases} [\mathbf{e}_i^T \mathbf{M}_i^-(\xi, \theta_0) \, \mathbf{e}_i]^{-1} & \text{if } \mathbf{e}_i \in \mathrm{Range}[\mathbf{M}_i(\xi, \theta_0)] \\ 0 & \text{otherwise} \end{cases}, \qquad i = 1, \ldots, p \quad (12)$$

and

$$\partial \Phi_i(\xi; \xi_x) = \frac{[\mathbf{e}_i^T \mathbf{M}_i^{-1}(\xi, \theta_0) \, \nabla \eta(x, \theta_0)]^2 - \mathbf{e}_i^T \mathbf{M}_i^{-1}(\xi, \theta_0) \, \mathbf{e}_i}{[\mathbf{e}_i^T \mathbf{M}_i^{-1}(\xi, \theta_0) \, \mathbf{e}_i]^2},$$

with $\nabla \eta(x, \theta) = [\frac{\partial \eta(x, \theta)}{\partial \theta_1}, \ldots, \frac{\partial \eta(x, \theta)}{\partial \theta_4}]^T$, we obtain $\pi_2^* = .574$ and $\pi_4^* = 1 - \pi_2^*$. Figure 1, which shows the sensitivity function on the left-hand side of (8), proves that the numerical solution ξ_s^* actually is a maxi-min efficiency design.

5 Conclusions and discussion

In practice, obtaining an optimal design that accounts for several goals or experimenter's interests, is a difficult task. There is much literature on different approaches, usually considering specific situations. In this study, we consider a quite general setting; we aims at finding a max-min efficiency design which maximizes the minimum of the efficiencies of several component-wise criteria (reflecting different tasks). This multi-objective criterion depends on some nominal values of the parameters, therefore a sensitivity analysis to assess this dependence is advisable.

We provide theoretical results, including an equivalence theorem which states that the maxi-min efficiency design is Bayesian optimal for a specific prior distribution on the set of the component-wise criteria. Furthermore, a method to identify this prior distribution is given. This is important for two reasons.

(i) It enables the application of the equivalence theorem in such a way the optimality of a particular design, e.g. found by the implementation of an algorithm, can be checked through the equivalence theorem since the prior probability can be determined.

(ii) This prior distribution tells the practitioner the weight the optimal design is assigning to each component-wise criterion. Notice that if a criterion does not receive any weight, this does not mean that the optimal design is going to be bad for that criterion. It is quite the opposite, as the efficiency of the optimal design with respect to that specific component-wise criterion will be higher than those with positive weights.

Appendix A

Lemma 4 *The maximum inefficiency criterion admits the following expression*

$$\Phi^{-1}(\xi) = \max_{i \in \{1,\ldots,k\}} \frac{1}{\text{Eff}_i(\xi)} = \max_{c \in \mathcal{C}} \Psi(\xi; c),$$

*where $\mathcal{C}$ is the convex hull
of the canonical vectors $\{e_1, \ldots, e_k, -e_1, \ldots, -e_k\}$ and*

$$\Psi(\xi; c) = c^T \begin{bmatrix} \frac{1}{\text{Eff}_1(\xi)} & \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & \frac{1}{\text{Eff}_k(\xi)} \end{bmatrix} c.$$

Proof The function $\Psi(\xi; c)$ is convex with respect to c; thus, it reaches its maximum at one (or more) of the vertexes of the set $\mathcal{C}$, and

$$\max_{c \in \mathcal{C}} \Psi(\xi; c) = \max_{c \in \mathcal{C}} c^T \begin{bmatrix} \frac{1}{\mathrm{Eff}_1(\xi)} & \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & \frac{1}{\mathrm{Eff}_k(\xi)} \end{bmatrix} c$$

