SORT 43 (1) January-June 2019, 163-190 DOI: 10.2436/20.8080.02.84

Efficient algorithms for constructing D- and I-optimal exact designs for linear and non-linear models in mixture experiments

Raúl Martín Martín[1], Irene García-Camacha Gutiérrez[1] and Bernard Torsney[2]

Abstract

The problem of finding optimal exact designs is more challenging than that of approximate optimal designs. In the present paper, we develop two efficient algorithms to numerically construct exact designs for mixture experiments. The first is a novel approach to the well-known multiplicative algorithm based on sets of permutation points, while the second uses genetic algorithms. Using (i) linear and non-linear models, (ii) D- and I-optimality criteria, and (iii) constraints on the ingredients, both approaches are explored through several practical problems arising in the chemical, pharmaceutical and oil industry.

MSC: 62-K05.

Keywords: Optimal experimental design, D-optimality, I-optimality, mixture experiments, multiplicative algorithm, genetic algorithm, exact designs.

1. Introduction

Applications of mixture problems can be found in several areas including the chemical, pharmaceutical and oil industries. Their main purpose is to identify the composition of different blends which optimally describe the characteristic-response of their products. Standard choice designs and models are typically applied in the literature. However, due to the benefits of the optimal experimental design (OED) theory, more attention is receiving the development of this theory for mixture experiments nowadays (Coetzer and Haines, 2017; García-Camacha Gutiérrez, 2017; Goos, Jones and Syafitri, 2016; Wong et al., 2015; Brown, Donev and Bissett, 2015). Many authors have worked on developing efficient algorithms for designing exact optimal experimental design. The limited number of theoretical results focuses on approximate optimal designs to precise parameter estimation (Cornell, 2002; Atkinson, Donev and Tobias, 2007). Kiefer (1961) analyt-

[1] Department of Mathematics, University of Castilla–La Mancha, Spain. Author for correspondence: Irene.GarciaCamacha@uclm.es, Avda. Carlos III s/n, Ed. 21 (21.1.04), 45071, Toledo, Spain

[2] School of Mathematics & Statistics, University of Glasgow, United Kingdom.

Received: February 2018
Accepted: February 2019

ically determined D-optimal designs for quadratic models. Galil and Kiefer (1977) extended these results for ϕ_p-optimization, while Mikaeili and Lim's works focused on cubic polynomials (Mikaeili, 1989; Lim, 1990). Nevertheless, no remarkable result exists for general-degree polynomials. Chan (1992) and Chan and Guan (1998) computed optimal designs for other classes of models such as log-contrast ones, with inverse terms or additive ones and Chan and Guan (2001) gave an extensive review about this topic. The book *Optimal Mixture Experiments* is an updated guide about both analytical and numerical results (Sinha et al., 2014). On the other hand, less attention has been paid in the statistical literature to seek I-optimal designs. Goos et al. (2016) provided a recent literature review on I-optimal designs and Coetzer and Haines (2017) introduced a new approach to the construction of D- and I- optimal designs when the mixture components are linearly constrained. Thus there is some space for exploiting the problem to develop more efficient numerical algorithms than the traditional ones.

The aim of this paper is to propose two novel design constructions algorithms for identifying exact D- and I-optimal designs in mixture experiments. The first one is based on a multiplicative algorithm (MA). This is a well known algorithm in OED (Torsney, 1977; Silvey, Titterington and Torsney, 1978). It consists of an update rule of probability measures and its convergence has been extensively studied for approximate design theory (Yu (2010)). However, the application of this methodology is not straightforward in exact mixture problems. In this work, we provide a new approach of the MA using a special class of designs known as exchangeable designs (Draper and Pukelsheim, 1999). The idea of these designs is to generate candidate points in the mixture designs using permutations of a fixed set of component values. In this paper, this class of designs are called *permutation mixture experimental designs* (PMEDs), where the use of MA takes advantage of exploiting the general equivalence theorem. On the other hand, an efficient genetic algorithm (GA) is provided as an heuristic alternative which is also valid in constrained mixture problems. Borkowski (2003) was a pioneer applying this numerical optimization tool to OED field and motivated its use for irregularly-shaped design regions. The nature of mixture experiments requires special conditions on the operators and even more if there are experimental limitations on the proportions. For that reason, although the basis of our algorithm is standard, adaptations of the operators have been carried out. GAs have been tested in a wide variety of contexts, in particular, they have been used as alternatives to exchange algorithms. Several modifications have already been developed to accelerate the convergence of these algorithms. Most of them are focused on the operators. Two new improvements are proposed in this paper. The first one based on the selection of the initial population and the second one is a new strategy based on a clusterization process around optimal points. Mixing laws for fluid viscosity, drug delivery systems, drug formulation and improvement of crude quality are some real examples suitable for computing optimal designs and for checking the goodness of the proposed algorithms.

The paper is organized as follows. Section 2 recalls the basis of mixture experimental design. We describe the existing designs for mixture experiments and an introduction

of the OED theory is presented. The proposed multiplicative and genetic algorithms for computing exact D- and I-optimal designs in mixture experiments are described in Section 3. Examples of applications to real problems are shown in Section 4. Finally, Section 5 provides a brief discussion and some future lines of research.

2. Background

2.1. Models and designs for mixture experiments

Controlled variables in a standard mixture problem are nonnegative, belonging to $[0,1]$ and dependent through the relationship $\mathbf{1}_q^\intercal p = 1$ where $\mathbf{1}_q = (1,\ldots,1)^\intercal \in \mathbb{R}^q$ and $p = (p_1,\ldots,p_q)^\intercal$ is the vector of relative proportions in a q-component mixture. These constraints define the design region χ as a $(q-1)$-dimensional simplex $\mathcal{S} = \{p \in [0,1]^q : \mathbf{1}_q^\intercal p = 1\}$. In addition, many real mixture problems are often constrained by lower and upper bounds on their proportions, $0 \leq L_i \leq p_i \leq U_i \leq 1$, $i = 1,\ldots,q$. This is mainly due to experimental limitations or ingredient availability considerations.

A suitable model must be selected a priori describing the composition-response relationship. Let $y = \boldsymbol{\eta}^\intercal(p)\boldsymbol{\theta} + \varepsilon(p)$ be the observed response, where $\boldsymbol{\eta}^\intercal(p) = (\eta_1(p),\ldots,\eta_k(p))$ is a vector of k linearly independent functions, $\boldsymbol{\theta} = (\theta_1,\ldots,\theta_k)^\intercal$ is the unknown parameter vector and $\varepsilon(p)$ is the error term. Additive uncorrelated random errors with common variance will be assumed. Because of the ordinary polynomials do not allow estimation of parameters due to collinearity between proportions, canonical polynomials introduced by Scheffé (1958) are the most commonly used for a large of practical situations. To illustrate, a third-order Scheffé polynomial (the full cubic model) is

$$E[y] = \sum_{i=1}^{q} \theta_i p_i + \sum_{i=1}^{q-1} \sum_{j=i+1}^{q} \theta_{ij} p_i p_j + \sum_{i=1}^{q-1} \sum_{j=i+1}^{q} \delta_{ij} p_i p_j (p_i - p_j) + \sum_{i=1}^{q-2} \sum_{j=i+1}^{q-1} \sum_{k=j+1}^{q} \theta_{ijk} p_i p_j p_k,$$

where δ_{ij} are reparametrizations of the parameters of an ordinary full third-order polynomial. In spite of being the most popular, other models have been proposed in the literature for data from mixture experiments with particular properties. Darroch and Waller's additive polynomials (Darroch and Waller, 1985), models with homogeneous functions (Becker, 1968), models with inverse terms (Draper and John, 1977), log-contrast models (Aitchison and Bacon-Shone, 1984) or Draper and Pukelsheim's *K-polynomials* (Draper and Pukelsheim, 1997) are some of them.

Although many aspects differ between experiments from different areas, standard designs are often used by practitioners in mixture problems. In general, standard mixture designs are adopted in the literature for fitting standard mixture models. If $m \geq 1$ is

an integer, the $\{q,m\}$-simplex lattice in $\mathcal{S}$ is defined as the collection of points whose coordinates are integer multiples of $1/m$, that is the set of points $\{p \in \mathcal{S},\ p_i = \frac{j}{m}, 0 \leq j \leq m,\ 1 \leq i \leq q\}$ (Scheffé, 1958). Thus a $\{q,m\}$-*simplex lattice* design describes a design that takes observations at the above set of points, the $\{q,m\}$-lattice. On the other hand, a $\{q,m\}$-*simplex centroid* ($1 \leq m \leq q$) is defined as a collection of points in $\mathcal{S}$ with $q-j$ coordinates equal to zero and j coordinates equal to $\frac{1}{j}$, $j = 1,\ldots,m$ (Scheffé, 1963). However, if interest is focused on exploring within the simplex, another class of designs named *axial designs* were suggested by Cornell (2002). Snee and McLean and Anderson (Snee, 1979; McLean and Anderson, 1966) proposed extreme-vertex designs for constrained mixture problems.

In summary, the analysis of mixture experiments has been developed using canonical polynomials models and other alternative linear models under standard designs. However, there are situations where models that are nonlinear in the parameters would be preferable and standard designs are not appropiated. The application of mixture experiments to nonlinear models appears to be a very interesing question which has been little explored (Coetzer and Focke, 2010; Brown et al., 2015). On the other hand, even considering linear models, if the design region is constrained, standard designs are not suitable 6 (Piepel, Cooley and Jones, 2005). In this paper we apply the OED theory to obtain optimal designs using both linear and nonlinear models and considering unconstrained and constrained regions. In the next subsection, we introduce the OED basis, which is used in what follows.