$$= \max_{i \in \{1,\dots,k\}} e_i^T \begin{bmatrix} \frac{1}{\mathrm{Eff}_1(\xi)} & \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & \frac{1}{\mathrm{Eff}_k(\xi)} \end{bmatrix} e_i = \max_{i \in \{1,\dots,k\}} \frac{1}{\mathrm{Eff}_i(\xi)} = \Phi^{-1}(\xi),$$

which proves the lemma. □

Proof of Proposition 2. The set $\mathcal{C} = \{c : c \in \mathbb{R}^k, |c|_1 \le 1\}$, where $|c|_1 = \max\{|c_1|, \dots, |c_k|\}$ is compact. Therefore, from Eq. (2.6.15) of Fedorov and Hackl (1997), the directional derivative of $\Phi^{-1}(\xi)$ evaluated at ξ in the direction of $\bar{\bar{\xi}} - \xi$ is

$$\partial \Phi^{-1}(\xi; \bar{\bar{\xi}}) = \max_{c \in \mathcal{C}(\xi)} \int_{\mathcal{X}} \psi(x, c, \xi) \bar{\bar{\xi}}(dx) = \max_{e_i \in \mathcal{C}(\xi)} \int_{\mathcal{X}} \psi(x, e_i, \xi) \bar{\bar{\xi}}(dx),$$

where $\psi(x, c, \xi)$ is the directional derivative of $\Psi(\xi; c)$ in the direction of $\xi_x - \xi$ and

$$\mathcal{C}(\xi) = \left\{ c : c(\xi) = \arg \max_{c \in \mathcal{C}} \Psi(\xi; c) \right\} = \left\{ e_i : e_i(\xi) = \arg \max_{i \in \{1,\dots k\}} \Psi(\xi; e_i) \right\}.$$

The last expression for $\mathcal{C}(\xi)$ is because $\Psi(\xi; c)$ always reaches its maximum at one or more points defined by the canonical vectors.

Moreover,

$$\psi(x, e_i, \xi) = \lim_{\lambda \to 0^+} \frac{\frac{\Phi_i(\xi_i^*)}{\Phi_i[(1-\lambda)\xi + \lambda\xi_x]} - \frac{\Phi_i(\xi_i^*)}{\Phi_i(\xi)}}{\lambda}$$

$$= -\frac{\Phi_i(\xi_i^*)}{\Phi_i(\xi)} \lim_{\lambda \to 0^+} \frac{\Phi_i[(1-\lambda)\xi + \lambda\xi_x] - \Phi_i(\xi)}{\lambda} \cdot \lim_{\lambda \to 0^+} \frac{1}{\Phi_i[(1-\lambda)\xi + \lambda\xi_x]}$$

$$= -\Phi_i(\xi_i^*) \frac{\partial \Phi_i(\xi, \xi_x)}{\Phi_i^2(\xi)}. \tag{13}$$

□

The following two lemmas are necessary to prove the Equivalence Theorem stated in Sect. 3.

Lemma 5 *Let ξ and $\bar{\bar{\xi}}$ be two designs and let η denote a discrete distribution on $\mathcal{C}(\xi)$, then*

$$\max_{\eta} \int_{\mathcal{X}} \int_{\mathcal{C}(\xi)} \psi(x, e_i, \xi) \, \eta(de_i) \, \bar{\bar{\xi}}(dx) = \max_{e_i \in \mathcal{C}(\xi)} \int_{\mathcal{X}} \psi(x, e_i, \xi) \, \bar{\bar{\xi}}(dx).$$

Proof The following inequality

$$\max_{\eta} \int_{\mathcal{X}} \int_{\mathcal{C}(\xi)} \psi(x, e_i, \xi)\, \eta(de_i)\, \bar{\xi}(dx) \geq \int_{\mathcal{X}} \int_{\mathcal{C}(\xi)} \psi(x, e_i, \xi)\, \eta(de_i)\, \bar{\xi}(dx)$$

holds for any η; thus, it is valid also for the measure η_i which puts the whole mass at the vector e_i,

$$\max_{\eta} \int_{\mathcal{X}} \int_{\mathcal{C}(\xi)} \psi(x, e_i, \xi)\, \eta(de_i)\, \bar{\xi}(dx) \geq \int_{\mathcal{X}} \psi(x, e_i, \xi)\bar{\xi}(dx).$$