2.2. Optimal experimental design background

Let a linear model $y = \boldsymbol{\eta}^{\mathrm{T}}(p)\boldsymbol{\theta} + \varepsilon(p)$ as defined above. A set of experimental conditions, p, must be determined in order to observe the outcome in an optimal manner (mainly to attain precise estimations of the parameters or to obtain accurate response predictions). An *exact design* will be a sequence of experimental conditions (mixture settings) $\xi_N = \{p^1, p^2, \ldots, p^N\}$ from a compact set (the $(q-1)$-dimensional simplex $\mathcal{S}$) which are not necessarily distinct. Assuming that only J of the points are different the design may be represented by a probability measure. Thus, if the point p^j appears n_j times in the design, $\omega_j = n_j/N$ will be the probability of p^j within the sample. Then the exact design problem can be viewed as one of determining these proportions optimally subject to them being rational. Using this idea Kiefer (1961) relaxed this condition, defining *an approximate design* as any probability measure ξ on χ with a finite support,

$$\xi = \begin{Bmatrix} p^1 & p^2 & \cdots & p^J \\ \omega_1 & \omega_2 & \cdots & \omega_J \end{Bmatrix},$$

where the ω_j values satisfy $0 \leq \omega_j \leq 1$ and $\sum_{j=1}^{J} \omega_j = 1,\ j = 1,\ldots,J$. From the Carathéodory theorem, an upper bound for the number of support points can be derived as $\frac{k(k+1)}{2} + 1$ (Chapter 8, Pukelsheim, 2006). For moderate and large numbers of runs,

the number of replicates of design points can be determined by integer appr oximation to the optimal measure.

The most important element for describing the quality of statistical inference that can be drawn from data collected with a design is the Fisher information matrix.

For an N-point exact design ξ_N we can assume $J = N$ and $\omega_j = 1/N$; so

$$M(\xi_N) = \sum_{j=1}^{N} \boldsymbol{\eta}(p^j)\boldsymbol{\eta}^{\intercal}(p^j)\omega_j = \frac{1}{N} VV^{\intercal} \propto VV^{\intercal},$$

where the i^{th} column of the matrix V is $\boldsymbol{\eta}(p^i)$ denoted by $v_i = v(p^i)$. The set of information matrices, $\mathcal{M}$, is convex and compact. The inverse of the information matrix is proportional to the covariance matrix of the least squares estimates. Thus, an experimental designing "optimizing", in some sense, the information matrix, should be found. Following convention, the ranking of alternative designs is based on a scalar-valued criterion function, $\psi[M(\xi_N)]$, so that, the problem becomes one of function optimization. A function ψ defined on the set of information matrices defines an optimality criterion if it is non decreasing in the Loewner sense ($\psi(M_1) \leq \psi(M_2)$ whenever $M_1 - M_2$ is non-negative definite). For notational issues, let us define two functions $\psi[\cdot]$ and $\phi(\cdot)$, both relative the criterion function whose use will depend on its argument, in particular $\psi[M(\xi_N)] = \phi(\xi_N)$. In this paper, we consider two optimality criteria: D- and I-optimality. The goal of D-optimality is connected to parameter estimation. This criterion seeks to minimize the volume of the confidence ellipsoid of the parameters and is formulated as $\phi_D(\xi_N) = \det[M(\xi_N)]^{-1/k}$. On the other hand, due to the importance of predictive capability of many mixture experiments, I-optimal designs were considered in this work too. This criterion focuses on precise prediction, and is defined by the following function: $\phi_I(\xi_N) = \frac{\int_{\mathcal{S}} \eta(p)^{\intercal} M^{-1}(\xi_N)\eta(p)dp}{\int_{\mathcal{S}} dp} = \Gamma(q) \cdot \text{trace}[M^{-1}(\xi_N)B]$, where B is the moment matrix given by $B = \int_{\mathcal{S}} \eta(p)\eta^{\intercal}(p)dp$ and $\int_{\mathcal{S}} dp = \frac{1}{\Gamma(q)}$ when the domain of the mixture settings is the simplex. Thus, I-optimal designs seek to minimize the average prediction variance over the design region.

A design optimizing the criterion function in the class Ξ_N of all exact designs of size N is referred to as an exact ϕ-optimal design, ξ_N^*. Thus we can compare the quality of two designs of the same size (N) through the ratio of the criterion values. When the optimal exact design is known, ξ_N^*, the efficiency of a design $\xi \in \Xi_N$ is defined as $\text{Eff}_\phi(\xi_N) = (\phi(\xi_N)/\phi(\xi_N^*))$.

However, finding an exact optimal design is not an easy task because it is a discrete optimization problem and there is no general analytical tool for confirming whether an exact design is optimal or not. On the contrary, approximate designs are easier to find. The most important advantage of searching approximate designs is the concavity (convexity) of the criterion functions. Under these conditions, an excellent tool to check whether a particular approximate design is optimal (especially for differentiable criteria) is the Equivalence Theorem. Even though finding approximate optimal design is

easier because of the above results, in practical settings, only exact designs can be implemented. So, when an optimal approximate design has been found, then it has to be rounded to obtain an exact design (Pukelsheim and Rieder, 1992). A weakness of this approach is that the final exact design obtained by rounding off an approximate design for implementation is not unique. In addition, a large sample size is needed to obtain a design close to the optimal exact design.

It is worth mentioning that in many real situations, mixing laws do not linearly respond as composition varies. For the linear case, optimal designs are independent of the value of θ. In the case where non-linear models are appropriate, the most common method for analyzing them is based on the use of the linear Taylor series approximation of the model. Under these conditions, the covariance matrix of the least squares estimator of $\boldsymbol{\theta}$ is asymptotically approximated by the inverse of the information matrix induced by the design

$$M(\xi_N, \boldsymbol{\theta}^0) = \frac{1}{N} \sum_{j=1}^{N} v(p^j, \boldsymbol{\theta}^0) v^{\mathsf{T}}(p^j, \boldsymbol{\theta}^0),$$

where $v(p^j, \boldsymbol{\theta}^0) = (\frac{\partial \eta(p^j, \boldsymbol{\theta})}{\partial \theta_1}, \ldots, \frac{\partial \eta(p^j, \boldsymbol{\theta})}{\partial \theta_k})^{\mathsf{T}}\big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^0}$ and $\boldsymbol{\theta}^0$ is a prior guess of $\boldsymbol{\theta}$ (Chernoff, 1953). In this sense, the computed designs are locally optimum.

3. Algorithms for solving mixture exact design problems

As it was defined in the previous section, finding a ϕ-optimal exact N-point design is a combinatorial problem, and it has been considered an NP-hard problem (Welch, 1982). Globally optimal exact designs usually cannot be established and, in most cases, we need to resort to heuristic algorithms to find good designs. Several algorithms are available in the literature, most of which can be only used to compute approximate designs. They can be categorised into two broad groups: *greedy algorithms* such as those based on Fedorov-type exchanges, candidate-free coordinate exchange and multiplicative updating of the weights, and *nature inspired algorithms* which include simulated annealing, genetic algorithms and swarm intelligence between others (Dean et al., 2015).

The first algorithms developed for dealing with exact designs are based on exchange methods and were proposed for the D-optimality criterion (Fedorov, 1972; Wynn, 1970). Some modifications of these procedures were suggested in order to speed up the original algorithms (DETMAX algorithm (Mitchell, 1974); KL-exchange algorithm (Atkinson and Donev, 1989); coordinate-exchange algorithm (Meyer and Nachtsheim, 1995). McLean and Anderson's method (McLean and Anderson, 1966), XVERT (Snee and Marquardt, 1974) and CONSIM (Snee, 1979) were specifically developed for obtaining designs on irregularly shaped experimental regions. The resulting designs are called *the extreme-vertex designs*. Most of these algorithms were later directly applied to mixture settings. Neither of the algorithms are guaranteed to find the globally optimum design

because the support points are chosen from a pre-specified grid points. This requirement implies an exhaustive search over all candidate points, which is time-consuming and inefficient. During the last few years, algorithms have been improved to avoid this drawbacks. In particular, for constructing approximate designs, hybrid algorithms have been developed for improving computational efficiency (Martín-Martín and García-Camacha Gutiérrez, 2015) for D-optimality, Saleh and Pan (2016) for G-optimality, and Coetzer and Haines (2017) for D- and I-optimality for mixture experiments with linear constraints). They are based on suitably adjusting the strategies followed by the standard algorithms so that the new proprieties were able to solve the arisen problems using these methods in an isolated way. Another class of algorithms, inside of the first group of algorithms, which has received much attention for finding optimal approximate designs is the class of multiplicative algorithms (Torsney, 1977; Silvey et al., 1978). In spite of the several improvements to this class of algorithms, only Torsney and Martín-Martín (2009) adapted the multiplicative algorithm to cope with exact designs. In the present paper, this numerical method will be adapted to the special nature of mixture design.

The second group of optimization techniques used in OED to compute optimal designs are the meta-heuristic optimization algorithms. Due to their flexibility and potential, they have become a common tool in computational statistics as alternatives to standard algorithms. One of the most popular ones is the GA. Borkowski (2003) was a pioneer applying this numerical optimization tool to OED field and motivated its use for irregularly-shaped design regions. Heredia-Langer et al. (2003) and Limmuun, Borkowski and Chomtee (2013) gave a substantial discussion about the relative merits of GAs for design of experiments and some of the potential pitfalls of the implementation. On the other hand, in a recent paper, Wong et al. (2015) proposed a modified particle swarm optimization (PSO) technique for computing D-optimal approximate designs for mixture linear models. It is important to highlight that these algorithms take the mixture proportions to be continuous over the design region. Variable-Neighbourhood Search (VNS) is also a metaheuristic strategy commonly used to escape from local optima. Several variants of VNS have been proposed in the literature (Vazquez, Goos and Schoen, 2018). In this work, two new improvements have been incorporated to the proposed GA. The first one is based on the selection of the initial population and the second one is a new strategy based on a clustering process around presumed optimal design points.

3.1. A novel approach of the MA to determining exact optimal design for mixture experiments

Symmetry and balancedness have always been a prime attribute of good experimental designs (Draper and Pukelsheim, 1999). Nevertheless, in the case of mixture experiments, symmetry cannot be conducted in the general geometrical sense since the simplex is not itself a symmetric region. The natural structure of symmetry in the simplex deals with the invariance under permutation of its coordinates, it means symmetry through

the centroid of the simplex. Following this idea and since the support points of most the optimal mixture designs obtained in the literature are permutations of proportions, we consider the use of permutations of fixed sets of q component values or proportions, say $p = (p_1, \ldots, p_q)$ where $\mathbf{1}_q^{\intercal} p = 1$, to generate candidate points for mixture designs. In this paper, this class of designs is called *Permutation Mixture Experimental Designs* (PMEDs).

Let $p = (p_1, \ldots, p_q)$ be a single mixture point in the $(q-1)$-dimensional simplex $\mathcal{S}$ and let

$$\mathcal{P}(p) = \{a = (a_1, \ldots, a_q) = \sigma(p_1, \ldots, p_q),\ \sum_{i=1}^{q} p_i = 1,\ p_i \geq 0, i = 1, \ldots, q\}$$

be the set of all possible permutations of its proportions, $\#\mathcal{P}(p) = q!$. A PMED of $p \in \mathcal{S}$ mixtures is an exact $N = q!$-design generated by one set of components $\xi_{\mathcal{P}(p)} = \{\mathcal{P}(p)\}$. From this definition, it is worth mentioning that the set of the permutation points of any point belonging to a linearly-constrained region into the simplex may not be entirely included in this region. Consequently, this new approach cannot be applied for solving constrained mixture problems. In this regard, new approaches are being explored for overcoming this situation.

Let us denote the PMED design $\xi_{\mathcal{P}(p)} \underset{Not}{\equiv} \mathbb{P}$. The corresponding information matrix will be written as

$$M(\mathbb{P}) = \frac{1}{q!} \sum_{j=1}^{q!} v(p^j) v^{\intercal}(p^j). \tag{1}$$

We are interested in finding $p^* = (p_1^*, \ldots, p_q^*)$ optimally to maximize a chosen design criterion, $\psi[M(\mathbb{P}^*)] = \phi(p^*) = \max_{p \in \mathcal{S}} \phi(p)$. This problem can be considered as special case of the general class optimization problem discussed by Torsney and Martín-Martín (2009). One advantage of this approach is that we can use calculus to determine first-order conditions of optimality for exact designs.

The first-order conditions for a local maximum (minimum) are:

$$F_i^* = F_\phi(p^*, e_i) = \begin{cases} = 0, & \text{for } p_i^* > 0 \\ \leq (\geq) 0 & \text{for } p_i^* = 0, \end{cases} \quad i = 1, \ldots, q \tag{2}$$

where $F_i = F_\phi(p, e_i)$ is the directional derivative of $\phi()$ at p in the direction of the extreme vertex $e_i \in \mathbb{R}^q$. It is noteworthy that the elements of the information matrix (1) in the mixture experiment context are not linear functions of p even for simple models such as higher first-order polynomials. Therefore, the criterion function $\phi(p) = \psi[M(\mathbb{P})]$ is probably a non-concave (non-convex) function, in which case (2) are necessary but not sufficient conditions for local maxima (minima). For illustrative purposes, directional derivatives for *D*-optimality are computed following the above considerations (see supplementary material A). The expression of the directional derivative in the case of non-

linear arguments is

$$F_\phi(p,e_i) = F_\psi\left[M(\mathbb{P}),\ M(\mathbb{P}) + \frac{\partial M(\mathbb{P})}{\partial p_i} - \sum_{l=1}^{q} p_l \frac{\partial M(\mathbb{P})}{\partial p_l}\right].$$

The directional derivatives for the D- and I-optimality criteria derived from these are

$$F_{\phi_D}(p,e_i) = \mathrm{Tr}\left[M^{-1}(\mathbb{P})\frac{\partial M(\mathbb{P})}{\partial p_i}\right] - \sum_{l=1}^{q} p_l \mathrm{Tr}\left[M^{-1}(\mathbb{P})\frac{\partial M(\mathbb{P})}{\partial p_l}\right], \tag{3}$$

and

$$F_{\phi_I}(p,e_i) = \mathrm{Tr}\left[LM^{-1}(\mathbb{P})\frac{\partial M(\mathbb{P})}{\partial p_i}M^{-1}(\mathbb{P})L^{\mathrm{T}}\right] - \sum_{l=1}^{q} p_l \mathrm{Tr}\left[LM^{-1}(\mathbb{P})\frac{\partial M(\mathbb{P})}{\partial p_l}M^{-1}(\mathbb{P})L^{\mathrm{T}}\right], \tag{4}$$

where L is the Cholesky factor of the moment matrix B.

To satisfy the constrains of this problem of maximizing a criterion function of proportions $p_1,\ldots,p_q$, we will use an iterative multiplicative algorithm. Thus, the n-th update corresponding to the i-th component of p is

$$p_i^{(n)} = \frac{p_i^{(n-1)} f(x_i^{(n-1)},\delta)}{\sum_{l=1}^{q} p_l^{(n-1)} f(x_l^{(n-1)},\delta)},\ \ i = 1,\ldots,q,$$

where $x_i^{(n-1)} = F_\phi(p^{(n-1)},e_i)$, $f(x_i^{(n-1)},\delta)$ is positive, $\partial f(x,\delta)/\partial x > 0$ and, if $\delta = 0$, $f(x,\delta)$ is constant; $n = 1,2,\ldots$ is the iteration number and $p^{(0)} = (p_1^0,\ldots,p_q^0)$ a starting point such that $M(\mathbb{P}^{(0)})$ is not a singular matrix. The choice of f plays an important role in the convergence of the algorithm. δ is a small positive constant whose choice must be suitably made for the monotoniciy of the algorithm. Since the criterion function can have negative derivatives, two appropriate choices of $f(x,\delta)$ are $f(x,\delta) = \Phi(\delta x)$, where Φ is the c.d.f. of the standard normal distribution, and $f(x,\delta) = \exp(\delta x)/(1+\exp(\delta x))$, i.e., the logistic c.d.f. evaluated at δx. An iteration of the algorithm will be completed when all components have been updated. It is important to note that the application of the standard version of the MA for computing the optimal approximate design with $q!$ points, will imply $q!\cdot q$ updates in each iteration, while it will be only q in the case of considering a permutation design due to only one set of proportions needs to be computed. The stopping rule will comprise checking if the first-order conditions (4) are satisfied up to a certain tolerance.

One of the limitations of considering one set of permutations is that, in many mixture systems, it is not sufficient to estimate all model parameters. This is mainly due to singularity occurring in the information matrix by the repetition of its elements (permutations

of blends with repeated coordinates) or simply because the number of design-points ($q!$) is lower than the number of parameters. In order to solve these problems, we provide a natural extension of the algorithm presented above. This approach consists of the simultaneous calculation of more than one set of permutations, say t sets,

$$p_{(h)} = (p_{h1}, \ldots, p_{hq}),\ h = 1, \ldots, t,$$

where

$$\sum_{i=1}^{q} p_{hi} = 1 \ \ \forall h = 1, \ldots, t \ \text{ and } \ p_{hi} \geq 0, \forall h = 1, \ldots, t,\ i = 1, \ldots, q.$$

Thus, a greater variety of designs points can be included in the designs,

$$\begin{aligned} \mathcal{P}(p_{(1)}, \ldots p_{(t)}) = \Bigg\{ & a_{(h)} = (a_{h1}, \ldots, a_{hq}) = \sigma(p_{h1}, \ldots, p_{hq}) : \\ & \sum_{i=1}^{q} p_{hi} = 1, p_{hi} \geq 0, h = 1, \ldots, t,\ i = 1, \ldots, q \Bigg\} \\ & = \mathcal{P}(p_{(1)}) \cup \ldots \cup \mathcal{P}(p_{(t)}). \end{aligned}$$

A PMED of $p_{(1)}, \ldots, p_{(t)} \in \mathcal{S}$ mixtures is an exact $N = t \cdot q!$ - design,

$$\xi_{\mathcal{P}(p_{(1)}, \ldots, p_{(t)})} = \Big\{ \mathcal{P}(p_{(1)}), \quad \mathcal{P}(p_{(2)}), \quad \ldots, \quad \mathcal{P}(p_{(t)}) \Big\}$$

consisting of all possible points formed by permutation of the coordinates of $(p_{(1)}, \ldots, p_{(t)}) \in \mathcal{S}$. Then, according to (1), the information matrix is

$$M(\mathbb{P}_{(1)}, \ldots, \mathbb{P}_{(t)}) = \sum_{h=1}^{t} M(\mathbb{P}_{(h)}) = \frac{1}{t \cdot q!} \sum_{h=1}^{t} \sum_{j=1}^{q!} v(p_{(h)}^{j}) v^{\intercal}(p_{(h)}^{j}).$$

Thus we are facing to the following optimization problem: optimize $\phi(p_{(1)}, \ldots, p_{(t)})$ over $p_{(1)}, \ldots, p_{(t)} \in \mathcal{S}$. Then the following ($h$-sets) simultaneous approaches are used

$$p_{hi}^{(n)} = \frac{p_{hi}^{(n-1)} f_h(x_{hi}^{(n-1)}, \delta_h)}{\sum_{l=1}^{q} p_{hl}^{(n-1)} f_h(x_{hl}^{(n-1)}, \delta_h)}, h = 1, \ldots, t, i = 1, \ldots, q$$

where n is the iteration number, $f_h(x_{hi}^{(n-1)}, \delta_h)$ are positive increasing functions and $x_{hi}^{(n-1)} = F_\phi(p_{(h)}^{(n)}, e_i) \underset{Not}{\equiv} F_{hi}$ are the directional derivatives defined as above. There are necessary optimality conditions equivalent to those in Eq. (2). Therefore, the algorithm stops when the following conditions

$$F^*_{hi} = F_\phi(p^*_{(h)}, e_i) = \begin{cases} = 0, & \text{for } p^*_{hi} > 0 \\ \leq (\geq) 0 & \text{for } p^*_{hi} = 0, \end{cases} \quad i = 1,\ldots,q, \;\; h = 1,\ldots,t \qquad (5)$$

are simultaneously satisfied.