The last inequality is satisfied for any $e_i \in \mathcal{C}(\xi)$, and this means that

$$\max_{\eta} \int_{\mathcal{X}} \int_{\mathcal{C}(\xi)} \psi(x, e_i, \xi)\, \eta(de_i)\, \bar{\xi}(dx) \geq \max_{e_i \in \mathcal{C}(\xi)} \int_{\mathcal{X}} \psi(x, e_i, \xi)\, \bar{\xi}(dx). \tag{14}$$

On the other hand,

$$\int_{\mathcal{C}(\xi)} \int_{\mathcal{X}} \psi(x, e_i, \xi)\, \bar{\xi}(dx)\, \eta(de_i) \leq \max_{e_i \in \mathcal{C}(\xi)} \int_{\mathcal{X}} \psi(x, e_i, \xi)\, \bar{\xi}(dx).$$

This inequality is obtained by replacing each term $\int_{\mathcal{X}} \psi(x, e_i, \xi)\, \bar{\xi}(dx)$ by $\max_{e_i \in \mathcal{C}(\xi)} \int_{\mathcal{X}} \psi(x, e_i, \xi)\, \bar{\xi}(dx)$, and it is satisfied for any measure η; thus,

$$\max_{\eta} \int_{\mathcal{C}(\xi)} \int_{\mathcal{X}} \psi(x, e_i, \xi)\, \bar{\xi}(dx)\, \eta(de_i) \leq \max_{e_i \in \mathcal{C}(\xi)} \int_{\mathcal{X}} \psi(x, e_i, \xi)\, \bar{\xi}(dx). \tag{15}$$

The lemma follows from inequalities (14) and (15). □

Lemma 6 *For any design ξ, the following equality is verified:*

$$\min_{\bar{\xi}} \max_{\eta} \int_{\mathcal{X}} \int_{\mathcal{C}(\xi)} \psi(x, e_i, \xi)\, \eta(de_i)\, \bar{\xi}(dx) = \max_{\eta} \min_{x \in \mathcal{X}} \int_{\mathcal{C}(\xi)} \psi(x, e_i, \xi)\, \eta(de_i).$$

Proof Given $\bar{\xi}$, the inequality

$$\int_{\mathcal{X}} \int_{\mathcal{C}(\xi)} \psi(x, e_i, \xi)\, \eta(de_i)\, \bar{\xi}(dx) \geq \min_{x \in \mathcal{X}} \int_{\mathcal{C}(\xi)} \psi(x, e_i, \xi)\, \eta(de_i)$$

holds for any η; therefore,

$$\max_{\eta} \int_{\mathcal{X}} \int_{\mathcal{C}(\xi)} \psi(x, e_i, \xi)\, \eta(de_i)\, \bar{\xi}(dx) \geq \max_{\eta} \min_{x \in \mathcal{X}} \int_{\mathcal{C}(\xi)} \psi(x, e_i, \xi)\, \eta(de_i). \tag{16}$$

Since inequality (16) holds for any $\bar{\xi}$, for the measure $\bar{\xi}$ that minimizes the left-hand side of (16),

$$\min_{\bar{\xi}} \max_{\eta} \int_{\mathcal{X}} \int_{\mathcal{C}(\xi)} \psi(x, e_i, \xi)\, \eta(de_i)\, \bar{\xi}(dx) \geq \max_{\eta} \min_{x \in \mathcal{X}} \int_{\mathcal{C}(\xi)} \psi(x, e_i, \xi) \eta(de_i). \tag{17}$$