Multiplicative algorithm for ϕ-optimal mixture design

Step 0. Input q, $p^{(0)}_{(1)}$, $\xi^{(0)}_{\mathcal{P}} = \xi_{\mathcal{P}(p^{(0)}_{(1)})}$, $\boldsymbol{\delta}_1$, tol. Set $n = 1, t = 1$.

Step 1. Update the proportions for each mixture point generator,$(p^{(0)}_{(1)},\ldots,p^{(0)}_{(t)})$

For $h = 1,\ldots,t$, do,

- For $i = 1,\ldots,q$, do $p^{(n+1)}_{hi} = \frac{p^{(n)}_{hi} f_h(x^{(n)}_{hi},\delta_h)}{\sum_{l=1}^{q} p^{(n)}_{hl} f_h(x^{(n)}_{hl},\delta_h)}$

with $x^{(n)}_{hi} = F_\phi(p^{(n)}_{(h)}, e_i)$ calculated as in (3), (4).

Step 2. Construct the design $\xi^{(n+1)}_{\mathcal{P}} = \xi_{\mathcal{P}(p^{(n+1)}_{(1)},\ldots,p^{(n+1)}_{(t)})}$.

Step 3. If $|M(\mathbb{P}^{(n+1)}_{(1)},\ldots,\mathbb{P}^{(n+1)}_{(t)})| \approx 0$, then repeat from step 1 to step 3 adding a new group of permutation, $\xi^{(0)}_{\mathcal{P}} = \xi_{\mathcal{P}(p^{(0)}_{(1)},\ldots,p^{(0)}_{(t)},p^{(0)}_{(t+1)})}$, $t = t+1$. Otherwise, go to step 4.

Step 4. Stopping rule: If

$$\min_{\substack{h=1,\ldots,t \\ i=1,\ldots,q}} \{F_\phi(p^{(1)}_{(hi)}, e_i)\} \leq 10^{-tol}$$

where tol is a number specified by the user, then STOP.
Else update $\xi^{(n)}_{\mathcal{P}}$ by $\xi^{(n+1)}_{\mathcal{P}}$, $n = n+1$, and return to Step 1.

In the next section we explore the potential of this method in a variety of examples encompassing both linear and non-linear models for D-optimality and I-optimality.

3.2. Genetic algorithm

When the design space is regular and conventional mathematics can be applied the NP-hard combinatorial optimization problem of finding a ϕ-optimal exact N-point design can be solved using traditional optimization techniques. However many difficulties such as the irregular structure of the design spaces, the non-linear and non-differentiable objective functions, etc. make that optimization techniques break down in many optimization problems. For this reason, metaheuristic strategies have been developed to solve these difficulties. The goal is to explore the design space in a smart way to get near-optimal solutions.

One of these algorithms is the genetic algorithm (GA). GAs are population based stochastic search algorithms inspired by Darwin's Theory of Evolution and the survival-of the fittest. The weakest individuals will disappear while the best ones will survive and be able to reproduce themselves for generating the next population. Although there is no metaheuristic algorithm that will be universally the winner, it should be pointed out that GAs are robust, flexible and easy to implement. As other metaheuristics strategies, the two main features of the algorithm are the locally and intensively exploring/searching around the best solutions (intensification) and the generation of diverse solutions to make sure the algorithm explores the design space globally (diversification).

It is common to find in the literature related to this class of algorithms a specific terminology based on Genetics. P denotes the population of M initial N-point exact designs. Potential solutions of the problem (designs) are named chromosomes, whereas support points (blends) are labelled genes.

GAs start to search from an initial population. The information provided for each exact design is measured in terms of the criterion function value relative to the population. This value is a probability measure of the design goodness known as the fitness function. At each iteration a number of operators is applied to the designs of the current population to generate the designs of the population of the next generation (iteration). The most popular genetic operators are (1) selection (certain elitism is used to ensure the monotonicity of the algorithm. Also, designs with higher fitness have higher probabilities of being selected for successive processes); (2) crossover, also called the recombination operator (new designs, called offspring, are generated from two designs, called parents with a crossover probability, P_C); (3) mutation (to avoid premature convergence toward local optimal, with a mutation probability, P_M). Applying this process iteratively, new generations of designs are created until some stopping rule is reached. In this work, the algorithm stops after performing a prefixed maximum number of consecutive iterations (N_{max}) without improvement of the best fitness function value.

In the first step of a GA an initial population of designs, which are created from a set of points, is needed. As in Heredia-Langer et al. (2003) we use a population size of $M = 40$ exact designs. It is reasonable to believe that if the set contains *good* points to create designs, then we will have more possibilities to find the near-optimal design. Thus, if some information about the optimal solutions is available it will be convenient to use. On the other hand, if no information about the solution is available, it would be expected that the more diverse the initial population is, the greater the possibility to find a solution (Diaz-Gomez and Hougen, 2007). With this in mind, several scenarios were considered in this work. Basically they are distinguished by the fact that they include just randomly points or also contain vertices, the overall centroid and the centroid of all lower dimensional simplices of a $(q-1)$-simplex. A detailed explanation of different frameworks can be found in the supplementary material B. Through numerical examples we study the effect of the initial populations in the convergence of the algorithm.

The choice of the algorithm operators and parameters is a hard problem that will determine whether the algorithm will find a near-optimum solution and whether it will

find such a solution efficiently (Eiben, Hinterding and Michalewicz, 1999). Although the proposed algorithm is based on the presented one in Limmuun et al. (2013), new modifications was needed to avoid infeasible solutions. In particular, solutions out of the feasible region were penalized during the recombination, whereas suitable replacements were carried out during mutation.

Finally, a new intensification strategy to improve the fitness of designs was applied when the fitness function was based on D-optimality. Due to exact D-optimal designs for Scheffé mixture models are $\{q,m\}$ simplex-lattice designs, $\{q,m\}$ simplex-centroid designs, and replications of points of them, that points can be viewed as consisting of *clusters* of points. It suggests that if the points of the designs are near of this *cluster* points (points in the open balls centred at cluster points with radius tol_{clu}), they will be reached in some iterations so an appropriate strategy consists of moving nearby points to them with certain frequency (n_{it}^{clu} iterations).

The step-by-step implementation of GA is explained as follows:

Genetic algorithm for D- and I-optimal mixture design

Step 0. Input M, N_{elite}, P_{elite}, P_C, P_M, N_{max}, tol, n_{it}^{clu}, tol_{clu}.

Step 1. Initialize $counter = 1$ and select an scenario to generate $\mathcal{P}^{(1)} = \{\xi_1^{(1)}, \xi_2^{(1)}, \ldots, \xi_M^{(1)}\}$:

- Unrestricted mixture experiments: RD, RUD or VD.
- Restricted mixture experiments: RRD, EVD or SEVD.

Step 2. For each $j = 1, \ldots, M$, calculate the fitness

$$fit_j^D = \frac{\Phi_D(\xi_j)}{\sum_{i=1}^{M} \Phi_D(\xi_i)} \text{ or } fit_j^I = \frac{1}{\sqrt{i_j} \cdot \sum_{j=k}^{M} \left(\frac{1}{\sqrt{i_k}}\right)},$$

according to the chosen optimality criterion. The subscripts $i_1, \ldots, i_M$ are refereed to the position of $\xi_1, \ldots, \xi_M$ increasingly sorted according to their criterion function values.

Step 3. Selection:

(i) *Selection with elitism.* Select the $N_{elite} = P_{elite} \cdot M$ designs with the highest fitness values.

(ii) *Probabilistic selection.* Select the i_1^*-th and i_2^*-th parent designs, being

$$i_1^* = min\{i : \sum_{s=1}^{i} fit_s^{\phi} \geq \gamma_1\} \quad \text{and} \quad i_2^* = min\{i : \sum_{s=1}^{i} fit_s^{\phi} \geq \gamma_2\},$$

where $\gamma_1, \gamma_2 \sim U(0,1)$. The superscript ϕ is taken to be ϕ_D or ϕ_I to denote D- or I-optimality respectively

Step 4. Crossover:

(i) *Arithmetic blending*. For each $p_j^{i_1^*} \in \xi_{i_1^*}$, $j = 1,\ldots,n$, generate $\gamma \sim U(0,1)$. If $\gamma < P_C$, then

$$p_{\text{off}_j} = \lambda p_j^{i_1^*} + (1-\lambda) p_j^{i_2^*} \quad \text{and} \quad p_{\text{off}_j} = (1-\lambda) p_j^{i_1^*} + \lambda p_j^{i_2^*}$$

where $\lambda \sim U(0,1)$. Otherwise, remain unchanged. ξ_{off_1} and ξ_{off_2} denote the new created offsprings.