On the other hand, the inequality

$$\min_{\bar{\xi}} \max_{\eta} \int_{\mathcal{X}} \int_{\mathcal{C}(\xi)} \psi(x, e_i, \xi)\, \eta(de_i)\, \bar{\xi}(dx) \leq \max_{\eta} \int_{\mathcal{X}} \int_{\mathcal{C}(\xi)} \psi(x, e_i, \xi) \eta(de_i)\, \bar{\xi}(dx)$$

holds for any $\bar{\xi}$; thus, it is also valid for any measure $\bar{\xi} = \xi_x$, that is,

$$\min_{\bar{\xi}} \max_{\eta} \int_{\mathcal{X}} \int_{\mathcal{C}(\xi)} \psi(s, e_i, \xi)\, \eta(de_i)\, \bar{\xi}(ds) \leq \max_{\eta} \int_{\mathcal{C}(\xi)} \psi(x, e_i, \xi) \eta(de_i), \quad x \in \mathcal{X}. \tag{18}$$

Since inequality (18) holds for any $x \in \mathcal{X}$, it is also valid for the value of x which minimizes the quantity $\int_{\mathcal{C}(\xi)} \psi(x, e_i, \xi)\eta(de_i)$, that is,

$$\min_{\bar{\xi}} \max_{\eta} \int_{\mathcal{X}} \int_{\mathcal{C}(\xi)} \psi(s, e_i, \xi)\, \eta(de_i)\, \bar{\xi}(ds) \leq \max_{\eta} \min_{x \in \mathcal{X}} \int_{\mathcal{C}(\xi)} \psi(x, e_i, \xi) \eta(de_i). \tag{19}$$

The lemma follows from inequalities (17) and (19).   □

Proof of Theorem 3. Since $\Phi^{-1}(\xi)$ is a convex function of ξ, a necessary and sufficient condition for ξ^* to be an optimum design is that

$$\partial \Phi^{-1}(\xi^*, \bar{\xi}) = \max_{e_i \in \mathcal{C}(\xi^*)} \int_{\mathcal{X}} \psi(x, e_i, \xi^*)\bar{\xi}(dx) \geq 0, \forall\ \bar{\xi},$$

that is,

$$\min_{\bar{\xi}} \max_{e_i \in \mathcal{C}(\xi^*)} \int_{\mathcal{X}} \psi(x, e_i, \xi^*)\bar{\xi}(dx) \geq 0.$$

From Lemma 5 the above is equivalent to

$$\min_{\bar{\xi}} \max_{\eta} \int_{\mathcal{X}} \int_{\mathcal{C}(\xi^*)} \psi(x, e_i, \xi^*)\, \eta(de_i)\, \bar{\xi}(dx) \geq 0. \tag{20}$$

On the other hand, from Lemma 6 inequality (20) is equivalent to

$$\max_{\eta} \min_{x} \int_{\mathcal{C}(\xi^*)} \psi(x, e_i, \xi^*)\, \eta(de_i) \geq 0.$$

Thus, there exists a measure $\overline{\eta}$ (the maximum of the previous one) satisfying

$$\min_x \int_{\mathcal{C}(\xi^*)} \psi(x, e_i, \xi^*)\, \overline{\eta}(de_i) \geq 0,$$

that is,

$$\int_{\mathcal{C}(\xi^*)} \psi(x, e_i, \xi^*)\, \overline{\eta}(de_i) \geq 0, \qquad x \in \mathcal{X}.$$

From Equation (13) and setting $\overline{\eta}_i = \overline{\eta}(e_i)$, the previous inequality becomes

$$\sum_{e_i \in \mathcal{C}(\xi^*)} \Phi_i(\xi_i^*) \frac{\partial \Phi_i(\xi^*, \xi_x)}{\Phi_i^2(\xi^*)} \overline{\eta}_i \leq 0, \qquad x \in \mathcal{X}. \tag{21}$$