(ii) *Single-crossover point*. Let $p_j = (p_1,\ldots,p_q)$ be the j-th gen of ξ_{off_1} from (i). Thus, p_j can be written as

$$p_j = (0.abc_1^j \mid def_1^j, ..., 0.abc_q^j \mid def_q^j),$$

being abc_k^j and def_k^j are the the decimal figures corresponding to the head and tail respectively. Let us consider third decimal position to divide for illustrating. For each $j = 1,\ldots,n$, if $\gamma < P_C$, then keep $abc_k^j \ \forall k = 1,\ldots,q$ and replace the tails by a random permutation $\sigma(def_1^j,\ldots,def_q^j)$. Otherwise, remain unchanged. Repeat the same operation with ξ_{off_2} genes. If there are constrains over the ingredients, remain unchanged cross points out of the feasible region.

Step 5. Mutation: Let ζ be a randomly selected $U(0,1)$. For each p_j, $j = 1,\ldots,n$, of ξ_{off_1} from (ii), if $\zeta < P_M$, then replace p_j by other randomly selected gen in the feasible region. Otherwise, remain unchanged. Repeat the same operation with ξ_{off_2} genes.

Step 6. Repeat step 3(ii)-5 until having obtained a new generation $\mathsf{P}^{(2)}$ of M new designs.

Step 7. Let ξ_1^{best} and ξ_2^{best} be the designs with highest (lowest) D- (I-)criterion function value in $\mathsf{P}^{(1)}$ and $\mathsf{P}^{(2)}$ respectively. If

$$\frac{\Phi_D(\xi_{best}^{(2)}) - \Phi_D(\xi_{best}^{(1)})}{\Phi_D(\xi_{best}^{(2)})} \leq 10^{-tol} \text{ or } \frac{\Phi_I(\xi_{best}^{(1)}) - \Phi_I(\xi_{best}^{(2)})}{\Phi_I(\xi_{best}^{(2)})} \leq 10^{-tol} \tag{6}$$

is satisfied, where tol is a number specified by the user, then $counter++$. Otherwise, $counter = 1$.

Step 8. If $\phi = \phi_D$ and $counter \equiv 0 \pmod{n_{it}^{clu}}$, then clusterize:

(i) Construct a distance matrix D, where $d_{ij} = \| p_i - p_j \|_2$, $p_j \in \xi_{best}^{(k+1)}$, $j = 1,\ldots,n$, $p_i \in C$ or V, $i = 1,\ldots,\#(C)$ or $\#(V)$, depending on whether it is a unrestricted or restricted mixture problem, respectively.

(ii) Define a new design ξ_{clu} to store the clusterized version of $\xi_{best}^{(k+1)}$ and initialize $\xi_{clu} = \xi_{best}^{(k+1)}$. Let

$$i_j^* = \underset{1 \le i \le \#(C) \text{ or } \#(V)}{arg\,\min} \; d_{ij}$$

be the position of the i_j^*-th point belonging to C or V nearest the j-th point of $\xi_{best}^{(k+1)}$. For each $p_j \in \xi_{clu}$, $j = 1, ...n$, if $d_{i_j^* j} < tol_{clu}$, then $p_j = p_i$.

iii) Replace $\xi_{worst}^{(k+1)}$ by ξ_{clu}, where $\xi_{worst}^{(k+1)}$ is the design with lowest D-criterion function value in $\mathcal{P}^{(k)}$.

(iv) If $\Phi_D(\xi_{clu}) > \Phi_D(\xi_{best}^{(k+1)})$, then $counter = 1$.

Step 9. Stopping rule: If $counter = N_{max}$, then STOP. Else update $\mathsf{P}^{(1)}$ by $\mathsf{P}^{(2)}$ and repeat from step 2.

4. Numerical Examples

Several real problems in the chemical, pharmaceutical and oil industry were used to demonstrate the effectiveness of the proposed algorithms. The selected models were set for three or four-ingredient blends since they were the most commonly used in the literature for data from mixture experiments. For illustrative purposes, D- and I-exact optimal designs were also computed for more ingredients and different numbers of points.

Both algorithms were developed in *R 3.6.0* software (R Core Team, 2018). The tolerance level considered with GA was 10^{-10} whereas it was 10^{-5} with MA since it has a more stringent stopping rule. It was established $N_{max} = 200$ in the stopping rule of the GA and the cumulative distribution function (CDF) of the standard normal distribution or logistic distribution were taken as the $f(x, \delta)$ function with $\delta = 1$ in the MA.

In all examples, we compared results from MA and GA with one of the most popular algorithm in the literature to compute exact designs, the KL-exchange algorithm (KLA), implemented in the R package OptimalDesign. As usual, it is recommended to verify the quality of the designs obtained by other heuristic methods. The application of this method is not direct since it is necessary to provide a set of candidate points. The type of initial mesh strongly affects the finding of the optimal designs. In this work, we propose several procedures for generating sets of candidate points (see supplementary material B) in order to improve its yield. On the other hand, we used the coordinate-exchange algorithm (CEA) of Piepel et al. (2005), which does not require specification of a candidate set. Other comparisons were made with other algorithms such as the cocktail algorithm, but they were not include in this paper for space considerations. A brief discussion of these algorithms will be provided in the last section.

In order to compute I-optimal exact designs with KLA, we found some computational problems considering the IV-optimality criterion provided in OptimalDesign package. Then, the corresponding problem of A-optimality was set such as imple-

menters suggest. The entries of the moment matrix for calculating I-optimal designs, $\mathbf{B}$, were obtained directly from the moments of a Dirichlet distribution (DeGroot, 1970, p. 51) when the experimental region was the $(q-1)$-dimensional simplex and linear models were considered (Goos and Syafitri, 2014). In other cases, that is, when the experimental region was a constrained space or the model was non-linear, the moment matrix was obtained by numerical calculations generating a large candidate set o points uniformly on the region. This was very important because a poor approximation could lead to suboptimal designs (Goos et al., 2016).

4.1. Real applications of the proposed algorithms

4.1.1. Tramadol matrix tablets formulation

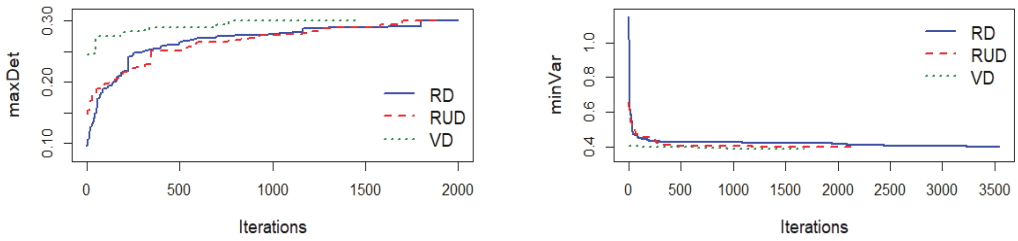
Polynomial models have been widely used in pharmacology, particularly in optimizing drug delivery systems. The following example is motivated by a real problem in which the aim was to determine the release-modifying effect of carboxymethyl xyloglucan for oral drug delivery (Madgulkar et al., 2013). A special cubic polynomial (7) was used to explain the percentage of drug release after a few hours in terms of the drug formulation. The mixture comprised three ingredients: p_1 =carboxymethyl xyloglucan, p_2 =gelling agent (HPMC K100M) and p_3 =dicalcium phosphate (DCP).

$$E[y(p)] = \theta_1 p_1 + \theta_2 p_2 + \theta_3 p_3 + \theta_{12} p_1 p_2 + \theta_{13} p_1 p_3 + \theta_{23} p_2 p_3 + \theta_{123} p_1 p_2 p_3. \quad (7)$$

Various softwares are often employed by practitioners to obtain designs on which must be carried out by experimenters. Classical designs such as simplex-lattice or simplex centroid are the most common choice suggested by these programs. In this simple case, there are analytical results about the D-optimal design. Uranisi (1964) showed that the $\{3,3\}$-simplex centroid was the D-optimal exact design of size 7. Indeed, when the size of the exact design N is proportional to the number of parameters, m, then the D-optimal exact design is the continuous one replicating N/m times each point. In other case, the design points should be as equireplicated as possible regardless which points are replicated most frequently (Goos et al., 2016). Our algorithms produced the same optimal designs for many common models. So, the validation of both techniques is especially interesting for our purposes.

We computed D-optimal designs with $N = 7, 14$ and 18 runs for $q = 3$, and $N = 25$ and 50 for $q = 5$ ingredients. MA was used considering three groups of permutations and GA algorithm was applied under the different scenarios (see supplementary material B). In order to compare our results with the KLA, the exact optimal designs were calculated taking into account that the initial candidate set of points were obtained through the same scenarios than GA. CEA was also run for the same study cases.

Table 1 (supplementary material C) collects the D-efficiencies of the obtained designs with regard to the optimal designs available in the literature or, otherwise, the best design achieved from the algorithms used in this work. They will be named as relative efficiencies. It is noteworthy the robustness of GA under different approaches we considered to construct the initial population in the case of three ingredients. Although this behaviour did not hold for five-ingredient mixtures, the optimum was always achieved under VD scenario (supplementary material B). Regarding the performance of the KLA, it is remarkable to say that this algorithm found difficulties to obtain the optimal design when a random grid of initial points was considered. This situation got worse when a bigger number of ingredients was considered. As it could be expected, the CEA achieved the optimum in all frameworks since it is not based on a set of candidate designs points and it was specially designed to tackle problems with large number of mixture components. On the other hand, when the optimum was a permutation design, such was the case of $N = 18$ runs, the MA quickly achieved the optimum. The convergence speed of the GA is shown in the Figure 1 (a). Despite the fact that the optimum was obtained in all scenarios, the initial population constructed from VD led to the solution faster than the others because it started from designs nearer optimum. Owing to space considerations, these figures are only presented for one case in each example.



(a) (b)

Figure 1: *Values of the D- and I-optimality criteria for GA applied to example 4.1.1 with three ingredients and $N = 18$ runs under different scenarios (RD, RUD and VD), (a) and (b) respectively.*

Regarding I-optimality, exact designs were computed for the second-order Scheffé model in order to compare our results with the presented ones in Goos et al. (2016). I-optimal designs with $N = 6,7,8,18$ and 30 runs for $q = 3$, $N = 15,16$ and 17 for $q = 4$, and $N = 15$ and 30 runs for $q = 5$ ingredients were calculated. From Table 2 (supplementary material C) it is deduced that, differently from D-optimality, there is no a strong dependence of the initial scenario for achieving the optimum regardless the number of ingredients considered in the problem. Designs achieved with GA are highly efficient in all study cases. This behaviour is also observed in the Figure 1 (b) in which it is shown that the optimum is practically obtained in 500 iterations for all scenarios. Optimal designs were obtained in all samples using the CEA. Again, when MA could be used, the optimal design was nimbly achieved. As in Goos et al. (2016) for

$q = 3$-ingredient mixtures, the $\{3,2\}$-simplex-lattice was found for $N = 6$ runs, whereas the $\{3,3\}$-simplex-centroid was obtained for $N = 7$. Nevertheless, some interior points appear in the optimum in many cases for I-optimality (Goos et al., 2016). These points cannot be obtained with the KLA unless they are included in the initial set of points. The higher the number of ingredients is, the lower the probability of being contained in the initial grid is. In spite of providing a thin grid, poorly efficient designs were obtained.

4.1.2. Mixing laws for fluid viscosity

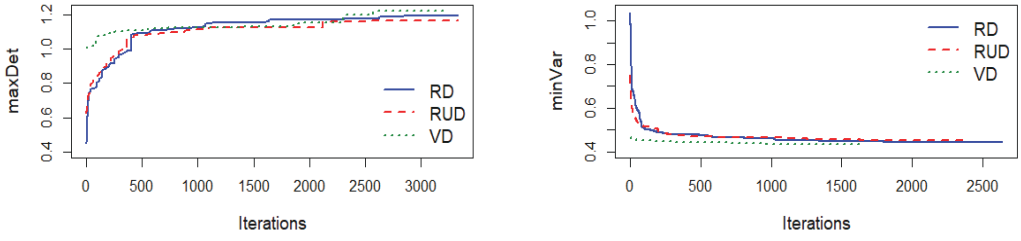
Another usual application of mixture models is found in chemistry and chemical engineering. When the purpose of the study is to analyse the kinematic viscosity of a fluid blend, optimal design tools are used to achieve the best parameter estimation in mixing laws. Most fluid viscosities do not linearly change as formulation varies. Therefore, researchers have developed complex mixture models for their prediction. The selected model in this example is a popular mixing law (8) provided by Grunberg and Nissan (1949). It is a particular case of a wide class of models named *power-mean-mixture models* (Focke, Sandrock and Kok, 2007). They used (8) to explain the viscosity as a function of the three components namely p_1 =acetone, p_2 =methanol and p_3 =water. We assume $\theta_{ij} = \theta_{ji}$ and the nominal values as in Focke et al. (2007),

$$E[y(p,\boldsymbol{\theta})] = \boldsymbol{\eta}(p,\boldsymbol{\theta}) = Exp\left(\sum_{i=1}^{3}\sum_{j=1}^{3} Ln(\theta_{ij})p_i p_j\right). \tag{8}$$

Coetzer and Focke (2010) computed a six-point D-optimal design for this model using a non-linear constrained optimization technique. Variations in the location of the design points caused a significant increase in the criterion function value. The design provided in Coetzer and Focke (2010) is 86.69% efficient relative to the six-point optimal design obtained with the GA and KLA as we can observe from Table 3 (supplementary material D). One set of permutations provided three different support points (permutations of $(1,0,0)$) which was not enough to estimate the model parameters with MA. Therefore, new groups of permutations were considered in the problem, although this involved adding $q!$ new design points for each group. We will compute the optimal designs with $N = 6, 12$, and 18 runs for ternary blends, and $N = 15$ and 30 for five-ingredient samples for both D- and I-optimality criteria.

Similar performances of the algorithms were found to those observed in the previous example (see Table 3, supplementary material D). The I-optimal designs obtained with GA and KLA for five ingredients and $N = 15$ runs are shown in Table 4 (supplementary material D). This table illustrates the GA searchability when optimal design points are located in the interior of the design region. This situation is frequently found when response is not linear in the parameters. Figure 2 shows that the speed of convergence is less dependent on the initial scenario for I-optimality than it is for D-optimality. The CEA cannot be directly applied on non-linear mixture models so that it is not imple-

mented in the most popular commercial softwares. Designs cannot be calculated using this method since a new adaptation is necessary to tackle the non-linearity of the model at the same time that mixture coordinates cannot be independently exchanged without violating the constraint that proportions must sum to one.



(a) (b)

Figure 2: *Values of the D- and I-optimality criteria for GA applied to example 4.1.2 with three ingredients and $N = 18$ runs under different scenarios (RD, RUD and VD), (a) and (b) respectively.*

The following examples are constrained mixture problems. As we mentioned in section 3.1, the extension of the MA proposed in this work does not allow to tackle such kind of problems. New approaches are being investigated for overcoming this situation. Nevertheless, if the initial population of designs is randomly generated over the constrained region, new solutions will remain in this region by construction of the operators proposed in the GA.

4.1.3. Size control of amphiphilic cyclodextrin nanoparticles

Natural or modified cyclodextrins are important excipients used in the pharmaceutical industry to reduce toxicity while improving stability, solubility and bioavailability of hydrophobic drugs (Choisnard et al., 2005). The nanoparticle capacity associated with a drug is expected to be partially influenced by nanoparticle size. This study was focused on controlling the size of amphiphilic β-cyclodextrin (βCDa) nanoparticles using a nano-precipitaciton procedure which strongly depends on solvent formulation. The influence of p_1 =water, p_2 =acetone and p_3 =ethanol proportions involved in this technique was investigated through an experimental design methodology using the full second-degree polynomial to estimate the nanoparticle size. Due to difficulties found in preliminary studies, the experimental region was limited to $0.4 \leq p_1 \leq 0.7$ and $0 \leq p_2, p_3 \leq 0.6$. These limitations are necessary to control the high solubility of βCDa in organic solvent and to avoid the low limit of scattering intensity.

The model chosen in Choisnard et al. (2005) is not an appropriate model for this kind of settings. It is not a canonical polynomial so that the parameters associated with its terms are not unique. Consequently, the design used by the experimenters with that model led to a singular determinant of the information matrix. Thus, a reparametrization

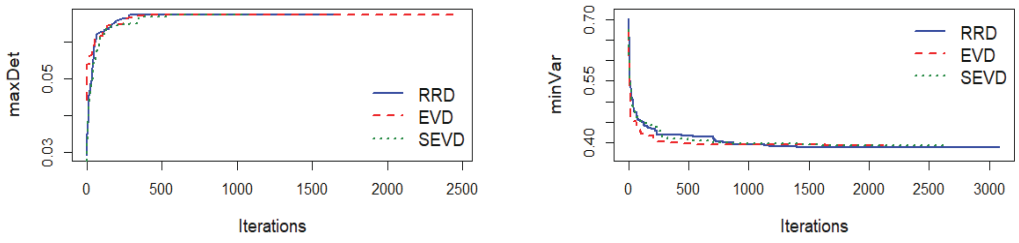
of the full second-degree polynomial was used in this work

$$E[y(p)] = \theta_1 p_1 + \theta_2 p_2 + \theta_3 p_3 + \theta_{12} p_1 p_2 + \theta_{13} p_1 p_3 + \theta_{23} p_2 p_3. \quad (9)$$

It does not only avoid the singularity of the information matrix but also it involves a reduction in the number of model parameters. Thus, fewer runs are needed to estimate the parameters.

Table 5 (supplementary material E) collects the D-efficiencies obtained with GA, KLA and CEA with $N = 6$ and 12 runs in the case of ternary blends, and $N = 15$ runs for five-ingredient mixtures. Both samples in this latter case have different complexity. In the first case $^{(\star)}$, fourth and fifth ingredient can be freely allocated into the simplex, whereas all ingredients are constrained in the second case $^{(\star\star)}$. D-optimal designs achieved with GA and CEA were quite robust, while KLA showed difficulty to find the optimum for RRD and SEVD scenarios for ternary blends and it was unable to achieve them for five ingredients.

Problems of numerical accuracy were found with KLA in the calculus of the I-optimal exact designs despite being recommended in the literature to verify the quality of the designs obtained by other heuristic methods (Harman, Bachrata and Filová, 2016). I-optimal designs cannot be calculated by using this algorithm. Table 6 (supplementary material E) contains the I-efficiencies obtained with GA and CEA in several examples. The CEA applicable for constrained mixture experiments was designed to D-optimally select design points without candidate points (Piepel et al., 2005). In view of the results, this strategy does not seem adequate to achieve I-optimal constrained mixture designs. On the contrary, GA results seem quite robust. It is noteworthy from Figure 3 that both D- and I-optimal design are quickly achieved in a few iterations which reveals the good GA performance in constrained problems.