Let $\mathcal{I}(\xi^*) = \left\{ i : i = \arg\min_{j \in \{1,\dots,k\}} \mathrm{Eff}_j(\xi^*) \right\}$, setting

$$\pi_i^* = \frac{\overline{\eta}_i}{\mathrm{Eff}_i^2(\xi^*)} \bigg/ \sum_{i \in \mathcal{C}(\xi^*)} \frac{\overline{\eta}_i}{\mathrm{Eff}_i^2(\xi^*)}, \qquad i \in \mathcal{I}(\xi^*)$$

inequality (21) becomes

$$\sum_{i \in \mathcal{I}(\xi^*)} \pi_i^* \frac{\partial \Phi_i(\xi^*, \xi_x)}{\Phi_i(\xi_i^*)} \leq 0, \qquad x \in \mathcal{X}. \tag{22}$$

The thesis follows from comparing inequality (22) with (4). □

References

Atkinson A, Fedorov VV (1975) Optimal design: experiments for discriminating between several models. Biometrika 62:289–303

Belmiro PMD, Wong WK, Atkinson AC (2015) A semi-infinite programming based algorithm for determining T-optimum designs for model discrimination. J Multivar Anal 135:11–24

Berger MPF, Joy KCY, Wong WK (2000) Minimax d-optimal designs for item response theory models. Psychometrika 65(3):377–390

Chen R, Chang S, Wang W, Tung H, Wong WK (2015) Minimax optimal designs via particle swarm optimization methods. Stat Comput 25:975–988

Chen P, Chen R, Tung H, Wong WK (2017) Standardized maxmin d-optimal designs for enzyme kinetic inhibition models. Chem Intell Lab Syst 169:79–86

Chen R, Chen P, Chen L, Wong WK (2020) Hybrid algorithms for generating optimal designs for discriminating multiple nonlinear models under various error distributional assumptions. PLoS ONE 15(10):1–30

Dette H (1997) Designing experiments with respect to 'standardized' optimality criteria. J R Stat Soc Ser B 59:97–110

Dette H, Biedermann S (2003) Robust and efficient designs for the Michaelis–Menten model. J Am Stat Assoc 98(463):679–686

Dette H, Melas VB, Pepelyshev A, Strigul N (2003) Efficient design of experiments in the Monod model. J R Stat Soc Ser B 65:725–742

Dette H, Haines LM, Imhof LA (2007) Maxmin and Bayesian optimal designs for regression models. Stat Sin 17:463–480

Dette H, Pepelyshev P, Shpilev P, Wong WK (2010) Optimal designs for discriminating between dose-response models in toxicology studies. Bernoulli 16(4):1164–1176

Fackle-Fornius E, Miller F, Nyquist H (2015) Implementation of maximin efficient designs in dose-finding studies. Pharm Stat 14:63–73

Fedorov VV, Hackl P (1997) Model-oriented design of experiments. Lecture notes in statistics. Springer, New York

López-Fidalgo J, Tommasi C (2004) Construction of MV- and SMV-optimum designs for binary response models. Comput Stat Data Anal 44(3):465–475

Müller CH, Pazman A (1998) Applications of necessary and sufficient conditions tor maximin efficient designs. Metrika 48:1–19

Nyquist H (2013) Convergence of an algorithm for constructing minimax designs. In: Ucinski D, Atkinson A, Patan M (eds) mODa 10-advances in model-oriented design and analysis. Springer, Heidelberg, pp 187–194

Schervish MJ (1995) Theory of Statistics. Springer, New York

Tommasi C, López-Fidalgo J (2010) Bayesian optimum designs for discriminating between models with any distribution. Comput Stat Data Anal 54(1):143–150

Tommasi C, Martín-Martín R, López-Fidalgo J (2016) Max-min optimal discriminating designs for several statistical models. Stat Comput 26:1163–1172

Wong WK (1992) A unified approach to the construction of minimax designs. Biometrika 79:611–619