(a) (b)

Figure 3: *Values of the D- and I-optimality criteria for GA applied to example 4.1.3 with three ingredients and $N = 12$ runs under different scenarios (RRD, EVD and SEVD), (a) and (b) respectively.*

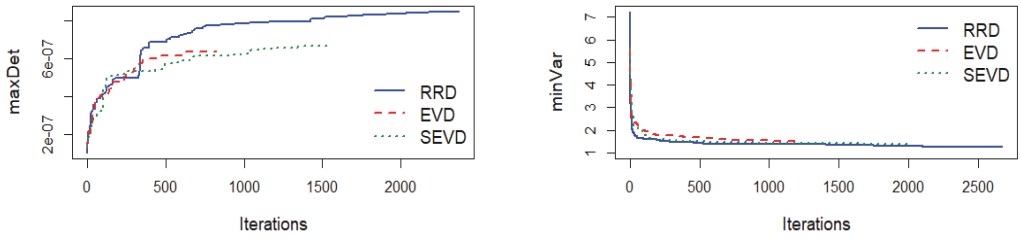
4.1.4. Aqueous phase composition of a microemulsion

Enhanced oil recovery process is obtained determining the optimal formulation of a microemulsion system. Water and oil are not miscible substances at ambient temperatures. The mixture needs to be made under critical conditions due to the existing incompatibility between these fluids. However, a small amount of surfactant, co-surfactant, brine and water may render them compatible to form a structure called microemulsion. This desirable effect is produced due to the properties of these substances. Jerirani et al. (2012) modeled this behaviour using the special cubic polynomial (10) for predicting IFT (interfacial tension) as a measure of energy at the interface of two immiscible fluids. Lower IFT is expected to produce a more effective microemulsion system. Providing a suitable model is essential to finding the formulation which yields its minimum value. Four components are involved in this experiment: p_1 =isopropyl alcohol (IPA), p_2 =sodium chloride (NaCl), p_3 =polysorbate 80 (Tween80) and p_4 =water. A relevant issue arises in the construction of valid formulations under which a microemulsion system is effective. A large amount of water is involved in this process and the rest of the components are practically negligible in spite of their significant positive effect. Particularly, the constraints are $0.01 \leq p_1 \leq 0.04,\ 0 \leq p_2 \leq 0.03,\ 0.002 \leq p_3 \leq 0.02$, and $0.91 \leq p_4 \leq 0.98998$. This fact implies an extreme difficulty in the search for the optimum.

$$E[y(p)] = \theta_1 p_1 + \theta_2 p_2 + \theta_3 p_3 + \theta_4 p_4 + \theta_{12} p_1 p_2 + \theta_{13} p_1 p_3 + \theta_{14} p_1 p_4 + \theta_{23} p_2 p_3 + \\ + \theta_{24} p_2 p_4 + \theta_{34} p_3 p_4 + \theta_{123} p_1 p_2 p_3 + \theta_{124} p_1 p_2 p_4 + \theta_{134} p_1 p_3 p_4 + \theta_{234} p_2 p_3 p_4 \tag{10}$$

Table 7 (supplementary material F) shows the GA power to seek a D-optimum over a severely constrained region, whereas the KLA is even less D-efficient than in the previous case. Unlike 4.1.3 example, CEA was unable to achieve the D-optimum. Figure 4 shows that a random or "semi-random" scenario is preferable to any other for D-optimality. This matter demonstrates that KLA and CEA are inefficient in samples where the optimum is not allocated on extreme-vertex points.

The same drawback than in the previous example was found considering KLA for I-optimality, so that we can only compare with the designs provided by the experimenters and the CEA for this criterion. A 20-point I-optimal design was selected in Jerirani et al. (2012) for IFT modelization. In view of the I-efficiencies shown in Table 8 (supplementary material F), we have that the design obtained by the experimenters is 4.51% efficient in comparison with the design obtained with GA. This fact implies the methodology used by them to carry the optimization out is not adequate. The robustness of the GA for I-optimality can be observed in Figure 4, whereas the CEA inefficiently performs again.



(a) (b)

Figure 4: *Values of the D- and I-optimality criteria for GA applied to example 4.1.4 with four ingredients and $N = 20$ runs under different scenarios (RRD, EVD and SEVD), (a) and (b) respectively.*

5. Discussion

This paper presents two new optimization tools for constructing D- and I-optimal exact designs, when the variables controlled by the experimenter are proportions, and then discusses their properties.

The MA is well known in OED and its convergence has been extensively studied in approximate design theory. However, its application to the solution of exact mixture problems is not straightforward and a new approach based on a class of permutation designs is proposed in this paper. Since symmetry and balancedness have always been a prime attribute of good experimental designs (Draper and Pukelsheim, 1999), and in view of the results obtained, considering PMED seems to be a suitable strategy to generate candidate points for mixture design. The new definition of the multiplicative iteration has a substantial advantage over the other algorithms: first order conditions can be obtained by exploiting the equivalence theorems, whereas stopping rules in the other methods are based on the idea of not finding a better exchange or a better solution. Another advantage the MA offers is that it does not need to anticipate the number of design points, unlike the other methods. The optimal number of permutation groups is automatically determined by the algorithm. However, disadvantages include the fact that it cannot be used when the design space has constraints beyond the natural one and the fact that the sample size has to be a multiple of q!. While this may not be too restrictive in a small q, in other cases it can become a difficulty.

GAs are a class of stochastic optimization methods, easy to implement and computationally powerful. We provided an efficient GA as a heuristic alternative when additional constraints over the experimental region appeared in real problems. One common feature of the GAs is that their computational time is relative. This situation has led to the development of a number of modifications to accelerate their convergence. Most of them focus mainly on the operators. Nevertheless, another interesting but much less studied option relates to initial populations. Several scenarios were proposed in this paper and substantial differences were observed in the speed of convergence rather than

the quality of final solution. A new strategy based on a clustering process around optimal points was also incorporated into the algorithm for this purpose. The number of iterations required to achieve the optimum is much lower than when clustering is not considered. This approach helps operators to explore and quickly reach potential solutions. Moreover, it prevents suboptimal designs from being obtained, in the sense of generating near-optimal points. Many algorithms have the disadvantage of achieving support points close to the vertices, the overall centroid and the centroid of all lower dimensional simplices of a $(q-1)$-dimensional simplex. This intensification strategy gives further guarantees of reaching the optimum. It is also noteworthy that if the optimum does not lie on the extreme vertex points, this new mechanism does not force their inclusion in the optimal design considering all possible cases. On the other hand, changes in the operators were made in order to hold the solutions within the feasible regions.

Genetic algorithms were seen as robust problem solvers that exhibit approximately the same accuracy over the different scenarios considered for constructing the initial populations (supplementary material B) in a wide range of problems. This property is even more evident when I-optimal designs are sought. In this regard, the MA and the CEA do not depend on an initial set of candidate points. However, the strong dependency of KLA on the initial set of points means it is a good choice when the interest is in selecting rather than finding solutions. As may be deduced from the examples, a GA does not offer significant benefits over exchange algorithms when the designs spaces are regular in the case of D-optimality. Unlike point-exchange algorithms, the CEA performs successfully when the optimal design points are located in the interior of the design region (I-optimality) in unrestricted regions. In spite of these advantages, this algorithm cannot be directly applied to non-linear mixture models. Due to the CEA efficiency, it could be interesting to explore a new approach to this algorithm in this kind of situation. On the other hand, when there is no evidence of potential candidate points as, for instance, in severely constrained design regions, the designs generated by exchange algorithms are not frequently optimal under any scenario. On the other hand, the GA and the MA (when possible) converged in all examples and showed excellent searchability.

Other algorithms were also used in this paper apart from KLA and CEA for comparison purposes. In the examples where the cocktail algorithm could be applied, the efficiencies of the designs obtained were the lowest due to rounding effects. Rounding methods take neither the model nor the criteria into account. As a consequence, they are guaranteed to produce efficient results only if the number of trials is high compared to the dimension of the unknown parameter (Harman and Filová, 2014). Results could not be obtained for a predetermined number of runs N since the approximation rule will depend on the weight assigned to each point of the discretized space. These results were omitted due to their poor performance and for considerations of space.

Particular attention is drawn to the successful performance of the proposed algorithms when non-linear mixture models are considered. We can recommend to practi-

tioners more efficient designs than those used in their experiments. They provided better results than general optimization solvers and the algorithms implemented in commercial software.

Finally, although at this stage the use of the multiplicative method seems to be limited, this approach offers the advantages previously noted. In regard to limitations, we are exploring other alternatives as a line of future research. In particular, we are looking at a partition of the simplex into symmetrical regions to simplify the research as the number of proportions increases, and we are working on imposing order constraints on the proportions so the sample size need not be a multiple of q!. In addition, it would be interesting to use the MA proposed here to construct *D*- and *I*-optimal designs for mixture experiments in which linear constraints are imposed on the components. A new adaptation of MA for tackling this kind of practical situation is also being explored. Moreover, we expect that these algorithms can be applied to find optimal designs for a much broader class of optimality criteria. All these studies will be aimed at solving real situations in other fields of study where OED with mixtures plays an essential role.

Acknowledgements

The authors have been sponsored by Ministerio de Economía, Industria y Competitividad, Agencia Estatal de Investigación (MTM2016-80539-C2-1-R), Consejería de Educación, Cultura y Deportes de la Junta de Comunidades de Castilla-La Mancha (SBPLY/17/180501/000380) and Fondo Europeo de Desarrollo Regional (FEDER).

References

Aitchison, J. and Bacon-Shone, J. (1984). Log contrast models for experiments with mixtures. *Biometrika*, 2, 323–330.

Atkinson, A. and Donev, A. (1989). The construction of exact D-optimum experimental designs with application to blocking response suface designs. *Biometrika*, 3, 515–526.

Atkinson, A., Donev, A. and Tobias, R. (2007). *Optimum Experimental Designs with SAS*. Oxford Statistical Science Series. United Kingdom: Oxford University Press.

Becker, N. (1968). Models for the response of a mixture. *Journal of the Royal Statistical Society*, 1, 107–112.

Borkowski, J. (2003). Using genetic algorithm to generate small exact response surface designs. *Journal of Probability and Statistical Science*, 1, 65–88.

Brown, L., Donev, A. and Bissett, A.C. (2015). General blending models for data from mixtures experiments. *Technometrics*, 4, 449–456.

Chan, L. (1992). D-optimal design for a quadratic log contrast model for experiments with mixtures. *Communications in Statistics. Theory and Methods*, 10, 2909–2930.

Chan, L. and Guan, Y. (1998). Design in mixture models with inverse terms for two components. *Private comunication*.

Chan, L. and Guan, Y. (2001). A- and D-optimal designs for a log contrast model for experiments with mixtures. *Journal of Applied Statistics*, 28, 537–546.

Chernoff, H. (1953). Locally optimal designs for estimating parameters. *The Annals of Mathematical Statistics*, 24, 582–602.

Choisnard, L., Géze, A., Bigan, M., Putaux, J. and Wouessidjewe, D. (2005). Efficient size control of amphiphilic cyclodextrin nanoparticles through a statistical mixture design methodology. *Journal of Pharmacetical Sciences*, 3, 593–600.

Coetzer, R. and Focke, W. (2010). Optimal designs for estimating the parameters in weighted power-mean-mixture models. *Chemometrics*, 24, 34–42.

Coetzer, R. and Haines, L. (2017). The construction of D- and I-optimal designs for mixture experiments with linear constraints on the components. *Chemometrics and Intelligent Laboratory Systems*, 171, 112–124.

Cornell, J. (2002). In *Experiments with Mixtures*, New York. Wiley.

Darroch, J. and Waller, J. (1985). Additivity and interaction in three component experiments with mixtures. *Biometrika*, 1, 153–163.

Dean, A., Morris, M., Stufken, J. and Binghman (2015). *Handbook of Design and Analysis of Experiments*.

DeGroot, M. (1970). *Optimal Statistical Decisions*. New York: McGraw Hill.

Diaz-Gomez, P. A. and Hougen, D.F. (2007). Initial population for genetic algorithms: A metric approach. In H.R. Arabnia, J.Y. Yang, and M.Q. Yang (Eds.), *GEM*, pp. 43–49. CSREA Press.

Draper, N. and John, R. (1977). A mixtures models with inverse term. *Technometrics*, 1B, 37–46.

Draper, N. and Pukelsheim, F. (1997). Mixture models based on homogeneous polynomials. *Journal of Statististical Planing and Inference*, 71, 303–311.

Draper, N. and Pukelsheim, F. (1999). Kiefer ordering of simplex designs for first- and second-degree mixture models. *Journal of Statistical Planning and Inference*, 79, 325–348.

Eiben, A., Hinterding, R. and Michalewicz, Z. (1999). Parameter control in evolutionary algorithms. *IEEE Transactions on Evolutionary Computation*, 3, 124–141.

Fedorov, V. (1972). In *Theory of optimal experiments*, New York. Academic Press.

Focke, W., Sandrock, C. and Kok, S. (2007). Weighted-power-mean mixture model: Empirical mixing laws for liquid viscosity. *Industrial and Engineering Chemistry Research*, 46, 4660–4666.

Galil, Z. and Kiefer, J. (1977). Comparison of simplex designs for quadratic mixture models. *Technometrics*, 4, 445–453.

García-Camacha Gutiérrez, I. (2017). *Diseño óptimo de experimentos para modelos de mezclas aplicados en la ingeniería y las ciencias experimentales*. Ph. D. thesis, Departamento de Matemáticas. Área de Estadística e Investigación Operativa. Universidad de Castilla-La Mancha.

Goos, P., Jones, B. and Syafitri, U. (2016). I-optimal design of mixture experiments. *Journal of the American Statistical Association*, 111, 899–911.

Goos, P. and Syafitri, U. (2014). V-optimal mixture designs for the qth degree model. *Chemometrics and Intellingent Laboratory Systems*, 136, 173–178.

Grunberg, L. and Nissan, A. (1949). Mixing law for viscosity. *Nature*, 164, 799.

Harman, R., Bachrata, A. and Filová, L. (2016). Heuristic construction of exact experimental designs under multiple resource constraints. *Applied Stochastic Models in Business and Industry*, 32, 3–17.

Harman, R. and Filová, L. (2014). Computing efficient exact designs of experiments using integer quadratic programming. *Computational Statistics and Data Analysis*, 71, 1159–1167.

Heredia-Langer, A., Carlyle, W., Montgomery, D., Borror, C. and Runger, G. (2003). Genetic algorithms for the construction of D-optimal designs. *Journal of Quality Technology*, 35, 28–46.

Jerirani, Z., Jan, B., Ali, B., Noor, I., Hwa, S. and Saphanuchart, W. (2012). The optimal mixture design of experiments: Alternative method in optimizing the aqueous phase composition of a microemulsion. *Chemometrics*, 112, 1–7.

Kiefer, J. (1961). Optimal design in regression problems. *The Annals of Mathematical Statististics*, 2, 298–325.

Lim, Y. (1990). D-optimal design for cubic polynomial regression on the q-simplex. *Journal of Statistical Planning and Inference*, 25, 141–152.

Limmuun, W., Borkowski, J. and Chomtee, B. (2013). Using a genetic algorithm to generate D-optimal design for mixture experiments. *Quality and Reliability Engineering International*, 29, 1055–1068.

Madgulkar, A., Bhalekar, M., Padalkar, R. and Shaikh, M. (2013). Optimization of carboxymethyl-xyloglucan-based tramadol matrix tablets using simplex centroid mixture design. *Journal of Pharmaceutics*.

Martín-Martín, R. and García-Camacha Gutiérrez, I. (2015). Combined algorithm to compute D-optimal designs. *Journal of Computational and Applied Mathematics*, 278, 248–257.

McLean, R. and Anderson, V. (1966). Extreme vertices design of mixture experiments. *Technometric*, 3, 447–454.

Meyer, R. and Nachtsheim, C. (1995). The coordinate-exchange algorithm for constriction exact optimal experimental designs. *Technometrics*, 37, 60–69.

Mikaeili, F. (1989). D-optimal design for cubic without 3-way effect on the simplex. *Journal of Statistical Planning and Inference*, 21, 107–115.

Mitchell, T. (1974). An algorithm for construction of D-optimal experimental designs. *Technometrics*, 16, 203–210.

Piepel, G., Cooley, S.K. and Jones, B. (2005). Construction of a 21-component layered mixture experiment design using a new mixture coordinate-exchange algorithm. *Quality Engeniering*, 17, 579–594.

Pukelsheim, F. (2006). *Optimal Design of Experiments*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics.

Pukelsheim, F. and Rieder, F. (1992). Efficient rounding of approximate designs. *Biometrika*, 79, 763 – 770.

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Saleh, M. and Pan, R. (2016). A clustering-based coordinate exchange algorithm for generating g-optimal experimental designs. *Journal of Statistical Computation and Simulation*, 86, 1582–1604.

Scheffé, H. (1958). Experiments with mixtures. *Journal of the Royal Statistical Society*, B, 344–369.

Scheffé, H. (1963). The simplex-centroid design for experiments with mixtures. *Journal of the Royal Statistical Society*, 2B, 235–263.

Silvey, S., Titterington, D. and Torsney, B. (1978). An algorithm for optimal designs on a finite design space. *Communications in Statistic*, A, 1379–1389.

Sinha, B., Mandal, N., Manisha, P. and Das, P. (2014). In *Optimal mixture experiments*, Lecture Notes in Statistics. Springer.

Snee, R. (1979). Experimental designs for mixture systems with multicomponent constrains. *Communications in Statictics. Theory and Methods*, 4, 306–326.

Snee, R. and Marquardt, D. (1974). Extreme vertices designs for linear mixture models. *Technometrics*, 16, 399–408.

Torsney, B. (1977). Contribution to discussion of "maximum likelihood estimation via the em algorithm" by dempster et al. *Journal of Royal Statistical Society*, B, 26–27.

Torsney, B. and Martín-Martín, R. (2009). Multiplicative algorithms for computing optimum designs. *Journal of Statistical Planning and Inference*, 139, 3947–3961.

Uranisi, H. (1964). *Optimal Design for the Special Cubic Regression Model on the q-simplex*. Mathematical Report 1, Kyushu University, General Education Department.

Vazquez, A.R., Goos, P. and Schoen, E.D. (2018). Constructing two-level designs by concatenation of strength-3 orthogonal arrays. *Technometrics*.

Welch, W.J. (1982). Three NP-hard problems in computational statistics. *Journal of Computation and Simulation*, 1, 41–48.

Wong, W., Chen, R.-B., Huang, C.-C. and Wang, W. (2015). A modified particle swarm optimization technique for finding optimal designs for mixture models. *PLoS ONE 10, e0124720*, 6.

Wynn, H. (1970). The sequential generation of D-optimum experimental designs. *The Annals of Statistics*, 6, 1273–1285.

Yu, Y. (2010). Monotonic convergence of a general algorithm for computing optimal designs. *The Annals of Statistics*, 38, 1593–1606.